

Characteristic Classification of Walkers via Underfloor Accelerometer Gait Measurements through Machine Learning

Dustin Bales

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Mechanical Engineering

Pablo Tarazaga
Mary Kasarda
Serkan Gugercin

April 20th, 2016
Blacksburg, Virginia

Keywords: Internet of Things, Goodwin Hall, Accelerometer gait measurement, walker classification, machine learning, data reduction for machine learning, rank aggregation

Copyright Dustin Bales 2016

Characteristic Classification of Walkers via Underfloor Accelerometer Gait Measurements through Machine Learning

Dustin Bales

ABSTRACT

The ability to classify occupants in a building has far-reaching applications in security, monitoring human health, and managing energy resources effectively. In this work, gender and weight of walkers are classified via machine learning or pattern recognition techniques. Accelerometers mounted beneath the floor of Virginia Tech's Goodwin Hall measured walkers' gait. These acceleration measurements serve as the inputs to machine learning techniques allowing for classification. For this work, the gait of fifteen individual walkers was recorded via fourteen accelerometers as they, alone, walked down the instrumented hallway, in multiple trials. These machine learning algorithms produce an 88 % accurate model for gender classification. The machine learning algorithms included are Bagged Decision Trees, Boosted Decision Trees, Support Vector Machines (SVMs), and Neural Networks. Data reduction techniques achieve a higher gender classification accuracy of 93 % and classify weight with 64% accuracy. The data reduction techniques are Discrete Empirical Interpolation Method (DEIM), Q-DEIM, and Projection Coefficients. A two-part methodology is proposed to implement the approach completed in this thesis work. The first step validates the algorithm design choices, i.e. using bagged or boosted decision trees for classification. The second step reduces the walking data measured to truncate accelerometers which do not aid in increasing characteristic classification.

Characteristic Classification of Walkers via Underfloor Accelerometer Gait Measurements through Machine Learning

Dustin Bales

GENERAL PUBLIC ABSTRACT

The ability to classify occupants in buildings has far reaching affects in many applications, such as security and retail space design. This work specifically classifies gender and weight of an individual based on their walking signature. Understanding a walker's gender and weight can aid in their re-identification in the event of a crime, for example. This thesis lays the frame work for the classification of walker characteristics using underfloor mounted accelerometers through a machine learning approach.

Accelerometers are mounted throughout Goodwin Hall (a 155,000 sq. ft. five-story building) on the Virginia Tech campus for the purpose of measuring walking signatures. These accelerometers are mounted to structural members of the building and enable measurement of walker's footsteps. These acceleration measurements are used as the inputs to machine learning approaches. Machine learning is a set of algorithms used to comb through data to find patterns which may not be understood without these techniques. This approach allows for prediction of a walker's gender and weight based on previously measured walkers with known gender and weight.

Through preliminary efforts walker gender was classified to an 88 % accuracy. Through data reduction techniques, which eliminate superfluous walking data, the gender classification accuracy was raised to 97 % and weight was classified to 64 %. The underfloor accelerometers are never in the line of walking and do not record details which people may want to keep private. This technology will lead the way for the future of smart buildings to better aid their occupants.

ACKNOWLEDGEMENTS

Thank you to my entire committee Dr. Pablo A. Tarazaga, Dr. Mary Kasarda, and Dr. Serkan Gugercin for their guidance and patience. This project would not be possible without my advisors Dr. Tarazaga and Dr. Kasarda who have asked the most of me and have challenged me to deliver. I have walked away from this experience not only a better engineer but a better person.

I wish to acknowledge the support as well as the collaborative efforts provided by the Virginia Tech Smart Infrastructure Laboratory (VTSIL) sponsors, VTI Instruments, PCB Piezotronics, Inc.; Dytran Instruments, Inc.; and Oregano Systems. Thanks for the support provided by the College of Engineering at Virginia Tech through Dean Richard Benson and Associate Dean Ed Nelson as well as VT Capital Project Manager, Todd Shelton, and VT University Building Official, William Hinson. I would also like to acknowledge Gilbane, Inc. and in particular, David Childress and Eric Hotek. I especially thankful to the Student Engineering Council (SEC) at Virginia Tech and their financial commitment to this project. The work was conducted through the support of the Virginia Tech Smart Infrastructure Laboratory and its members.

Special thanks to collaborators Dr. Bryan Joyce, Sriram Malladi, Mico Woolard, Jeff Poston, Austin Phoenix, Rodrigo Sarlo, and Patrick Musgrave.

Thank you to my family and my wife, Libby Bales, for their help throughout the process of completing my thesis, but more so for the help that I received in getting the opportunity to graduate with my Master's Degree. In the words of Mark Twain, "Don't let school get in the way of your education," which has been particularly true for my time completing my Masters.

Table of Contents

Table of Figures.....	vii
List of Tables.....	viii
Chapter 1. Introduction.....	1
Chapter 1.1. Virginia Tech Smart Infrastructure Laboratory and Goodwin Hall	4
Chapter 1.2. Gait.....	6
Chapter 1.3. Potential Applications of the Technology	7
Chapter 1.4. Gait Measurements Technologies	10
Chapter 1.4.1. Visual Based Methods	11
Chapter 1.4.2. Non-visual Based Methods	12
Chapter 1.5. Most Pertinent Literature for Gait Classification and Data Reduction.....	17
Chapter 1.5.1. Gait Classification	17
Chapter 1.5.1. Gait Measurement Data Reduction.....	18
Chapter 2. Experimental setup	20
Chapter 2.1. Testing Facility	20
Chapter 2.2. Walkers Investigated.....	25
Chapter 2.3. Experimental Procedure	27
Chapter 2.4. Processing Gait Measurement.....	29
Chapter 2.5. Data Representation	32
Chapter 3. Machine Learning.....	35
Chapter 3.1. Machine Learning Algorithms Used	37
Chapter 3.2. Examples of Gait Classification in Literature	44
Chapter 4. Data Reduction Approaches	49
Chapter 4.1. Feature Selection.....	49
Chapter 4.2. Instance Selection	56
Chapter 4.2.1. Sample Rate Selection – Instance Selection Method (SRS-ISM).....	57
Chapter 4.2.2. Accelerometer Selection – Instance Selection Methods (AS-ISM).....	58
Chapter 4.3. Rank Aggregation	66
Chapter 4.3.1. Current Methods	66
Chapter 4.3.1. Singular Value Decomposition – Count Rank Method	67
Chapter 4.4. Examples of Gait Measurement Reduction in Literature	69
Chapter 5. Gender classification of walkers via underfloor accelerometer measurement	72
Chapter 5.1. Results and Discussion.....	75
Chapter 5.2. Conclusion	82
Chapter 6. data reduction via sensor truncation	84
Chapter 6.1. Results and Discussion.....	89
Chapter 6.1.1. Accelerometer Selection – Instance Reduction Methods (AS-ISMs).....	90
Chapter 6.1.2. Sampling Rate Reduction	111

Chapter 6.2. Conclusion	117
Chapter 7. Conclusion	120
Chapter 7.1. Project Summary.....	120
Chapter 7.2. Project Aims.....	121
Chapter 7.3. Contributions.....	127
Chapter 7.4. Future Work	128
References	131
Appendix A: Sample Coding.....	135
Appendix B: formulation of svd-cr rank aggregation method	139
Appendix C: Analysis for statistical significance	142
Appendix D: Full Gender Classification Results Table.....	144
Appendix E: Full Instance Selection Results.....	147
Appendix F: Figure and IRB Figure Permissions.....	155

TABLE OF FIGURES

Figure 1. Goodwin Hall on the Virginia Tech campus.	4
Figure 2. Accelerometer mount.	6
Figure 3. The gait cycle of human movement, showing both the stance and swing phases. [2].....	7
Figure 4. An example sensitive floor from [1].	15
Figure 5. Aerial view of Goodwin Hall.	20
Figure 6. Tri-axis accelerometer mount.	21
Figure 7. Goodwin Hall Representation.	22
Figure 8. Testing Hallway.	23
Figure 9. Accelerometer Locations.	24
Figure 10. Accelerometer Warmup.	28
Figure 11. Accelerometer Detrending.	31
Figure 12. A typical walking signal.	32
Figure 13. Walking trial schematic.	33
Figure 14. Example supervised learning data set.	35
Figure 15. Ten-Fold-Cross-Validation Method.	36
Figure 16. Illustration of a Decision Tree.	38
Figure 17. Neural network illustration.	41
Figure 18. Support Vector Machines illustration.	42
Figure 19. Representation of the peak aligning and averaging.	51
Figure 20. SAS10 compared to the maximum peak footstep.	52
Figure 21. Illustration of subset of sensor arrangement.	53
Figure 22. Effects of Downsampling.	58
Figure 23. Rank independence relationship.	59
Figure 24. Average normalized singular value decay of all walking trial.	61
Figure 25. Typical Projection Coefficients.	64
Figure 26. Chapter 5 Methodology Overview.	73
Figure 27. Representation of the gender classification error.	76
Figure 28. SVM and NN parameter study.	82
Figure 29. The generic methodology for completing ADC and ST.	86
Figure 30. Algorithm Design Confirmation Methodology.	93
Figure 31. Statistical analysis results.	94
Figure 32. General Gender and Weight Results.	95
Figure 33. Six weight bin normalized confusion matrix.	97
Figure 34. ηp error for all possible rankings of five sensors.	99
Figure 35. Sweeping for Truncation Methodology.	105
Figure 36. Sweeping for Truncation Results.	106
Figure 37. Sweeping for Truncation with various rank aggregation methods.	108
Figure 38. Singular Value Decay Comparison.	110
Figure 39. Down Sampling Methodology.	113
Figure 40. Minimum Sampling Rate Results.	114
Figure 41. Entire Methodology Summary.	116
Figure 42. Classification Error as a Function of Sampling Rate.	126
Figure 43. Classification Error for all Data Treatments.	148

LIST OF TABLES

Table 1. Summary of Data Types in the structured database.....	34
Table 2. Gait classification further reading.....	45
Table 3. Illustration of the feature types construction.	54
Table 4. The error metrics comparing the SVD-CR rank aggregation method to other popular techniques.	102
Table 5. The best eight sensors ranked by the various ranking aggregation methods.	102
Table 6. Statistical Testing Summary.	118
Table 7. Statistical Analysis Results for all Data Treatments.	123
Table 8. Rank Aggregation Results.	124
Table 9. Algorithm and Feature Set Error Summary in Time Domain.	145
Table 10. Algorithm and Feature Set Error Summary in Frequency Domain.	146
Table 11. The tabulated form of the classification error for all data treatments.	149

Characteristic Classification of Walkers via Underfloor Accelerometer Gait Measurements through Machine Learning

CHAPTER 1. INTRODUCTION

The instrumentation of Virginia Tech's Goodwin Hall (GH) with underfloor accelerometers allows for the development of smart building technologies. The underfloor accelerometers enable the non-invasive measurement of gait or human walking. This work queries these gait measurements for the purpose of computing characteristics of walkers in a smart building. There are many potential applications associated with measuring a walker's characteristics (e.g., gender, weight, etc.,) including security and energy efficiency.

The overall aim of this thesis work is to examine the effectiveness of characteristic classifications via a machine learning approach. Machine learning is a branch of artificial intelligence, developed for the purpose of studying patterns using input data to tune models for making predictions on new, never before seen data. In this work, the input to these machine learning algorithms is the gait measurements acquired from the underfloor mounted accelerometers.

The experiment completed for this work included fifteen walkers. Each was recorded with the underfloor mounted accelerometers as he or she traversed a 98-foot-long hallway on the fourth floor of GH. The developed database of walking trials allows for the appending of future walking trials for continued research in making buildings smarter.

From the database, specific features, or inputs to machine learning algorithms, are constructed. The features used in the machine learning algorithms greatly affect the classification accuracy, but determining a feature's value a priori is difficult. Thus, multiple features sets are used to understand which is most appropriate for high classification accuracy. Multiple machine learning techniques are also used to ensure the robustness of a specific feature set regardless of the algorithm used.

Combining the developed features with various machine learning algorithms led to a classifier capable of 88 % accuracy in gender classification. The work in Chapter 5 is based on the work in the academic journal [\[3\]](#) and conference paper [\[4\]](#).

The specific questions answered in Chapter 5 are:

- (1) which of the proposed feature sets are most useful for gender classification,
- (2) which of the proposed machine learning algorithms performed best, and
- (3) which domain, time or frequency, contains the most readily useful information for gender classification.

Next, this thesis work details a new methodology to validate algorithm design choices and truncate the accelerometers which do not contribute to successful characteristic classification. This methodology has allowed for an improvement in gender classification to 93 % accurate while reducing the amount of data stored and processed by approximately half. The weight classifier predicts with a 64% accuracy. Chapter 6 is based on the work in an academic journal, Bales et al. [5].

The proposed methodology is split into two steps: Algorithm Design Confirmation (ACD) and Sweeping for Truncation (ST). The ACD step provides a method to prove a parameter in a machine learning analysis is statistically better than another parameter. For example, proving that machine learning algorithm A produces better classification results than algorithm B. The ST step determines and truncates the accelerometers which do not aid in classification.

The specific questions examined in Chapter 6 are:

- (A) what is the most appropriate type of accelerometer selection instance selection method,
- (B) what is the effect of the various data treatment types (e.g. type of data set used, walking direction, machine learning algorithm type, etc.),
- (C) what is the minimum number of accelerometers required for effective classification,
- (D) how effective is the proposed Singular Value Decomposition – Count Rank, rank aggregation method,
- (E) what effect does rank aggregation method have on the Sweep for Truncation step, and
- (F) what relationship exists between sampling rate and classification accuracy.

Contribution

In summary, this thesis work makes contributions in three specific categories:

- Gait Measurement – The underfloor mounted accelerometers potentially offer many advantages not plausible using other gait measurements techniques. The underfloor accelerometers can be developed to see through a

burglar's mask by measuring the walker's characteristics, i.e., gender, weight, height, etc. Instrumenting a building rather than a walker, like some previous works, could ensure that a burglar does not need to be compliant with a worn technology. These accelerometers also potentially allow walkers to retain a higher degree of privacy when compared to other gait measurement technologies like video or audio recordings. The current efforts do not identify individuals thus the labeling of a walking measurement as male allows that walker to retain their anonymity. The gait measurement method is believed to be novel.

- Characteristic Classification Via Gait – Many state-of-the-art classification examples in literature are interested only in the individual identification, but there are many applications where characteristic classification or soft biometrics can be just as important. Attributing characteristics to an individual could aid in their re-identification. Many of the previous works overlook the value in labeling a walker with specific characteristics like weight and gender, discussed in Chapter 1.3. The gender and weight classification using gait measured from underfloor accelerometers are believed novel.
- Data Reduction – The reduction methods (DEIM, Q-DEIM, and Projection Coefficients) can reduce the large data storage and computational time to classify an individual. The size of the reduction is application specific, but this work reduces by nearly 50% of data stored. Previous works that use data reduction methods do not examine the effect of their reduction on classification accuracy, so a direct comparison is difficult. The use of the included data reduction methods as an instance selection method is novel.

The remainder of the Introduction chapter goes into further detail the above overview of the thesis work. First, this introduction discussed Goodwin Hall and the Virginia Tech Smart Infrastructure Laboratory. Second, an overview of gait provides the reader with insight into the complexities of human walking. Multiple potential applications of this technology are then discussed to motivate these advancements in smart building instrumentation. More specifically, potential applications of the characteristic classification are discussed. The last two major sections of the Introduction examine the literature of gait measurement methods and a few of the most closely related works to the classification and data reduction efforts with respective discussion.

Chapter 1.1. Virginia Tech Smart Infrastructure Laboratory and Goodwin Hall

The 5-story, 155,000 square foot GH situated on Virginia Tech's campus is an active building laboratory. Figure 1 shows the main entry of GH. The VTSIL has envisioned GH as a model 'smart building,' allowing for new technological developments. These technologies could aid building occupants in ways once thought impossible. Imagine walking into a building that acknowledges you by name at the front door and calls for an elevator. The building turns on the lights to your office after unlocking the door for you. The temperature is appropriate, but the air conditioning has not been running all night to conserve energy. You have an email from the building's security system reporting that an unknown person was apprehended by authorities last night for suspected burglary. The individual made repeated office visits at night over the last two months and entered an office that was not his. This futuristic situation is being made closer to a realization through the work in VTSIL and this thesis. VTSIL hopes to inform the design of all future buildings aimed at being 'smart.'

The instrumentation of an operational building allows GH to be an ideal real-world test bed. The measurements recorded in GH will contain difficulties not present in a controlled laboratory setup. The models developed using a controlled laboratory setup may be oversimplified and may not generalize into real buildings easily. The noise of an active building facilitates the development of models that are robust to the complexities of an actual building.



Figure 1. Goodwin Hall on the Virginia Tech campus.

GH houses a large number of rooms that serve many functions such as classrooms, laboratories, auditoriums, and staff offices. The varied nature of these spaces potentially allows for the development of models which are robust to building type. The models developed in GH aim to be robust allowing for easy adaptation into new smart buildings.

VTSIL has instrumented GH with a distributed data acquisition (DAQ) system. This distributed DAQ system can be expanded to record many types of sensors with relatively little installation difficulty. This ease of installation facilitates further instrumentation of various sensors capable of querying various measurements such as air flow, temperature, floor dynamics, magnetic fields, etc. These various measurements allow for exploration of models that may contribute to a better human-building interaction.

Other instrumented buildings aided in the design of the data acquisition capability in GH. The most pertinent buildings are the Millikan Library at Caltech and the Factor Building at the University of California, Los Angeles. The Millikan Library was one of the first buildings instrumented with accelerometers for the purpose of examining building dynamics in the 1960's. Thirty-six accelerometers have measured every major earthquake in California since 1971. The Factor Building, 198,000 square foot building, was instrumented with 72 accelerometers for similar purposes to the Millikan Library. Hamilton [\[6\]](#) discusses the instrumentation of these buildings as well as others. These buildings informed the instrumentation of GH.

GH is, at the time of this writing, instrumented with 212 accelerometers making it the most instrumented public building for vibrations in the world. With 140 mounting locations further instrumentation is possible. Where each mount potentially houses up to three accelerometers. More accelerometers are in various stages of installation and commissioning.

The accelerometers instrumented in GH mount directly onto structural members of the building via steel mounting beams. Figure 2 shows an example mounting of an accelerometer. The underfloor mounting of the accelerometers ensures non-invasive gait measurements. The gait measurements remain non-invasive as walkers never need to alter their gait to avoid an accelerometer. On-top-of-floor mounted accelerometers would require gait altering to avoid these tripping hazards. These sensors preserve an individual's privacy because walkers retain their anonymity. More intrusive methods like video or audio could record sensitive information such as a person's direct activities or their conversations. Jain et al. [7] show that gait is a highly accepted measurement by society. The underfloor accelerometers are used to measure gait of walkers, and the subsequent section discusses the complexities of gait.



Figure 2. Accelerometer mount.
Vertical, single-axis sensor shown.
The mount is welded to a structural member with the fire retardant stripped to expose steel.

Chapter 1.2. Gait

Gait is the manner in which humans walk. It consists of the interaction between hundreds of muscles and joints in the body [8]. A gait cycle is the reoccurring, nearly identical events in the walking process [9]. Two phases comprise the gait cycle: stance and swing. Approximately 60% of the gait cycle occurs during the stance phase and the remaining 40% during the swing phase [10]. There are also two periods in a gait cycle which are double supported: where both feet are in contact with the ground. Footsteps occur during the stance phase beginning with a heel strike and ending with toe off [2]. The swing phase occurs from toe off until the subsequent heel strike. Figure 3 visualizes the phases of walking. The three functional tasks associated with walking are weight acceptance, single limb support, and limb advancement. Walkers generate forces via these functional tasks [2]. While neither the stance/swing phase nor the functional tasks are explicitly examined using the machine learning algorithms, the background on gait allow for an

appreciation of the complexity of this daily task. These complexities offer some insight as to why differentiation of walkers and their characteristics is difficult.

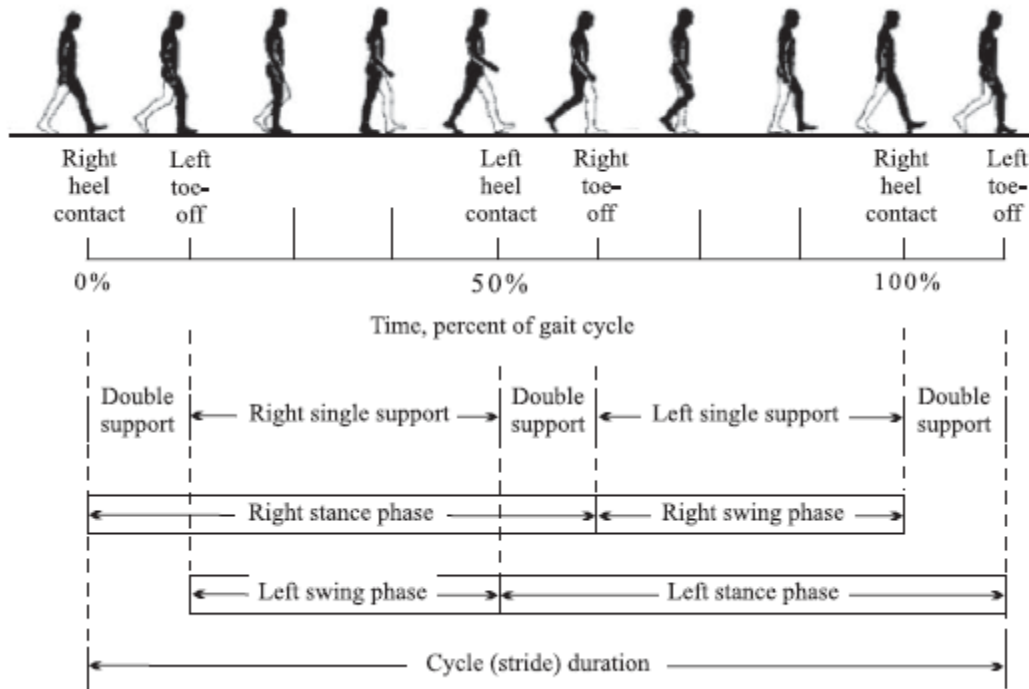


Figure 3. The gait cycle of human movement, showing both the stance and swing phases. [2]

The cyclic patterns in gait are similar between humans, but not identical [8]. The differences in gait between walkers lead to the assumption that gait analysis can decide characteristics of a walker. Nixon and Carter [11] support classifying these differences suggesting that gender classification is achievable by observation of gait with cameras. Physiological differences between the genders suggested a larger hip swing for women and shoulder swing for men, allowing for classification. Other physical differences might allow for the classification of other walker characteristics.

There are multiple potential applications where gait classification via underfloor accelerometers could overcome some of the difficulties associated with other gait measurement technologies.

Chapter 1.3. Potential Applications of the Technology

Occupant characteristic classification has the potential to be of prime importance in numerous applications such as security, improving sales in a retail space, improving building occupant health monitoring, and energy resource management. VTSIL continuously develops further applications, so this list is not comprehensive. The above four

example applications are described to illustrate the necessity of characteristic classification via underfloor mounted accelerometers.

Security

The ability to identify characteristics of a walker plays an important role in re-identifying a suspect in the event of a crime. Providing authorities with the weight, gender, height, shoe size, etc., of a suspect can lead to quicker re-identification and apprehension. Producing these characteristics, known as soft biometrics, has been extensively studied for their ability to aid in the identification of a person. A biometric is a unique attribute of a person, for example, their fingerprint. Biometrics allow for identification of an individual, and soft biometrics are used to aid in that identification [12]. Soft biometrics contribute in two major ways: increasing the speed of existing identification techniques by removing less likely individuals and improving the accuracy of recognition. For example, the time to comb through an entire fingerprint database could be reduced by knowing the fingerprints belong to a male.

Video surveillance systems have traditionally been used in security settings to identify individuals, but these face limitations. Video falls short in two ways (1) potentially violating an individual's privacy and (2) not being able to identify an individual who has altered their appearance. This thesis work shows that underfloor accelerometers allow for extraction of walker characteristics while maintaining their privacy because gait is measured not their appearance. These accelerometers could even monitor sensitive rooms of a building (i.e. restrooms) for safety while maintaining privacy. The characteristics classified by the underfloor mounted accelerometers can aid in security settings while overcoming limitations associated with current solutions.

Improving Sales in a Retail Setting

Providing retailers with information about their customers can allow them to match more effectively people with products. Efficiently laying out a retail space can generate an increase in sales [13]. Supplying a retailer with a walker's characteristics, without compromising that customer's privacy, can better inform store layouts. Information such as how women move through a department store may inform store layout design in efforts to increase sales. Retail was a \$3.19 trillion a year industry in 2014 and increasing sales through this type of technology would be pursued [14].

There are already companies that specialize in video surveillance to examine who enters a store and how they move through the store. As previously claimed, these video methods can violate the privacy of a shopper. Moreover, there can also be issues with spatial aliasing without surplus equipment as unknowingly exemplified in Clifford and Gavett [15, 16]. Underfloor mounted accelerometers can maintain privacy while ‘seeing through’ floor displays which block proper video measurement. The characteristics classified by the underfloor mounted accelerometers could inform a store layout design while overcoming the limitations of the current technologies.

Occupant Health Monitoring

Monitoring the health of individuals who occupy a building is of great importance to society. Classifying characteristics of an individual via gait measurement can be useful in specific diagnosis, for example, ALS and MLS [17]. Diagnosing disease via these characteristic classifiers can aid in occupants seeking medical treatment as early as possible.

The ability to classify characteristics leads to the belief that event classification is possible. Visual identification of doors closing in GH measurements validates event classification on a superficial level with further research ongoing in the VTSIL. Classifying events in conjunction with characteristics can lead to better help for people in need of assistance.

Falls are a specific event worth classifying. The United States spends an estimated \$34,000,000,000 on medical treatment due to falls annually [18]. Being able to classify a fall and seek help is paramount for safety if a faller is unable to seek help themselves. Recording and documenting fall events are of prime importance in the prediction of the next fall [19]. Providing physicians with fall information can allow for more effective treatment of a patient [20]. The underfloor mounted accelerometers allow for privacy while keeping a watchful eye, even when a person might not be able to do so for themselves. The current methods of monitoring falls, either violate older adults’ privacy or require self-reporting which would not be the case with the underfloor mounted accelerometers.

Energy Resource Management

Underfloor mounted accelerometers can improve the energy resource management of a building. Effectively counting occupants can aid in managing building heating and cooling systems to operate only when necessary. In 2014, buildings consumed approximately 41% of the total energy in the United States [21]. The resources required to light, heat, and condition account for a sizeable percentage of our total energy consumption. Buildings can have a positive impact on the environment ensuring lights are only on, and the air only conditioned when a room is occupied. The Environmental Protection Agency (EPA) estimates that smart buildings will reduce the amount of energy used in those buildings by approximately 20% [22].

Current methods of energy resource management are lacking. Automatic lighting is ineffective often leaving individuals in the dark and air, conditioned regardless of who is in the space. Underfloor accelerometers provide the ability to know a person is in a room while not requiring movement to reaffirm their presence such as many lighting sensors. Other occupancy detectors are often video-based which yet again may violate the privacy of occupants.

The ability to classify characteristics of individuals has far-reaching effects for those occupants, their communities, and society. Throughout the above applications, walkers have been described in many ways from shopper to the patient because people assume all of these roles. Regardless of their title there remains the constant need for buildings to monitor occupants for their benefit. The technology and methods presented in this work can provide this monitoring while overcoming the limitations of current methods. The next section discusses the literature of gait measurement, specifically organized by the technology used for the measurement.

Chapter 1.4. Gait Measurements Technologies

There are various methods of gait measurement, and many have been used to classify a variety of walker characteristics. The focus of this section is on the actual techniques. The references given below do not encapsulate all gait measurement studies as they are many, but cover the range of techniques to measure gait. One approach to organizing existing gait measurement methods is to label an approach as either visual or non-visual based.

Chapter 1.4.1. Visual Based Methods

Visual based methods involve video recording and associated analysis to measure the movement of a walker. Visual based approaches are a popular measurement method with large amounts of ongoing research. Since these methods are not similar to the technology used in this work they are only introduced, and the size of writing contained here is not proportional to a large amount of ongoing visual based gait measurement research. Active and passive methods comprise visual-based methods.

The active method, also known as structure from motion, tracks specific points on the human body to characterize its motion [23]. The underlying assumption of structure from motion is that recognition of a person's gait is made possible by measuring a finite number of characteristics. Recording these characteristics, therefore, allows for their classification [24]. Passive methods are not interested in the underlying physical structure of a body's motion. These methods treat each frame of the video as a unique pose made by a walker [25]. Querying these unique poses allows for the extraction of gait features such as body measurements or velocities.

While visual methods are popular for gait measurement, there are limitations associated. When using video to record people, there is always a potential violation of privacy. Monitoring sensitive areas (i.e. restrooms) exacerbates this potential violation. Underfloor accelerometers currently do not record specific activities mitigating these privacy concerns. Other limitations of the video processing often include assumptions of constant lighting, static backgrounds, subjects remain in the workspace at all times, special color clothes, and tight-fitting clothing, among others [26]. These limitations mean that visual-based methods require very specific settings to function correctly. These limitations could be used to beat a visual based gait measurement system in a security application.

One limitation of the underfloor mounted accelerometers overcome by the use of video is the intuitive understanding of the measurement. Humans are capable of identifying gait quickly from the video but understanding the same phenomena from accelerometer measurement is more difficult. A challenge for underfloor mounted accelerometers moving forward will be to enable visualization of these measurements for easy understanding.

Chapter 1.4.2. Non-visual Based Methods

Non-visual based methods broadly measure either the walker's or the floor's reaction due to gait. These methods can provide information that is not necessarily available in the visual-based methods such as the force exerted by the walker. This additional gait information makes non-visual methods a popular measurement technique. Non-visual methods of gait measurement are achieved by force plates, wearables, microphones, smart floors, and floor measurement of acceleration and velocity.

Force Plate

Force plate systems use load cells to measure the loading of a plate as walkers step directly on them. The load cell measures the force of the walker inputs to the plate. The force plate is the gold standard for measuring ground reaction forces according to Sinclair et al. [27]. Addlesee et al. [28] is one example that uses force plates mounted beneath the walking surface to measure gait. The limitations of this measurement method are the relatively small recording area and installation of sensors in an existing building. The number of sensors required to measure a large walking area is often large reducing the active area of this measurement method because of cost. Addlesee et al. themselves raise concerns about the price required to instrument the floor fully. This method also requires the installation of sensors below a floor which might be extremely costly to install in an existing building. Underfloor mounted accelerometers allow for gait measurement across a large distance, even with a single sensor. The accelerometers are also relatively easy to install assuming that the ceiling below the walking surface is accessible.

The force plate does allow for the direct measurement of force which, not the case with accelerometers. In applications where researchers desire the force of a walker, the underfloor accelerometers might not be the most appropriate method. Further study is ongoing by VTISL comparing the force plate and underfloor mounted accelerometers.

Wearables

There are a variety of wearable sensors including force plate shoes and worn accelerometer-based measurement systems. A wearable system is one that is contained by the walker, i.e., the walker carries all of the system with them. These wearable sensors are becoming more ubiquitous with the rise of step counters and the like.

Liu et al. [29] discusses the development of a mobile force plate housed in a walker's shoes for the purpose of unobstructed assessment of gait. They compare the resulting mobile force plate to stationary force plates and the XSSENS motion tracking system. Their results show the mobile force plate produces comparable results to the other measurement methods.

The accelerometer-based measurement methods function in a variety of ways such as commercially available step counters, accelerometer instrumented vests, walker mounted accelerometers, and smartphones. Godfrey et al. [30] offer a survey of many examples measuring gait with accelerometers, including many commercially available products: the RT3 Tri-axial Research Tracker Kit and activPAL™ to name a few. Orendurff et al. [31] study human walking patterns during regular activity over a period of two weeks using StepWatch™ Activity Monitors. Nyan et al. [32] studies gait recorded by an accelerometer instrumented vest worn by walkers. The accelerometer instrumented vest has sensors arranged in anteroposterior and vertical directions mounted on the shoulder of the vest. Sinclair et al. [27] mounted an accelerometer to the distal tibia of the human leg to measure axial accelerations from gait. Non-stretch tape was used to mount the accelerometer to the tibia to measure the longitudinal acceleration. They find a strong relationship between the shank-mounted accelerometer measurements and the force plate to which they correlate. The last discussed wearable system are smartphones. Hoang and Choi [33] recorded gait via an accelerometer in a smartphone. The ubiquitous technology of smartphones means that widespread gait measurement is a possibility, but there are limitations with all wearable technologies in certain applications.

While wearable gait measurement systems allow for the recording of gait, there are applications where this technology is not plausible. In a security situation assuming that a criminal would choose to wear technology that would identify them in a crime is unrealistic. In a less malicious example, if a person forgets to wear their device they will also not be identifiable by their building's system which could effectively lock them out. The underfloor mounted accelerometers overcome walker non-compliance to ensure gait measurement.

One limitation of the underfloor mounted accelerometers overcome by a wearable is the installation of the technology. As mentioned above the wearable technology is becoming ubiquitous meaning that instrumentation of a building is not necessary.

Microphones

Another method of measuring gait is through recording sound. In one example, Sabatier and Ekimov [34] complete an active and passive method of measuring gait using sound waves. The active method uses continuous ultrasonic waves directed at a walker; then use the Doppler effect to analyze the reflected waves for gait measurement. The passive method uses a microphone to measure the audible friction interaction between a walker's shoe and the floor. In a separate work, Ekimov and Sabatier [35] ask walkers to traverse various surface types to develop an understanding of the frequency content of a footstep via microphone recording.

While microphones have been shown as potential gait measurement systems, they share a pitfall with video recording. They potentially violate an individual's privacy. Recording audio for gait measurements will inevitably record conversations meant to be private. The underfloor accelerometers, again potentially allow for maintained privacy.

Also similar to video, audio recordings of gait are intuitive to interpret. Underfloor mounted accelerometer measurements are difficult to understand, requiring software to effectively interpret and analyze the data.

Sensitive Floors

Various studies have shown that sensitive floors are a viable way of measuring gait. Various studies have been completed using capacitors to measure occupancy of a specific area enabling coarse gait measurement [1, 36, 37]. Yun [38] describes the UbiFloorII, which measures using photo-interrupters. A tile instrumented with many photo-interrupters measure changes in light due to the occupancy of portions of the tile. Qian et al. [39] also discuss a pressure sensitive floor for the purpose of recognizing the gait of a specific individual. They use a commercially available pressure sensing mat, the Tekscan 5315, to record the gait of walkers. Figure 8 shows the pressure sensitive floor used Middleton et al. [1]. They use capacitors to measure cell occupancy.

While these sensitive floors allow for underfloor mounting of sensors similar to the underfloor mounted accelerometers, they still face limitations. Many of these methods require re-installation of a floor which may not be feasible for existing buildings. Sensitive floors also require a relatively large number of sensors to have an adequate spatial resolution to determine if a footstep took place. For instance, the system in Figure 4 houses 1536 individual sensors. Underfloor accelerometers provide relatively easy installation if the ceiling below the walking surface is accessible. These accelerometers enable gait measurement at relatively large distances (~ 70 feet) as shown in this thesis work. This result effectively reduces the number of sensors required for classification.

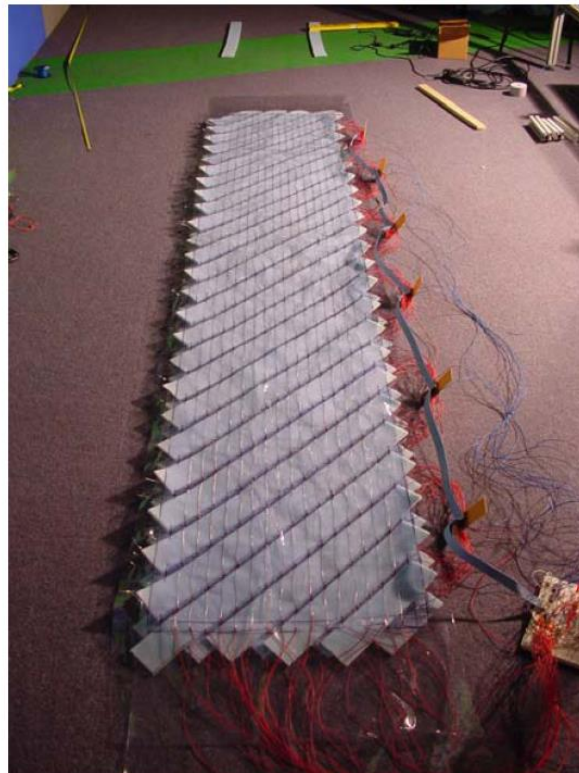


Figure 4. An example sensitive floor from [1]. This type requires large amounts of equipment. Reproduced with permission.

The sensitive floors overcome a spatial resolution issue possible with underfloor accelerometers. Using a higher sensor density like those in most sensitive floors allows for accurate footstep location. Ongoing work in the VTSIL is investigating the spatial accuracy of footstep localization.

Ground Measurement

The gait measurement methods most closely related to the technology in this thesis work measure the floor dynamics caused by a walker. Pan et al. [40] use geophones to measure the vertical velocity of the floor caused by a walker.

They discuss the ability to use these sensors in a sparsely instrumented manner overcoming some of the issues associated with the previously discussed gait measurement types, specifically sensitive floors. Sabatier and Ekimov [34] stake geophones into the ground outdoors to measure gait. In another work, Ekimov and Sabatier [41] use stake mounted accelerometers to measure the response of footsteps.

These gait measurement systems measure very similar signals when compared to the underfloor mounted accelerometers, but they face different limitations because of their mounting location. Mounting these sensors on the walking surface causes them to become a tripping hazard. Filling a hallway with obstacles would make walking unnatural. In a security setting, these sensors could be avoided or destroyed. Underfloor mounted accelerometers have locations beneath the walking surface so walkers can move naturally, and would be difficult to destroy in a security setting.

These on top of floor sensors do overcome a challenge faced by the underfloor mounted accelerometers: reconfiguration. Due to the welding of the accelerometer mounts to the structural members in GH, sensors cannot be easily moved. This lack of reconfiguration means that system design (i.e. what is the best way to arrange sensors) is difficult in GH.

In summary, there are many methods currently used to measure gait, but all of them face limitations overcome by the underfloor accelerometers discussed in this work. Issues ranging from privacy concerns to retrofitting sensors in an existing building are all potentially solved by underfloor mounted accelerometers. While the underfloor mounted accelerometers face limitations themselves, the challenges they overcome outweigh their problems. Using underfloor accelerometers in this gait measurement manner is believed to be novel. There are many instances of classifications based on gait measurements.

The next section provides the most pertinent literature review for the gait classification and data reduction. Chapter 3.2 and Chapter 4.4 detail further examples for gait classification and data reduction respectively. This further literature review is not included until the formal introduction of machine learning, and data reduction as the specifics of each do not help better place this work in context but use technical terms not yet discussed.

Chapter 1.5. Most Pertinent Literature for Gait Classification and Data Reduction

A major contribution of this thesis work is the gait classification of walker's characteristics and the data reduction methodologies that improve the classification rates through the removal of 'excess' measurements. This section discusses the most pertinent examples of gait classification and gait data reduction to place this work in context without diving too deep into the details of the techniques. The examples from the literature will be discussed in further detail in Chapter 3.2, but are presented here before the formal introduction of machine learning and data reduction.

Chapter 1.5.1. Gait Classification

The walker characteristics classified in this thesis work are gender and weight. As stated above, Chapter 5 achieves a gender classification accuracy of 88 %, which rises to 97 % through the data reduction efforts in Chapter 6. Chapter 6 establishes a weight classifier of 64 % accuracy for 20 lbs. weight bins. Below discusses the most closely related works in gender classification and weight classification.

Gender Classification

The most similar efforts in gender classification as completed in this thesis work are by DeLoney [42] and Li et al. [43]. Both works use acoustic gait measurements of walkers to classify gender to accuracies not report (NR) and 75%, respectively. Li et al. [43], which reports a 75% accuracy in gender classification, uses a principal component analysis to transform the acoustic signal. Then a discriminant analysis to classify gender.

This thesis work uses machine learning techniques to interpret the floor acceleration measured by the underfloor mounted accelerometers to classify the gender of a walker. As further discussed Chapter 6, this work is capable of gender classification accuracies of 97 %. This classification accuracy is the largest in known literature.

Weight Classification

Current state-of-the-art weight classification techniques often use video recordings of individuals to extract anthropomorphic measurements. Other methods directly measure anthropomorphic features to classify weight. This section presents literature that classifiers based on both non-gait measurements and gait measurements. First, non-gait measurements are presented.

Velardo and Dugelay [12] provided a weight estimation classifier after examining a large database of pictures taken of study participants. A thresholding algorithm classifies weight using various measurements of the participants' bodies. They claim weight estimation to an accuracy of $\pm 5\%$. Verlaro et al. [44] provided an interesting application of weight estimation in that of weighing astronauts while in space. They developed a technique using an RGBD camera (i.e. Kinect) to measure study participants. From the extracted measurements they can classify weight to within ± 2.7 kg by using a thresholding algorithm. Cao et al. [45] used the CAESAR ID database containing anthropometry measurements of 1000 individuals to estimate body weight. They use support vector machine (a later discussed form of machine learning) to achieve 0.0108 mean absolute error in weight classification. The above examples of weight classification do not classify using gait measurements but help to form the picture of how well current methods of weight classification perform. The remaining weight classifiers use gait measurements as the input to determining weight.

Krishan [46] examines the relationship between measurements of a footprint and body weight. The study included 50 individuals. Each participant had their footprint measured before and after holding weights. Their results showed that when a subject held 20 kg weights, their footprint length and breadth tended to increase. These differences allow for estimation of body weight. Labati et al. [47] used two cameras to record a front and side view of walkers. They used computer vision techniques to extract features of the walkers. The extracted features are input to a neural network (a later discussed form of machine learning) algorithm to estimate the weight of 20 study participants. Their mean error was 0.07 kg (standard deviation of 2.30 kg). These studies demonstrate that other gait measurement techniques have the potential to complete weight estimation.

The best weight classifier produced in this thesis work is 64 % accurate when the weight bins were 20 lbs. large. Where a single bin is anyone weighing between 90-110 lbs. for example. The weight bin classifier makes a direct comparison to the previously discussed weight classifiers difficult. The contribution of weight classification is through the novel method of gait measurement.

Chapter 1.5.1. Gait Measurement Data Reduction

There are several examples of managing large amounts of gait data through various manipulation methods. The literature sources provided below give readers a source for further information. While the details each type of data

manipulation method discussed below are not introduced until Chapter 4. At this point, it is important for the reader to understand that these manipulation methods can be used to reduce large datasets. The importance of data size reduction will be investigated further in this work, but overall, reducing the data size reduces both the time to make predictions of characteristics and increases the accuracy of the classifying models.

Chau [48, 49] provides a survey paper of the many methods most often used in transforming gait measurements. These two sources highlight several different methodologies such as the use of fuzzy logic clustering, multivariate statistics, fractal dynamics, neural networks, and wavelet analysis. Further reading on topics in Chau but not included in the survey paper are provided here: principal component analysis [24, 38, 43], canonical analysis [50], neural networks [51], and wavelet analysis [32]. Many of the methods overviewed in the above survey papers [48, 49] are introduced to familiarize the reader with gait manipulation techniques having the potential to be data reduction methods as well. Neural networks are discussed in-depth in Chapter 3.1.

CHAPTER 2. EXPERIMENTAL SETUP

The experimental setup chapter details the development of the gait measurement database. The gait database serves as the input to the machine learning algorithms. The machine learning algorithms search this data to find patterns and make predictions (discussed further in Chapter 3). This section provides more detail on the instrumentation of the testing laboratory: Goodwin Hall. This section discusses sensor locations and the accelerometers themselves. The next section describes the walkers included in the investigation. Description of the experiment itself and the procedures follows. These steps conclude the collection of the gait measurements. Then the recorded data must be processed and effectively stored; thus, the last two sections detail the processing and storage of the data.

Chapter 2.1. Testing Facility

Virginia Tech's Goodwin Hall instrumentation with underfloor mounted accelerometers enables a novel method of gait measurement. The 155,000 square foot building facilitates the collection of active building measurements. Figure 5 shows Goodwin Hall (GH) for reference. These real world measurements allow for the development of characteristic classifiers which non-invasively records the ground acceleration due to footsteps.



Figure 5. Aerial view of Goodwin Hall. Camera facing West.

The accelerometers mount to steel mounting pegs, and these pegs are welded directly to the structural steel members of the building. Figure 6 shows a tri-axis accelerometer mount which is capable of measuring the three principal directions. The accelerometers are located throughout GH for the purpose of recording many types of measurements. This analysis uses only the vertical measuring accelerometers. Figure 7 depicts GH for reference of the building layout.

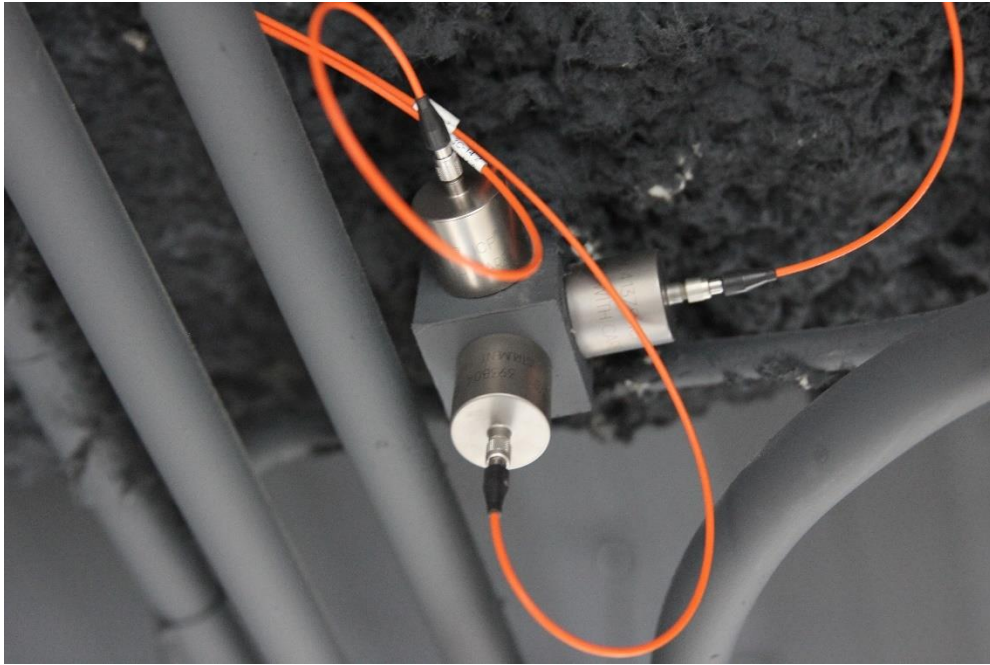


Figure 6. Tri-axis accelerometer mount. The steel mounting peg is welded to the structural steel beam (shown with fire proofing). The accelerometers are threaded onto the mounting peg.

The accelerometers enable gait measurement. Machine learning approaches use these measurements as input. This thesis uses machine learning approaches to find patterns in walking data. Understanding these patterns allows for the prediction of walker characteristics. Specifically, the characteristic classifiers of walker gender and weight are investigated in this thesis using underfloor mounted accelerometer for gait measurement.

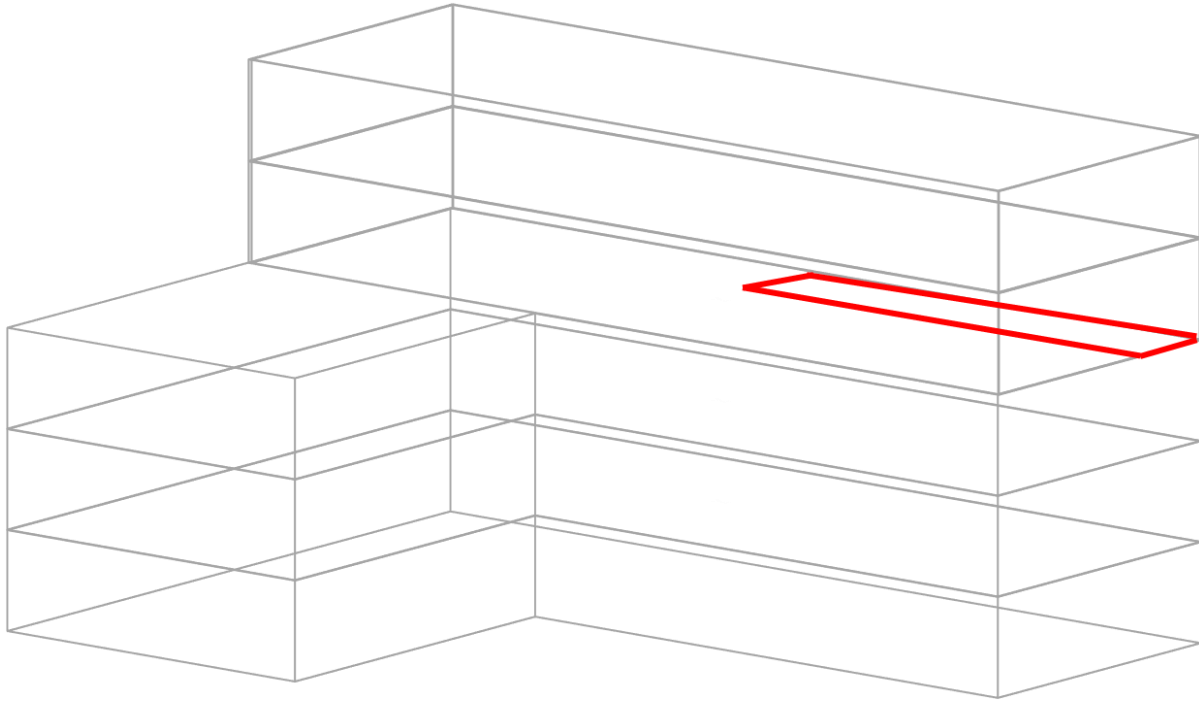


Figure 7. Goodwin Hall Representation. Schematic of floor layout of the entire five-story building. The red highlighted area is the hallway of the building where the testing was completed. Reproduced with Dr. Bryan Joyce's permission.

A single hallway in GH was used to collect all of the walking data used in this thesis. The hallway is 98-foot-long and located on the fourth floor, North side of the building (highlighted in Figure 7). The hallway is one of the most densely instrumented areas in the building facilitating rich gait measurement. Figure 8 shows the building plans for the entire fourth floor. Each of the circled numbers in the figure represents qualitatively the location of a sensor mount and the exact number of accelerometers mounted at that point. Also, highlighted in the figure is the specific testing hallway with a picture of the hallway for scale.

Fourteen accelerometers were used to record gait from in and adjacent to the testing hallway. Researchers verified the accelerometer locations before testing, but locations were not reproducible after the experiment due to a simultaneous software and hardware transition in the VTSIL. After the transition, accelerometers no longer corresponded to the previously recorded locations. Reproducing the transition is implausible meaning that the original location of the accelerometers may always remain unconfirmed. The current location of all sensors is known.

Not knowing accelerometer locations does slightly affect the creation of the walking database. Each sensor has a specific sensitivity factor used to convert from the measured voltage to the corresponding acceleration. This difficulty is mitigated because the accelerometers have a low variance in the factor used to convert to acceleration. The accelerometers used in this study have a nominal factor of $1000 \text{ mV/g} \pm 10\%$ [52]. The small variance in the nominal sensitivity factor reduces the error associated with attributing the incorrect factor to an accelerometer.

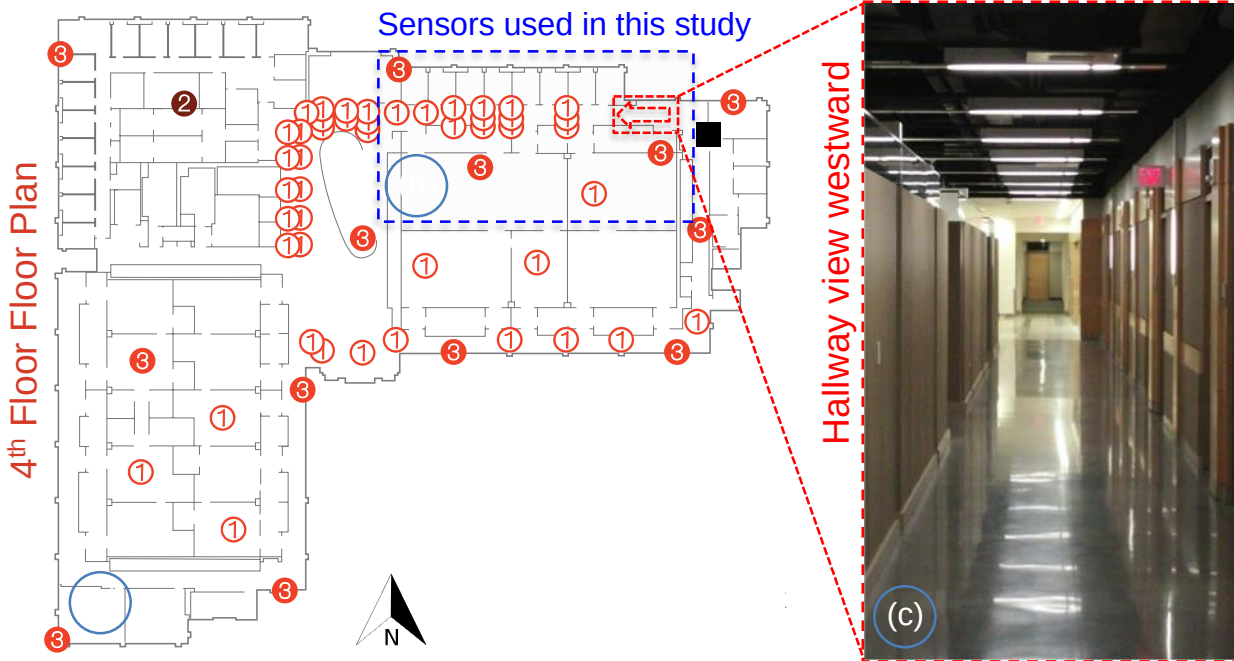


Figure 8. Testing Hallway. The hallway used for testing: (a) shows a floor plan of the 4th floor depicting the location of the accelerometers, (b) the blue dashed area shows the section used in this study, and (c) shows the view from the east side of the building facing westward. Building north is labeled. The experiment used only a subset of the total accelerometers in (b).

The lack of reproducible accelerometer locations also potentially creates an issue with the development of two of the feature sets. The machine learning algorithms use the feature sets as the inputs. The feature types affected are the multiple sensor step (MSS), and single sensor step (SSS) feature types, introduced in Chapter 4. These features assume spatial relations for their formulation which requires knowing the location of the sensors. Figure 9 houses the supposed accelerometer locations below for reference. The supposed locations of accelerometers were used in the development of these features, and each produces varied results for gender classification in Chapter 5.

The last potential effect of not knowing the sensor locations is in the methodology in Chapter 6. The Sweeping for Truncation step in the methodology discusses truncating not useful accelerometers for consideration in the classification algorithms. The examination of the sensors to be truncated, determined by the methodology, is not completed in this thesis because the sensor locations were unreproducible. This methodology can be completed with known sensor locations to try to infer the relationship between ‘good’ and ‘bad’ sensors. The next section discusses the hardware, both sensors and data acquisition, used to record the gait information.

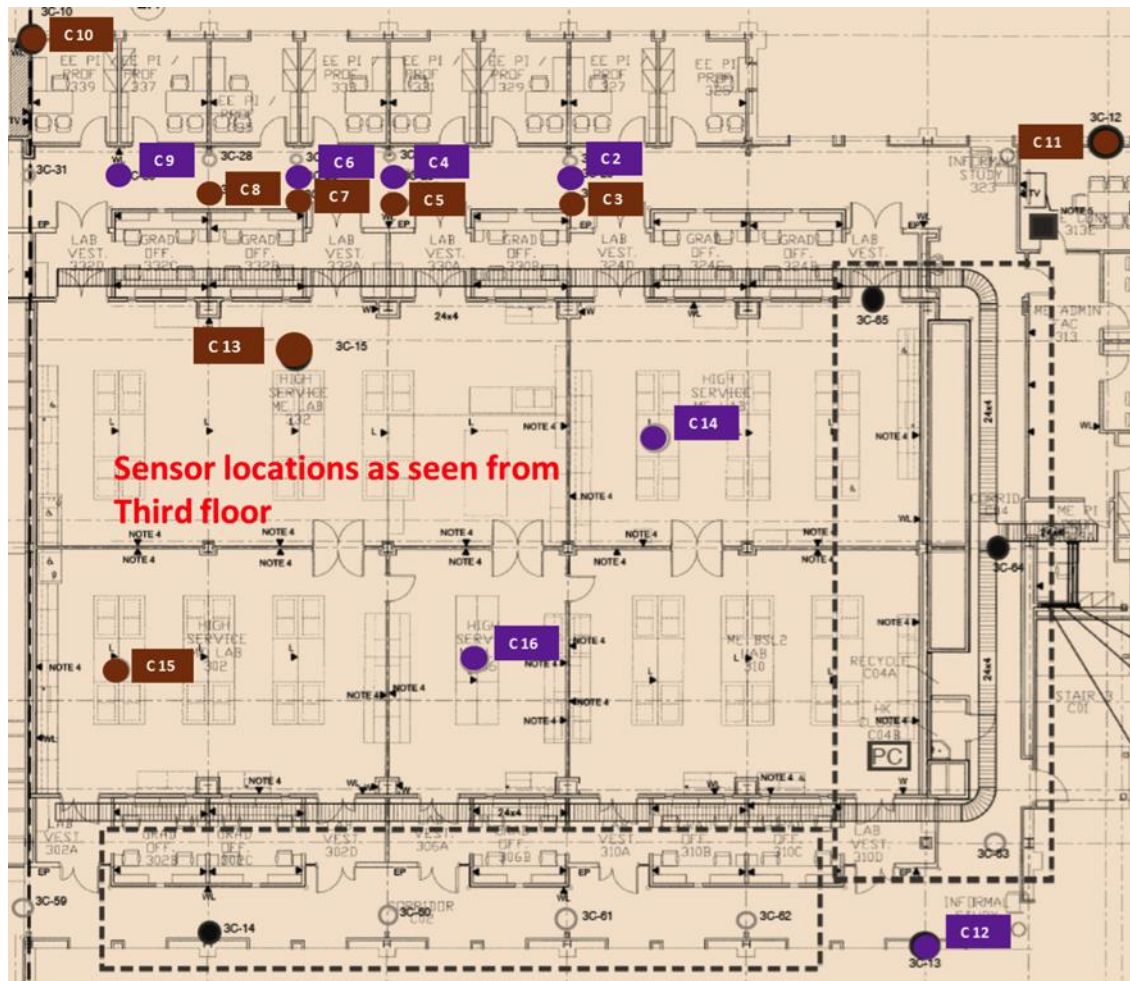


Figure 9. Accelerometer Locations. The accelerometer locations as verified before the walking experiment. The sensors used in this study are those listed as C – 1 through 14. This diagram is of the third floor of Goodwin Hall. The sensors are mounted to the ceiling and therefore are under the walking surface of the fourth floor test hallway.

Hardware

Fourteen of the 212 accelerometers (model PCB393B04) positioned in and around the testing hallway recorded gait in the walking experiment. A VTI Instruments data acquisition system (DAQ) captured data from the accelerometers. Multiple of these DAQs are positioned throughout GH to limit the length of cabling for each accelerometer, but a single DAQ was used to record all of the data in this work. A short description of the DAQ components follows.

A DAQ is composed of multiple components from various companies, and all five DAQs are interconnected on a private network in GH. In a single DAQ, a VTI CMX-09 chassis houses multiple EMX-4250 measurement cards. These cards record measurements and connect to the accelerometers through a breakout board (VTI model EMX-4016). The cards then transfer the acceleration measurements to the CMX-09 chassis. The chassis, card, and break out board encompass a single DAQ and the interconnectivity the five DAQs in GH is discussed next.

The DAQs network to each other via Ethernet cable using multiple Oregano switches (model syn1588 Gbit switch). These so-called transparent (due to very high throughput speed) switches connect to each DAQ and then to a data acquisition computer on the third floor of GH. A further explanation of the instrumentation installed in GH is available in Tarazaga et al. [53] and Hamilton [6].

Chapter 2.2. Walkers Investigated

Collection of the gait measurements took place in a single hallway while each study participant walked alone. The title ‘walkers’ now refers to these study participants. Two characteristics of each participant were collected: gender and weight. Fifteen walkers, eight males and seven females, completed this study. The range of the walkers’ weight was 109 – 240 lbs. The below sections detail other notes about the walkers.

Occupant Diversity – The walkers were mainly students attending Virginia Tech with some faculty also walking. While not only students walked in the study, future care must be taken to ensure the database of walkers is diverse. A more diverse database will enable the development of robust characteristic classifiers because it was created ‘seeing’ the gait of many types of people. While this experiment included an equal number of men and women as well as a wide range of body weight, this study is potentially not representative of the average age of any building.

While the experimental procedure is not collect walker age, it would be a safe assumption the average age of all walkers was near 25 with little variance. This tight grouping in age makes understanding how the developed classifying models would generalize to different age groups difficult. Jain et al. [7] have shown that gait varies with age, so the classifiers developed in this thesis may only reproduce the reported accuracies for new walkers in a similar age group. It is important that the walker database is varied because developing classification models on a more diverse database will ensure robustness in classifying anyone.

Shoe Type – Although, footwear has been shown to have a small effect on gait measurements [43, 54], walkers were asked to state whether they considered their shoes hard or soft soled. Three females and two males were asked to walk in both hard and soft soled shoes, and the other walkers wore either hard or soft soled shoes. This strategy was used to account for differences in footwear that may be worn by any ‘typical’ walker in GH. In summary, thirteen soft soled and seven hard soled walkers were tested, totaling 20 walkers. Each of the 20 walkers completed six walking trials: three walking Eastward, East to West, and three Westward. In total, 120 walking trials are included in the database at the time of this writing. The next section mentions the review process that oversees the collect of human data at Virginia Tech.

Internal Review Board

The Internal Review Board (IRB) oversees the collection of all human data at Virginia Tech. The experiment was conducted by best practices and Virginia Tech requirements. IRB Protocol # 15-681 warranted the collection of the data for this thesis. In this Protocol, consent from each walker was implied through their participation in the experiment. The only information collected during this study was the gender and weight of each walker. All future human walking tests will require adhering to IRB protocol. A second IRB Protocol was approved (# 15-838) which can include up to 100 individual walkers, and allows for the collection of more potentially useful characteristics of walkers for further classification studies.

The next section discusses the experimental procedures used in recording every walking trial. The section includes both an experimental description to provide perspective on the insight gained during the experimentation and the procedures themselves.

Chapter 2.3. Experimental Procedure

This section provides perspective to the experimental setup and the specific procedures used. The experimental description provides background about the specific testing to facilitate proper understanding of the procedures.

Experimental Description

The description of the experiment is divided into subsections that provide a deeper understanding of the experiment and its requirements. This section is presented before the procedures so the reader will have a background of the experiment before the presentation of the specific instructions.

Testing Environment – The experiment was completed on a Saturday (morning to evening) to limit unintended gait measurement. Still non-study participants moved throughout GH. The walking trial required repeating if research scientists noted the movement of non-study participants in the testing hallway. Research scientists did not monitor the movement of non-study participants in labs and offices adjacent to the hallway during the experiment. Research scientists did not move during the experiment to ensure the gait measurements were that of the walker alone.

Walker Velocity – Walkers chose their walking speed as the procedure does not stipulate this value. Not controlling this experimental parameter was decided by researchers to ensure walkers moved naturally. Recording a walkers' natural gait is important so that the measurements are truly representative of them. Dictating walker velocity could bias their natural gait and may adversely affect the developed classifiers.

Characteristic Collection – The classifiers of interest, gender and weight, were collected after the walking trials were completed to avoid any biasing that might result from asking a walker before walking. On a similar note, the weight of each participant was asked for by the research scientist potentially causing issues. The reported weight could be different from the actual value for multiple reasons, i.e. a person might choose to be misleading, or they might not know their weight. Future experiments should quantify whenever possible the characteristics collected from walkers.

Required Sensor Warmup – There is a required waiting period between the start of data acquisition and the start of recording meaningful data. When started, the accelerometers initialize to a steady state voltage. The measurements

collected during this transient period are non-representative of the walking signal. Thus, before each walker was signaled to begin, the accelerometers recorded for approximately 10 seconds for appropriate measurement of gait. Figure 10 shows a walking signal which includes the accelerometer warmup for illustration of the issue.

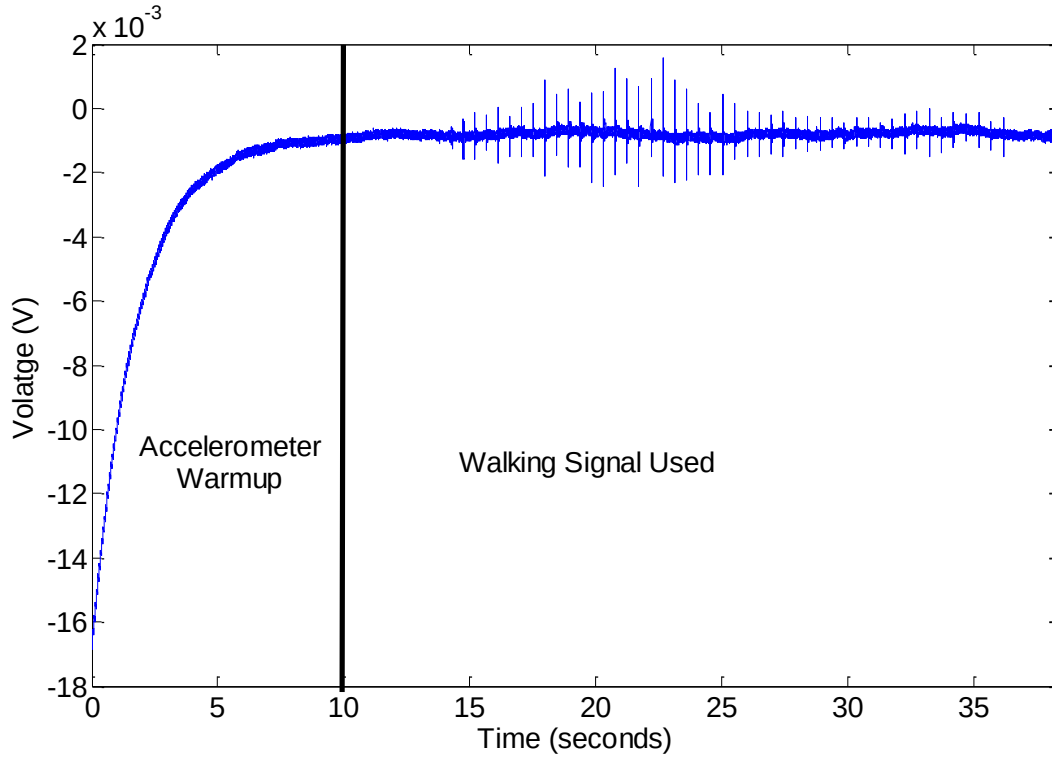


Figure 10. Accelerometer Warmup. The depiction of the required accelerometer warmup period. After approximately 10 seconds the signal reaches a constant voltage.

Data Acquisition Code – Joe Hamilton developed the code used to acquire the gait measurements in MATLAB [6]. This version of the code was developed to take preliminary data from the DAQs while the full acquisition software was in development. Hamilton [6] contains the developed code. Between the recording of gait measurements and the analysis for this thesis work a transition in hardware and software occurred at the VTSIL. This transition caused an issue with accelerometer locations used during the experiment as discussed above.

Sampling Rate – The sampling rate used throughout this experiment was 51200 Hz. The relatively high sampling rate allows for high fidelity time resolution but requires accounting for resonant effects by filtering in post-processing. There are two resonant effects inherent to the DAQ system requiring removal: the accelerometer itself and the

accelerometer mount. The manufacturer reports a natural frequency ≥ 2500 Hz [52]. Hamilton et al. [55] found the natural frequency of the mount to be 3400 Hz. Thus, the minimum resonant frequency is 2500 Hz. The filtering discussed below removes the resonant effects of the DAQ system. Chapter 2.4 discusses this filtering. The next section discusses the procedure used to collect each trial of walking data.

Procedure

The experiment required three research scientists to effectively record a walking trial. The three research scientists are named 'A', 'B', and 'C' below. As discussed above the data acquisition computer is on the third floor of GH, and the walking experiment was completed on the fourth floor. Research scientist 'A' controlled the data acquisition computer and research scientist 'B' and 'C' were positioned at the start and finish line of the testing hallway respectively. The ordered list below details the experimental procedures used for all walkers.

- a. Position the walker facing Westward at the start line marked by tape.
- b. Begin acquiring accelerometer measurements. (Completed by researcher A)
- c. Signal the walker to begin. (Completed by researcher B)
- d. Signal research scientist A to stop data acquisition. (Completed by researcher C)
- e. Repeat steps a. through c. for the walker moving Eastward.
- f. Repeat steps a. through d. twice more to record six walking trials: three in each direction.
- g. Request the weight of the walker and note their apparent gender.

The next section details the methods used after the collection of the gait measurements through the above procedure. The raw gait measurements were processed to reflect typical data processing practices and to filter the measurements for the known physicality of gait.

Chapter 2.4. Processing Gait Measurement

The recorded acceleration measurements were processed using typical digital signal processing such as:

- DC offset removal from of each accelerometer – The mean of each accelerometer signal is subtracted from each respective signal to center the measurements around 0.0.

- Initialize a zero value for each accelerometer – The initial value of each accelerometer signal was translated to zero and with all subsequent data points moved accordingly.
- Apply a low pass filter – A Butterworth filter of order three with a 500 Hz cutoff frequency was used to filter the measurements. This filtering removed the resonant effects of the DAQ system while retaining the expected frequency content due to the normal force of a footstep in accordance with Ekimov and Sabatier [35].

Further data processing was required when a drift in the accelerometer signals was noted after the creation of the walking database. Chronologically this drift was noted after the analysis for Chapter 5 and before Chapter 6 . During laboratory testing, Dr. Bryan Joyce discovered that the accelerometers have a large temperature dependence. Sizable changes in acceleration measurement (on the order of 0.5 G) were induced by blowing on or even cupping warm hands around the accelerometers.

The experiment procedures did not monitor the temperature at each accelerometer, so this drift is present in the walking database but is correctable. To remove the drift every walking trial in the database was linearly de-trended. Figure 11 illustrates this effect of de-trending the signals. Two sensors, one requiring de-trending and one not, are shown. Also shown, is the effect of de-trending on each signal. The de-trending of the sensor with no apparent drift has little effect. While the sensor requiring drift removal now behaves similarly to the non-drifting accelerometer.

In addition to the 'detrend' function, a low cutoff frequency was used to remove the drift. The filter used a low cutoff frequency of 10 Hz. This low cutoff frequency is to ensure the removal of the low-frequency effects of the drift. Sabatier and Ekimov [41] support this low cutoff frequency as they found that 10 Hz is an appropriate low cutoff frequency for walking measurements.

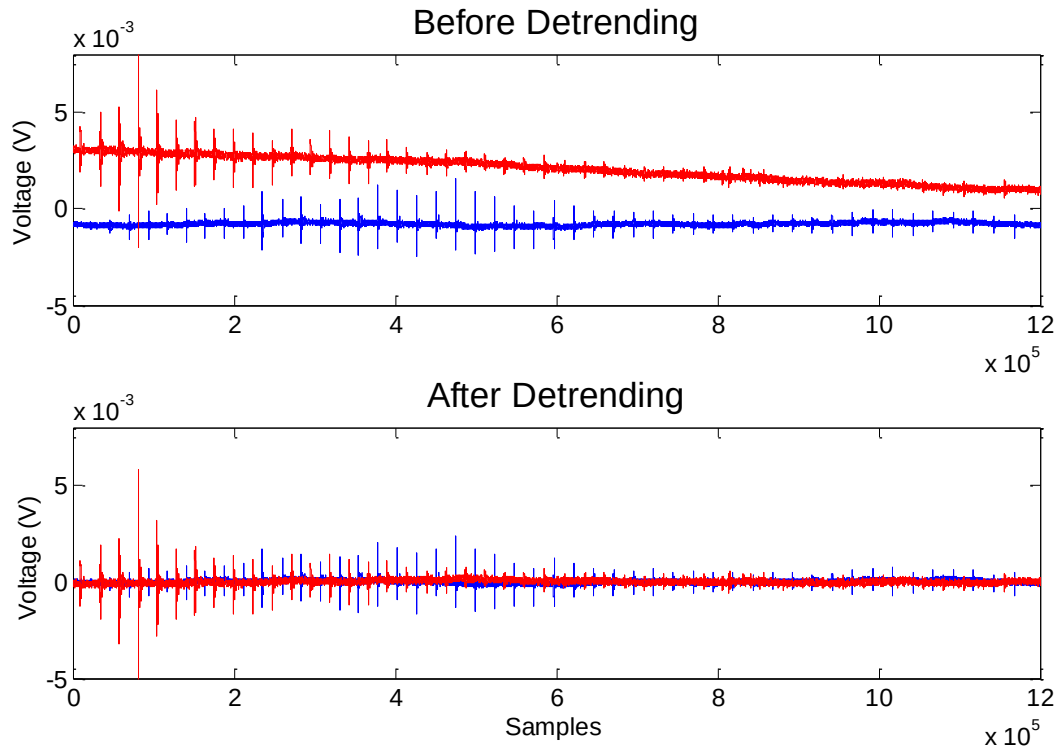


Figure 11. Accelerometer Detrending. The walking signal of two accelerometers: one requiring de-trending and the other not. A before and after of the de-trending shows that there is little difference in the signal not requiring de-trending, but there is a difference between the before and after of the accelerometer requiring de-trending.

Although the database used in Chapter 5 (where the aim is to classify gender of walkers) does not account for the drift, it does not affect the legitimacy of the results because the machine learning algorithms were able to produce 88 % gender classification in the presence of the drift. Chapter 6 use the de-trended database for all analysis. As previously stated Chapter 5 reports an 88 % gender classification accuracy while the data reduction methods in Chapter 6 lead to a 93 % accuracy. Although, de-trending of the database took place between the two chapters the improvement in classification accuracy increase is believed to be due to the data reduction methodology and not to the removal of the drift. This assumption is supported by the fact that a small portion of the sensors required de-trending. Only

approximately 7 % of the accelerometer recordings required de-trending. The representation for future analysis and storage of the database is discussed next.

Chapter 2.5. Data Representation

The dimensionality of the raw data for a single walking trial is $\mathbb{R}^{m \times n}$. Where m is a function of the number of samples and time required to complete a walking trial and n is the number of accelerometers used (i.e. fourteen). A typical walking signal recorded by one accelerometer for is shown in Figure 12 for reference, each peak represents a footstep. The detail of the figure shows a single step as used in the formulation of the feature types discussed in Chapter 4.1.

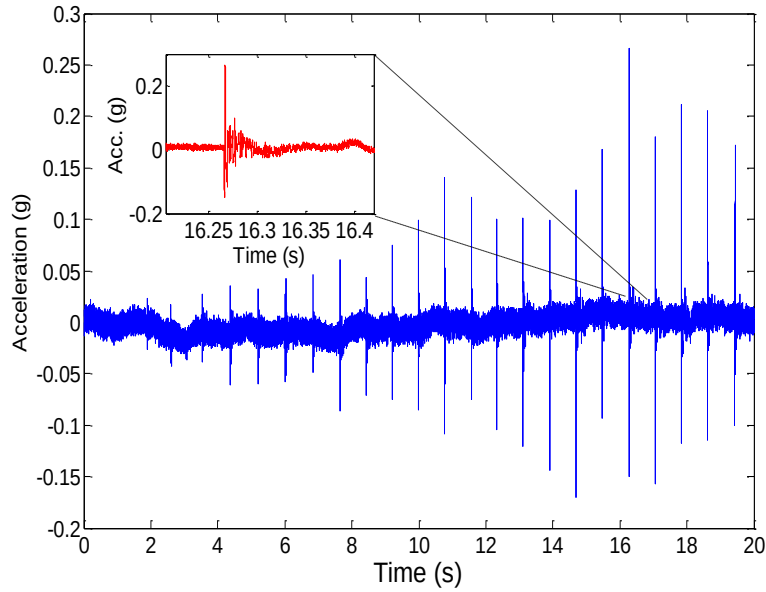


Figure 12. A typical walking signal. This particular signal is trial one of Female A wearing hard soled shoes. The peaks of the signal represent footsteps. The inlaid plot shows a portion of the same signal.

Database Variable Structuring – Since the walking speeds were not constant for each walker, there is a variance in length of gait signal measured. Due to the variation in signal length, the gait measurements requiring structuring for effective storage in the database as simple concatenation is not possible. The details of this structure are discussed below for future use of the data. The data recorded in the development of the database is stored as a (.mat) file following the structure in Equation 1.

$$Data.(Direction).(Data Type).(Reduction Type).(Walker Name).trial(number).value \quad \text{Equation 1}$$

Where:

- *(Direction)* is *(Left)* or *(Right)* (i.e. (Eastward) or (Westward) walking trials)
- *(Data Type)* is the name of the data type summarized in Table 1.
- *(Reduction Type)* is not included in this database set, thus this value will always remain NoRedux, representing no reduction method.
- *(Walker Name)* is the name of the walker participating in the study which follow the format of:

$$(Gender)_{(Letter)}_{(Shoe Type)}$$

Where *(Gender)* is either Male or Female, *(Letter)* is a capital letter ranging from A-H for males and A-G for females, and *(Shoe Type)* is either hard soled shoes, HS, or Soft Soled Shoes, SS.

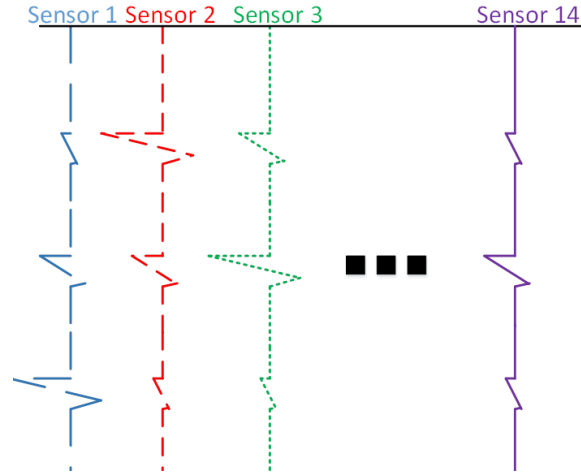


Figure 13. Walking trial schematic. This illustration shows a caricature representation of a single walking trial. The representation is used to illustrate the construction of the feature types used in the machine learning algorithms in the feature selection section below. The fourteen sensors used in the experiment manifest as columns, n , and the row dimension is m , with each deviation from zero representing a footstep.

Table 1. Summary of Data Types in the structured database.

Name	Name of Data Type
Avg_10	SAS10
Avg_5	SAS5
Avg_3	SAS3
Avg_10_FFT	SAS10 Freq
Avg_5_FFT	SAS5 Freq
Avg_3_FFT	SAS3 Freq
Multistep	MSS
Multistep_FFT	MSS Freq
Single_1	SSS
Single_1_FFT	SSS Freq
Raw	Raw Data

- *trial(number)* is the walking trial number ranging from 1-3. Each number is in chronological order, i.e. *trial(1)* occurred before *trial(2)*.
- *value* is the $\mathbb{R}^{m \times n}$ walking trial measurements.

With the completion of the experimental setup, the reader should now understand the walking database to include the limitations associated with the data acquisition. The subsequent chapter introduces machine learning. This broad field of artificial intelligence enables searching data for complex patterns. It is assumed that these patterns hold the gait information that enables characteristic classification. The results of this approach, shown in Chapter 5 and Chapter 6, show the usefulness of this approach to classification.

CHAPTER 3. MACHINE LEARNING

Machine learning essentially is a field that finds complex patterns in data. Machine learning algorithms function by using input data to ‘learn’ appropriate parameters that produce a predictive or descriptive model. Supervised and unsupervised learning comprise the two major categories of machine learning. In Supervised learning, the models are developed to predict the outcomes of never before seen observations of data. While in unsupervised learning tries to make sense of the data in a more general way; a specific prediction may not be possible for these techniques.

Supervised Learning

Supervised learning finds patterns in data using labeling for each observation in that data. Each observation in a dataset has multiple features and a corresponding label. A supervised learning algorithm tunes algorithm specific parameters to optimize the outputted model’s ability to predict the label for observations. These developed models then predict labels for new/never-before-seen observations. The goal of these developed models is to have a low classification error (few mislabels) for known and new observations.

Figure 14 shows a representative supervised learning data set. The entire figure represents a dataset built from multiple observations of specific features. The machine learning algorithms develop a model to predict the label for a new observation based on the value of each feature. The discrete values of the label in this figure make this a classification model similar to those included in this thesis work. A continuous value for a label would produce a regression model.

Observation	Feature 1	Feature 2	...	Feature N	Label (Class 0 or 1)
I	1	0	...	1	0
II	1	0	...	0	0
III	0	1	...	0	1
...	...	...	...	...	...
N	0	1	...	0	1

Figure 14. Example supervised learning data set. Supervised learning requires the label of a specific observation to develop a model to effectively label a new observation based on the provided features. Features are not required to be binary.

The supervised learning algorithms in this work require three stages: training, validation, and testing. The training stage uses a certain percentage of observations to develop a general predictive model. Then this general model is used

to classify a certain percentage of the observations in the validation stage. This stage sets ‘hyper-parameters’ or model design choices to optimize the performance of the generalized model from the training stage. Finally, the test stage determines the classification error of the tuned model in classifying new data observations. The observations for each stage are not shared to reduce the bias of the algorithm as training on the testing percentage of observations is learning with the answer and is ‘cheating.’ In other words, having unique observations for each stage ensures that the model can predict on new, unseen data.

The error (mislabel percentage) in the test stage is the most important error value as it represents an estimation of how well the model will generalize to a new observation. In a broader sense, this test error represents how an algorithm should function for any new dataset.

The percentage of observations in each stage can vary. The initialization of observations for use in each stage can affect the developed model’s accuracy in classifying test observations. Use of a ten-fold cross-validation (TFCV) method reduces the effect of this feature initialization. TFCV randomly partitions the observations into ten, nearly equal parts, known as folds. The testing stage uses a single fold, and the training and validation stages use the remaining nine folds. The model developed using the training and validation folds then classifies the observations in the held out tenth fold and records the error. Then, a different fold is held out, repeating this process until all folds have been held out a single time. The test error from each fold is averaged together. This average test error serves as a good representation of the expected error when classifying a never before seen observation [56]. Figure 15 illustrates the cross-validation method.

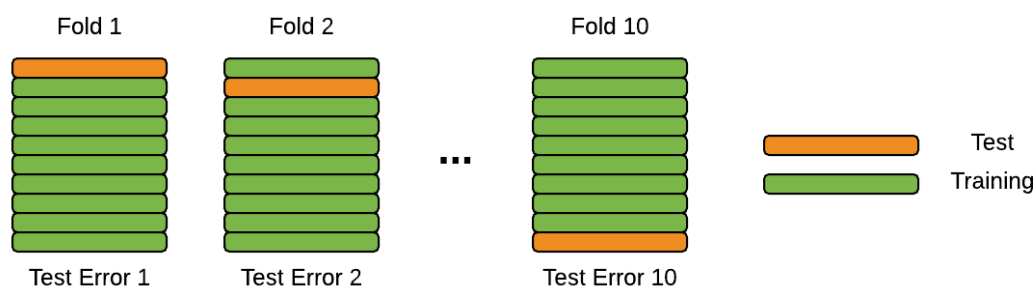


Figure 15. Ten-Fold-Cross-Validation Method. Ten folds divide the total number of observations. The training and validation stages are represented in the training stage. The ten test errors are then averaged to estimate how a model will generalize to never before seen data.

Unsupervised Learning

Unsupervised learning is used to organize data without observation labels. These types of algorithms are useful in many applications and can even be used to examine data where labels are unknown. Since labels for the data are not present, there is no concept of classification error as there is no ‘correct’ answer. Unsupervised learning is presented here for clarification purposes only, and it is beyond the scope of this thesis.

Chapter 3.1. Machine Learning Algorithms Used

One aim of this thesis work is to understand the effect of choosing a specific machine learning algorithm on the characteristic classification accuracy. The machine learning techniques examined are Bagged Decision Trees, Boosted Decision Trees, Support Vector Machines (SVMs), and Neural Networks (NNs). Selection of these algorithms took place because of their general applications and acceptance in the machine learning field. Below each algorithm type is discussed with references given for further reading.

Decision Trees – Bagged and Boosted

Decision trees involve splitting on an observation’s features until there is a high probability of a specific label. A node represents the splitting or querying point of a feature, and a branch represents the flow of an observation’s features to further nodes. Classification takes place at leafs when labeling an observation is likely to be correct. Figure 16 shows an example decision tree for reference.

Exploring the example in Figure 16, the initial splitting examines the feature x . In the case that $0 < x < 1$, the unlabeled observation’s features is carried down the middle branch. The next node queries another, not shown, feature of the same observation, branching to one of two leafs. The observation will be assigned a predicted labeled at either of the two leafs. This simple example shows how decision trees work and can easily be scaled up to accommodate many more features.

The training and validation stages of decision trees determine features the tree should split on and the appropriate depth of the tree. The depth of the tree is the number of nodes necessary for an accurate classification. To test a tree’s accuracy, each feature of a new observation is evaluated at each node until a label is assigned. Further reading on decision trees is found in Quinlan [\[57\]](#).

According to Martin and Hirschberg [58], the computational complexity of a simple decision tree is $\theta(A^2N^3)$ where A is the number of observations and N is the number of features. This nomenclature is known as the Big θ (theta) notation which estimates the lower bound for computation time. This notation is useful for understanding how computational time changes based on the input size [59]. This computational efficiency is introduced because it motivates the need to perform data reduction to diminish computational loads because more observations and features require more computation. The data reductions are discussed in Chapter 4.

Choosing the most appropriate features on which to split is difficult a priori and motivates the combination of many decision trees. Ensemble learning methods combine multiple learners to solve complex problems. Ensemble methods include bagging and boosting machine learning algorithms. These types of decision trees manipulate the observations used in the training stage to generate multiple models. The performance of these multiple models combines to develop a better performing overall model [60]. Using ensemble methods requires the assumption that each of the weak learners (e.g., decision tree) performs better than 50% accurate to combine into a useful ensemble model [61].

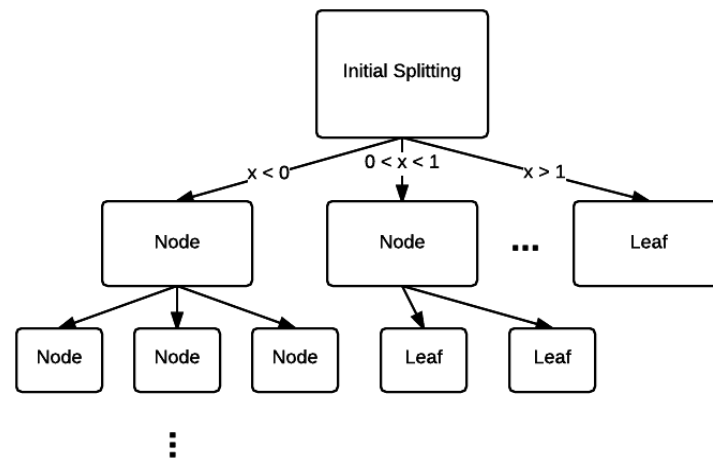


Figure 16. Illustration of a Decision Tree. Each node poses another ‘question’ of a specific feature and each branch carries an observation’s remaining features to the next node. Leafs terminate the flow of features and label or classify an observation.

As a background on the types of errors associated with supervised learning, there are two: bias and variance. Bias is the error resulting from assumptions made about the learning algorithm. Large bias error can lead to underfitting, where a model neglects apparent relationships between features and labels. Variance is the error resulting from the

variance in the observations allotted to the training stage. Models with high variance are susceptible to overfitting, or over-weighting the noise present in an observation's features [62]. Ideally, a model has low bias and variance error, but an inherent balance between bias and variance error exists. Reducing bias error almost always increases variance error [63].

Bagged Decision Trees

Bagging machine learning algorithms reduces the variance associated with high variance, low bias algorithms [60, 64]. The process of bagging takes the original observation set and constructs a new data set by taking examples of observations with replacement from the original observation set. The dataset is developed to have approximately 63% unique features, while the remainder are duplicates of the original observations [64]. A model, a decision tree, is developed using the labels and this dataset developed through duplication. This process repeats for a specific number of iterations, and the test error of each model averages for a representative test error.

Boosted Decision Trees

Boosted machine learning algorithms reduce the bias associated with high bias, low variance systems [65]. In the process of boosted learning, every simple model (i.e. a decision tree) is weighted based on its ability to classify accurately. The final model (i.e. boosted decision tree) is a weighted combination of these simple models. This work uses the AdaBoost method. This method is an adaptive boosting method leveraging the results of many weak learners [66]. Adaptive in the sense that the later developed trees focus on the difficult to classify observations as determined by the other previous tree classifiers. This method uses optimally-weighted, majority vote of the weak classifiers to determine an appropriate label for a new observation. Further reading found in Dietterich [60].

The decision tree models were developed in MATLAB using the 'fitensemble' command. The 'Kfold' cross-validation value was set to 10 folds by the above-discussed ten-fold cross-validation. One hundred trees make up the weak learners. For the boosted trees the 'AdaBoostM1' setting was selected, and for the bagged trees the 'Bag' setting. The 'kfoldLoss' command then estimated the generalization error associated with the developed ensemble model. Appendix A shows an example of the coding used.

The ensemble decision trees prove to be some of the best performing in Chapter 5 for accurate classification. For this reason, they are the only included machine learning algorithms in Chapter 6. The next machine learning algorithm discussed is neural networks, which were inspired by the neurons in the brain.

Neural Networks

Neural networks were inspired by the neurons in the human brain that essentially break down complex inputs and uses a large number of neurons to make a choice of the appropriate response to a given situation. Neural networks are machine learning algorithms which map an observation's features to 'hidden' neurons and then to a label producing output node via a series of non-linear functions. Figure 17 shows the type of neural network used throughout this thesis work, where a non-linear function represents a neuron. For each feature, a weighting must be learned to map to a neuron in the hidden layer. So-called "hidden" because these weights and neurons are not the focus of the analysis: the observation label is the focus. The weights learned for a single neuron in the hidden layer feed through a non-linear function. Any non-linear function could be used with common functions being Sigmoid or step functions. The outputs from the neurons in the hidden layer are weighted and summed into the output layer. The non-linear function in the output layer determines the labeling for a specific observation. The training and validation stages learn these weights.

This work uses linear mappings to neurons in a single hidden layer. Using more than a single hidden layer can add unnecessary computational complexity and require further parameter sweeping to optimize the network [67]. More hidden layers could approximate more complex patterns. This work uses a single layer neural network because they can approximate any continuous function [48]. The number of neurons in the hidden layer does, however, require optimization. Chapter 5.1 investigates the number of neurons.

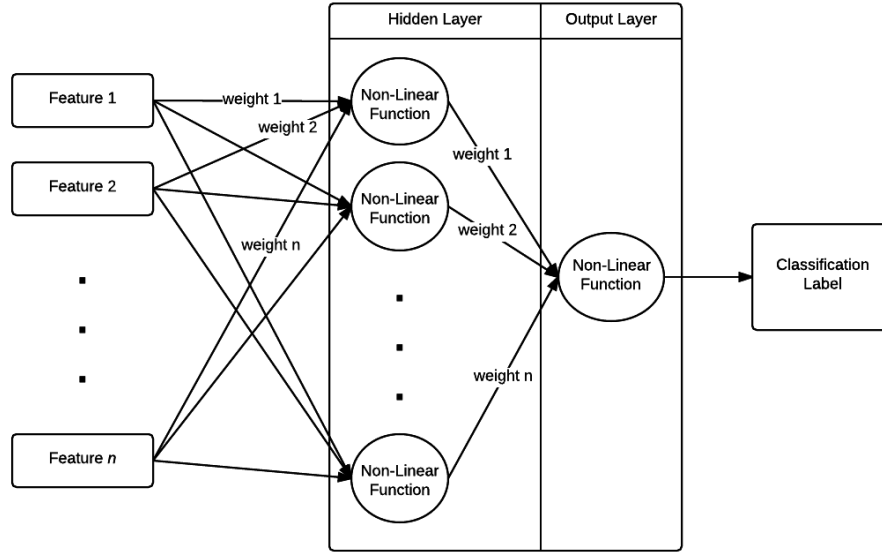


Figure 17. Neural network illustration. The number of neurons in the hidden layer is a parameter that is chosen a priori, thus requiring optimization via a parameter sweep. Every connection represents a weight used as an input to the following non-linear function.

MATLAB's 'patternnet' command was used to develop the neural networks in this thesis. The 'patternnet' command allows for the generation of an empty neural network. Then the 'train' command develops the empty neural net into a useful model using training observations and corresponding labels. The training method used was the scaled conjugate gradient backpropagation because it does not require further user inputs to operate successfully. This method essentially eliminates a further optimization requirement by other training methods [68]. Finally, the developed net is used to classify the test set. The ten-fold-cross-validation coded was completed with a developed function. This developed function divides the walking observations into the appropriate folds. Appendix A houses example coding.

These neural network models produced statistically poor results when compared to the other machine learning algorithms in Chapter 5. Thus, the computation time for typical neural networks is not examined because Chapter 6 does not include them in the analysis. Neural networks are still included in this work despite the poor performance because of the popularity of the algorithm type. Haykin [69] provides further reading on the immense field of neural networks.

Support Vector Machines

Support Vector Machines (SVMs) ‘learn’ a hyperplane which divides the classes. The hyperplane specifically maximizes the margin between the two classes. Figure 18 shows an example SVM. This hyperplane is defined by the observations closest to the boundary dividing the classes: those which are most difficult to classify. Those difficult observations serve as the support vectors effectively defining the hyperplane, hence, the support vector machine name.

Solving for the hyperplane requires writing an SVM as a convex optimization. Then a quadratic program solves the convex optimization problem [70]. Figure 18 depicts linearly separable classes, but classes are often not linearly separable necessitating higher-order hyperplanes. Various types of hyperplanes are made possible by what is called the kernel trick.

The kernel trick allows for the replacement of inner products (i.e. measures of similarity) which are inherently present in the formulation of SVMs, with kernel functions [71]. Kernels map an observation’s features into a higher dimensional space through a chosen non-linear mapping. The non-linear mapping (i.e. polynomial) is chosen a priori to develop an optimal hyperplane that separates the classes [72, 73]. A suitable Kernel must be a proper inner product, meaning that it adheres to Mercer’s Theorem [74].

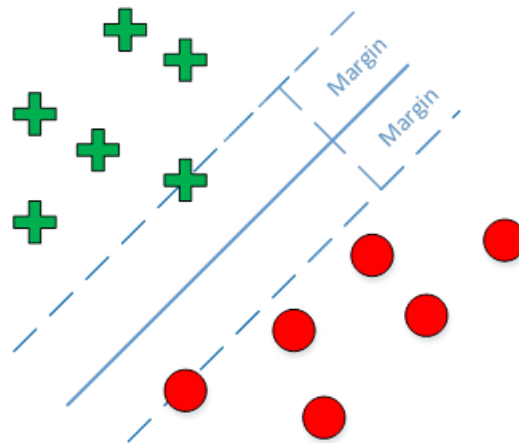


Figure 18. Support Vector Machines illustration. There are many lines that can separate the two classes, but the SVM solves for the line that gives the maximum margin between them.

Specifically, this thesis work studied the Radial Basis Function (RBF) and polynomials of order 1-25 SVMs to understand how the type of SVM affects classification accuracy. Equation 2 and Equation 3 show the polynomial and radial basis functions for reference.

Polynomial

$$K(x, y) = (x^T y)^d$$

Equation 2

Kernel

Where x and y are the observations (specific walking trials) and d is the order of the polynomial.

Radial

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right)$$

Equation 3

Basis

Kernel

Where, x and y are the observations (specific walking trials) and σ is an optimized free parameter.

The computational complexity for typical modern SVMs are $O(n^d l)$ [75], adapted for a polynomial kernel function. Where $O()$ represents the Big O notation, estimating an upper bound on the computational time. Above n is the number of observations, d is the dimensionality of the polynomial kernel used, and l is the number of features. As discussed with the decision trees, this estimate of computation time is given to understand how the inputs to these algorithms can increase computation time. Limiting the number of features and observations plays a key part in reducing computational time, potentially allowing for real-time classification. This informs the use of the data reduction techniques in Chapter 4.

The ‘fitsvm’ command in MATLAB was used to develop the SVMs. The ‘KFold’ setting was chosen as 10 for the above-mentioned ten-fold cross-validation, resulting in each fold containing 12 walking trials. The ‘kfoldLoss’ command is then used on the developed SVM to estimate the test error of the algorithm. The quadratic program used to solve the quadratic program was the Sequential Minimal Optimization. The Sequential Minimal Optimization (SMO) reduces computation time making computation of large dataset possible based on removing the matrix computation [76]. Appendix A houses example coding.

As previously discussed, SVM kernels for polynomial and radial basis functions were used to classify the gender of walkers in Chapter 5. Chapter 5 shows no statistical difference between SVMs and the two types of decision trees. However, SVMs are not included in the Chapter 6 analysis because they require further optimization through parameter sweeps (i.e., polynomial order and σ for RBFs). Optimally choosing the kernel would require 26 (25 polynomials and RBF) models to be completed, greatly increasing computation time, so only the just-as-well performing decision trees were carried over to the data reduction work in Chapter 6.

In summary, these four machine learning algorithms were combined with the ten developed feature types (see Chapter 4.1) to investigate the effectiveness of each algorithm for characteristic classification accuracy. Machine Learning has a tremendous number of algorithm types that have varied success in numerous applications; other algorithms may prove more appropriate for this type of analysis. The next section examines successful classifications that have taken place in literature. Some of these works use some of the previously mentioned machine learning algorithms, and others use different approaches. All classification methods use some form of machine learning because even those only thresholding a single value are technically using a decision stump (one-level decision tree).

Chapter 3.2. Examples of Gait Classification in Literature

Many of the previously mentioned gait measurement techniques (Chapter 1) have used their recorded gait data in a classification context. Numerous studies have classified various attributes using only gait measurements as inputs, ranging from walker identification to diagnosing neurodegenerative diseases [17, 77]. Yun [38] contains a survey of gait classification examples. Table 2 below contains a small survey of other previous gait classification work not included in Yun. These classification algorithms range from thresholding a single term to complex machine learning strategies. The gait measurements serve as the inputs to these models, and the label depends on the developed type of classifier. The remainder of this section discusses the specifics of the classifications detailed in Table 2. The examples not closely related to the classification in this thesis work are organized by gait measurement technology below. Following this formal introduction of machine learning, the closely related gender and weight classifiers from Chapter 1.5.1 are reexamined.

Table 2. Gait classification further reading. The source, the gait measurement method, and what was characteristics were classified.

Category	Source	Gait Measurement Method	Classification Goal	Accuracy Achieved (%)	Machine Learning Algorithm	Special Notes
Gait Measurement Classification	[77]	Video Surveillance	Human Recognition	95	k-nearest neighbors	Generalized Symmetry Operator
	[24]	Video Surveillance	Individual Identification	93	k-nearest neighbors	Leave-one-out-cross-validation
	[51]	Video Surveillance	Individual Identification	89.33 - 99.07	Self-organizing- maps (SOM)	N/A
	[17]	Database/ Video Surveillance	Various Neurodegenerative Diseases	79.04 - 93.96	Support Vector Machines	10-fold cross-validation
	[32]	Walker Mounted Accelerometers	Walking upstairs/ downstairs/ level ground	97.72 / 93.18 / 93.93	Decision Trees	Discrete Dyadic Wavelet Transformation
	[1]	Pressure Sensors	Individual Identification	80	Decision Trees	N/A
	[40]	Walking Surface Mounted Geophones	Footsteps/ Occupancy	99.55 / 85	Decision Trees	N/A
Gender Classifications	[42]	Microphone Recording	Gender/ Individual Identification	NR / 60	Support Vector Machines	Modulation Features
	[43]	Microphone Recording	Gender	75	Discriminant Analysis	Principal Component Analysis
Weight Classifications	[12]	Picture Database	Weight	100 $\pm$ 5	Multiple Regression Analysis	N/A
	[44]	Video/ RGBD	Weight	96	Decision Trees	N/A
	[45]	Anthropomorphic Measurements	Weight	87.8 – 93.3	Decision Trees	N/A
	[46]	Footprint Measurements	Weight	N/A	Statistical Test	N/A
	[47]	Video	Weight	N/A	Neural Networks	10-fold-cross-validation

Video Surveillance

Hayfron-Acquah et al. [77] achieve 95% accuracy in classifying a human in images in video. They use the nearest neighbor approach to classify humans. The nearest neighbor algorithm is a supervised learning method that assigns the label of the ‘closest’ labeled point [78]. The distance between known observations’ features and a new observation’s features is calculated, and the new observation is labeled the same as its nearest neighbor.

Multiple video recording references classify an individual’s identity, such as BenAbdelkader et al. [24] and Lin et al. [51], to varying accuracies, 93%, and 90.40-99.07%. The former study uses the nearest neighbor method. The latter study identifies individuals via a self-organizing map. These maps are a specialized form of a neural network model that clusters “like” data together. This self-organizing map is an unsupervised learning method that requires the interpretation of results to correctly define groups.

The last video example uses an SVM to classify walkers with various neurodegenerative diseases such as ALS, MS, etc. The accuracies vary for each comparison from 79.04 – 93.96% [17]. These results are exciting and somewhat expected as physicians classically diagnosis these types of diseases in this visual manner.

These previous example classifications illustrate the capabilities of gait classification. Their results are exciting because they demonstrate that there are inherent patterns in gait allowing for a range of applications from identification of a person to diagnosing neurodegenerative diseases. These classifiers could potentially be developed using the underfloor mounted accelerometers to further the applications made possible in the VTSIL.

Wearable Technology

Wearable technology has been used to predict if a person is walking upstairs, downstairs, or on level ground. Nyan et al. [32] report 98, 93, 94% accuracy for each activity, respectively. The algorithm used in this classification is a decision tree. The lack of accelerometers in the stairwells of GH makes immediate implementation of this type of classifier difficult. More accelerometers could be added to the stairwells in GH to build the same type of classifier.

Sensitive Floors

Individual identification has been completed based on gait measurements from pressure sensitive floors by Middleton et al. [1]. They report an 80% accuracy in identification using decision trees. The features extracted from the sensitive floor are: stride length, stride cadence, and time on toe to time on heel ratio. These types of features are potentially extracted from the underfloor mounted accelerometers but these were not used in this work. The positive identifications may indicate that research can develop similar identification classifiers for use in GH.

Ground Measurement

One of the most similar methods of measuring gait to the GH underfloor mounted accelerometers, is that of Pan et al. [40]. They studied the ability to classify human occupancy via geophone measurements. Their occupancy methodology classifies based on using the spatial relationships of the geophones and ‘landmark’ steps of a person walking. The largest footstep recorded by a sensor is a landmark step, signifying that a walker is near that sensor. With approximate location, via the landmark steps, and knowledge of the building layout, logic can inform when a walker has left a hallway. They report 85 % accuracy in occupancy classification. The machine learning algorithm used in Pan et al. [40] work is a decision tree, or more specifically, a decision stump, which is a thresholding of a single feature. Geophones record the ground velocity while the underfloor accelerometers measure the acceleration. The similarity between these measurement types lead to the belief that a human occupancy classifier could be developed for use in GH.

Most Closely Related Classifications

The introduction chapter (specifically Chapter 1.5.1) examined the classifications of walkers most closely related to those in this thesis work. These references are briefly reexamined below now that machine learning has been introduced. These closely related classifiers are divided into gender and weight classifiers.

Gender Classifiers

The most similar efforts in gender classification as completed in this thesis work is by DeLoney [42] and Li et al. [43]. Li et al. [43], which reports a 75% accuracy in gender classification, uses a principal component analysis to transform the acoustic signal of gait. The principal component analysis is an unsupervised learning method that retains the

maximum amount of variance while reducing to a lower dimensional space. Then a discriminant analysis, or statistical analysis, is used on the transformed signal to classify gender.

Weight Classifiers

Krishan [46] examines the relationship between measurements of a footprint and body weight. Their results showed that when a subject held 20 kg weights, their footprint length and breadth tended to increase, allowing for estimation of body weight. Labati et al. [47] used two cameras to record a front and side view of walkers. They used computer vision techniques to extract features of the walkers. The extracted features are input to a neural network algorithm to estimate the weight of 20 study participants. Their mean error was 0.07 kg (standard deviation of 2.30 kg).

In summary, all of the classifications using gait measurements point to the ability to develop the same results regardless of gait measurement technology. Suggesting that research into the previously discussed classifiers could aid in the development of these technologies for use in GH using underfloor mounted accelerometers. The gender classification in this thesis is novel because of the gait measurement technology, and this work reports the highest accuracy found in the literature (75 %). The weight classification is novel because it is completed through a novel method of gait measurement. Future work may refine the weight classifier to achieve higher accuracies.

Summarizing the entire chapter, the machine learning algorithms discussed here enable gait pattern recognition. These approaches enable the classification of various walker characteristics without specific querying of features. These methods require little specific knowledge of the feature type and, as will be discussed, they perform well for characteristic classification.

The classifications completed in these previous works examine large amounts of gait measurements. Large data sizes lead to a need to understand appropriate reduction methods for successful characteristic classification. The following chapter discusses the methods of data reduction used in this work and methods used in literature currently.

CHAPTER 4. DATA REDUCTION APPROACHES

Data reduction enables faster training of machine learning algorithms and faster predictions by the developed models. The motivation for data reduction is to provide a contraction in data storage and computational time for classification – while often improving the generalization of an algorithm by removing unnecessary features [79, 80]. This chapter gives a broad introduction to data reduction techniques, discusses the methods used in this thesis work, and places these reduction methods in the context of previous works.

As a background to data reduction, methods are grouped into two categories: selecting “features” and selecting “instances” [81]. ‘Global’ feature selection methods initially reduce the walking database dimensionality. These feature selections create feature sets used as the input to the machine learning algorithms in Chapter 5 and Chapter 6. Chapter 6 uses instance selection methods presented in Chapter 4.2 to determine which accelerometers are the most useful for accurate classification models. The subsequent sections explain the two types of reductions.

Chapter 4.1. Feature Selection

Multiple features sets are created for use in machine learning algorithms to classify gender and weight. Ten feature sets are developed below and used in combination with the four machine learning algorithms presented in Chapter 3. Chapter 6 details the performance of each feature selection method.

Due to the large dimensionality of the walking database, raw data is often not a suitable feature type for machine learning algorithms. Large feature sets lead to large computation time and can result in low accuracies in classification [81]. There are countless transformations (or reductions) that can be used to reduce the database size, but determining the usefulness of a transformation a priori is difficult [48]. To truly determine how effective a feature type is the classification accuracy was used as a metric; requiring:

- a function must the feature must extract features from the raw data,
- a machine learning model must train with these features,
- testing must showcase the classification error using that feature, and
- statistical analysis must determine if this feature is truly better than another.

Because determining the worth of a feature is an involved process, initial features were developed to be ‘smart’ or rather to encompass a large amount of gait measurement.

The proposed feature types are ‘global,’ meaning that they encapsulate raw accelerometer measurements from an entire footstep. Figure 12 shows a walking trial recorded by a single accelerometer. A local feature is a more precise value that represents a very specific query of a footstep, like footstep amplitude, time between footsteps, the decay rate of a footstep, to name a few. Global features encapsulate many so-called local features, mitigating the risk associated with developing a ‘bad’ feature that must go through all of the above steps to be proven poor at classification.

The developed feature types are Sensor Averaged Step (SAS), Multiple Sensor Step (MSS), and Single Sensor Step (SSS). The SAS feature type averages multiple steps measured by a single accelerometer into one representative step for each of the fourteen accelerometers. The MSS feature type is composed of three steps from specific accelerometers with a known spatial relationship between the sensors and walker. The SSS feature type is a single footstep measured by a single sensor. A description of the construction of each feature set follows accompanied by the illustrations in Table 3. Table 3 constructions each feature type based on the original depiction of a walking trial in Figure 13.

Sensor Averaged Step (SAS)

The SAS feature type selects the n most prevalent footsteps, based on amplitude as recorded by each accelerometer during the entirety of a walking trial. These n steps were maximum value aligned and then averaged. Figure 19 depicts the peak aligning and averaging process. The most appropriate number of steps to average is unknown a priori, and therefore three SAS feature types are created where $n = 10, 5$, and 3 . These number of steps were chosen to understand the effect of increasing the number of steps for overall classification accuracy. Chapter 5 examines the results of the importance of the number of steps. These three feature sets are denoted SAS10, SAS5, and SAS3. Collectively these feature types are referred to as SAS.

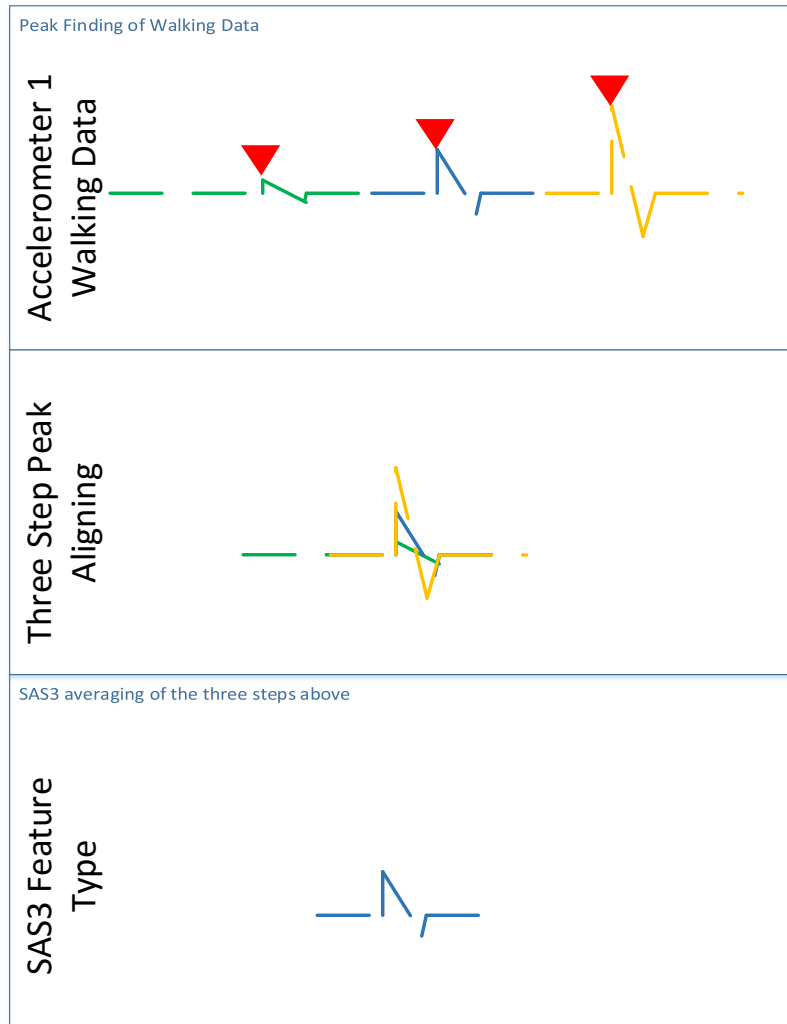


Figure 19. Representation of the peak aligning and averaging. The creation of all of the SAS feature types use this process.

There was a hypothesis that an averaged step would remove the variance in each accelerometer measurement warranting the development of the SAS feature type. By averaging multiple steps, the random error associated with each accelerometer is reduced. An averaged step is a global feature type that averages many local features.

Figure 20 shows an SAS10 footstep in comparison to the footstep at the maximum peak. It is hypothesized that the largest peak amplitude for a footstep occurs when the walker is stepping closely to the accelerometer, which is supported by Pan et al. [40]. As is seen in the figure, the largest peak footstep is reduced in magnitude by the averaging of multiple footsteps. Some of the peaks in the Maximum Peak plot are shifted or even removed by the averaging of

multiple steps leading to the hypothesis that the maximum peak footstep misses important gait information. The future work examines the assumptions made in this analysis.

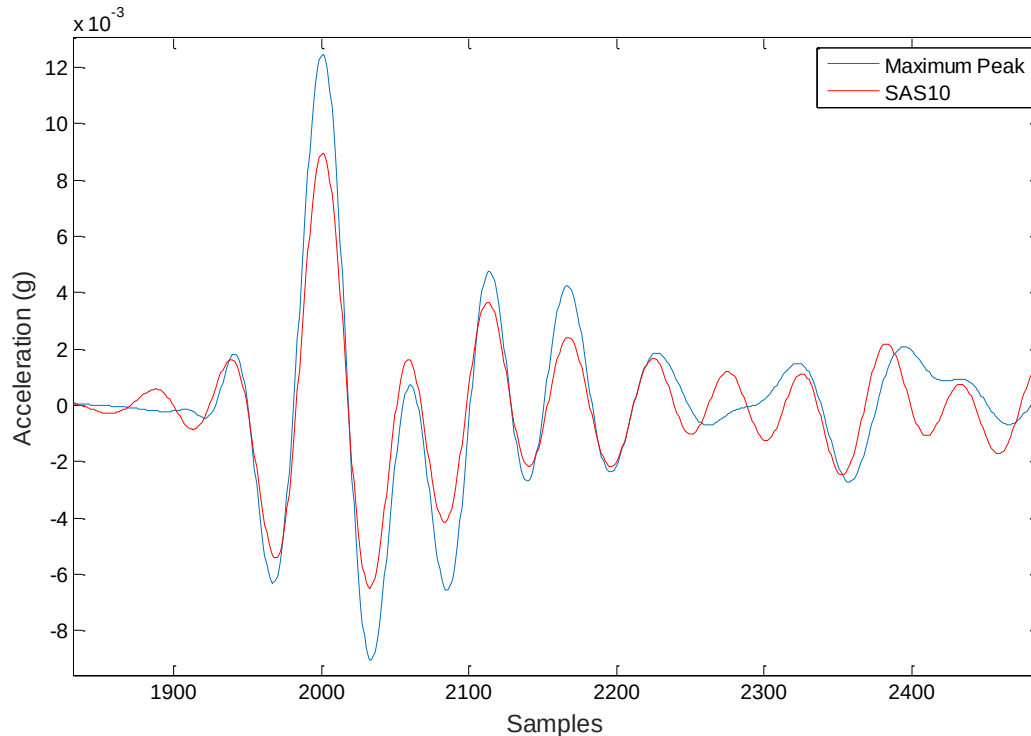


Figure 20. SAS10 compared to the maximum peak footstep. This step is taken from trial 1 sensor 1 for Female A Hard Sole.

The SAS feature type performs relatively well for classifying gender in Chapter 5. Chapter 6 includes the SAS10 and SAS5 feature types as these perform the best in Chapter 5. The choice of 3, 5, and 10 steps for averaging was to determine the effect of the number of steps averaged on the classification error. In some walking trials more than ten steps were difficult to find with the peak finder (explained below) as only up to ten steps were considered. The SAS feature type seems to be representative of the necessary gait information to classify walker characteristics as evident by their use in the 88 % and 97 % accurate gender classifiers in Chapter 5 and Chapter 6 respectively.

Multiple Sensor Step (MSS)

The MSS feature type is constructed by taking the most prevalent step, based on amplitude, recorded by accelerometer 4. Then documenting the same footstep in time as recorded by accelerometers 2 and 6. The feature type contains these three steps in chronological order of accelerometers as seen by the walker. That is, the Westward trials used

accelerometers 2, 4, 6 and Eastward 6, 4, 2 in those respective orders. Figure 21 details a representative layout of sensors 2, 4, and 6 in the testing hallway in GH. These three accelerometers were chosen for in the creation of the MSS type because these sensors were known to be in the middle of the walking trial. This arrangement of sensors is to the best of the author's knowledge and is the likely arrangement. A mislabeling of accelerometers was found after the creation of the walking database.



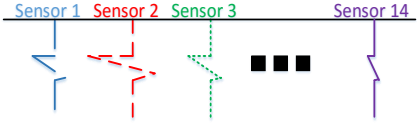
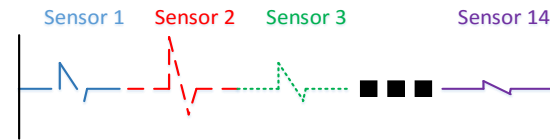
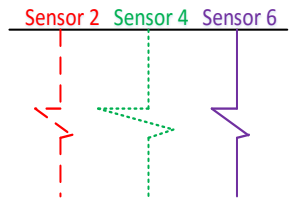
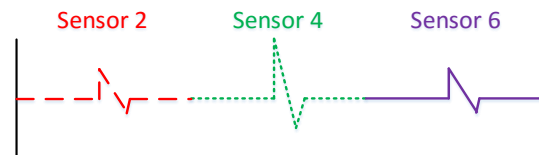
Figure 21. Illustration of subset of sensor arrangement. The spatial relation between sensor 1, 2, 4, and 6 in the testing hallway. These sensors are used in the construction of the MSS and SSS data types.

There was concern that averaging of SAS may adversely affect the classification accuracy because of the loss of spatial content warranting the development of the MSS feature type. This loss of spatial information results because the largest n steps are averaged in the SAS set; these averaged steps are not required to happen in chronological order. This lack of chronological information causes the loss of the spatial relationship between the walker and the accelerometers. These spatial relationships are retained with the MSS data as the largest peak recorded by accelerometer 4 meaning that the walker was nearly over the top of the sensor. Comparing the results of SAS and the MSS feature types allow for an understanding of the importance of spatial information to characteristic classification.

Single Sensor Step (SSS)

The SSS feature set takes the most prevalent single step recorded by accelerometer 1 as shown in Figure 21. Accelerometer 1 was chosen randomly as a control feature type with the most limited amount of gait measurement. This feature set is used as a control to determine the usefulness of a single accelerometer and single step to successfully classify walker characteristics.

Table 3. Illustration of the feature types construction. The steps required to develop the three feature types (SAS, MSS, and SSS) in the time domain are illustrated schematically for ease of explanation.

Feature Type	Step 1	Step 1 Representation	Step 2	Step 2 Representation (Final Feature)
SAS	Averaging of n steps of each sensor resulting in one average step for each accelerometer.	 Dimensionality: $\mathbb{R}^{7001 \times 14}$ 14 Accelerometers used	Concatenation of each average step.	 Dimensionality: $\mathbb{R}^{98014 \times 1}$
MSS	Selection of the most prominent step recorded by sensor 4 and documentation of the corresponding steps in sensors 2 and 6.	 Dimensionality: $\mathbb{R}^{7001 \times 3}$ 14 Accelerometers used	Concatenation of each step. The order of the concatenation is dependent on walking direction. Eastward: 2, 4, 6 Westward: 6, 4, 2	 Dimensionality: $\mathbb{R}^{21003 \times 1}$
SSS	Single most prominent step recorded from only sensor 1.	 Dimensionality: $\mathbb{R}^{7001 \times 1}$ 14 Accelerometers used		

The most prominent peaks for each feature set were computed using the MATLAB command 'findpeaks'. The settings were developed qualitatively to pick out footsteps from the database. The most important setting was the minimum number of samples between peaks. Research determined an appropriate value of 16000 samples or approximately 0.31 seconds for the minimum number of samples between peaks. This number of samples between footsteps is reasonable because it is smaller than the average time between footsteps. The average walking cadence is 115 steps per minute or one step (left or right step) every 0.51 seconds [82].

A single footstep in all feature sets is defined as the 61.0 *ms* (2000 samples) before a footstep peak and 152.6 *ms* (5000 samples) after the peak. So the total number of samples in a single footstep is 7001 samples of gait measurements, where the extra 1 sample is counting the most prominent peak itself. The time range was chosen to qualitatively encapsulate a full footstep. The larger dynamic response after the peak informs the biasing of the number of samples before and after the peak of a footstep (2000 compared to 5000). Figure 12 shows an example footstep in the inlaid plot.

The values of 16000 samples and a footstep range of 7001 samples proved plausible as they produce promising results for the feature developed in Chapter 5 and Chapter 6. The future work section discusses studying the effects of selecting these two numbers as a different peak finder and fewer samples defining a footstep may achieve similar results. Appendix A presents example MATLAB coding used to find the footsteps in this thesis work.

Development of the SAS, MSS, and SSS feature types took place in the time domain. A Fast Fourier Transform (FFT) transforms these feature types into the frequency domain. Specifically, the frequency domain is quantified by a two-sided power spectrum. The transformation allows for the examination of the importance of feature type domain on the classification accuracy. Chapter 5 shows the time domain features more useful for characteristic classification, for these included features sets only.

In summary, there are ten sets of features. The SAS-10, SAS-5, SAS-3, MSS, and SSS in the time domain and their frequency domain counterpart (denoted Freq). Comparison of these feature types in Chapter 5 allows for understanding the assumptions made in the creation of these specific features. The next section explains the instance

selection methods used and their importance in this work for removing from the analysis the accelerometers not useful for improved classification accuracy.

Chapter 4.2. Instance Selection

As opposed to feature selection, which reduces the number of features input to the machine learning algorithms, instance selection limits the number of observations and entire portions of feature sets. While this difference is nuanced, it is important to make this distinction as the former creates useful features and the latter limits from which data the features are created. This thesis examines multiple instance reduction method (ISM) types.

The dimensionality of a single walking trial is $\mathbb{R}^{m \times n}$, where m is based on the sampling rate and n is the number of accelerometers used (14). Sample rate selection methods reduce the m dimension by choosing which points within a feature are necessary for a high classification accuracy. Accelerometer selection methods reduce the n dimension by choosing the ‘best’ accelerometers. A successful instance selection method will decrease the computational load by reducing the number of points in each of the previously discussed features while retaining or enhancing the classification accuracy.

As previously discussed, increasing the number of observations provided to a machine learning algorithm greatly increases the computational cost and can reduce overall accuracy [79, 80]. The downfalls to having a large number of observations clearly motivate using data: from accelerometers that are most useful and at a minimum sampling rate producing an acceptable classification accuracy. Reducing the sampling rate by a factor of two will reduce the number of points in a feature by 50 % for example. Removing a single accelerometer from consideration will reduce the number of features by 7 % (if the feature examines every accelerometer used in recording the walking trials). The sample rate selection section details downsampling the original data to reduce the m dimension. The accelerometer selection section discusses the approaches to reducing the n dimension by choosing the best accelerometers and truncating measurements from insignificant sensors.

Chapter 4.2.1. Sample Rate Selection – Instance Selection Method (SRS-ISM)

Downsampling the walking trials measured by the fourteen accelerometers is a method reducing unnecessary features. Downsampling is the exclusion of specific data points to produce artificially a lower sampling rate, illustrated in Figure 22. For instance, reducing the sampling rate by a factor of two discards every other point from the original signal. The original sampling rate used in the experiment was 51200 Hz. This high sampling rate facilitated the downsampling study which potentially reduces a large number of features. Chapter 6 studies the effect of reducing these features on classification accuracy.

This Sample Rate Selection ISM (SRS-ISM) has a huge potential to reduce the number of features used in the machine learning algorithms. For example, a single footstep has 7001 values and reducing by a factor of 2 results in 3501 values which represent the ‘same’ information. SRS-ISMs are an effective method of reducing the number of samples used to represent time-series information. The goal of these SRS-ISMs is therefore to find the minimum sampling rate that produces acceptable classification accuracy. In this work, the minimum sampling rate is determined by downsampling the recorded walking trials and determining how effective each downsampled signal is at producing a high classification accuracy.

Figure 22 shows the effect of downsampling on a footstep. As can be seen, at some reduced sampling rate the footstep will effectively lose its inherent patterns. Thus, a balance must be established between reducing the sampling rate for computational efficiency while still representing the gait signal. Chapter 6 investigates this balance. The next section examines the accelerometer selection - instance selection methods (AS-ISMs) for the purpose of finding which accelerometers allow for the creation of the best features.

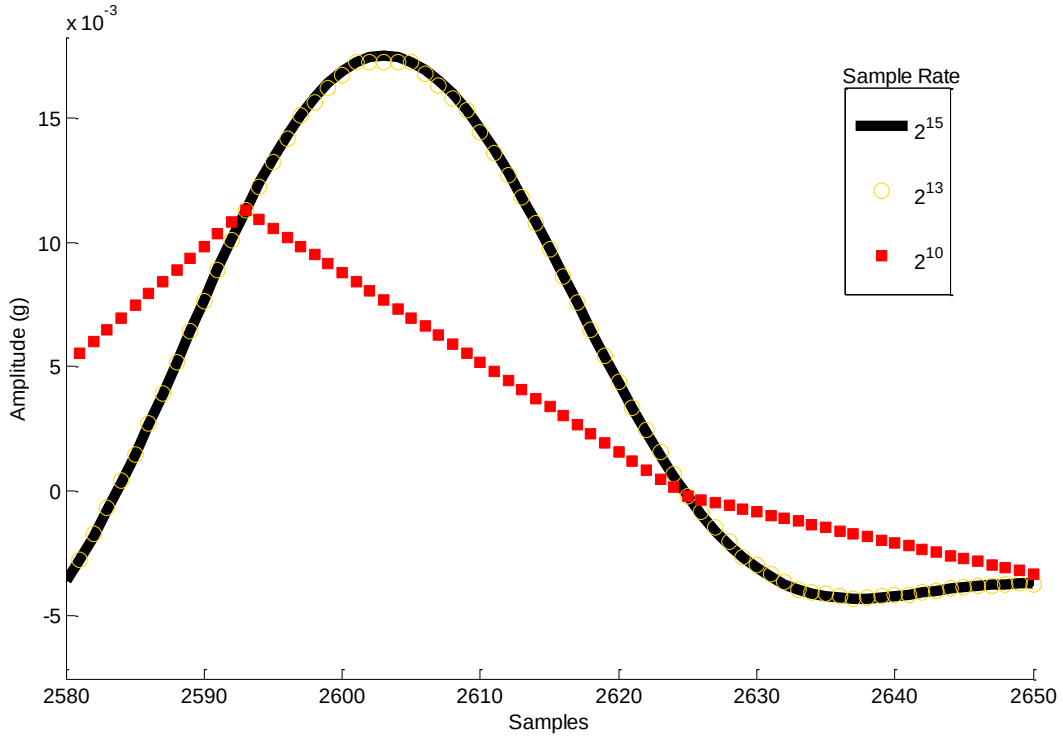


Figure 22. Effects of Downsampling. The first peak of a footstep is plotted at various sampling rates to show the effect of down sampling.

Chapter 4.2.2. Accelerometer Selection – Instance Selection Methods (AS-ISM)

The accelerometer selection - instance selection methods (AS-ISMs) investigated are Discrete Empirical Interpolation Method (DEIM), an alternate formulation of DEIM: Q-DEIM, and Projection Coefficients (PC). These AS-ISMs were chosen because they were novel approaches to reducing gait measurements with known potential for large reductions of dynamic data while maintaining acceptable accuracy. Each AS-ISM outputs a ranking of accelerometers from independent to redundant using the gait measurements of each walking trial as input. The produced ranking of accelerometers can inform which sensors do not improve the predictions of the developed characteristic classifier.

It is difficult to know a priori that either independent or redundant accelerometers will produce the best classifier accuracy. The importance of independence vs. redundancy is studied further in Chapter 6. Figure 23 illustrates the relationship between independence and redundancy for each ranking output by the various AS-ISMs.

Accelerometer Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Accelerometer Name	12	7	9	4	1	5	9	8	6	3	2	10	11	14

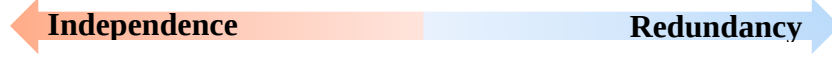


Figure 23. Rank independence relationship. Illustration between ranking and independence of accelerometer from the entire walking trial measurements. These accelerometers are used for example only.

The formulations of DEIM and PC require a Singular Value Decomposition (SVD), and Q-DEIM requires a QR transformation. The reduction background section introduces the SVD and QR transformation. Then the specific ISMs are described individually. Lastly, the η_p approximation metric is discussed for quantifying how close the ranking output by each ISM reduction is compared to optimum.

Reduction Background

The examination of the included ISMs first requires a background in a few mathematical concepts. Specifically, the Singular Value Decomposition (SVD) and QR matrix transformation. These two matrix transformations enable the formulation of the ISMs.

Singular Value Decomposition and Singular Value Decay

The SVD is the transformation of a matrix, $A \in \mathbb{R}^{m \times n}$ into the form below as shown in Golub and van Loan [83].

$$A = U\Sigma V^T \quad \text{Equation 4}$$

“If A is real, then there exists orthogonal matrices

$$U = [u_1, \dots, u_m] \in \mathbb{R}^{m \times m} \text{ and } V = [v_1, \dots, v_n] \in \mathbb{R}^{n \times n}$$

such that

$$U^T A V = \text{diag}(\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0)$$

where $p = \min(m, n)$.”

As discussed in Chaturantabut and Sorensen [84], the SVD “extracts the basis elements that contain characteristics of the space of the expected solutions.” In reducing the SVD basis dimension from r , where r is the rank of matrix A , to k (where $k < r$), a set of orthonormal vectors whose linear span that best approximates (in $\|\cdot\|_2$) the space of the matrix is sought [83]. This set of orthonormal vectors is provided by the left eigenvectors computed by the SVD of the matrix, A .

In this thesis work the rank, r , is the total number of accelerometers and k is the number of sensors desired in an optimum subset of sensors. The appropriate k is initially examined by the decay of singular values, Σ , and then through a sweep of the number of sensors as a function of classification accuracy. A singular value decay analysis can be completed after creation of a walking database to estimate reducibility of the system. This estimation can take place before the more time consuming proposed methodology in Chapter 6. While the full methodology of Chapter 6 allows for calculation of the minimum number of useful accelerometers.

The singular value decay is the decreasing magnitude of the singular values, Σ , from Equation 4. The rate at which these values decays gives insight into how much a give data set can be reduced. Figure 24 shows the normalized average singular value decay for 120 walking trials. If the singular values decayed rapidly then a good approximation of the original matrix, A , can be had with only a few of the vectors of the V matrix [85]. If only a few vectors of V are required, then only the first few accelerometers ranked by the ISMs would be required for accurate characteristic classification. These few ‘important’ sensors are expected to accurately represent of the full walking trial, A (the fourteen accelerometers recording for length of time required to walk the length of the hallway), and will producing the same classification accuracy as the full data set (if not better). Unfortunately, Figure 24 does not show a ‘rapid decay’ in singular values meaning that initial estimates of k might be near the full 14 sensors. A large k does not represent a large reduction in the walking data.

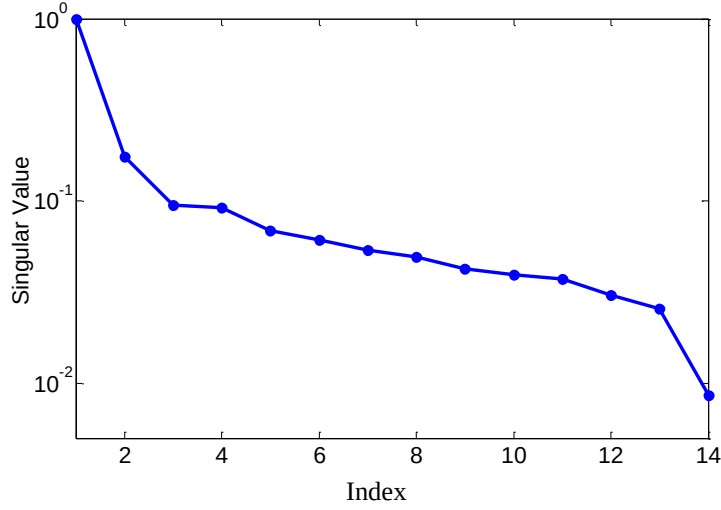


Figure 24. Average normalized singular value decay of all walking trial.

It is hypothesized that the classification accuracy has a correlation with the singular value decay. Meaning that the classification accuracy is expected to follow a similar trend to the singular values decay for an increase in the number of accelerometers. This hypothesis informs the development of the Singular Value Decomposition – Count Matrix rank aggregation method in Chapter 4.3.1. Chapter 6 examines the relationship between the classification rate and singular value decay.

The MATLAB command ‘svd’ is used to compute all SVDs in this work. The large dimensionality of the walking database requires the use of the economy version of the command. This economy specification causes only the first n columns of U to be calculated. A sample of the coding used to compute SVDs is included in Appendix A. The QR transformation enables to formulation of the Q-DEIM AS-ISM.

QR Matrix Transformation

The QR transformation further explained in Wilson and Martinez [86], is the factorization of a matrix, A , as:

$$A = QR \quad \text{Equation 5}$$

Where Q is an orthogonal matrix and R is an upper triangular matrix. The MATLAB command ‘qr’ is used to compute the QR transformation of A , where A is a walking trial (i.e. 14 accelerometers recording the time taken to traverse the hallway). A sample of the code used to calculate the QR transformation is shown in Appendix A.

With the mathematical background established for the SVD and QR transform the description and discussion of the ISMs can follow.

Chapter 4.2.2.1. Discrete Empirical Interpolation Method (DEIM) for Data Reduction

Chaturantabut and Sorensen [84] proposed the Discrete Empirical Interpolation Method (DEIM) for use in the reduction of dynamical systems; the work includes results showing little error between the full and reduced systems with multiple examples. Sorensen and Embree [87] is a previous work that uses a database search similar to the use of the algorithm in this work to find representative web pages from two geographic areas. This work uses an SVD of the gait measurements as opposed to the CUR factorization in Sorensen and Embree.

“DEIM is an efficient method to project high dimensional non-linear functions into a lower dimensional space for the purpose of data reduction” [84]. The DEIM algorithm essentially minimizes the error of the approximation, or projection onto a lower dimensional space iteratively. The largest singular value determines the first index in the algorithm. The remaining entries with the largest residual error serve as the next index. The residual is effectively the error between the input basis and its approximation due to interpolating the basis. This iterative process repeats until the number of singular values is reached. The DEIM methodology allows for the determination of the most independent accelerometer for every trial of the experiment. The output from this method is 120 rankings of the accelerometers from most independent gait measurements provided during a walking trial to most redundant.

Chapter 4.2.2.2. Q-DEIM for Data Reduction

The Q-DEIM method, which is an alternate formulation of the DEIM, is developed using a pivoted QR matrix factorization for the purpose of large-scale systems. The Q-DEIM method offers an improved theoretical bound on the error of the projection onto a lower dimensional space than DEIM [88]. Drmac and Gugercin [88] provide previous applications of this methodology. They reduce various dynamic systems to compare to the DEIM method. One advantage that is offered by Q-DEIM is the possibility of using a stochastic method in the above work. The stochastic methods may be necessary as the database grows as this method does not require the examination of every walking trial. Q-DEIM outputs a ranking of the independence of the gait measurement by each sensor just like DEIM.

Chapter 4.2.2.3. Projection Coefficients (PC) for Data Reduction

The Projection Coefficient (PC) method is often used in the dynamics modeling field to determine the degree in which experimental and analytical mode shapes match. Perinpanayagam and Ewins, and Chen [89, 90] show examples of using PCs.

The PC is calculated as the cosine of the angle between two vectors, in this case, a single left eigenvector from the SVD and a single original walking trial measured by each accelerometer as:

$$\text{Projection Coefficient} = PC \equiv \cos(\theta) = \frac{a \cdot b}{|a||b|} \quad \text{Equation 6}$$

Where θ is the angle between the two vectors, a is the left eigenvector, and b is the original measurement of each accelerometer. The coefficient varies between 1 and 0, where 1 represents a high degree of similarity between the two vectors and 0 represents orthogonality between them. Understanding how the first left eigenvector relates to the original database allows for an investigation of which parameters are most independent.

Figure 25 shows a typical projection coefficient matrix for a single walking trial. The measurements from each accelerometer (1-14), for one walking trial, are compared to each of the same walking trial's left eigenvectors. This process repeats for all 120 walking trials. In the figure, a colored spot indicates the two vectors (accelerometer measurement and left eigenvector) were similar to one another.

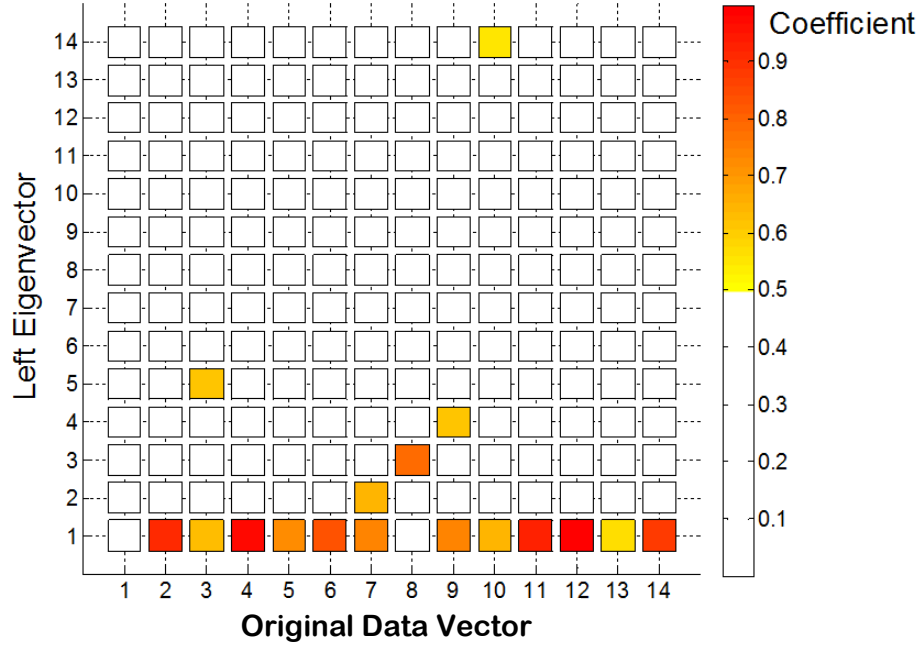


Figure 25. Typical Projection Coefficients. The resulting projection coefficient matrix for a typical walking trial. The values below 0.5 (low similarity) are neglected and plotted as white. The magnitude in the first left eigenvector row is used to create the ranking.

Selection of the first eigenvector is justified by the previously mentioned singular value decay, as on average the first singular value accounts for 63 % of the summation of all singular values. In other words, the first eigenvector is chosen as it is the single eigenvector that is most representative of the original database. The PC method creates, just like DEIM and Q-DEIM, a ranking of the most independent accelerometers (PC value closest to 1) to most redundant (PC value closest to 0).

In summary, three AS-ISMs are included that output a ranking of accelerometers for all 120 walking trials similar to that depicted in Figure 23. Three methods are included in this work as the rankings output by each method differ, and there is no simple method of determining which method will produce the best rankings. Understanding which types of AS-ISMs are most appropriate for selection of accelerometers in gait measure is one aim of this entire thesis. The next section introduces the η_p approximation metric which allows for comparison to an optimum ranking.

Chapter 4.2.2.4. η_p Approximation Metric

An estimation of the error between the optimum reduction and the output of the ISMs is the η_p approximation metric. This metric is later used to compare rankings of accelerometers. The error metric was defined in [87, 88, 91] and is defined as:

$$\eta_p = \|(P^T V)^{-1}\| \quad \text{Equation 7}$$

Where $V \in \mathbb{R}^{m \times k}$ contains the k leading left singular vectors of a matrix A ($A \in \mathbb{R}^{m \times n}$). $P = I(:, p)$ where I is the identity matrix and p = ranking of sensors as output by any of the AS-ISMs. This error metric gives a proportional constant with a minimum value of 1 representing the effectiveness of the approximation of the original matrix, A . An $\eta_p = 1$ represents exact approximation of the original data using an arbitrary reduction method. This value can be used to determine the ‘goodness’ of the approximation.

This metric is used in Chapter 6 to compare the best five accelerometers ($k = 5$, p = the five highest ranked sensors) as determined by the ISMs to all possible rankings of five accelerometers. This optimum determination is only completed using five sensors because completing this analysis on a larger number of accelerometers becomes intractable quickly. That is, using five sensors requires 2002 matrix η_p calculations, while using all fourteen sensors would require more than 87 billion.

In summary, the η_p approximation metric allows for the examination of how well an AS-ISM’s outputted ranking compares to an optimum ranking. The use of the ranking of the five best sensors enables the calculation of all possible rankings (2002) to determine how effective the AS-ISM performs compared to optimum.

The accelerometer and sample rate selection methods could be used in series or independently to reduce the amount of data used in the creation of a feature. Chapter 6 uses the AS-ISMs to investigate the minimum number of accelerometers necessary. That is, which of the fourteen accelerometers used provided the most insignificant data for classification. Then, Chapter 6 uses the SRS-ISMs (after the AS-ISMs) to determine the minimum sampling rate for effective classification.

Each AS-ISM produces a ranking of accelerometers for each walking trial (120 in total). Thus, each AS-ISM produces 120 rankings of 14 accelerometers. The next section focusses on combining the 120 rankings from respective AS-ISMs into a single ‘best’ ranking. This single best ranking represents the sensors that are most often representative of the walking trials. Chapter 6 compares multiple rank aggregation methods for which method produces the best overall ranking.

Chapter 4.3. Rank Aggregation

Combining multiple rankings into a single ‘best’ ranking is known as rank aggregation. There are many methods of completing such a combination, and the method producing the best ranking is application specific. This section introduces some of the current methods of rank aggregation and proposes the SVD-Count Ranking (SVD-CR) method applied in Chapter 6.

Chapter 4.3.1. Current Methods

Dwork et al. [92] serve as an introduction and general survey paper for many rank aggregation methods. The rank aggregation methods combining the rankings of the web pages output by multiple search engines into one most appropriate ranking. They show that aggregation allows for robustness to spam and reduces the bias of each search engine. These aggregations ultimately provide a better web page ranking than possible from a single search engine.

Lin [93] provides another survey introduction to rank aggregation methods. Lin work points out some of the advantages of using rank aggregation methods: they are robust to outliers and invariant to transformation. Lin work broadly divides algorithms into three categories: heuristic, stochastic, and distributed. Heuristic methods are approaches designed to be computationally efficient and do not optimize any criterion. Stochastic methods are optimization methods that do not circumvent the need to calculate all possible rankings. Distributed calculate all rankings using some form of distance metric. Lin discusses the result of various heuristic and stochastic algorithms applied to a provided long list. A long list is a few rankings containing many ranked entries (e.g., four people’s rankings of 150 ice cream flavors). The distribution methods are not used on the long lists as they are suited for short lists. A short list is a rank aggregation of many rankings of only a few ranked entries (e.g. 150 people’s rankings of 4 ice cream flavors).

Two distance metrics, Spearman footrule and Kendall tau, are discussed in the above works [92, 93]. These metrics quantify the errors between a computed best ranking and the aggregated rankings. The Spearman distance is the sum of the absolute difference between the ranks presented in two lists (i.e. best list and the aggregated list). The Kendall distance examines the number of “pairwise adjacent transpositions required to transform from one list to another” [92]. A low Spearman and Kendall distance represent a ranking which is representative of the many rankings, but as is shown in Lin [93], these two metrics are not guaranteed to agree on the best overall ranking.

This work proposes a rank aggregation method: Singular Value Decomposition – Count Ranking. Chapter 6 examines multiple rank aggregation methods to compare the effectiveness of the proposed method. The rank aggregation is completed using the RankAggreg analysis package in R [94]. Appendix A houses the code for used in R for the RankAggreg package. The included aggregation methods are Cross-Entropy Monte Carlo (CE) and Genetic Algorithm (GA), with either Spearman or Kendall distance optimization. The next paragraph details the CE and GA methods.

The CE algorithm works in two phases: generate random rankings and update parameters of the ranking randomizer to build ‘better’ overall rankings. By optimizing either the Spearman or Kendall distances, and iterating until convergence, the better ranking is achieved. The GA takes inspiration from genetics allowing for randomized cross-over of rankings during iterations and mutation to determine the overall best list. Again, the Spearman and Kendall distances are optimized in the creation of the algorithm [94]. The combinations of these algorithms and distance metrics result in four methods: CE Spearman, CE Kendall, GA Spearman, and GA Kendall. These algorithms allow for the comparison of the effectiveness of the proposed SVD-CR method to popular rank aggregation methods. The formulation of the proposed SVD-CR method is below.

Chapter 4.3.1. Singular Value Decomposition – Count Rank Method

The formulation of the Singular Value Decomposition – Count Ranking (SVD-CR) requires the number of times an accelerometer appears in a specific rank and the average singular value decay. The example used to describe the formulation of this method is the 120 rankings output by any AS-ISM. The formulation steps are:

- A count matrix is developed by arranging data so that the rows represent the rank, and the column represents the specific accelerometer. The entry of the matrix is the count of times that accelerometer (column) appears in that rank (row).
- A scaled count matrix is created by scaling the appropriate rank (row) by the corresponding singular value. For example, the first row of the count matrix is scaled by one, due to singular values normalization. The second singular value scales every entry in the second row, and so on.
- Summing the entries along specific columns allows for the creation of a summed vector. The summed vector is ordered from highest to lowest value. The column (accelerometer) corresponding to the highest sum is the first ranked accelerometer with this ranking continuing for all remaining sensors.

This SVD-CR method aggregates the 120 rankings from each walking trial into a single ranking representative of every walking trial. This rank aggregation method allows for a determination of which accelerometers are likely most important in the test hallway. Chapter 6 discusses the specific results of the SVD-CR method in comparison to other popular rank aggregation methods. Appendix B houses the coding and illustration of the step by step development of this rank aggregation method.

Similarity to Current Methods

The SVD-CR method is similar to a Borda method discussed in Lin and Dwork et al. [92, 93]. The Borda method linearly scales the importance of each rank. That is, the first ranked item receives the largest scaling value as it is the most important. A linear decrement, based on the number of items in the list, scales the remaining rankings. Then a summed vector, similar to the above, allows for the creation of the best overall ranking. The difference between SVD-CR and Borda is the scaling. SVD-CR scales based on the singular values not linearly like in Borda.

Limitations of SVD-CR

An issue with the SVD-CR method arises when the rankings provided are not of appropriate dimension. More specifically, when there are more items ranked than there are rankings to aggregate (e.g., four people's rankings of 150 ice cream flavors), as known as a short list. This issue results from the SVD. There will only be x singular values

with which to scale; where x is determined by the minimum dimension (m or n , samples in a walking trial or number of accelerometers respectively) of the matrix A in Equation 5.

For the use of SVD-CR in the setting of this thesis work, there must be more observations than the number of accelerometers to complete this analysis. This limitation means SVD-CR is not effective in providing an aggregated rank for so-called long lists but is acceptable for short lists. This limitation of the SVD-CR method precludes it from use on long list rank aggregation problems similar to the distributed methods in Lin [93].

In summary, the SVD-CR rank aggregation method is proposed, and other well-established rank aggregation methods are described. Chapter 6 examines the effectiveness of these five rank aggregation methods to determine the effectiveness of SVD-CR and what effect rank aggregation method has on the truncatable number of accelerometers.

Now that the feature selection and instance selection methods studied in this thesis have been introduced, the next section elaborates on the literature of gait manipulation literature originally discussed in Chapter 1. There are many examples of using data reduction methods on gait because the measurement techniques often produce large amounts of data. This section attempts to place the feature selection and instance selection methods in the context of other works. The comparison of feature selection methods is fairly straightforward as all methods reduce from the full data set to a reduced number of features. Direct comparison of the instance selection methods is difficult because the literature does not choose the most appropriate sensors or sampling rate for analysis. The reduction techniques of the previous work are more feature selection techniques than instance selection techniques.

Chapter 4.4. Examples of Gait Measurement Reduction in Literature

As previously stated Chau [48, 49] provides two survey papers of the many methods often used in transforming gait measurements. These two sources highlight several different methodologies such as the use of fuzzy logic clustering, fractal dynamics, principal component analysis, and wavelet analysis. These methods are discussed further after the formal introduction of the data reduction methods of feature and instance selection.

Fuzzy Logic Clustering

Fuzzy logic is a type of machine learning algorithm. It requires the assumption that membership in a class does not have to be sharp, but can be fuzzy. For example, a walker does not need to belong to a binary class of gender (male or female), as it can exist between these states. The vagueness introduced by allowing fuzzy membership facilitates the clustering of data that may otherwise be viewed as dissimilar when using rigid boundaries [48]. This clustering of data could allow for selecting the minimum number of features required to cluster “like” observations together. This choosing of the minimum number of features is feature selection for classification assuming that the clusters of data are interpretable. For example, if through experimentation, researchers define the clusters of male and female then a minimum number of features can be selected that describe that cluster. Those fewer features will reduce the computation time while retaining classification accuracy.

Fractal Dynamics

A fractal is a phenomenon that is self-similar regardless of scale [48]. Meaning a pattern exists at any observable scale of measurements. Many physiological systems exhibit this type of relationship (e.g. the human heartbeat). The ‘dynamics’ in fractal dynamics, in this case, refers to fractal characteristics exhibited during gait [95]. This methodology means that gait measurements can be reduced in size while retaining larger scale patterns. Fractals are another method of feature selection, allowing for the choice of the most appropriate feature types. In Chapter 6, the SSS feature type, representing a single footstep, is compared to the SAS type, representing multiple footsteps, allowing for examination of the fractal nature of gait.

Principal Component Analysis

Principal Component Analysis (PCA) allows for a reduction in dimensionality of data by reducing it to principal components. These principal components capture the dimensions incorporating the largest amounts of variance [48]. There is potential for labeling the principal components with physical meaning, but this analysis is often difficult and beyond the scope of this thesis [48].

PCA is a common method of data reduction in many fields. It allows for retained variance (i.e. similarity to the original data) and data reduction by truncating the less useful principal components. PCA could be directly used for feature selection as discussed in Chau [48], or it could be used in instance selection methods similar to those presented. This

instance selection is apparent because of the similarity of PCA to the SVD. The difference between PCA and SVD is through the computation of the covariance matrix in PCA while SVD does not take this step. PCA is probably the most applicable literature example that could be used to reduce the required number of sensors.

Wavelet Analysis

The wavelet analysis allows for the examination of the frequency domain of a signal similar to a Fourier transform. However, the wavelet transforms goes further than the Fourier transform, allowing for the study of the signal in time and frequency domains simultaneously [48]. There has been work to show that wavelet transformations are an acceptable method of feature selection reduction methods for classifying footsteps [96]. Wavelet coefficients could be used as an effective feature selection method for building feature sets that produce accurate classification models. These wavelet coefficients combine the time and frequency information potentially performing better than the feature types proposed in this work as they examine either time or frequency.

In summary of the above, the presented previous methods have the ability to be used to reduce the gait measurements while ideally increasing classification accuracy. With further development, researchers could use these reduction methods in either feature or instance selection techniques with further development. The methods presented here are chosen because they have been found to be some of the most well-established methods.

In summary to Chapter 4, these previous reduction methods provide data reduction possibilities but have not been used in a methodology like that proposed in Chapter 6, which confirms algorithm design choices and determines which sensors provide the most interesting data. The use of these SRS-ISM and AS-ISMs is believed to be novel for instance selection of gait analysis. The next chapter discusses the classification of a walker's gender using the machine learning algorithms of Chapter 3 and the features created in Chapter 4.

CHAPTER 5. GENDER CLASSIFICATION OF WALKERS VIA UNDERFLOOR ACCELEROMETER MEASUREMENT

The motivation for this chapter is to demonstrate the usefulness of machine learning techniques to classify the walker characteristic of gender. Gender is an example of a soft biometric. Soft biometrics are characteristics that are not unique to an individual but have been shown to aid in the re-identification of an individual [12]. Identifying the gender of walkers in a building could benefit multiple applications, such as security and improving retail sales. The manner in which a person walks, gait, is measured here via underfloor mounted accelerometers. These accelerometers mount to the structural members of Virginia Tech's Goodwin Hall (GH) and enable the measurement of the vibrational response of the floor due to a walker's gait. These gait measurements serve as the input features to machine learning algorithms that predict the gender of new walking trials.

Chapters 1-4 discuss the technical background and previous work for this chapter, allowing this chapter to focus on the results and discussion. This introduction examines the methodology, and the questions specifically addressed in this chapter. The results and discussion are presented then conclusions are offered. The results and discussion of this chapter are the basis for Bales et al. [3].

Methodology

The methodology for the analysis in this chapter has been discussed in detail in the previous chapters, but is summarized here:

- Gait is measured using underfloor accelerometers,
- Digital signal processing prepares the gait measurement,
- Peak finding algorithms select features from the gait measurements,
- Machine learning algorithms use the features to predict gender (via training, validation, and testing stages),
and
- Statistical analysis determines the best performing data treatments.

A data treatment is defined as any operation completed to the processed data. The types of data treatments in this chapter are feature type, walking direction, feature domain (time or frequency), and machine learning algorithm. The

idea of this data treatment term is presented as in the statistical analysis when a best performer in a data treatment type is tested all other data treatment types are considered. That is, when determining the best machine learning algorithm data treatment, the statistical test uses the error for all other types of data treatments (feature type, walking direction, and feature domain).

Each step of the bulleted methodology is discussed below briefly with a reference to the detailed description of the writing in the thesis. Figure 26 illustrates the above methodology bullets.

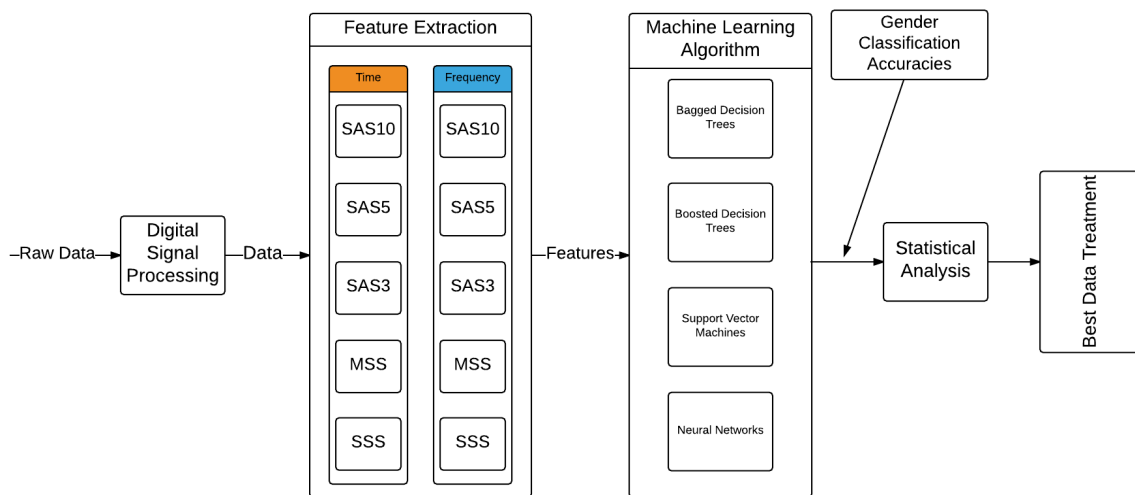


Figure 26. Chapter 5 Methodology Overview. The process required to go from the raw data recorded by the underfloor accelerometers to gender classification (and therefore gender classification accuracy) is depicted here.

A data acquisition system recorded the gait of 20 individuals as they, individually, walked down a 98-foot-long hallway in Virginia Tech's Goodwin Hall (GH). Fourteen underfloor mounted accelerometers recorded the floor acceleration caused by a walker's footsteps. The experimental setup is discussed completely in Chapter 2. Chapter 2 also discusses the instrumentation and digital signal processing methods completed on the raw data to form the walking database. Algorithms (peak finder, and peak alignment, etc.,) extract features from the walking database.

The feature selection methods discussed in Chapter 4 extract five feature types in the time domain: SAS10, SAS5, SAS3, MSS, and SSS. The SAS feature type is the Sensor Averaged Step, which takes either the 10, 5, or 3 most prevalent steps from every sensor respectively, peak aligns those steps and averages them. The MSS feature type is

the Multiple Sensor Step, where a single footstep measured by three specific location accelerometers are concatenated. The SSS feature type is the Single Sensor Step, which takes the most prominent step from a single sensor. These time domain features are then transformed into the frequency domain using a Fast Fourier Transform (FFT). These feature types represent a large data reduction from the original data size (approximately 0.1% of the original data size). These feature types serve as the inputs to the machine learning algorithms.

The machine learning algorithms use the developed features to tune model parameters to classify the gender for never before seen or new walking trials. This chapter addresses the use of bagged and boosted decision trees, support vector machines (SVMs), and neural networks as machine learning algorithms. Chapter 3 discusses the specifics of each algorithm and also provides a further reference for each. A ten-fold-cross-validation method is used to reduce the variance associated with the randomization of the features for the training, validation, and test stages of the supervised learning. Each combination of the four machine learning algorithms and ten feature type produces a single value of gender classification accuracy. These accuracies must be compared in a statistical manner to evaluate the best performing in each category.

The method of statistical analysis used in this chapter, and in Chapter 6, is an analysis of variance (ANOVA) and then a Tukey multiple comparison test. The ANOVA identifies if a difference in gender classification accuracy exists as a result of using different data treatment types. Then, the Tukey test allows for the determination of which data treatment performs better. During these analyzes, one data treatment is tested at a time, while considering all other data treatments in the analysis. The ANOVA and Tukey tests use a 95 % confidence interval for determining statistical significance. Appendix C houses further detail on these analysis methods. These methods allow for the determination of which data treatment is the best performing. The best performing data treatments have the statistically lowest classification error in the developed machine learning model.

Questions Addressed in this Chapter

This chapter examines broad questions to understand the possibility of completing characteristic classification using the methods discussed above. The questions are presented here and discussed in the results below. The addressing of these questions will lead future research into other characteristic classifiers (e.g., height, shoe size, etc.,) with a

preliminary understanding of which features to create and which machine learning algorithms to start with. The results and discussion follow the questions.

- (1) Which of the proposed feature sets (SAS10, SAS5, SAS3, MSS, SSS in both time and frequency domains) are most useful for gender classification?
- (2) Which of the proposed machine learning algorithms (bagged and boosted decision trees, SVMs, and Neural Networks) perform best?
- (3) Does the time or frequency domain house the most readily useful information for gender classification based on the methods used in this study?

Chapter 5.1. Results and Discussion

Each of the four machine learning algorithms use the ten previously discussed feature types (SAS10, SAS5, SAS3, MSS, and SSS in both time and frequency domains), respectively, as inputs. The overall range of accuracies for all combinations of feature and algorithm type was 42 – 95 %.

Figure 27 shows the classification error results for all combinations of feature types and machine learning algorithms in this study. In the figure, the two broad columns represent the time and frequency domains, and the further subdivided columns represent the feature types. The dependent y-axis shows the error for each algorithm – feature type combination, averaged over both walking directions (Eastward and Westward). The averaging over both directions is the reason the 95% accurate model listed in the above paragraph is not shown directly in the figure.

Figure 27 shows the frequency domain produces larger classification error than the time domain. Secondly, the neural networks produce a larger classification error than the other three types of machine learning algorithms. Further trends

are difficult to interpret directly from the figure motivating the use of the statistical analysis described above and in Appendix D.

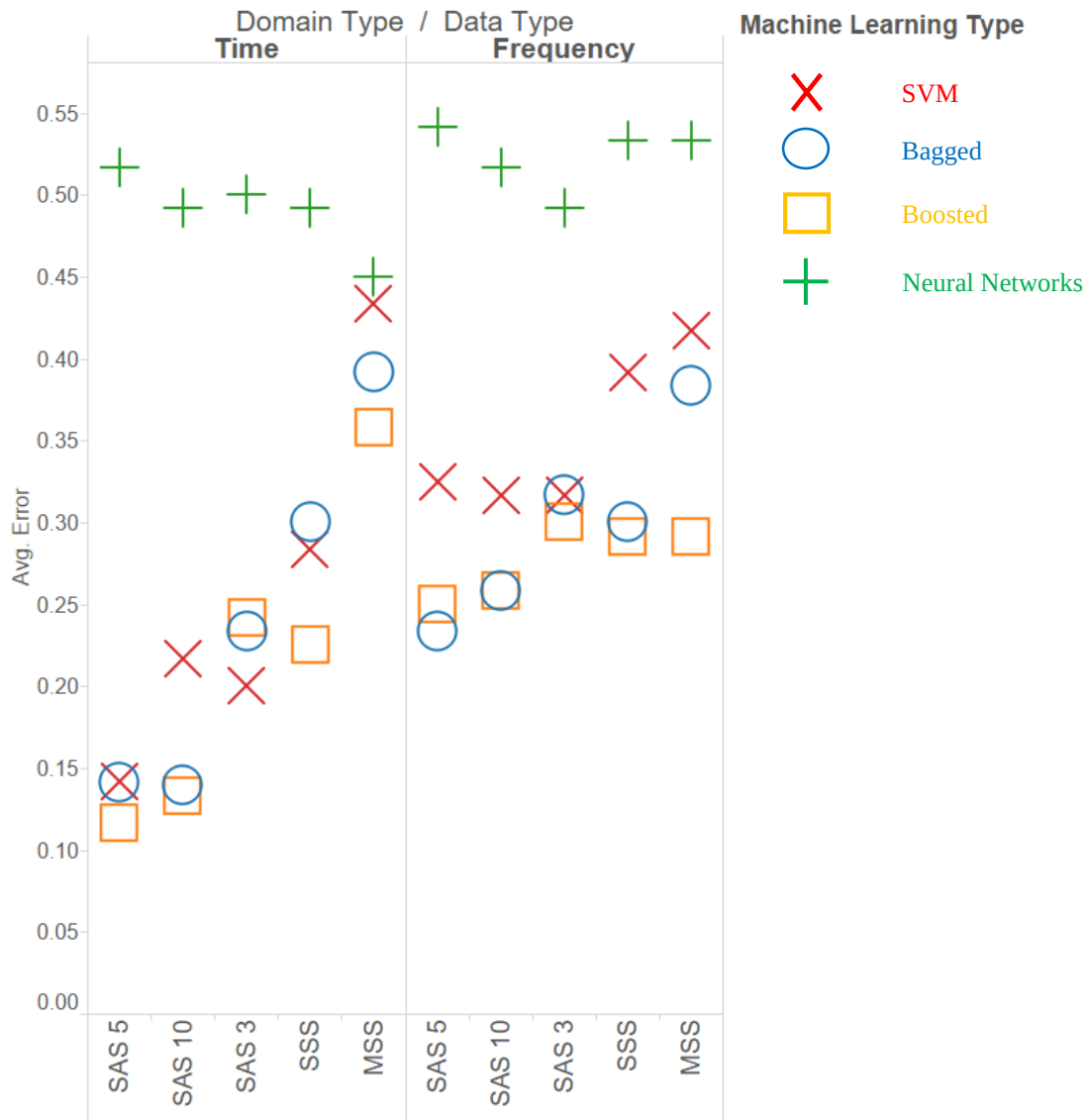


Figure 27. Representation of the gender classification error. The x-axis represents the five feature types in both the time and frequency domain (shown as columns). The y-axis shows the average classification error. This error is averaged for over both directions of walking (Eastward and Westward). The symbols represent the machine learning algorithm type.

Comparison of Time Domain vs. Frequency Domain Representation

As shown in Figure 27, the time domain data features give a higher accuracy than the features in the frequency domain. The overall accuracy of the time domain is 70 % while the frequency domain is 64 %. This result may be partly due to the relative richness of the input data in the time domain versus the frequency domain. Generally speaking,

dynamicists are looking for a way to “simplify” complex time series data to understand underlying phenomena by using the frequency domain. However, this type of transformation of the features into the frequency domain may prove adverse to the success of the machine learning algorithms.

More specifically, multiple peaks in the time domain may be represented only by a single dominant peak in the frequency domain. For example, walking at 1.2 Hz will produce many peaks in the time domain (multiple for each footstep) but will be represented by a single peak in the frequency domain at 1.2 Hz. Thus, the time domain has a larger number of potentially interesting (non-zero) features than the frequency domain. The machine learning algorithms may be missing these important peak frequencies due to the noise provided by a large number of other features.

Ekimov and Sabatier [35] support the hypothesis explaining the relatively poor results of the frequency domain features. Their work shows that 500 Hz bounds the energy in the frequency domain of a footstep for the normal force interaction between the foot and the ground. The frequency domain in this study contains higher frequencies terms than 500 Hz. Not truncating the frequency domain to below 500 Hz may result in many near zero features that only add noise to the machine learning algorithm. Based on these results, future works should investigate time domain features of gait unless there are specific features in the frequency domain of interest.

Comparison of Machine Learning Types

Bagged and boosted decision trees as well as SVMs are statistically equivalent machine learning algorithms for gender classification accuracy (see Appendix C for further discussion). While each of these three algorithm types have different average classification errors, statistically they are indistinguishable. On the other hand, neural networks perform statistically worse than the other machine learning algorithms.

The best machine learning algorithm, when examining only the average error is the bagged decision tree. Bagged decision trees produced a classification error of 21.5 % and 27.8 % across all time and frequency domain features, respectively. Bagged decision trees are suggested for gender classification because they produce the lowest average

error across all feature types, meaning they are robust to the feature type used. Based on the statistical analysis, to within 95 % confidence, both of the decision tree types and the SVMs are suitable algorithm types.

The good performance of both types of decision trees in comparison to the other machine learning algorithms is the reason they are the only algorithms considered in Chapter 6. While the SVMs perform equally as well, they require more parameter optimization and, therefore, more computational time to build the models. The Notable Observations discuss the parameter optimization for SVMs and NN below. Appendix D contains the fully tabulated results of this study. Chapter 6 will address the effects of instance selection methods on characteristic classification not to produce the best possible classifiers, so this analysis does not include SVMs.

Comparison of Feature Type

The statistical testing reveals that there is no distinguishable difference between the SAS5 and SAS10 feature types. Although, when examining only the mean error, the SAS5 feature type produces the lowest classification error of 22.9 % and 33.8 %, for time and frequency respectively. All feature types were averaged across all four machine learning algorithms. This averaging leads to the development of robust features independent of machine learning approach. Chapter 6 only considers the SAS5 and SAS10 feature types based their equivalent performance here.

The feature types of SAS3, MSS, and SSS perform statistically worse than the other two features. The notable observations section discusses the MSS, and SSS feature types. The strong performance of the SAS data type speaks to the use of an averaged footstep in the time domain as a good representation of an individual's gait.

Most Accurate Machine Learning – Feature Combination

The single best performing feature type and machine learning combination were the boosted decision trees using SAS5 time features. This model produced an 88.3 % accuracy in the classification of a walker's gender. Similar results may be possible using either the bagged decision tree or SVM as the learning algorithm in combination, and either the SAS5 or SAS10 feature type, as these all produce statistically equivalent results. The strong results of the boosted decision tree using SAS5 speaks to the potential of using underfloor mounted accelerometers and a machine learning approach to classifying the characteristics of walkers.

Notable Observations

Additional observations beyond the primary test goals were noted. The subcategories below present these results, and each is discussed to offer insight to inform future work. The notes are loosely formatted to discuss feature type results then machine learning algorithm trends.

SSS Feature Type

The investigation of the SSS feature type reveals that using data from a single footstep measured by one accelerometer provides classifiable information. The boosted decision trees produced the highest accuracy of 78 % with this SSS feature set. This result implies that a high density of the information required for gender classification lies within a single footstep. This work is exciting because it shows that in a security setting, taking one step into an instrumented building could provide information about a suspect.

If higher classification rates are desired than 78 %, multiple footsteps should be recorded as demonstrated by the improved performance of the SAS feature types. Recording more steps can increase the accuracy of the gender classification, at the cost of computational time. For instance, recording an additional four more steps (SAS5) can increase accuracy by 10 %, but in doing so, the algorithms must accommodate 14 times more data, which affects computation time. A delicate balance exists between the desired accuracy of a model and the required amount of data for a feature. This balance becomes even more important when an application needs to run in real-time, requiring minimum computation.

The relative success of the SSS data in comparison to the SAS supports the belief that gait is a fractal phenomenon. As discussed in Chapter 4, research has shown that gait exhibits fractal properties, meaning that there is an inherent pattern regardless of the observable scale [48, 95]. This result means patterns exist whether the observable scale is a single footstep or many footsteps. The SSS feature type retains a large portion of the inherent gait pattern supporting this fractal theory.

Spatial Importance in Features

Comparing the results of the SAS and MSS feature types can lead to an understanding of the importance of a spatial relationship between accelerometers and a walker. The SAS feature potentially loses spatial information because the averaged steps are chosen by the magnitude and not chronologically in order of sensor triggered by a walker. However, the MSS feature type explicitly retains the accelerometer/walker spatial information.

The MSS feature type produced the worst gender classification accuracy by a sizable margin (9 % points worse than SSS feature type in the time domain) of any of the features considered. The SAS type produces the best results: 26 % gender classification error for all SAS feature types. Even the SSS data type which has both fewer steps and no spatial relationships performs better than the MSS type. This result leads to the belief that the spatial relationship between a walker and the accelerometers does not necessarily improve characteristic classification. This result means that future constructed features do not require explicitly examining sensors location.

Neural Network Performance

The poor performance of the neural networks is not surprising due to the number of weights that must be learned for this type of algorithm. This particular form of the algorithm must learn $nd + n$ number of weights, where n is the number of neurons and d is the dimensionality of the feature set. The feature set size can be seen in Table 3 of Chapter 4, and the number of neurons is discussed below. It is hypothesized that the large number of features has led to the relatively poor results as a large number of weights must be learned. Successfully selecting features which have lower dimensionality is predicted to improve neural network performance [81]. This study is beyond the scope of this thesis work.

Machine Learning Parameter Optimization

The SVM and neural network algorithms require a free parameter optimization. The free parameters for the SVMs are the polynomial order of the kernel and for NNs are the number of neurons in the hidden layer. A parameter sweep completes this optimization. That is, sweeping through integer values to find in optimum in a specified range. The SVM polynomial order and the number of neurons in the neural network were optimized respectively. The optimum

polynomial degree for the SVMs (24) and the optimum number of neurons in the neural networks (2) were determined by the iterative search for the ranges considered here (i.e. 1st – 25th).

Figure 28 shows the results of this parameter sweep. For SVMs, the polynomial order has relatively little effect on a polynomial order of 12 but does have a minimum at 24th order. Further investigation of SVM kernel type should investigate these models as they have proven to be a useful algorithm as kernel type has been shown, in this work, to have a great effect on the classification accuracy. For NNs, the number of neurons has little effect on the accuracy of classification leaving little confidence in robust classifying with the neural networks single-layer scheme chosen for this study. This result does not eliminate the use of neural networks from consideration for characteristic classification but does point to the need to develop more complex networks to accurately predict characteristics and more tailored features for this type of learning algorithm.

While SVMs have quality results the extra computation time required to complete a parameter sweep with the increased number of data treatments in Chapter 6 causes them not to be considered in the Chapter 6 analysis. This result is not to say that SVMs are not a viable option for classifying walker characteristics. On the contrary, it is hypothesized that the improved accuracies for the decision trees using instance reduction methods in Chapter 6 would apply to the use of the SVMs as well.

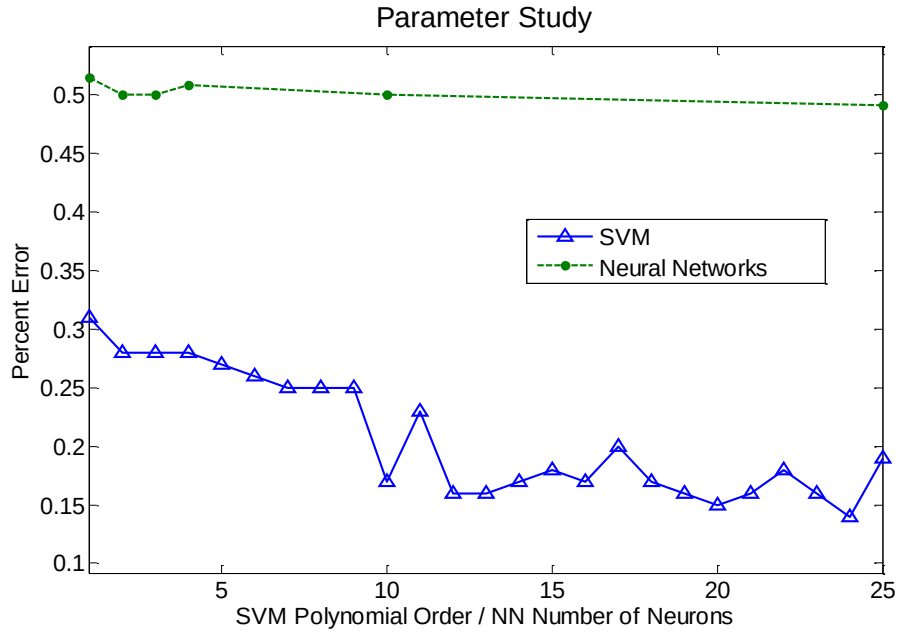


Figure 28. SVM and NN parameter study. The independent axis is either the order of the SVM polynomial or the number of neurons in the neural network each plotted against the percent error for gender classification.

Chapter 5.2. Conclusion

Using data from underfloor accelerometers as input features for machine learning techniques shows promising results for accurately classifying the gender of a walker. With 88 % accuracy achieved for SAS5 and boosted decision trees. The underfloor mounting of accelerometers potentially offers advantages over other gait measurement technologies. The sensors are out of sight and potentially retain a walker's privacy more than video or audio systems. The underfloor mounting of the sensors makes them tamper-resistant. These potential technological advantages in conjunction with the ability to accurately classify characteristics demonstrate the use of this technology and these approaches as effective in future smart building applications.

This paragraph directly addresses the questions studied in this chapter from Chapter 1.

- (1) The SAS feature types produce the most accurate classification models. The SAS5 and SAS10 features are used in the development of the most accurate models regardless of the machine learning method used. The SSS feature type shows encouraging possibilities for recording a small amount of gait and then using machine learning approaches.

- (2) The bagged and boosted decision trees as well as support vector machines produce statistically equivalent results and are robust to feature type.
- (3) The time domain provides more readily available features for successful gender classification than the frequency domain.

The single best classifier created was combining a boosted decision tree and the SAS5 time feature set to achieve an accuracy of 88 % in gender classification (averaged over both directions of walking). The best performing features, SAS5 and SAS10, and algorithms, bagged and boosted decision trees, are carried into the next chapter to examine the effect of choosing the ‘best’ accelerometers through instance selection methods on characteristic classification. The instance selection methods, discussed in Chapter 4, are studied to reduce computational requirements while retaining or improving classification performance reported in this chapter.

CHAPTER 6. DATA REDUCTION VIA SENSOR TRUNCATION

The underfloor mounted accelerometers instrumented throughout Virginia Tech's Goodwin Hall (GH) enable the measurement of a walker's gait. Algorithms create features from these gait measurements, and those features serve as inputs to machine learning algorithms. The machine learning algorithms classify the gender and weight of a walker.

The feature types used in this chapter are the best performing from previous findings in Chapter 5: SAS5 and SAS10. Likewise, bagged and boosted decision trees are the machine learning algorithms considered in this chapter. For this study, instance selection methods reduced computational requirements and improved classification accuracies when applied.

Instance selection methods (ISMs) enable selection of the appropriate amount of features to use as input to the machine learning algorithms. This chapter uses two types of ISMs: accelerometer selection and sample rate selection methods. The accelerometer selection – instance selection methods (AS-ISMs) determine which accelerometers are most useful for improving characteristic classification. Chapter 4 describes the formulation of the AS-ISMs used in this work (DEIM, Q-DEIM, PCs). Sample rate selection – instance selection methods (SRS-ISMs) study the minimum sampling rate via downsampling. Sampling rate refers to the number of data points collected in the walking experiment. Downsampling is the exclusion of specific gait measurements. Chapter 4.2.1 further describes down sampling. The types of ISMs applied in this work are believed to be novel for gait analysis. This chapter is the basis for the work completed in Bales et al. [5].

ISMs could have a great impact on potential applications requiring characteristic classification. Reducing the number of features can greatly reduce the computational expenses (i.e. data storage and computational time) while even improving accuracy [79, 80]. Providing more accurate classification results in a shorter amount of time would be imperative in a security setting, for instance. In a situation when algorithms must run in real-time, to identify a threat as soon as possible, reducing the number of features will likely be necessary. The methods in this chapter demonstrate the potential of a reduction in the number of features while improving the gender classification accuracy reported in Chapter 5.

The remainder of this introduction provides an overview of the methodology proposed, AS-ISM and SRS-ISM separately discussed. Then the project questions to be answered in this chapter are explicitly stated. Finally, the mapping for the remainder of the chapter is given.

Methodology

The methodology section further divides into accelerometer selection and sample rate selection sections. The DEIM, Q-DEIM, and PC algorithms comprise the AS-ISMs studied in this work, and Chapter 4 presents the formulation and discussion of each. The sample rate selection uses a downsampling process which reduces the total number of features examined.

Because Chapter 2 and Chapter 5 previously described the data acquisition methodology, the walking experiment is minimally summarized here. Underfloor accelerometers recorded the gait of 20 walkers. These measurements were then processed (see Chapter 3) to form a gait walking database. Future mention of “Data” refers to this walking database specifically. As the next sections detail specifically, the walking database is used to develop features which serve as inputs to machine learning algorithms. These machine learning algorithms allow for an estimation of the gender and weight classification accuracies expected if these machine learning algorithms were applied to new walker recordings.

Accelerometer Selection – Instance Selection Methods (AS-ISMs)

The AS-ISM methodology is comprised of two major steps: Algorithm Design Confirmation (ADC) and Sweeping for Truncation (ST). The ADC is similar to the completed work in Chapter 5, where features are extracted, these features serve as inputs to machine learning algorithms, and a statistical analysis determines which data treatments are the most appropriate. Where a data treatment is an operation to the walking data. This work uses the term data treatment because in the statistical analysis step a single data treatment is used to determine the best performing approaches while considering all other data treatment types. For example, when determining the best machine learning algorithm all feature types, walking directions, etc. are considered. However, the analysis in this chapter introduces more data treatments than in Chapter 5. The ST step uses the best data treatments from the ADC step to sweep through the number of accelerometers included in the analysis. The accelerometers not improving classification accuracy

should be truncated, or removed, from the analysis as they increase data storage, increase computational time, and decrease classification accuracy.

Figure 29 summarizes the general methodology for the ADC and ST. The bulleted points below overview each block in Figure 29 and latter sections elaborated on each as the discussion of each step of the methodology is detailed.

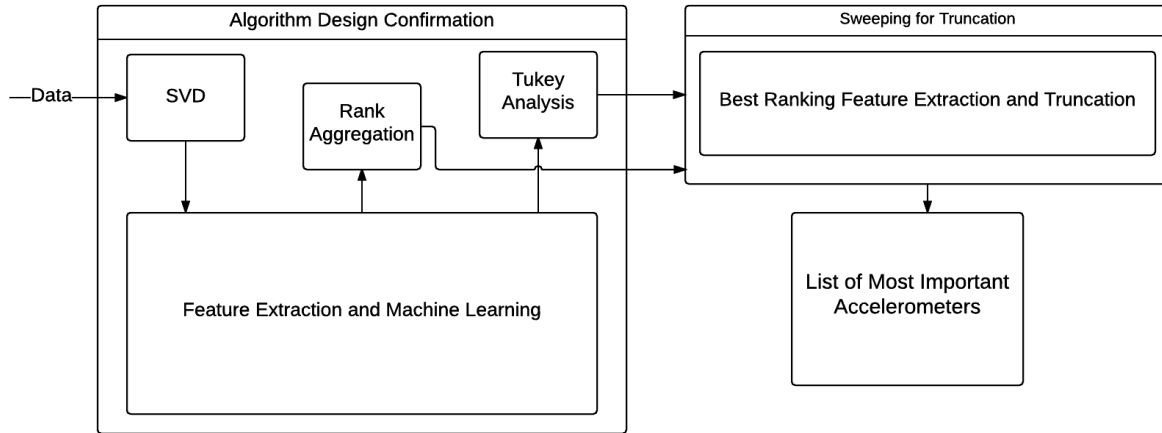


Figure 29. The generic methodology for completing ADC and ST.

- The singular value decay (SVD) [97] is completed on the walking database to inform the number of sensors to initially examine. In this work, five sensors are chosen for an initial number of sensors based on the relatively low fifth singular value on average for all walking trials. Figure 24 exhibits the average singular value decay for the walking trials in this study.

Examining an initial number of accelerometers merely reduces the computational complexity of the problem; this step could be neglected at the cost of computation time. Although as is shown in the Sweeping for Truncation section, considering all fourteen sensors does not necessarily lead to better results. In fact, in this study, considering more than ten sensors does not result in better characteristic classification. So reducing the number of sensors initially examined will reduce computation time and may improve classification error. The symbol X now represents the number of initially considered accelerometers.

- The “Feature Extraction and Machine Learning” step represent multiple components.
 - (1) The AS-ISMs examine all walking trials and rank the importance of each accelerometer. This approach results in a single ranking of all fourteen accelerometers for all 120 walking trials recorded. The full set of rankings (14×120) is one output of the “Feature Extraction and Machine Learning” block which then feeds into the “Rank Aggregation” block in Figure 29.
 - (2) The extraction of features (i.e., SAS10 and SAS5) from the full walking data. Chapter 4 discusses the feature extraction.
 - (3) The selection of the features specifically from the X best sensors as ranked by the AS-ISMs in (1).
 - (4) Use the developed features from (3), in the machine learning algorithms (i.e., bagged and boosted decision trees). The classification errors for each model are the second output of the “Feature Extraction and Machine Learning” block. These classification errors are then inputted to the “Tukey Analysis” block. Chapter 3 examines the machine learning algorithms used in this work.

- The “Rank Aggregation” block represents the combining of all of the 120 rankings of fourteen sensors. The block creates a best overall ranking for a single AS-ISM. The output of this blocks is a single ranking of the fourteen sensors used in the experiment. This single ranking is used to reduce the amount of computation completed in the Sweeping for Truncation step. All 120 rankings of sensors (for each AS-ISM) could be computed, but would *greatly* increase computation time ($\times 120$). The rank aggregation methods are discussed in Chapter 4.3.

- The “Tukey Analysis” block takes in the classification errors for each data treatment and determines the best data treatment types. Hence the name of the Algorithm Design Confirmation (ADC) step. The Tukey analysis allows for a fair comparison of the data treatments to understand, for example, which machine learning algorithm performed best. This block reduces the number of algorithm design choices needed to be made in the Sweep for Truncation step by allowing for the choice of the best performers (i.e., machine learning algorithm, feature type, etc.). The “Sweeping for Truncation” block could analyze all data treatments at the cost of computation time (dependent on the number of features, machine learning algorithms examined). Appendix C presents the Tukey analysis.

- The “Best Ranking Feature Extraction and Truncation” step uses the best performing data treatments from the ADC block in a sweep of the number of accelerometers included in the analysis from one sensor to all fourteen sensors. The “Rank Aggregation” block determines the order of subsequent sensor addition in the “Best Ranking Feature Extraction and Truncation.” The “Sweeping for Truncation” block outputs a list of accelerometers that provide the best features for effective characteristic classification. During this step, the methodology truncates sensors which do not contribute to improved characteristic classification.

In this work, the proposed methodology reduces the amount of data stored and improves the overall characteristic classification. The best combination of algorithms truncate the use of 6 accelerometers (43 % of the original walking data) and reaches 95 % accuracy in gender classification. These results show the importance of using the most important accelerometers as Chapter 5 only achieves 88 % accuracy while considering all fourteen accelerometers.

The details of the “Feature Extraction and Machine Learning” and “Best Ranking Feature Extraction and Truncation” are explained below. The discussion section which re-presents the methodology with the added detail to the methodology presented in Figure 29. The next section discusses the methodology for the sample rate selection – instance selection method.

Sampling Rate Selection – Instance Selection Method (SRS-ISM)

The original sampling rate used in the development of the walking data was 51200 Hz. The relatively high sampling rate allowed for the study of the minimum necessary sampling rate. Minimizing the sampling rate reduces the number of features, in turn reducing computation time, while maintaining or improving the classification accuracy. The SRS-ISM used in this work is downsampling. Chapter 4.2.1 discusses downsampling.

The SRS-ISM functions similarly to the “Feature Extraction and Machine Learning” block as shown in Figure 29 with the addition of downsampling by various factors. The factors tested in this work were 1 (Full data set), 2, 4, 6, 8, 16, 32, 64, 128, and 256. These factors allow for the study of sampling rates from 51200 to 200 Hz. Chapter 6.1.2 contains the details of methodology.

Questions Addressed in this Chapter

The questions answered in this chapter broadly relate to determining the best ISMs, the best overall data treatments, and the overall effectiveness of the ISMs on classification accuracy. More specifically the questions answered in this chapter are:

- (A) What is the most appropriate type of accelerometer selection – instance selection method (AS-ISM)?
- (B) What is the effect of the various data treatment types (e.g. feature type, walking direction, machine learning algorithm type, etc.)?
- (C) What is the minimum number of accelerometers required for the highest accuracy characteristic classification?
- (D) How effective is the proposed Singular Value Decomposition – Count Rank, rank aggregation method in comparison to other well-established rank aggregation methods?
- (E) What effect does rank aggregation method have on the number of truncatable accelerometers in the Sweeping for Truncation (ST) step?
- (F) What relationship exists between sampling rate and classification accuracy?

The tools and arguments necessary to answer these questions are discussed in the results and discussion sections and largely serve as the contribution made in this chapter. The conclusion of this chapter explicitly answers the above questions.

Chapter 6.1. Results and Discussion

This section presents results while providing further detail of the methodologies discussed above. The full methodology described in Figure 29 is re-presented at the end of the discussion section and shown in Figure 41.

Chapter 6.1.1 explains the full ADC and ST methodologies, in addition to the rank aggregation discussion. The Algorithm Design Confirmation section (Chapter 6.1.1.1) addresses questions (A) and (B). The rank aggregation section (Chapter 6.1.1.2) addresses question (C). The Sweeping for Truncation section (Chapter 6.1.1.3) examines questions (D) and (E).

Chapter 6.1.2 examines the relationship between the sampling rate and characteristic classification. A methodology for examining the sampling reduction rate is similar to the steps in the Sweeping for Truncation algorithm. This section examines question (F).

Chapter 6.1.1. Accelerometer Selection – Instance Reduction Methods (AS-ISMs)

A single walking trial is of the dimension $\mathbb{R}^{m \times n}$. The accelerometer selection – instance selection methods (AS-ISMs) work to reduce the number of accelerometers considered, n . This section discusses the heart of proposed methodology for reducing the number of instance while improving classification accuracy: Algorithm Design Confirmation (ADC) and Sweeping for Truncation (ST). The ADC step contains all of the components discussed in the “Feature Extraction and Machine Learning” block in Figure 29. The ST step contains the components of the “Best Ranking Feature Extraction and Truncation” block in Figure 29. Secondary to the ADC and ST steps is the rank aggregation section which compiles a best overall ranking of sensors to reduce the computation requirements in the ST step. This section discusses ADC, rank aggregation, and ST respectively. Within the ADC section, the specific methodology and results are presented.

Chapter 6.1.1.1. Algorithm Design Confirmation

The Algorithm Design Confirmation step provides a systematic approach to effectively choose appropriate data treatments for large classification accuracy. Generally speaking, the data from various feature types are input into the accelerometer selection ISMs. Then groupings of five sensors are defined which select the corresponding features for use in machine learning algorithms. The singular value decay discussion Chapter 4 in forms the use of five sensors in this chapter.

Figure 30 depicts the ADC step, and a description follows. The full walking data pass into: (a) all of the AS-ISMs (DEIM, Q-DEIM, and PC) and (b) a feature extractor.

The “AS-ISM block” outputs 360 rankings of accelerometers (120 walking trials through the three methods). A higher ranking by any of the AS-ISMs does not necessarily represent the importance of an accelerometer; these methods only rank sensors from most independent to most redundant. Knowing a priori that the most independent accelerometers

will provide the most important features for effective classification is difficult. This difficulty requires defining multiple groupings to determine if the independence of an accelerometer is important for effective classification. Understanding the results of the best grouping type can reduce computation time in future analyzes because heuristically a single grouping does produce high classification accuracy (80 % for gender).

Each ranking of sensors resulting from the AS-ISMs block in Figure 30 is further divided into three groupings. The first grouping is the five most independent sensors (labeled IND), the second is the five most redundant sensors (labeled RED), and the final group is the middle five sensors (labeled MID). The walking experiment used fourteen accelerometers to record the gait of walkers, resulting in the MID grouping containing only four sensors. The sensor ranked fifth or tenth by the accelerometer selection ISM was randomly assigned to be the fifth sensor of the MID group. In summary, the MID group contained sensors ranked 6th, 7th, 8th, 9th and either 5th or 10th. Splitting the 360 sensor rankings into the three groupings results in 1080 data treatments to this point.

In addition to the three groupings (IND, MID, and RED), a random ranking of sensors (labeled RAND) was analyzed as well. This RAND grouping allows for comment on the effectiveness of the other grouping types. If a random ranking of five sensors performs as well as any of the other groupings, then a stochastic method may prove to be just as effective at characteristic classification while reducing the computation time required to run the AS-ISMs. One-hundred RAND rankings were analyzed and are discussed below as they perform unexpectedly well.

The “Feature Extraction Algorithms” block extracts SAS10 and SAS5 features in the same manner used to create features in Chapter 5 and described in Chapter 4. The SAS10 and SAS5 are used because they are the best-performing algorithms from the analysis in Chapter 5. Both walking directions (Eastward and Westward) are included in this analysis and examined independently to determine if walking direction has an effect on characteristic classification. The output of the “Feature Extraction Algorithms” block in Figure 30 only shows as SAS10 feature type for the ease of illustration, but both feature types were analyzed. The SAS10 and SAS5 feature types house all fourteen accelerometers at the output of this block.

The feature types and the sensor groupings meet in the “Feature Creation” block. Here, features to be used in the machine learning algorithms are created by selecting the features from the “Feature Extraction Algorithms” block from the appropriate grouping for the “Group Split” block. The output of this block is the true features used in this study and are selected based on the specific accelerometer ranking. Accounting for the 1080 data treatments from the “Grouping Split” block with two feature types and two walking directions results in 4320 sets of features to be used in the machine learning algorithms.

The “Machine Learning Algorithm” block uses each of these features output by the “Feature Creation” block as inputs to bagged and boosted decision trees. Bagged and boosted decision trees were chosen in this study because of their good classification performance in Chapter 5 (22 % and 24 % error for gender classification of time domain features). Using two methods of machine learning provides 8640 classification errors for gender and weight.

The 8640 classification errors allow for the determination of the most appropriate data treatment types similar to the method used in Chapter 5. The Tukey test discussed in Appendix C makes this statistical determination possible. This analysis allows for the examination of the best data treatments (e.g., feature type, machine learning algorithm type, grouping type, etc.,). Understanding the best data treatments confirms the algorithm design choices which are made a priori (e.g., which machine learning algorithms to use). Then using these best performers reduces the computation time required to determine which accelerometers are truncatable from future analysis.

The ADC method can be adapted to include more feature types or machine learning algorithms to determine which perform best for a given data set and characteristic to be classified. The purpose of completing this analysis is to determine some of the best performing methods for future analysis. That is, the Sweeping for Truncation (ST) step could have analyzed both bagged and boosted decision trees, but this would double the computation time. While ADC is not necessary for ST, it improves computation time and in this work produced better results than those presented in Chapter 5 for any combination of data treatments. Following the methodology figure is the results of using this described method.

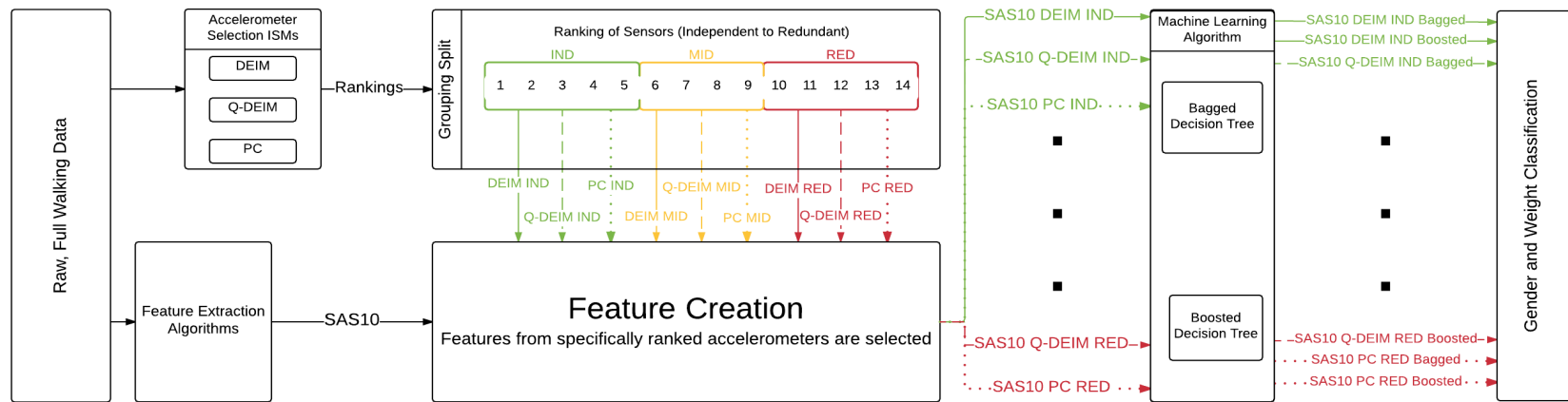


Figure 30. Algorithm Design Confirmation Methodology. The containers show the processes that act on the data and the arrows represent the flow of different data sets. While only the SAS10 feature type is shown in this figure both the SAS10 and SAS5 features types were considered.

Figure 31 shows the average gender and weight classification error for each of the data treatments. All data treatments are represented in this figure except AS-ISM. Figure 31 only examines the DEIM AS-ISM as this is the ‘best’ AS-ISM studied (discussion below) and the full figure showing all of the AS-ISMs is large and difficult to understanding. The discussion from Figure 31 applies to the full figure as well. The full figure and the accompanying tabulated data is found in Appendix E. As Figure 31 shows there is a much higher classification error for weight than there is gender. This result is somewhat expected as choosing gender correctly by guessing is correct 50 % of occurrences, while guessing the weight bin (20 lbs. intervals of body weight) correct is only 17 % probable. Also, concluded are the similarity between in results between the walking direction. It seems as though bagged decision trees perform better from the figure. The figure also shows that SAS10 probably performs better than the SAS5 feature type. While this figure allows for a general understanding of trends, more specific results require statistical analysis to truly understand if SAS10 is better than SAS5. The statistical analysis was used in Chapter 5 and discussed in Appendix C.

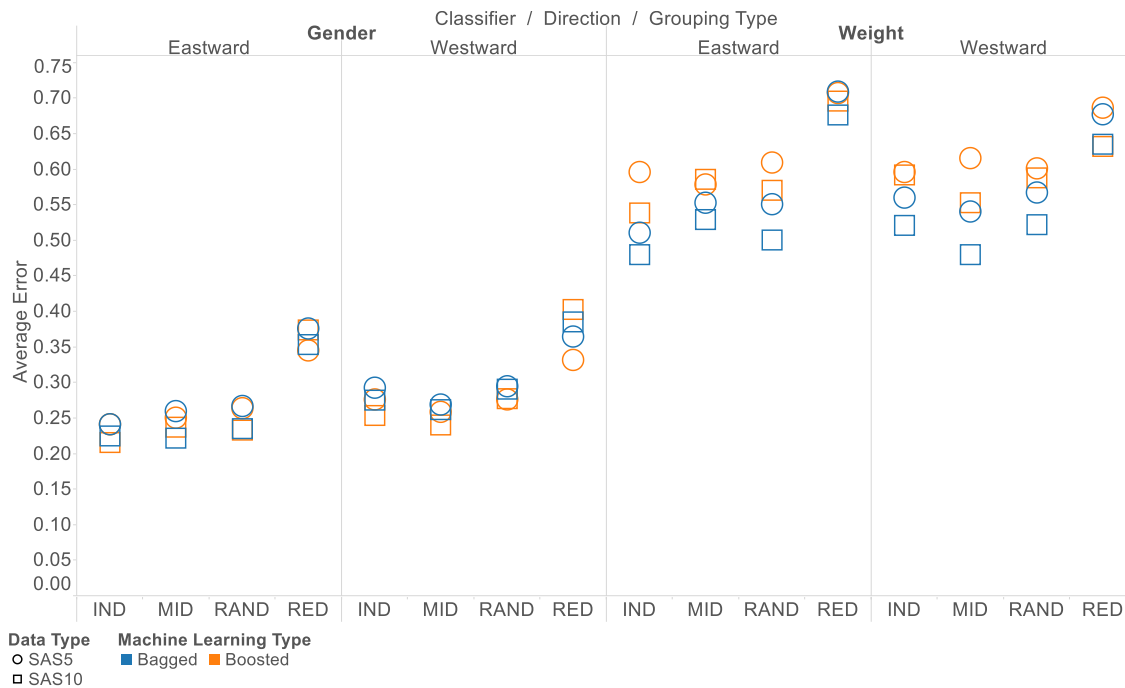


Figure 31. Statistical analysis results. Each row represents a specific data treatment and the x-axis represents the average classification error. Each data treatment is labeled with a marker for mean and 95 % confidence interval. Type I data treatments are statistically different than Type II. For statistical difference the 95% intervals must not be overlapping.

Figure 32 represents the results of this statistical analysis. The rows of the figure show the various data treatment types and the x-axis represents the classification error. The vertical placement within a row has no physical meaning. The figure labels each data treatment and shows an accompanying mean and 95 % confidence interval (plotted as a symbol and line). The mean marker and line type show the statistical difference between the types. That is, the data treatments labeled Type I by the legend are statistically different from those labeled Type II. On the other hand, all data treatments that are Type I or Type II are statistically indistinguishable from one another.

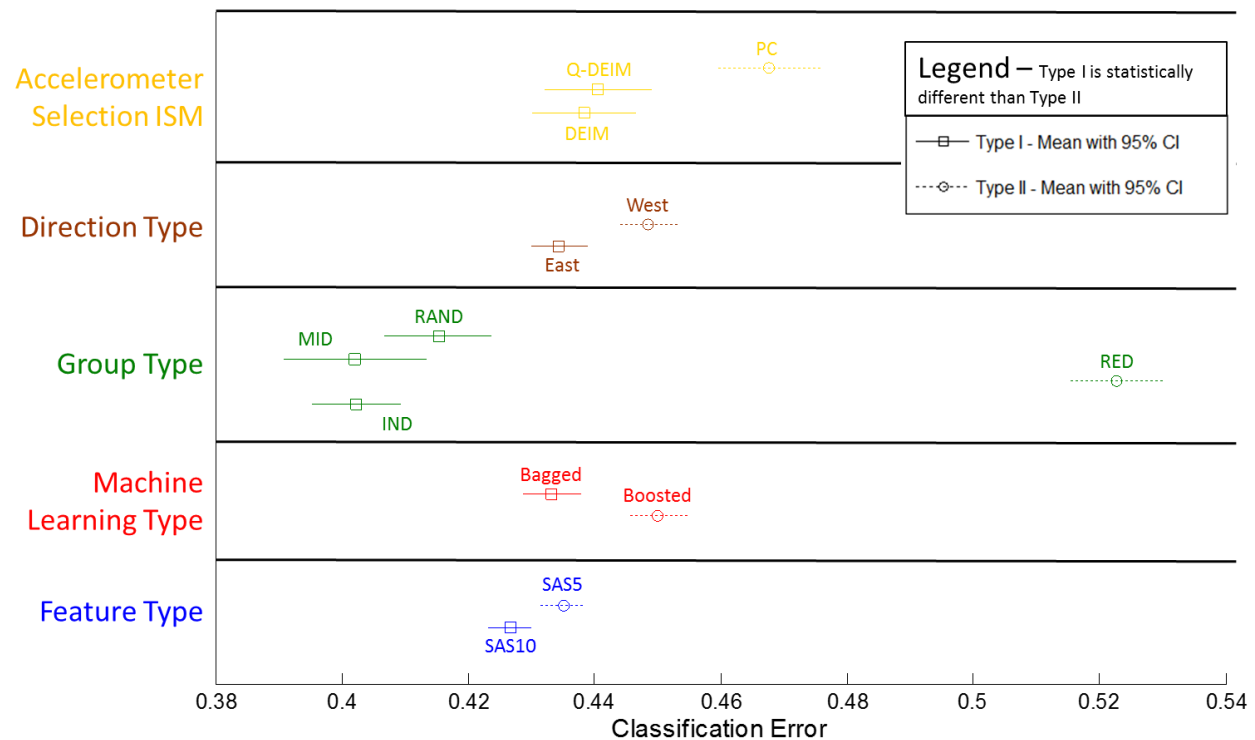


Figure 32. General Gender and Weight Results. The gender and weight classifiers are shown in separate columns and are further broken down into respective walking directions. The x-axis is the various grouping types, and the y-axis is the average classification error. The legend details the machine learning type with color and the data type with the shape.

The results shown in Figure 32 are averaged across all other data treatment types respectively. That is, in testing for the best machine learning type, for instance, all classification errors for gender, weight, accelerometer selection ISM, direction type, group type, and feature type are considered. This type of comparison is the crucial assessment made in the ADC methodology. This step allows for the determination of best machine learning algorithm type without computation of multiple trials of the same data treatment. That is, with this methodology the best machine learning algorithm can be determined while varying other variables without rerunning only the difference of machine learning multiple times with all other variables constant. This secondary method, holding all other variables constant, would

require much further computation than the proposed methodology which can vary multiple data treatment types all at once to find the best data treatment in a specific category (i.e. machine learning type).

Because the methodology considers all other treatment types the classification errors listed in Figure 32 should be carefully examined. The figure is truly trying to illustrate that, for instance, the RED grouping type will perform statically worse than IND, MID, or RAND; not that RED will produce on average $\sim 52\%$ error. To understand the full potential of the RED grouping type only the best feature type, machine learning type, etc. should be considered. The assumption made in this work, and shown in Figure 32, is that if the IND, MID, RAND, and RED grouping sets were each combined with the best remaining data treatment types, RED would still perform worse on average than the others. For this reason, in the discussion of the best data treatments in each category is qualitative rather than quantitative because the quantifications would be misleading unless all of the data treatment categories were specified.

Characteristic Classified

The gender classification model produces a better classification rate than that of the weight classification model, as shown in Figure 32. It is important to note that randomly choosing gender correctly is 1 in 2 (50 %) and 1 in 6 for weight (17 %), so this direct comparison may be unfair. The comparison does show an interesting result: the relatively poor performance of the weight classifier leads to the belief that the gender classifier does not classify based on the relationship between gender and weight. Statistically speaking males weigh more than females [98], so it is logical to think that the gender classifier might effectively be a weight classifier. The strong performance of gender classification means that the classification models are finding separate patterns for gender and weight, respectively.

This work uses weight bins to formulate walker weight estimation as a classification problem. The bins range from 90 to 210 lbs. in 20 lbs. increments. Each walker belongs to one of the six weight bins. There is a possibility that the weight bins themselves are masking the effectiveness of a potential higher accuracy weight classifier. The classifier outputs a discrete variable for the predicted weight bin. These discrete weight bins penalize the machine learning algorithm if the weight bin is not labeled correctly, which can be problematic. For example, if an 110 lbs. person (bin 2) is labeled bin 1 the learning algorithm adjusts to try to learn better the bin boundaries, but if this person was at the high end of the bin (e.g., 109 lbs.), then the algorithm was performing well all along.

To test this hypothesis, a confusion matrix for the weight classifier was created, shown in Figure 33. A confusion matrix examines which bins were correctly and incorrectly classified. A perfect classifier would result in a confusion matrix that is a height of 1 along the diagonal. As seen, the first four weight bins are often mislabeled as one another (there are many off-diagonal values which are not zero). Weight bin 4 is correctly classified most often while weight bin 5 is frequently labeled bin 4. This result suggests there were not enough bin 5 walkers entries in this study, as the learning algorithms were not capable of determining the appropriate boundaries. Future work examines further the results of the weight classifier.

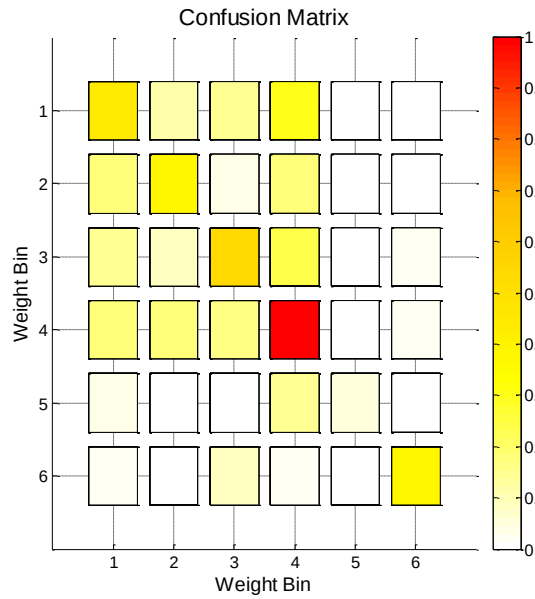


Figure 33. Six weight bin normalized confusion matrix. A large value represents classification in that coordinate. The confusion matrix for a 100 % accurate classifier is the identity matrix.

Accelerometer Selection Instance Selection Method

The most effective ISM was both DEIM and Q-DEIM, as these methods were statistically indistinguishable. DEIM has a slightly lower average classification error than Q-DEIM, but their 95 % confidence intervals overlap making them equivalent in performance. Either method could be used to effectively rank sensors. Future work might examine the use of the Q-DEIM method because of the stochastic formulation discussed in Dramc and Gugercin [88]. The stochastic methods may be more computationally effective than DEIM. See Dramc and Gugercin [88] for a discussion on the computational reduction, essentially having to do with the number of rows in a matrix that requires calculation. The PCs produced statistically inferior classification accuracies and are not suggested as a future AS-ISM.

Direction Type

The walking direction (Eastward or Westward) does have an effect on the ability to classify gender and weight. Eastward trials produce better classification results. A potential explanation for this result is that every walker started the walking experiment completing a Westward trial. The bias associated with the experiment could explain the difference in classification accuracy. Because it is each walker's first trial, it is hypothesized that participants walk in a biased fashion. This bias is due to each walker's inexperience in walking 'naturally' while being recorded.

This bias could make the Westward trials more difficult to classify because instead of three trials in this direction, there is one trial of heightened awareness. Another possible explanation for this result includes the non-symmetry of the sensors placement along the hallway. Visual cues seen by the walkers could also influence walking results. In the Westward trials, walkers moved toward an open space (as seen in Figure 8), while in Eastward trials walkers moved toward a wall. The effect of walking direction is not studied further in this work.

Grouping Type

The IND and MID accelerometers sets are the most successful grouping for characteristic classification. Choosing the most redundant sensors (RED) will result in ineffective classification. This result shows the importance of a DEIM like ranking strategy because independence results in lower classification error.

The relatively low classification error recorded in 100 RAND rankings is interesting but potentially misleading. The result means that stochastic methods may produce quality classification results if the random ranking contains sensors in the IND and MID groupings. On the other hand, the RAND grouping could select sensors included in the RED ranking resulting in poor classification performance. The RAND solution space holds the ranking of IND, MID, and RED and, therefore, could produce a wide array of classification error rates. Using the RAND method for selecting sensors requires 'enough' averaging so the effect of poorly performing sets of five sensors can be reduced. Future work should investigate stochastic sensor selection methods. The IND or MID sensor sets allow for consistent performance in classification without excess averaging.

To provide insight into how effective the IND and RED groupings are, and possibly why the RAND set performs well, the η_p approximation metric introduced in Chapter 4 can be used. When examining the best possible ranking of five sensors (from all fourteen) the computation of η_p becomes tractable. There are 2002 permutations of 14 sensors when 5 are chosen (nchoosek where $n = 14, k = 5$). ‘nchoosek’ is a MATLAB function used to compute this value. All 2002 possible rankings allows for an understanding of the ‘best’ possible ranking. Figure 34 shows the η_p metric for all 2002 rankings of five sensors possible.

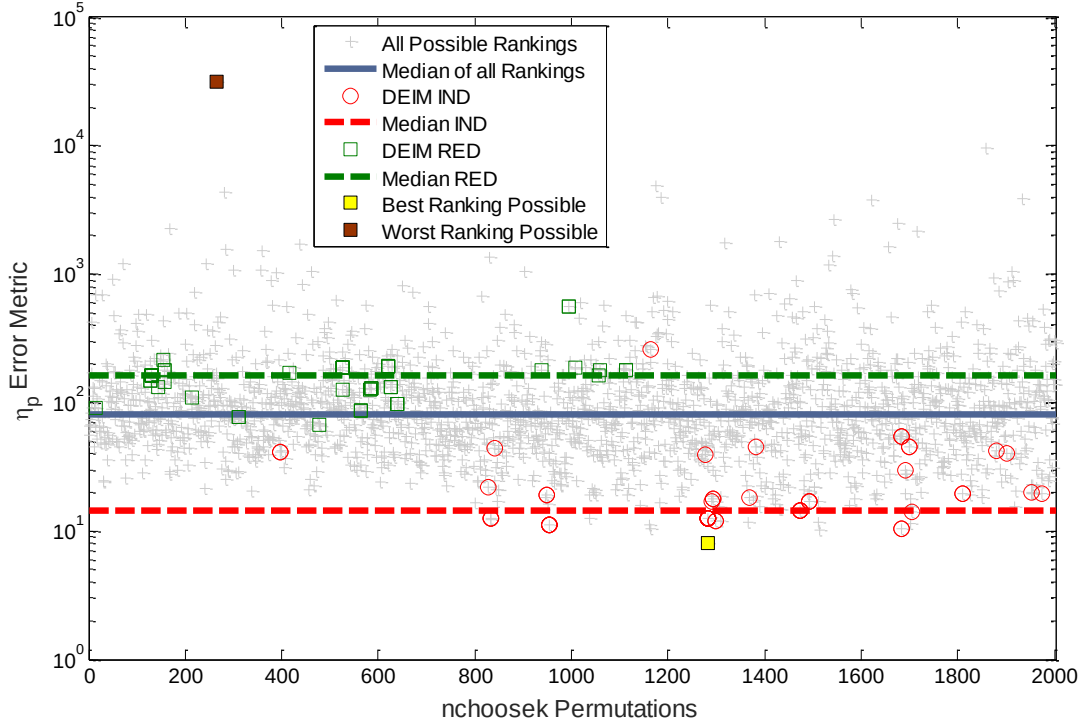


Figure 34. η_p error for all possible rankings of five sensors. The median of all rankings as well as for the IND and RED grouping sets are shown. Note the log-y scale.

Figure 34 shows that the IND grouping chooses very well performing rankings on this walking database at choosing the best rankings. Also shown, is the RED rankings which perform worse than the median of all possible rankings. This result confirms the result from the hypothesis testing; IND is more effective than RED groupings for choosing a useful sensor for classification. One further note on the grouping type, the relatively tight grouping for most of the possible permutations might be the reason that the RAND grouping performs well. RAND methods for selection sensors may have adverse effects if the database has a different distribution of the η_p approximation metric.

When completing this ADC methodology on a larger problem (i.e., more accelerometers), selecting the IND grouping type is an effective heuristic method for choosing near optimal rankings of reduced numbers of sensors. Choosing the IND grouping would effectively eliminate the need for the grouping split in the ADC analysis, but blindly choosing the IND grouping type may not lead to accurate results. The IND grouping houses the most independent accelerometers as determined by DEIM, but because the signal is the most ‘different’ than the others does not mean it represents the recorded phenomenon (gait). Care must be taken to ensure that the most independent sensors are independent not because they inherently measure something different than the intended gait of a walker. The possibility of removing the first ranked sensors because they did not record gait is examined further in the rank aggregation methods of Chapter 6.1.1.3.

Machine Learning Type

The bagged decision trees outperform the boosted decision trees. The analysis in Chapter 5 is unable to determine if bagged or boosted decision trees are most effective for gender classification. This work includes more data treatment types and, therefore, more averaged classification errors allowing for the determination of the statistical difference between the two learning method types. Bagged decision trees are suggested for characteristic classification of walker’s when using gait measurements based on their higher classification accuracies than the boosted decision trees.

Feature Type

The SAS10 data type outperforms SAS5. This result is different than presented in Chapter 5. Chapter 5 showed that SAS10 and SAS5 features were statistically indistinguishable. There are more data treatments (i.e., more blocks in the methodology in this chapter than there were in Chapter 5) producing more numerous classification errors. The number of samples included in the Tukey analysis has a large effect on the results. The more classification errors in this chapter allows for the Tukey test to reveal the statistical difference between the two features types even though this difference was not seen in Chapter 5. SAS10 is suggested for future gait classification.

In summary, ADC allows for the reduction in the parameters that need to be considered in the second half of the methodology: Sweeping for Truncation. The ADC analysis allows for the creation of a best overall performing classifier. The best classifier uses the best performing combinations of data treatments from the ADC analysis above.

The reduction of computation time required to investigate further the questions posed in the introduction exemplifies the benefits of completing the ADC analysis. The best classifier used in the remainder of this chapter is DEIM, Eastward walking, bagged decision trees, and SAS10 features. Before the ST analysis, the best overall ranking of sensors must be established to reduce further computation time, as 120 rankings of accelerometers exist from the output of the DEIM AS-ISM.

Chapter 6.1.1.2. Rank Aggregation Comparison of SVD-CR Method

There are 60 rankings of sensors, one for each Eastward walking trial, output by the DEIM ISM. Combining these 60 rankings into one overall ranking is known as rank aggregation, as discussed in Chapter 4. The SVD-Count Ranking (SVD-CR) allows for an objective scaling of the importance of a specific rank and ultimately combines the 60 DEIM rankings into a single best rank.

Chapter 4 discusses the SVD-CR rank aggregation method and the four ranking aggregation methods used to evaluate the effectiveness of the SVD-CR. The four rank aggregation methods used in the comparison are CE Spearman, CE Kendall, GA Spearman, and GA Kendall.

In this chapter, the rank aggregation methods combine the 60 Eastward trials' rankings to produce a best overall ranking. Each aggregation method produces a best overall ranking of the fourteen accelerometers. To evaluate the effectiveness of the overall rankings the rank aggregation distance metrics and η_p approximation metric for the first five ranked sensors are compared, shown in Table 4.

Table 4 shows the distance metrics for the Eastward and Westward trials. The Spearman and Kendall distances are similar for both walking directions, although the aggregate rankings were developed only using the 60 rankings from the Eastward trials. This result means that the rank aggregation methods are robust to walking direction. Further research should study walker direction effect on classification error as this result contradicts that results of Figure 31. The rank aggregation metrics show that a ranking of sensors for one direction is good for the other walking direction as well while the analysis above has shown that walking direction does produce statistically different results.

Table 4. The error metrics comparing the SVD-CR rank aggregation method to other popular techniques. Minimums for each column are bolded.

Aggregation Method	η_p Error Metric	% Error from Optimum η_p	Median Distance from 60 Eastward Rankings			
			Eastward Trials		Westward Trials	
			Spearman	Kendall	Spearman	Kendall
SVD-CR DEIM	12.5	41.72	0.24	0.18	0.27	0.21
CE Spearman	12.5	41.72	0.23	0.19	0.23	0.21
CE Kendall	12.5	41.72	0.23	0.19	0.24	0.20
GA Spearman	45.52	416.10	-0.06	-0.05	-0.10	-0.10
GA Kendall	14.75	67.23	0.23	0.18	0.28	0.26

It is important to note the GA Spearman aggregation because it produced the lowest (best) Spearman and Kendall distances. Meaning that this algorithm may be effective for choosing the optimum ranking. This aggregation method also produces the largest η_p approximation error meaning that its top five sensors are different from the optimum calculated through the nchoosek brute force approach in Chapter 4.2.2.4.

To further examine this trend,

Table 5 shows the top eight ranked sensors for all methods and the η_p optimum five sensors. The GA Spearman is the only method to house the five η_p optimum accelerometers in the top eight sensors. If sensor 12 were discarded, then the GA Spearman method has the correct η_p optimum sensors. This result means that poor performance in η_p error may not mean the entire ranking will result in poor classification ability. The success of the GA Spearman algorithm may indicate that the Spearman and Kendall distances are more representative of the best overall ranking than η_p . Chapter 6.1.1.3 examines the success of the GA Spearman algorithm used instead of SVD-CR in the Sweeping for Truncation methodology.

Table 5. The best eight sensors ranked by the various ranking aggregation methods. The optimal η_p ranking (not ordered) is also presented in a shifted format to show the alignment between the η_p optimum ranking and the GA Spearman algorithm.

Method	Top 8 Sensors							
SVD-CR DEIM RANK	12	7	4	9	3	8	2	11
CE Spearman	12	7	9	4	3	8	13	11
CE Kendall	12	7	9	3	4	8	13	11
GA Spearman	12	7	11	9	4	3	13	8
GA Kendall	12	7	3	9	8	4	13	11
Optimal η_p Ranking (Order is insignificant)	-	3	4	7	9	11	-	-

In summary, the proposed SVD-CR method performs well in Spearman, Kendall, and η_p approximation metrics when compared to other popular ranking aggregation methods. It is difficult to make any further conclusions about the proposed method's effectiveness other than it seems acceptable for continued use. The remainder of this chapter heavily uses the results of the SVD-CR analysis. Chapter 6.1.1.3 directly compares the SVD-CR ranking to the GA Spearman to determine the effect on rank aggregation method in the Sweeping for Truncation step.

Chapter 6.1.1.3. Sweeping for Truncation

The Sweeping for Truncation (ST) step provides a methodology to reduce the number of sensors considered in the analysis of the walking database developed in this work. This methodology could be used to reduce the number of accelerometers used in future studies as well by truncation of sensors which do not improve classification accuracy. Accelerometers not aiding in effective classification should be truncated from the analysis as they only increase the computational costs of the analysis without the benefit of improved classification abilities.

Figure 35 details the ST methodology. The full walking data is passed through the feature extractor which produces the SAS10 feature type for all fourteen accelerometers (this step does require completion if the feature types from the ADC analysis shown in Figure 30 were saved). From the "Rank Aggregation" block of Figure 29, the best aggregated DEIM ranking of sensors is used as input to the "Feature Creation" block. DEIM and SAS10 are chosen as these are the best performing data treatments, in their respective categories, found by the ADC analysis.

The "Grouping Split" is no longer IND, MID, or RED, but rather the subsequent addition of an additional sensor to the analysis. For instance, the first grouping is only the highest ranking sensors; the second is the highest and second highest ranked sensors, and so on. Any rank aggregation method could be used to produce this best overall ranking of sensors, but unless otherwise noted this work uses the proposed SVD-CR method.

The "Feature Creation" block represents the extraction of the appropriate SAS10 features from the appropriate sensors in the order based on the best overall ranking. The resulting features pass into the bagged decision tree algorithm. The bagged decision tree was the best performing machine learning algorithm in the ADC analysis, so it is included here.

Since the rank aggregation method produces a single best ranking and the feature type and machine learning algorithm are defined, multiple trials must be completed and average to use the Tukey analysis. The running of the sweeping for truncation methodology one time will produce a classification accuracy for each number of sensors included in the ranking (14). This work completed this process 30 times, and the error for the corresponding number of sensors is averaged. This averaging allows for the development of multiple samples. Multiple samples allow for the assumption that the distribution of these errors is approximately normal [99]. Each time the analysis is complete there is a different classification error because of the ten-fold-cross-validation method used. The randomization of the method, i.e., which walking trials are assigned to which fold, results in variability in the classification error

Figure 36 shows the results of the Sweeping for Truncation step. Two different random rankings of sensors enable comparison between the SVD-CR rank aggregation method and stochastic rankings. The first random set (RAND) is a random ranking of sensors. The second random set (DEIM RAND) is a random selection from the 120 DEIM rankings. The success of either the RAND or DEIM RAND ranking may show that the rank aggregation is not a necessary step of this methodology. Thirty RAND and thirty DEIM RAND rankings were analyzed through the ST analysis, and the resulting classification errors were averaged respectively.

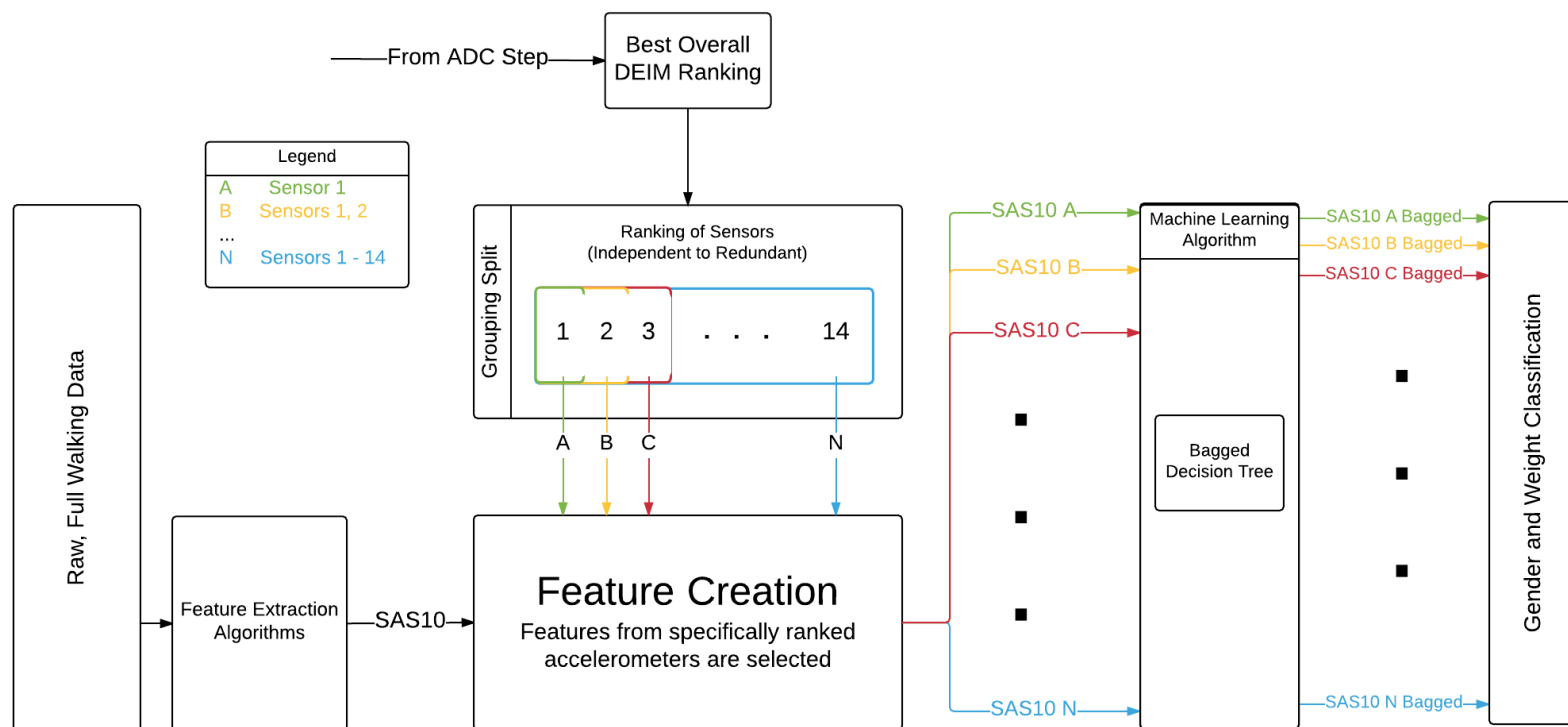


Figure 35. Sweeping for Truncation Methodology. The SAS10 feature types and the best overall ranking from the ADC analysis combine to form the useful features in the “Feature Creation” block. These features are then used in the “machine learning algorithm” block. The output of this process is the classification error for gender and weight. The data treatments of SAS10, DEIM, and bagged decision trees were used based on their best performance in each category of the ADC analysis.

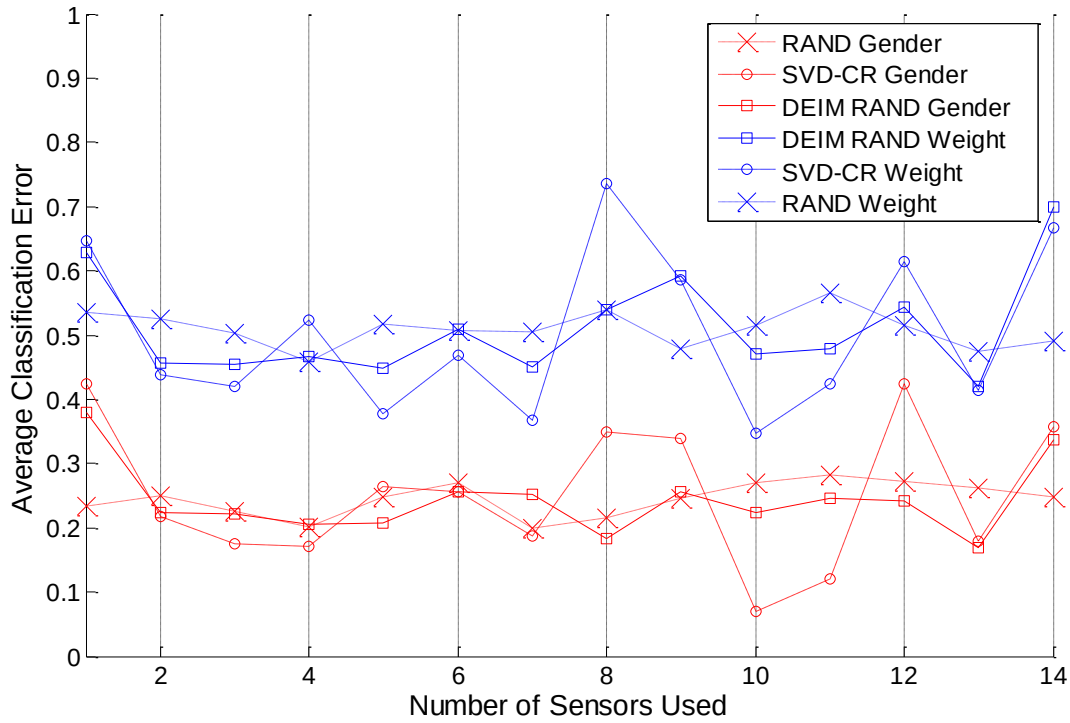


Figure 36. Sweeping for Truncation Results. The classification error when adding subsequently ranked sensors. Various rankings of sensors are included: SVD-CR, random ranking (RAND), and a random selection of one of the 120 rankings not aggregate from DEIM (DEIM RAND).

The usefulness of the ST step is shown in Figure 36 as there is now an objective method of determining which accelerometers are most useful. Just as important, this method shows which sensors are truncatable. In this case, the last four ranked sensors should be truncated as they never yield lower classification error and consume larger computational resources. There is no benefit to considering sensors ranked 11th through 14th.

Figure 36 shows that SVD-CR is more robust rank aggregation method than either of the RAND methods. While RAND and DEIM RAND produce a lower classification rate at some number of sensors included, the SVD-CR method produces the lowest overall classification error of 7 % for gender and 36 % for weight. The SVD-CR rank aggregation method also produces better classification accuracies than the DEIM RAND method. This result shows the effectiveness of the rank aggregation method: SVD-CR.

Both minimum classification errors occur at ten sensors included in the analysis. This result leads to the belief that the ST method would produce low classification error for any characteristic classifier at the same number of

accelerometers. In this case, these ten best sensors will result in good characteristic classification regardless of the classifier (potentially for height, shoe size, etc.). More generally, the sensors determined to be most important are truly the best sensors for recording gait to be used in a characteristic classification approach.

The above results show that rank aggregation method is important for the overall classification accuracy of a model. How important choosing an effective rank aggregation method is for overall algorithm effectiveness is examined further next.

Gender Classification using GA Rank Aggregation

Chapter 6.1.1.2 discusses the success of the GA Spearman rank aggregation method in both the Kendall and Spearman distances compared to the other rank aggregation methods. In that section, there is also discussion about the removal of the most independent accelerometer resulting in the selection of the η_p optimum sensors. These insights have led to the definition of the GA Spearman - η_p optimum rank aggregation method, which moves the most independent sensor from first ranking to last ranking. For this work this effectively moves sensor 12 listed in Table 5 from ranked first to fourteenth.

This section reproduces the analysis completed in the ST step now using the GA Spearman and the GA Spearman - η_p optimum rank aggregation methods. These other rank aggregation methods are compared to the SVD-CR results shown in Figure 36 to understand the importance of minimizing the rank aggregation error metrics (Spearman, Kendall, and η_p). More importantly, this analysis allows for comment on the importance of finding the ‘correct’ rank aggregation method. Figure 37 shows the result of the sweeping for truncation analysis for these rank aggregation methods.

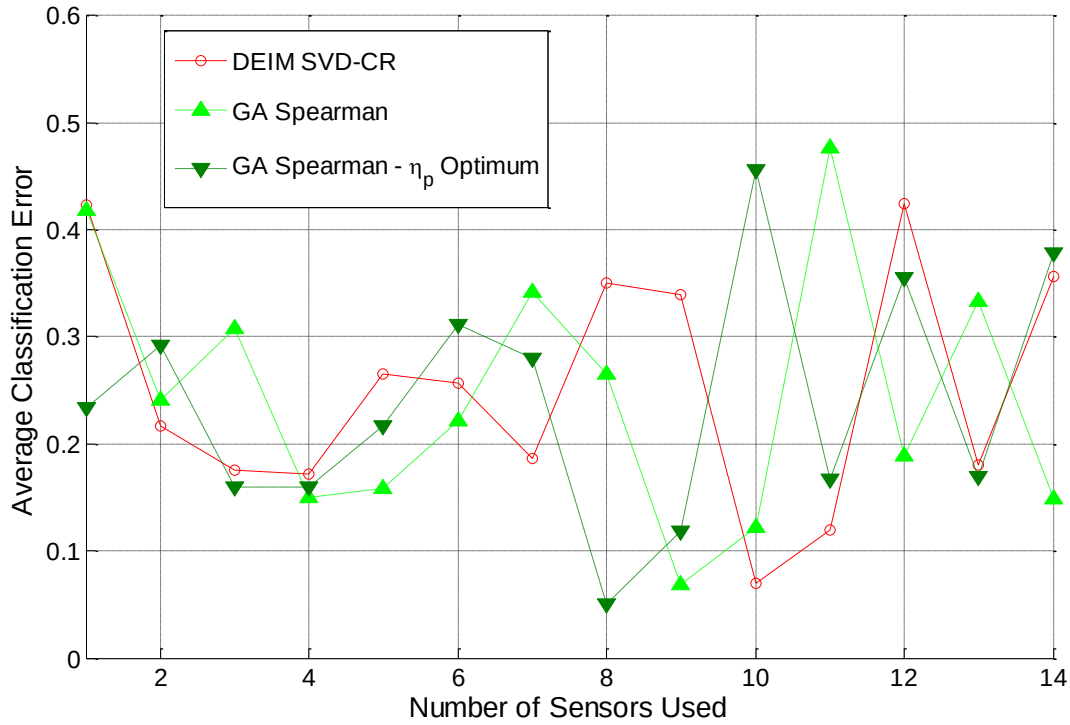


Figure 37. Sweeping for Truncation with various rank aggregation methods. The figure only shows the classification error for gender.

There are three interesting results of this study. First, the rank aggregation methods seem to have similar trends with a sort of phase difference. Second, the GA Spearman - η_p optimum rank aggregation method has a lower classification error at one sensor. Third, the GA Spearman and the GA Spearman - η_p optimum rank aggregation methods produce lower overall classification error than the proposed SVD-CR method.

The results of each aggregation method are similar until the inclusion of five sensors in the analysis. At five sensors the SVD-CR error increases, then at six sensors, the GA Spearman - η_p optimum increases in error, and at seven sensors the GA Spearman method spikes. These type of patterns are seen in multiple places in Figure 37. The trends in classification error as a function of the number of sensors included seem to be present in all of the rank aggregation methods. This result leads to the belief that these patterns may be inherent to the data acquisition system and maybe more so the sensor locations. Further examination of this hypothesis would require the methodology in this thesis to be completed where the sensors are arranged in a different configuration.

The GA Spearman - η_p optimum rank aggregation method produces the lowest classification error at one sensor. This confirms the choice to move the most independent sensor to the last position of the ranking. This result shows the importance of developing sensor rankings that are η_p optimum for better classification accuracy.

Both the GA Spearman and the GA Spearman - η_p optimum ranking methods produce lower overall classification error than the proposed SVD-CR method. Meaning the GA Spearman and the GA Spearman - η_p optimum methods are better accelerometer selection ISMs. Another advantage of these rank aggregation methods is that each increase the number of truncatable sensors. The SVD-CR method encourages truncating four sensors as the classification error cannot get better than this value. While the GA Spearman would truncate five sensors and the GA Spearman - η_p optimum ranking would truncate six sensors. These truncations represent a 36 % and 43 % decrease from the original size of the walking database respectively, while SVD-CR is only 29 %. These results speak to the importance of using the correct rank aggregation method as this choice has a large effect on the outcome of the ST analysis. In this work the GA Spearman - η_p optimum method performed the best for reducing sensors and improving classification accuracy.

In summary, choosing the most appropriate rank aggregation method is imperative for the best possible outcome of the ST analysis. The GA Spearman and GA Spearman - η_p optimum rank aggregation methods produce a lower overall classification error than the SVD-CR method, at a lower number of sensors (6.83 %, 5 %, and 6.99 % respectively). The lowest overall classification error was recorded by the GA Spearman - η_p optimum method leading to the hypothesis that rank aggregation methods should minimize Spearman and Kendall distances while optimizing the η_p approximation metric. The relationship between classification error and the singular value decay is examined below to validate a previously discussed hypothesis.

Sweeping for Truncation Compared to Singular Value Decay

Chapter 4 hypothesizes that there is a relationship between the classification error and the singular value decay. This section examines this relationship by comparing the sweep of included sensors using the SVD-CR and GA Spearman

- η_p optimum ranking aggregation methods to the singular value decay. Figure 38 allows for the direct comparison of the singular value decay and average classification error for both rank aggregation methods.

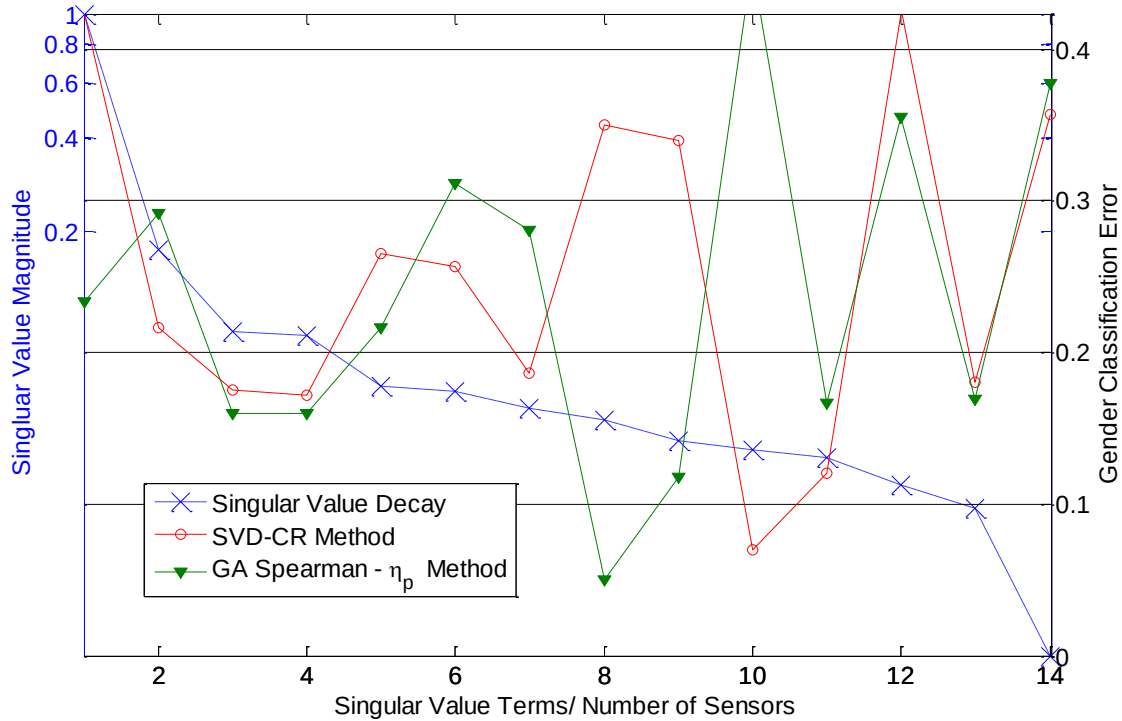


Figure 38. Singular Value Decay Comparison. The comparison of the Singular Value Decay and the average classification error for the SVD-CR and GA Spearman η_p rank aggregation methods. Note the two different y-axes. The left axis is for the singular value decay and the right axis is the gender classification error. The error for 10 and 12 sensors of the SVD-CR and GA Spearman η_p methods, respectively, are slightly off this plot for the purpose of matching the data point at one accelerometer included.

Figure 38 shows little relationship between the singular value decay and the two rank aggregation methods (SVD-CR and GA Spearman - η_p optimum). Although, there is an interesting relationship between the singular value decay and the first four points of the SVD-CD Method. A similar trend exists between the singular values and the SVD-CR sweep for these first four points, but then the algorithms do not correlate when considering more sensors. When examining the average singular values, the first four singular values are the largest and then they begin to level out, until another large drop between 13 and 14 accelerometers included. It is during this relatively flat section in singular values (after the first four) that the SVD-CR method begins to operate in a different manner. In this region where the singular values are relatively unchanged from one value to the next, the SVD-CR method may begin to behave similarly to the other rank aggregation methods because the scaling of the SVD-CR is no longer the dominating factor.

Chapter 4.2.2 hypothesizes that the classification accuracy will have a correlation with the singular value decay. The results presented above, little correlation between the classification accuracy and singular value decay prove this hypothesis incorrect.

In summary, the sweeping for truncation section provides a methodology for removing or truncating sensors that do not contribute to the success of characteristic classification. Not only are sensors truncated, reducing the computation time, but the overall classification accuracy can be improved. Further studies of the ST, show the importance of choosing the ‘correct’ rank aggregation method by minimizing rank aggregation metrics. The correct rank aggregation method may be intractable due to the number of possible permutations of sensors, but the methods detailed provide relatively good results. The other instance selection method, sampling rate reduction, is examined in the next section to improve the classification accuracy of the SVD-CR aggregation method.

Chapter 6.1.2. Sampling Rate Reduction

The accelerometer selection – ISMs above deal with reducing the n^{th} dimension of the walking database, $\mathbb{R}^{m \times n}$. On the other hand, reducing the m^{th} dimension is possible through a sample rate selection – ISM (SRS-ISM). The SRS-ISM considered in this work is downsampling which is explained in Chapter 4. A successful SRS-ISM will reduce the required sampling rate, while maintaining or improving the classification results listed above. Reducing the sampling rate can reduce the amount of features input to the machine learning algorithms. For example, the SAS data type will now need fewer data points to describe an entire average footstep. If the signal is downsampled by 2 then the SAS feature type of a single footstep contains 3501 data points rather than the 7001 of the full feature.

Understanding the minimum sampling rate for successful classification may also address plausibility concerns associated with real-world applications. Knowledge of the minimum sampling rate for effective classification informs the required quality of both sensors and the data acquisition system required to record gait using underfloor mounted accelerometers.

This SRS-ISM study uses a methodology similar to that in the ST step with the addition of a downsampling step. Figure 39 details the methodology to complete this analysis. The downsampling factors examined in this analysis were

1, 2, 4, 6, 8, 16, 32, 64, 128, and 256 resulting in sampling rates ranging from 51200 to 200 Hz. The grouping split is still the best overall ranking as determined by the SVD-CR ranking.

Figure 39 shows the methodology for the SRS-ISM study. The methodology is identical to that in the Sweeping for Truncation methodology with the addition of an SRS-ISM block. The various downsampling factors alter the feature. The new output features are then used in the bagged decision trees. Only gender classification results are produced in the SRS-ISM section.

Figure 40 shows the minimum sampling rate for the minimum classification error rate as a function of the number of included sensors. Each number of included sensors (1-14) results in a classification error at each sampling rate, but for ease of understanding the figure, only the minimum classification error at the minimum sampling rate is shown. The minimum sampling rate is the only value shown because even if the same classification error is possible at a higher sampling rate, there would be little reason to include more information than necessary.

Current results show that there is a complex relationship between the number of accelerometers included in the analysis and the minimum sampling rate. This complexity means recording gait at a high sampling rate and then conducting a sweep of sample rate selection factors to determine the most appropriate rate on a case by case basis.

Although no general trend is clear from this analysis, Figure 40 does show one interesting result; the minimum gender classification error in this entire thesis is achieved using an SRS-ISM. This model which combines the best data from the ADC step, the SVD-CR rank aggregation method, and a reduced sampling rate achieves a 96.7 % accuracy. Thus, the effort in completing a case by case sampling reduction sweep may be worth the effort as it has led to the best overall results.

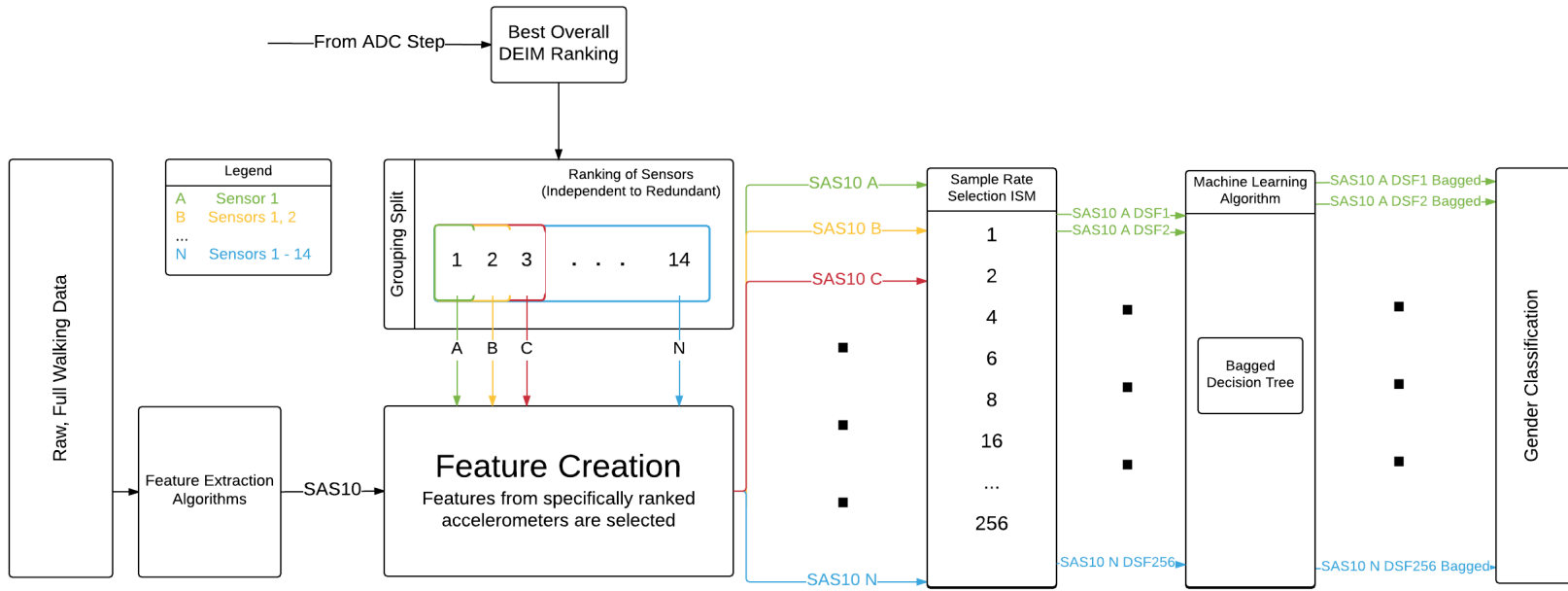


Figure 39. Down Sampling Methodology. The best performing data treatments from the ADC section are again used as input to the accelerometer selection ISM. Then the ranking as decided by the SVD-CR rank aggregation method is used to sweep through the number of sensors included. Then the sample rate selection ISM was completed by down sampling based on a down sampling factor (i.e., down sampling by two is a factor of two). Then the bagged decision tree algorithm type is used to generate the gender classification error.

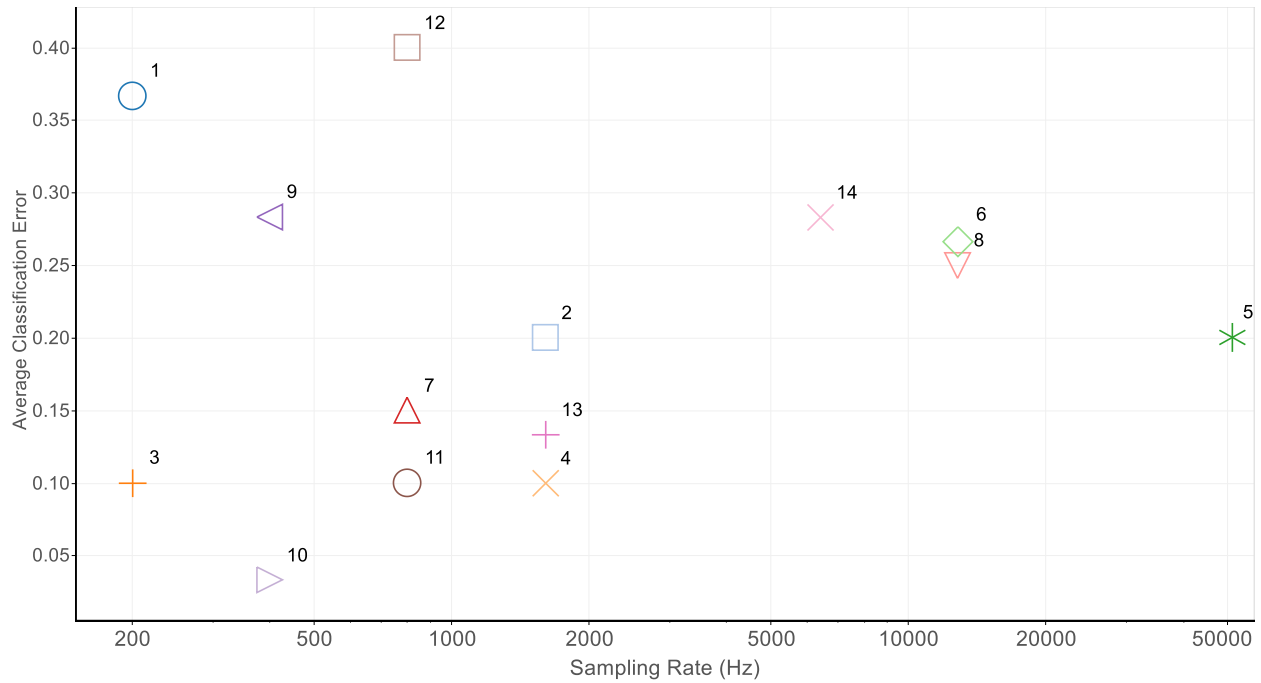


Figure 40. Minimum Sampling Rate Results. Sampling rate that results in the minimum average classification error for 1 – 14 sensors included in the analysis. For example, the minimum sampling rate using the three best SVD-CR ranked sensors occurs at 200 Hz resulting in 10 % error. Note the log scaling for the x-axis.

It is also interesting to note that the minimum classification error occurs at ten sensors, the same as the results of the SVD-CR rank aggregation method in the ST step. This result implies that the other rank aggregation methods GA Spearman and GA Spearman - η_p optimum would yield reduced classification error at their respective values for minimum classification error (9 and 8 sensors included).

In summary, the sample rate selection ISM of down sampling is an effective method for improving the accuracy of characteristic classifiers. A complex relationship exists between the minimum sampling rate, the number of sensors included in the analysis, and the minimum classification error. This study is unable to provide specific trends in this relationship, but the computation of this analysis after the completion of the ADC and ST steps may be worth the added accuracy of the classification models dependent on the application.

Figure 41 shows the overall methodology discussed in the introduction to this chapter. The ADC allows for querying of any data treatment of interest to reduce the required computation understand which sensor are truncatable. As has been shown above the grouping split of the ADC is removable as the IND set heuristically performs near optimum.

The ST step allows for the determination of which sensors to truncate. The selection of the appropriate rank aggregation method is imperative. The sample rate selection step can improve the overall characteristic classification.

The methods in this chapter enable an examination of taking appropriate gait measurements in future experiments. For instance, the ADC and ST steps result in a list of truncatable sensor locations. Knowing these locations can improve future sensor arrangements. Understanding the minimum sampling rate, through the sample rate selection process, can inform data acquisition requirements of new systems.

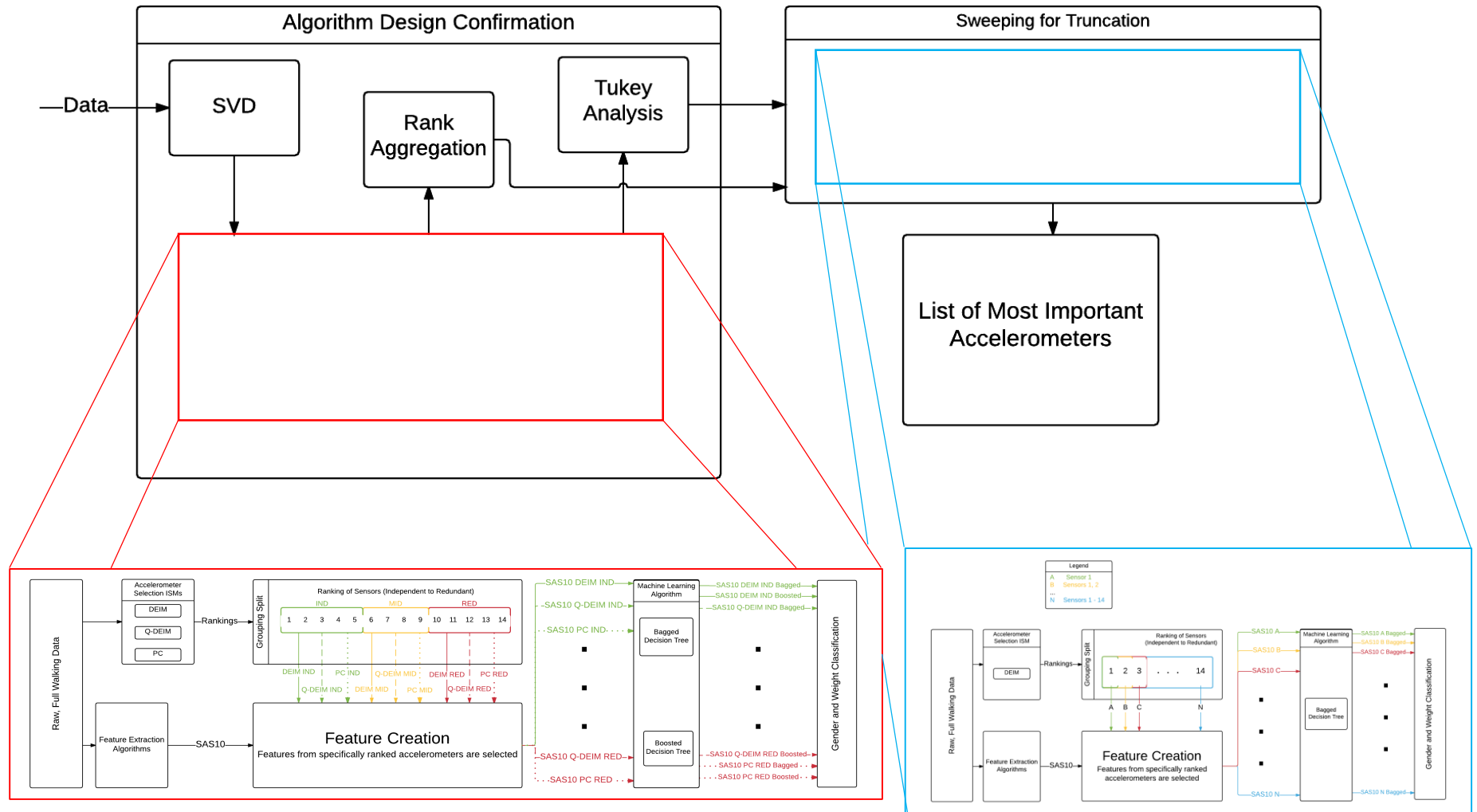


Figure 41. Entire Methodology Summary. The entire methodology shows both the Algorithm Confirmation and Sweeping for Truncation. The initial Singular Value Decay determines a suitable number of sensors to complete the ADC. The process shown in Figure 30 is applied, allowing for both a best overall ranking to be generated via SVD-CR and a Tukey analysis to determine the most appropriate design choices. The Sweeping for Truncation is then applied, as shown in Figure 35, facilitating a truncation in the number of sensors required. This methodology effectively allows for a systematic reduction of the number of sensors required for effective classification.

Chapter 6.2. **Conclusion**

The broad sections of this chapter have independent conclusions in subsections below. Overall, the use of instance selection methods is a tangible method of improving characteristic classification accuracy while reducing the amount of data required to make a classification. These methods will lead to faster classifiers that potentially make multiple applications possible. The title of each subsection includes the letter of the specific question from Chapter 1 addressed. The specific questions are presented below for reference.

- (A) What is the most appropriate type of accelerometer selection – instance selection method (AS-ISM)?
- (B) What is the effect of the various data treatment types (e.g. feature type, walking direction, machine learning algorithm type, etc.)?
- (C) What is the minimum number of accelerometers required for the highest accuracy characteristic classification?
- (D) How effective is the proposed Singular Value Decomposition – Count Rank, rank aggregation method in comparison to other well-established rank aggregation methods?
- (E) What effect does rank aggregation method have on the number of truncatable accelerometers in the Sweeping for Truncation (ST) step?
- (F) What relationship exists between sampling rate and classification accuracy?

Algorithm Design Confirmation (Questions A and B)

A proposed Algorithm Design Confirmation (ADC) aids in choosing the most appropriate data treatments. The singular value decay can be completed before the analysis of the data to determine an appropriate initial number of sensors to consider in determining the best data treatment. A statistical analysis (Tukey test) determines which of the data treatments result in the most accurate classification. The SVD-CR rank aggregation method reduces computational complexity allowing for a sweep of sensors. The table below summarizes the best data treatments.

Table 6. Statistical Testing Summary. The best and worst performing of each data treatment type are noted.

Data Treatment Type	Best Performing	Worst Performing
Accelerometer Selection ISM	DEIM, Q-DEIM	PCs
Walking Direction	Eastward	Westward
Grouping Type	IND, MID, RAND	RED
Machine Learning Algorithm	Bagged Decision Trees	Boosted Decision Trees
Feature Type	SAS10	SAS5

Rank Aggregation Comparison (Question C)

The proposed SVD-CR method produces a best overall ranking from the 60 DEIM rankings of accelerometers. This ranking orders sensors from most independent to most redundant. The first five sensors ranked by the SVD-CR are nearly η_p optimum meaning the approximation of DEIM is relatively close to optimum. SVD-CR performs as well or better in η_p approximation metric as the other rank aggregation methods presented. SVD-CR has similar Spearman and Kendall distance measures to the other methods with the exception of the GA Spearman algorithm. The GA Spearman method is a viable method of rank aggregation as well as the proposed SVD-CR method.

Sweeping for Truncation Questions D and E)

This chapter proposed the ST methodology for reducing the number of accelerometers used to measure gait in post-processing. The positive results of the ST step allow for the reduction of the number of sensors considered in analysis while improving the classification accuracies presented in Chapter 5. When using the SVD-CR ranking, the number of sensors included for a minimum classification error was ten sensors for both the gender and weight classifier.

The GA Spearman rank aggregation method produced better overall gender classification error then the SVD-CR ranking (6.8 % compared to 7.0 % error). The creation of the GA Spearman - η_p optimum ranking produced the best overall gender classification accuracy (5 % error). Both of the GA Spearman rank aggregation methods produce better classification accuracy while increasing the number of truncatable sensors. Quality rank aggregation methods are imperative to the success of this Sweeping for Truncation methodology because different rank aggregation methods were able to produce better classification error while increasing the number of truncatable sensors.

Sampling Rate Reduction (Question F)

This study produced a thesis best minimum classification error of 3.3 %. However, the sampling rate selection provided no general trend that might inform a minimum sampling rate for future experiments. This result means that future experiments should use a high sampling rate and then sweep for the minimum sampling rate to reduce the classification error.

In summary, the proposed methodology allows for an increase in gender classification accuracy of walkers over Bales et al. [3] at 97 % and a 64 % weight classifier accuracy. The instance selection methods not only increase gender classification accuracy but reduce the amount of data storage and computation time required. These methods can lead to an understanding of the spatial relationships between truncatable accelerometers and the minimum sampling rate. Understanding these results can inform how to effectively take data in future experiments.

CHAPTER 7. CONCLUSION

Project summary, project aims, contributions, and future work sections comprise the conclusions section. The project summary frames the project background to showcase the results and contributions of the work. The project aims present the questions posed in the thesis introduction, answers those questions and establishes the impact of the results. The contributions are a bulleted list of the novelty and the role filled by this thesis. Finally, the future work section gives suggestion on how to continue this characteristic classification research.

Chapter 7.1. Project Summary

Goodwin Hall is a five-story, 155,000 square foot living lab. Goodwin Hall is the most instrumented public building in the world for recording vibrations. The instrumentation of the building by the Virginia Tech Smart Infrastructure Lab has enabled investigation into multiple studies aimed at improving the human-building interaction. This thesis work focuses on the characteristic classification of walkers. The outcomes of this research lay the groundwork for developing methods of re-identification of a suspect using the soft biometrics of gender and weight.

There is a wide variety of gait measurement techniques in the literature ranging from video recording to ground response measurement due to footsteps. The latter method is most similarly related to the underfloor mounted accelerometers used in this thesis. Various previous works have examined gender and weight classification using gait measurements, but none from the measurement of gait through underfloor mounted accelerometers. The highest found gender classification accuracy using gait measurements was 75 %, and the highest for weight was ± 0.07 kg. In this thesis, the highest gender classification accuracy is 97 % and 64 % for weight classification.

An experiment was conducted in Goodwin Hall (GH) to provide gait measurements as input data for machine learning algorithms. Twenty study participants walked along a 98-foot-long instrumented hallway in Goodwin Hall. Fourteen of the underfloor mounted accelerometers surrounding and in the test hallway recorded gait. The gait measurement from the underfloor accelerometers and the gender and weight for 120 walking trials facilitated the creation of a walking database. From the walking database features were created for use in the machine learning algorithms which predict the class of never-before-seen walking trials.

Supervised learning methods are used throughout this work to generate models to predict the gender and weight of walkers. These machine learning algorithms enable the finding of complex patterns inherent in gait in potentially real-time algorithms. These algorithms facilitate characteristic classification from the underfloor mounted accelerometers. Bagged and boosted decision trees, support vector machines, and neural networks are the machine learning algorithms used in this work. The success of these algorithms is dependent upon the quality of the input features.

Data reduction methods produce successful features, but computationally smaller, used in the machine learning algorithms. Instance selection methods further reduce the quantity of features, while improving their classification accuracy. These methods reduce the number of features, decreasing the computation time required, and therefore, making more technological applications possible. In addition to the reduction in the number of features, these methods improve the classification of both gender and weight. The instance reduction methods used in this work are DEIM, Q-DEIM, Projection Coefficients (for accelerometer selection) and downsampling (for sample rate selection).

Chapter 7.2. Project Aims

The project aims are the questions originally posed in the introduction and discussed throughout this thesis. The answers to these questions can inform future research in the field of characteristic classification using both underfloor mounted accelerometers or any multiple sensor measuring system. Each question is stated, answered, and discussed.

Chapter 5 Conclusions

(1) Which of the proposed feature sets are most useful for gender classification?

Of the ten feature types used in Chapter 5, the SAS feature types produced models that had the best gender classification accuracy. Within the SAS type, the SAS10 and SAS5 feature types were determined to have statistically equivalent classification accuracy when considering all four machine learning algorithms. The results of the SAS type suggest the need to record multiple steps and that peak aligning and averaging them builds a feature containing classifiable characteristic information. The MSS feature type was developed to understand the importance of spatial relations between walkers and accelerometers produced the lowest accuracy of any feature. This result suggests that developing features incorporating spatial relationships are not necessary for successful features. The SSS feature type performed relatively well in gender classification

suggesting that measurement of a single footstep with a single sensor contains classifiable characteristic information. The classification error for the SAS10, SAS5, SAS3, MSS, and SSS were 25, 23, 29, 41, and 32 % in the time domain respectively. The frequency domain results were 34, 34, 36, 41, 38 % respectively.

(2) Which of the proposed machine learning algorithms performed best?

The bagged and boosted decision trees as well as the support vector machines (24th order polynomial) produced statistically equivalent classification accuracy independent of feature type. The support vector machines require further optimization than the decision trees but potentially run computationally faster. Any of these three algorithms can be used in further gait characteristic classification. The neural network algorithm produced the lowest accuracy classification model. The poor performance of the neural networks may be a result of the large number of features or the relatively low complexity of the model (i.e., difficult patterns are difficult to learn with simple models).

(3) Which domain, time or frequency, contains the most readily useful information for gender classification?

The time domain contains statistically better features than the frequency for gender classification. The frequency domain summarizes many of the peaks in the time domain into a single non-zero feature. The resulting relatively few interesting features in the frequency domain, are hypothesized to be overshadowed by the majority of noisy features. These results suggest that developing features in the time domain will yield the high classification accuracies.

Chapter 6 only includes the best performing feature types and machine learning algorithms from Chapter 5, greatly reducing the number of iterations required in the Algorithm Design Confirmation step of the methodology.

Chapter 6 Conclusions

(A) What is the most appropriate type of accelerometer selection – instance selection method (AS-ISM)?

The Discrete Empirical Interpolation Method (DEIM) and its alternate formulation Q-DEIM produce statistically equivalent gender classification accuracies. The grouping splits of IND, MID, and RED based on the rankings output by DEIM and Q-DEIM produce better classification accuracy models than the

Projection Coefficients (PCs). Either the DEIM or Q-DEIM method should be used to rank sensors from independent to redundant. There is a potential to use Q-DEIM in a stochastic formulation which might tip the scales for this algorithm if the database examined is relatively large. Drmac and Gugercin [88] present the formulation and discussion for the stochastic Q-DEIM.

(B) What is the effect of the various data treatment types (e.g. feature type, walking direction, machine learning algorithm type, etc.)?

The table below summarizes the best and worst performing data treatments for gender classification accuracies. Following the data treatments not discussed in other project, aim questions are addressed specifically.

Table 7. Statistical Analysis Results for all Data Treatments.

Data Treatment Type	Best Performing	Worst Performing
Accelerometer Selection ISM	DEIM, Q-DEIM	PCs
Walking Direction	Eastward	Westward
Grouping Type	IND, MID, RAND	RED
Machine Learning Algorithm	Bagged Decision Trees	Boosted Decision Trees
Feature Type	SAS10	SAS5

There is an interesting result showing that the Eastward walking direction produced better classification accuracy models than Westward trials. A bias in the walking experiment may explain this result; as each walker began each experimental trial walking Westward. This first trial is potentially biased because each walker is self-conscious of their walking with this trend subsiding as walkers continue the experimentation. The direction dependence could also be explained by the non-symmetric arrangement of sensors or the visual cues that the walkers are walking toward in each trial.

The singular value decay informed using five sensors in the Algorithm Design Confirmation step in Chapter 5. Therefore, four grouping types divide the ranking of all fourteen sensors. The grouping types of IND, MID, and RAND produce higher classification accuracies than the RED set. Further study shows that the IND ranking of sensors enables the removal of the grouping split from the ADC methodology. The IND set serves

as a heuristic set in place of calculating the best possible set of sensors, which quickly becomes intractable with an increase in the number of sensors examined. The RAND, or randomly created, set produces relatively high classification results, but the results must be averaged over multiple iterations to ensure the ranking produces good classification accuracies.

The best machine learning algorithm is the bagged decision trees. This result is different than the answer to question (2) from Chapter 5. This previous question stated the boosted and bagged decision trees were statistically equivalent. An increase in the number of data treatments included in Chapter 6 allows for the statistical determination that bagged decision trees are robust to data treatment type. Future characteristic classification studies should include bagged decision trees as these algorithms produce the best results of the studied machine learning algorithms.

(C) What is the minimum number of accelerometers required for the highest accuracy characteristic classification?

The minimum number of accelerometers for effective gender and weight classification is dependent on the desired classification error. Relatively few accelerometers can be considered to achieve poor classification accuracy, but more sensors are needed to improve the accuracy to a point. The minimum number of accelerometers is also dependent on the rank aggregation method. The table below shows the minimum number of sensors required and the gender classification error for the three examined rank aggregation methods.

Table 8. Rank Aggregation Results. The minimum classification error for gender and weight and the corresponding number of accelerometers included in the analysis.

Rank Aggregation Method	Minimum Classification Error Gender (Weight)	Number of Accelerometers at the Minimum Classification Error
SVD-CR	7 % (36 %)	10
GA Spearman	7 % (NA)	9
GA Spearman - η_p Optimum	5 % (NA)	8

The study of a minimum number of accelerometers required for successful classification shows that some accelerometers do not aid in reducing the classification error, and future analysis should truncate these sensors from analysis. This methodology allows for a reduction in the number of sensors analyzed in these characteristic classification problems reducing computation time while improving classification accuracy.

(D) How effective is the proposed Singular Value Decomposition – Count Rank, rank aggregation method in comparison to other well-established rank aggregation methods?

The SVD-CR rank aggregation method performs as well or better than most of the other rank aggregation methods compared in three metrics: Spearman and Kendall distances as well as η_p approximation error. This establishes this method as a potential rank aggregation method warranting future use, but it is not the best performing algorithm found in this work. The GA Spearman and GA Spearman - η_p Optimum methods produce lower classification error at a lower number of considered sensors (shown in Table 8). These other rank aggregation methods are suggested for use in this database over the developed SVD-CR method.

(E) What effect does rank aggregation method have on the number of truncatable accelerometers in the Sweeping for Truncation (ST) step?

As is slightly touched on in the previous two questions the choice of rank aggregation method greatly affects the outcome of the sweeping for truncation step of the proposed methodology. Not only can choosing the appropriate rank aggregation method reduce the characteristic classification error but it can do so with a reduced number of accelerometers. For instance, GA Spearman - η_p optimum method achieves a 5 % gender classification error while reducing the amount of data used in the feature development by 43 %. Further research into the ‘best’ rank aggregation method should be completed, but this thesis has established that a lower Kendall, Spearman, and η_p error metrics produce higher classification accuracies.

(F) What relationship exists between sampling rate and classification accuracy?

The relation between sampling rate and classification accuracy is complex, and this work can offer no general trends as shown in Figure 42. No straightforward tendency exists as the minimum classification error does not occur at any particular sampling rate regardless of the number of sensors included in the analysis (from

1 to 14). The sampling rate selection - instance selection method is useful for reducing the overall classification error. Thus, the downsampling methods used may warrant their relatively large computation time in applications that require minimum classification error. The large computation time results from the need to sweep through many downsampling factors (as shown in Figure 42) because no heuristic, general trend, is readily available from this analysis. Based on this lack of result, to gain improve the classification error results requires a full sweep of many sampling rates.

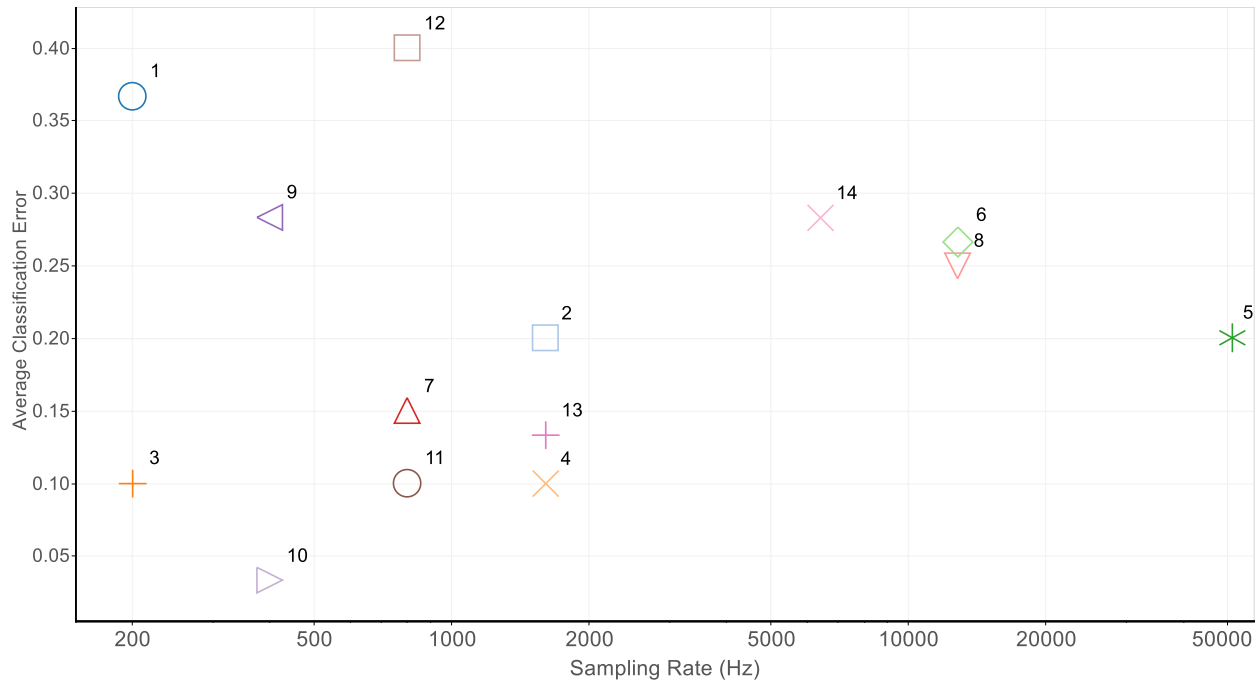


Figure 42. Classification Error as a Function of Sampling Rate. Each symbol and marker color corresponds with the number accelerometers included in the analysis. There seems to be no relationship between classification error and neither the sampling rate nor the number of sensors included in the analysis.

The project aims answered in Chapter 6 demonstrate the success of the proposed accelerometer selection process: Algorithm Design Confirmation and Sweeping for Truncation. These steps together aim to reduce the computation time required to determine the best data treatments and the accelerometers that are truncatable. These steps can be used in future analysis to determine which sensors should be examined for the purpose of improving the classification accuracies found using a simpler analysis, as the analysis in Chapter 5. The next section states the contributions of this thesis work.

Chapter 7.3. Contributions

- Walking Database – A walking database filled with the gait measurements of 20 individuals (120 trials total) was created and is available for future analysis by VTSIL. This database can also be appended easily for the continuation of this characteristic classification research.
- Gender Classification from Gait Measurement – The algorithms from Chapter 5 produce an 88 % accurate model for predicting gender using boosted decision trees and SAS5 features. Chapter 5 uses all fourteen accelerometers in the classification process while the instance selection methods of Chapter 6 reduce the number of accelerometers examined. Chapter 6 uses data reduction methods to increase the gender classification accuracy to 97 %. The Chapter 6 model uses SAS10 features, bagged decision trees, Eastward trials, the DEIM accelerometer selection method, the SVD-CR rank aggregation method, including ten sensors in the analysis, and a downsampling factor of 128 (sampling rate of 400 Hz) to achieve this 97 % accuracy in classification. This accuracy is the highest known for gender classification using any gait measurement technology.
- Weight Classification from Gait Measurement – The approaches in Chapter 6 enables the development of a model that is 64 % accurate in weight classification. The algorithm which achieves 64 % accuracy is SAS10 features, bagged decision trees, Eastward trials, the DEIM accelerometers selection method, the SVD-CR rank aggregation method, and considering ten accelerometers in the analysis. The weight bin classification of this thesis makes a direct comparison the other literature difficult as other works report standard deviations in the number of pounds they were able to predict. This work uses a classification approach rather than regression, so a continuous variable is not output, but rather a bin to which a walker belongs. The accuracy is the highest known for an underfloor mounted accelerometer classification.
- Rank Aggregation – The SVD-CR method proposed is as good as most of the algorithms compared to in this work warranting its future use in many rank aggregation methods. Using the η_p approximation error to

remove poorly performing, independently ranked sensors is believed novel. This process improved the overall gender classification error while reducing the number of accelerometers that must be considered.

- Instance Selection Methods – The use of the instance selection methods (DEIM, Q-DEIM, and PCs) to determine the most useful accelerometers is novel and enables the reduction of those accelerometers from analysis.

These contributions will positively impact the emerging field of gait analysis, and underfloor mounted accelerometer approaches. Hopefully these contributions aid in the further development of smart buildings and enable better human-building interactions. The next section offers a suggestion on the continuation of this research.

Chapter 7.4. Future Work

The bulleted list below details suggestions for future work in this area of research.

- Regression, not Classification – An alternative method for predicting the weight of a walker is to treat the problem as a regression rather than a classification. Classification methods were used to understand if this technology is viable for predicting weight, but may be hiding the effectiveness of the ability of the learning algorithms.

Regression is a supervised machine learning algorithm which outputs a continuous variable rather than a discrete output. All of the machine learning algorithms used in this work (bagged and boosted decision trees, SVMs, and neural networks) can be formulated for use in a regression problem. Then a regression analysis is possible by changing the observation labels to individual walker weights rather than a discrete weight bin label.

- Feature Types – This work investigates many feature types (SAS10, SAS5, SAS3, MSS, and SSS in both the time and frequency domain) as an initial attempt at characterizing an individual's gait. More representative features may exist to characterize a specific class, for example, weight, but the importance of these very specific feature types and how they relate to a broader classification still must be investigated. The best

possible feature is one that allows for classification, or regression, of many characteristics (i.e. height gender weight, etc.) not only a single one.

- Machine Learning Algorithms – There are countless algorithms and methodologies associated with machine learning that may perform better in classifying walker characteristics. The results of this work demonstrate that machine learning is a viable approach for characteristic classification. The algorithms studied in this work are meant to point aim future research to algorithms which may achieve a better overall methodology.
- Rank Aggregation Methods – The proposed SVD-CR method for combining multiple rankings is an initial effort into this field. There are many ‘appropriate’ methods to provide an overall rank that is representative of all the ranking aggregated together. Since examining all 87 billion possible rankings of sensors is not possible, rank aggregation methods must be well understood and used.
- Increasing Database Size – There are always questions as to how robust a model is to generalization. That is, correctly predicting the gender of an individual which not included in the original walking data set. The ten-fold-cross-validation method ensures the generalization of the developed models; more walking trials could further validate the models developed in this work.
- Investigating the Digital Signal Processing – There are remaining questions to be further investigated about the digital signal processing used in this thesis. There remains question about the effects of the:
 - Do the filtering order and cutoff frequency used affect classification accuracy?
 - Do the 7001 samples which define a step represent all the key characteristics of a footstep?
 - Is the peak alignment and then averaging the most appropriate method for combining footsteps?
- Testing Multiple Walkers – All of the experiments conducted for this thesis work have been done using a single walker at a time. The addition of multiple walkers would add complexity, but would be more representative of the environments where these technologies would be used.

- Physically Understanding the Classification Outcomes – Another important question to investigate is why a machine learning algorithm labels a walker in a specific characteristic, for example, gender. Understanding the physical nature of the gait patterns that point to male or female can allow for the development of machine learning algorithms that probe more directly the underlying difference.

The results in this thesis demonstrate the potential use of underfloor mounted accelerometers in characteristic classification. Understanding a walker's characteristics can lead to applications which require soft biometrics. Such applications are security and individual re-identification. The future work will further aid in the understanding of the limitations of this technology while delving deeper into the successful classifications presented in this thesis. This research field potentially leads to exciting applications that will be far reaching for how we use buildings, and more so how we make them work for us.

REFERENCES

- [1] L. Middleton, A. A. Buss, A. Bazin, and M. S. Nixon, "A floor sensor system for gait recognition," in *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, 2005, pp. 171-176.
- [2] V. Racic, A. Pavic, and J. M. W. Brownjohn, "Experimental identification and analytical modelling of human walking forces: Literature review," *Journal of Sound and Vibration*, vol. 326, pp. 1-49, September 25, 2009 2009.
- [3] D. B. Bales, P. A. Tarazaga, M. E. Kasarda, D. Batra, A. G. Woolard, J. D. Poston, *et al.*, "Gender Classification of Walkers Via Underfloor Accelerometer Measurements," *IEEE Internet of Things Journal*, 2/11/2016 2016.
- [4] D. Bales, P. Tarazaga, M. Kasarda, and D. Batra, "Gender Classification using Vibration Measurements," in *IMAC 34*, 2016.
- [5] D. B. Bales, P. A. Tarazaga, M. Embree, S. Gugercin, and M. E. Kasarda, "Sensor Truncation via Gait Classification using Underfloor Accelerometers," *IEEE Internet of Things Journal*, 2016 (In Preparation).
- [6] J. M. Hamilton. (2015). *Design and implementation of vibration data acquisition in Goodwin Hall for structural health monitoring, human motion, and energy harvesting research*.
- [7] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 4-20, January 2004 2004.
- [8] C. BenAbdelkader, R. Cutler, and L. Davis, "Stride and cadence as a biometric in automatic person identification and verification," in *Fifth IEEE International Conference on Automatic Face and Gesture Recognition. Proceedings*, 2002, pp. 372-377.
- [9] E. Ayyappa, "Words about words: the terminology of human walking, bipedal exchange."
- [10] *Clinical Assessment of Human Gait. : JPO: Journal of Prosthetics and Orthotics*.
- [11] M. S. Nixon and J. N. Carter, "Automatic Recognition by Gait," *Proceedings of the IEEE*, vol. 94, pp. 2013-2024, 2006.
- [12] C. Velardo and J. Dugelay, "Weight estimation from visual body appearance," in *2010 Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, 2010, pp. 1-6.
- [13] S.-M. Aghazadeh, ""Layout strategies for retail operations: A case study," *Management Research News*, vol. 28, pp. 31-46, 2005 2005.
- [14] *Annual retail industry sales in the United States 2000-2014 | Statistic*.
- [15] S. Clifford and Q. Hardy, "Attention, Shoppers: Store Is Tracking Your Cell," in *The New York Times*, ed, 2013.
- [16] G. Gavett. (2014). *How Data Visualization Answered One of Retail's Most Vexing Questions*.
- [17] M. Yang, H. Zheng, H. Wang, and S. McClean, "Feature selection and construction for the discrimination of neurodegenerative diseases based on gait analysis," in *3rd International Conference on Pervasive Computing Technologies for Healthcare, 2009. PervasiveHealth 2009*, 2009, pp. 1-7.
- [18] "Important Facts about Falls," Centers for Disease Control and Prevention January 20, 2016 2016.
- [19] M. E. Tinetti, M. Speechley, and S. F. Ginter, "Risk factors for falls among elderly persons living in the community," *The New England Journal of Medicine*, vol. 319, pp. 1701-1707, Dec 29, 1988 1988.
- [20] C. Sherrington, J. C. Whitney, S. R. Lord, R. D. Herbert, R. G. Cumming, and J. C. T. Close, "Effective exercise for the prevention of falls: a systematic review and meta-analysis," *Journal of the American Geriatrics Society*, vol. 56, pp. 2234-2243, Dec 2008 2008.
- [21] "How much energy is consumed in residential and commercial buildings in the United States?," U.S. Energy Information Administration April 3, 2015 2015.
- [22] J. Zhang, R. Lutes, G. Liu, and M. Brambley, "Energy Savings for OccupancyBased Control (OBC) of VariableAir-Volume (VAV) Systems," U.S. Department of Energy, Pacific Northwest National Laboratory PNNL-22072, January 2013 2013.
- [23] D. D. Hoffman and B. E. Flinchbaugh, "The interpretation of biological motion," *Biological Cybernetics*, vol. 42, pp. 195-204, 1982.
- [24] C. BenAbdelkader, R. Cutler, H. Nanda, and L. Davis, "EigenGait: Motion-Based Recognition of People Using Image Self-Similarity," in *Audio- and Video-Based Biometric Person Authentication*, J. Bigun and F. Smeraldi, Eds., ed: Springer Berlin Heidelberg, 2001, pp. 284-294.
- [25] C. Cedras and M. Shah, "A survey of motion analysis from moving light displays," in *1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94*, 1994, pp. 214-221.

- [26] T. B. Moeslund and E. Granum, "A Survey of Computer Vision-Based Human Motion Capture," *Computer Vision and Image Understanding*, vol. 81, pp. 231-268, March 2001 2001.
- [27] J. Sinclair, S. J. Hobbs, L. Protheroe, C. J. Edmundson, and A. Greenhalgh, "Determination of gait events using an externally mounted shank accelerometer," *Journal of Applied Biomechanics*, vol. 29, pp. 118-122, Feb 2013 2013.
- [28] M. D. Addlesee, A. Jones, F. Livesey, and F. Samaria, "The ORL active floor [sensor system]," *IEEE Personal Communications*, vol. 4, pp. 35-41, October 1997 1997.
- [29] T. Liu, Y. Inoue, K. Shibata, and K. Shiojima, "A Mobile Force Plate and Three-Dimensional Motion Analysis System for Three-Dimensional Gait Assessment," *IEEE Sensors Journal*, vol. 12, pp. 1461-1467, 2012.
- [30] A. Godfrey, R. Conway, D. Meagher, and G. ÓLaighin, "Direct measurement of human movement by accelerometry," *Medical Engineering & Physics*, vol. 30, pp. 1364-1386, 2008.
- [31] M. Orendurff, J. Schoen, G. Bernatz, A. Segal, and G. Klute, "How humans walk: Bout duration, steps per bout, and rest duration," *Journal of Rehabilitation Research & Development*, vol. 45, pp. 1077-1090, 2008 2008.
- [32] M. N. Nyan, F. E. H. Tay, K. H. W. Seah, and Y. Y. Sitoh, "Classification of gait patterns in the time–frequency domain," *Journal of Biomechanics*, vol. 39, pp. 2647-2656, 2006.
- [33] T. Hoang and D. Choi, "Secure and privacy enhanced gait authentication on smart phone," *TheScientificWorldJournal*, vol. 2014, p. 438254, 2014 2014.
- [34] J. M. Sabatier and A. E. Ekimov, "A Review of Human Signatures in Urban Environments Using Seismic and Acoustic Methods," in *2008 IEEE Conference on Technologies for Homeland Security*, 2008, pp. 215-220.
- [35] A. Ekimov and J. M. Sabatier, "Vibration and sound signatures of human footsteps in buildings," *The Journal of the Acoustical Society of America*, vol. 120, pp. 762-768, 2006.
- [36] N. A, P. M, and N. R, "Cost Effective Immersive Room with Pressure Sensing Floor," in *IMEKO 20th TC3, 3rd TC16 and 1st TC22 International Conference*, 2007.
- [37] M. Sousa, A. Techmer, A. Steinhage, C. Lauterbach, and P. Lukowicz, "Human Tracking and Identification using a Sensitive Floor and Wearable Accelerometers," *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 166-171, 2013.
- [38] J. Yun, "User Identification Using Gait Patterns on UbiFloorII," *Sensors (Basel, Switzerland)*, vol. 11, pp. 2611-2639, 2011.
- [39] G. Qian, J. Zhang, and A. Kidané, "People Identification Using Floor Pressure Sensing and Analysis," *IEEE Sensors Journal*, vol. 10, pp. 1447-1460, 2010.
- [40] S. Pan, A. Bonde, J. Jing, L. Zhang, P. Zhang, and H. Y. Noh, "BOES: Building Occupancy Estimation System using sparse ambient vibration monitoring," 2014, p. NA.
- [41] J. M. Sabatier and A. E. Ekimov, "Range limitation for seismic footstep detection," 2008, pp. 69630V-69630V-12.
- [42] C. DeLoney, "Person Identification and Gender Recognition from Footstep Sound using Modulation Analysis," p. NA, 2008.
- [43] X. Li, R. J. Logan, and R. E. Pastore, "Perception of acoustic source characteristics: Walking sounds," *The Journal of the Acoustical Society of America*, vol. 90, pp. 3036-3049, 1991.
- [44] C. Velardo, J. Dugelay, M. Pleari, and P. Ariano, "Building the space scale or how to weigh a person with no gravity," in *2012 IEEE International Conference on Emerging Signal Processing Applications (ESPA)*, 2012, pp. 67-70.
- [45] D. Cao, C. Chen, D. Adjeroh, and A. Ross, "Predicting gender and weight from human metrology using a copula model," in *2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2012, pp. 162-169.
- [46] K. Krishan, "Establishing correlation of footprints with body weight—Forensic aspects," *Forensic Science International*, vol. 179, pp. 63-69, July 18, 2008 2008.
- [47] R. D. Labati, A. Genovese, V. Piuri, and F. Scotti, "Weight estimation from frame sequences using computational intelligence techniques," in *2012 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA)*, 2012, pp. 29-34.
- [48] T. Chau, "A review of analytical techniques for gait data. Part 1: fuzzy, statistical and fractal methods," *Gait & Posture*, vol. 13, pp. 49-66, 2001.
- [49] T. Chau, "A review of analytical techniques for gait data. Part 2: neural network and wavelet methods," *Gait & Posture*, vol. 13, pp. 102-120, April 2001 2001.

- [50] P. S. Huang, C. J. Harris, and M. S. Nixon, "Human gait recognition in canonical space using temporal templates," *Vision, Image and Signal Processing, IEE Proceedings -*, vol. 146, pp. 93-100, August 1999 1999.
- [51] Y.-C. Lin, B.-S. Yang, Y.-T. Lin, and Y.-T. Yang, "Human Recognition Based on Kinematics and Kinetics of Gait," *Journal of Medical and Biological Engineering*, vol. 31, pp. 225-263, 2010.
- [52] "Model: 393B04.," ed: PCB Piezotronics Inc., 2016.
- [53] P. A. Tarazaga, M. E. Kasarda, J. E. Hamilton, B. S. Joyce, R. Sarlo, D. B. Bales, *et al.*, "Design and Implementation of a Distributed Data Acquisition System for Vibration Monitoring in Virginia Tech's Goodwin Hall," *Experimental Techniques*, 2016.
- [54] R. J. Orr and G. D. Abowd, "The Smart Floor: A Mechanism for Natural User Identification and Tracking," 2000, pp. 275-276.
- [55] J. M. Hamilton, B. S. Joyce, M. E. Kasarda, and P. A. Tarazaga, "Characterization of Human Motion Through Floor Vibration," in *Dynamics of Civil Structures, Volume 4*, F. N. Catbas, Ed., ed: Springer International Publishing, 2014, pp. 163-170.
- [56] B. Parmanto, P. W. Munro, and H. R. Doyle, "Reducing Variance of Committee Prediction with Resampling Techniques," *Connection Science*, vol. 8, pp. 405-426, 1996.
- [57] J. R. Quinlan, "Induction of Decision Trees," *Machine Learning*, vol. 1, pp. 81-106, 1986.
- [58] K. J. Martin and D. S. Hirschberg, "On the Complexity of Learning Decision Trees," Department of Information and Computer Science, University of California Irvine.
- [59] D. E. Knuth, "Big Omicron and Big Omega and Big Theta," *SIGACT News*, vol. 8, pp. 18-24, April 1976 1976.
- [60] T. G. Dietterich, "Ensemble Methods in Machine Learning," in *Multiple Classifier Systems*, ed: Springer Berlin Heidelberg, 2000, pp. 1-15.
- [61] "BOOSTING (ADABOOST ALGORITHM)," MIT.
- [62] D. M. Hawkins, "The Problem of Overfitting," *Journal of Chemical Information and Computer Sciences*, vol. 44, pp. 1-12, January 1, 2004 2004.
- [63] *Understanding the Bias-Variance Tradeoff*.
- [64] L. Breiman, "Bagging Predictors," *Machine Learning*, vol. 24, pp. 123-140, 1996.
- [65] R. E. Schapire, "The strength of weak learnability," *Machine Learning*, vol. 5, pp. 197-227, 1990.
- [66] Y. Freund and R. E. Schapire, *A Short Introduction to Boosting*, 1999.
- [67] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals and Systems*, vol. 2, pp. 303-314, 1989.
- [68] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks*, vol. 6, pp. 525-533, 1993 1993.
- [69] S. Haykin, *A comprehensive foundation: Neural Networks*, 2004.
- [70] "SVM as a Convex Optimization Problem," Computer Science CMU.
- [71] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*: MIT Press, 2012.
- [72] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," 1992, pp. 144-152.
- [73] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY: Springer New York, 2000.
- [74] H. Q. Minh, P. Niyogi, and Y. Yao, "Mercer's Theorem, Feature Maps, and Smoothing."
- [75] C. Chih-Chung and L. Chich-Jen, "LIBSVM: A Library for Support Vector Machines," Department of Computer Science, National Taiwan University, Taipei, Taiwan March 4, 2013 2013.
- [76] J. Platt, "Sequential minimal optimization: A fast algorithm for training support vector machines," 1998.
- [77] J. B. Hayfron-Acquah, M. S. Nixon, and J. N. Carter, "Recognising human and animal movement by symmetry," in *International Conference on Image Processing. Proceedings*, 2001, pp. 290-293.
- [78] N. Bhatia and Vandana, "Survey of Nearest Neighbor Techniques," *arXiv:1007.0085 [cs]*, 2010-07-01 2010.
- [79] J. R. Cano, F. Herrera, and M. Lozano, "On the combination of evolutionary algorithms and stratified strategies for training set selection in data mining," *Applied Soft Computing*, vol. 6, pp. 323-332, March 2006 2006.
- [80] D. R. Wilson and T. R. Martinez, "Reduction Techniques for Instance-Based Learning Algorithms," *Machine Learning*, vol. 38, pp. 257-286, 2000-03 2000.
- [81] I. Czarnowski and P. Jędrzejowicz, "Data Reduction Algorithm for Machine Learning and Data Mining," in *New Frontiers in Applied Artificial Intelligence*, N. T. Nguyen, L. Borzowski, A. Grzech, and M. Ali, Eds., ed: Springer Berlin Heidelberg, 2008, pp. 276-285.
- [82] M. Kitagawa, "Walking in Graphs," ed: University of Texas, Dallas.

- [83] G. H. Golub and C. F. Van Loan, *Matrix computations* vol. 3: JHU Press, 2012.
- [84] S. Chaturantabut and D. C. Sorensen, "Discrete Empirical Interpolation for nonlinear model reduction," in *Proceedings of the 48th IEEE Conference on Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009*, 2009, pp. 4316-4321.
- [85] R. L. Grossman, C. Kamath, P. Kegelmeyer, V. Kumar, and R. Namburu, *Data Mining for Scientific and Engineering Applications*: Springer Science & Business Media, 2013.
- [86] J. G. F. Francis, "The QR Transformation—Part 2," *The Computer Journal*, vol. 4, pp. 332-345, 01/01/1962 1962.
- [87] D. C. Sorensen and M. Embree, "A DEIM Induced CUR Factorization," *arXiv:1407.5516 [cs, math]*, 2014-07-21 2014.
- [88] Z. Drmac and S. Gugercin, "A New Selection Operator for the Discrete Empirical Interpolation Method -- improved a priori error bound and extensions," *arXiv:1505.00370 [cs, math]*, 2015-05-02 2015.
- [89] Perinpanayagam, S and D. J. Ewins, "Test Strategy for Component Modal Test for Model Validation," 2004.
- [90] G. Chen, "FE MODEL VALIDATION FOR STRUCTURAL DYNAMICS," University of London, Imperial College of Science, Technology and Medicine.
- [91] N. Halko, P. G. Martinsson, and J. A. Tropp, "Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions," *SIAM Review*, vol. 53, pp. 217-288, Jun 2011 2011.
- [92] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar, "Rank Aggregation Methods for the Web," 2001, pp. 613–622.
- [93] S. Lin, "Rank aggregation methods," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, pp. 555-570, September 1, 2010 2010.
- [94] *RankAggreg.*
- [95] J. M. Hausdorff, Y. Ashkenazy, C.-K. Peng, P. C. Ivanov, H. E. Stanley, and A. L. Goldberger, "When human walking becomes random walking: fractal analysis and modeling of gait rhythm fluctuations," *Physica A: Statistical Mechanics and its Applications*, vol. 302, pp. 138-147, December 15, 2001 2001.
- [96] A. Itai and H. Yasukawa, "Footstep classification using wavelet decomposition," in *International Symposium on Communications and Information Technologies, 2007. ISCIT '07*, 2007, pp. 551-556.
- [97] P. G. H. Golub and D. C. Reinsch, "Singular value decomposition and least squares solutions," *Numerische Mathematik*, vol. 14, pp. 403-420, 1970-04 1970.
- [98] *Mean Body Weight, Height, and Body Mass Index, United States 1960–2002.*
- [99] A. Hyvärinen, J. Hurri, and P. O. Hoyer, "Independent Component Analysis," in *Natural Image Statistics*, ed: Springer London, 2009, pp. 151-175.
- [100] R. A. Fisher, "On the "Probable Error" of a Coefficient of Correlation Deduced from a Small Sample.," 1921 1921.
- [101] J. W. Tukey, "Comparing Individual Means in the Analysis of Variance," *Biometrics*, vol. 5, pp. 99-114, 1949 1949.
- [102] C. Eisenhart, "The Assumptions Underlying the Analysis of Variance," *Biometrics*, vol. 3, pp. 1-21, 1947 1947.

APPENDIX A: SAMPLE CODING

This appendix houses sample coding for the various algorithms used throughout this work. All of the codes are in MATLAB with the exception of the Rank Aggregation subtitle which is written using 'R'. The subtitles describe what is coded and the following code is explained to the level necessary for future of these methods.

Decision Trees

Bagged Decision Trees

```
TreeBAG_Model_L_Avg_10_Time_R=fitensemble("Observations",...
    "Labels",'Bag',num_learning_cycles,...
    'Tree','type','classification','Kfold',num_fold);
L_Avg_10_Time_R_Error=kfoldLoss(TreeBAG_Model_L_Avg_10_Time_R);
```

Boosted Decision Trees

```
TreeBoost_Model_L_Avg_10_Time_R=fitensemble("Observations",...
    "Labels",Boost_Method,num_learning_cycles,...
    'Tree','Kfold',num_fold);
L_Avg_10_Time_R_Error=kfoldLoss(TreeBoost_Model_L_Avg_10_Time_R);
```

SVMs

```
SVM_Model_L_Avg_10_Time_R=fitcsvm("Observations",...
    "Labels",'Kfold',num_fold,'KernelFunction',svm_kernel_name);
Error=kfoldLoss(SVM_Model_L_Avg_10_Time_R);
```

Neural Networks

```
net=patternnet(neurons_num); % Initialization of an empty net
net=train(net,"Observations","Labels"); % Training of the empty net
y=net("Observations withheld for testing"); % Testing of the net on
observations withheld specifically for testing
```

Peak Finder

The peak finder function used to find footsteps within this work is shown below. The inputs to this function are a walking trial y. The left and right step tolerances represent the number of samples to the left and right of the peak of the footstep which are included in the footstep. The samps_between_steps variable is the number of samples defined between each footstep (16000 in this work). The num_of_steps is the number of steps which should be defined.

```
%% Feature Extraction Secondary Functions
function
mean_step=Avg_Step_Finder(y,left_step_tol,right_step_tol,samps_between_steps,
num_of_steps)

% Finding Global Peaks and Valleys
% The below code finds peaks and valleys.

% PEAKS
[peak_mag,peak_loc]=findpeaks(y,'MINPEAKDISTANCE',...
    samps_between_steps,'SORTSTR','descend');
[~,I] = sort(peak_loc);
peak_loc=peak_loc(I);
```

```

peak_mag=peak_mag(I);
[~,b]=max(peak_mag);
[~,I] = sort(abs(peak_loc-peak_loc(b)));
peak_loc=peak_loc(I(1:num_of_steps));
peak_mag=peak_mag(I(1:num_of_steps));

[~,I] = sort(peak_loc);
peak_loc=peak_loc(I);
peak_mag=peak_mag(I);

% Local Step Data
for index=1:num_of_steps;

    if peak_loc(index)-left_step_tol>0 &&
peak_loc(index)+right_step_tol<peak_loc(end)
        step_data(:,index)=y(peak_loc(index)-left_step_tol:peak_loc(index)...
            +right_step_tol);
    else
        step_data(:,index)=zeros(right_step_tol+left_step_tol+1,1);
    end
end

% Rejecting of zero step_data column vectors (This arises when a peak is
% chosen that does not have a full step on the right or the left of it
% because of when it occurs, it also helps eliminate steps that are
% artificial at 0 samples created by the Butterworth filtering.

[~,I]=find(step_data==0);
I_remove=unique(I);
step_data(:,I_remove)=[];

% Mean of all selected steps

mean_step=mean(step_data');
mean_step=mean_step-mean_step(1);
end

```

SVD

```

[U,S,V]=svd(A,0); % Economy definition of the svd command operating on a
                  % matrix A

```

QR Transformation

```

[~,~,E]=qr(A,'vector'); % Ranking calculation through the qr ranking
                        % method of a matrix A

```

Rank Aggregation

The below code calculates the best overall rank (A) using the RankAggreg command. The ranking matrix is named 'G', there are 14 items to be ranked, the method is set to 'GA' for genetic algorithm. The optimized distance metric is set to 'spearman'. The rho is a step size parameter, N is the maximum number of iterations, and convIn is the number of iteration after which optimization stops when the error is unchanged.

```
A <-RankAggreg(G,14,method="GA",distance="Spearman",rho=.1,N=100,convIn=5)
```

Tukey Testing

The below MATLAB code was used to determine the statistical differences between the machine learning algorithm types for the work in Chapter 5. The MLType variable contains the classification errors for the four machine learning algorithm types when considering all other variables.

```
%% Machine Learning Type

MLType=[Bagged, Boosted, SVM, NN];
MLType_name={'Bagged','Boosted','SVM','NN'};

[~,~,stats]=anova1(MLType,MLType_name);

c=multcompare(stats);
```

Projection Coefficient

The function used to calculate the projection coefficients is given below.

```
function
[Proj_Coeff_mat,Proj_opt_sensors]=projection_coefficient(data,left_eig_used,i
ndependence)
%% Description
% This code returns the projection coefficient matrix for a specified
% number of trials for a provided data set with the structure of:
% data(trial #).value(:,sensor #)
% This function also plots the matrix in a 3d bar plot for the purpose of
% determining which sensors provide the most redundant information.

%% Definition of Projection Coefficient Matrix

y=data;
[U,S,~]=svd(y,0);

Sing=diag(S./S(1,1))

for i=1:size(U,2)
    for q = 1:size(U,2)
        a=y(:,i);
        b=U(:,q);
        Proj_Coeff_mat(i,q)=abs(dot(a,b))/(norm(a)*norm(b));
    end
end

% Scaling by the Singular Value
for k=1:size(U,2)
    A(:,k)=Proj_Coeff_mat(:,k)*Sing(k);
end

% Summation
A=sum(A,2);
```

```

% Sensor Optimization
% left_eig_used - is the number of which vector
% of the left eigenvector set
% will be used (most often the first vector)

if left_eig_used==0
    [~,index]=sort(A,independence);
else
    [~,index]=sort(Proj_Coeff_mat(:,left_eig_used),independence);
end

Proj_opt_sensors=index;
clear index

```

APPENDIX B: FORMULATION OF SVD-CR RANK AGGREGATION METHOD

This appendix describes in detail and example the formulation of the Singular Value Decomposition – Count Ranking method proposed in this thesis. The rank aggregation method is used to combine multiple ranking of accelerometers into a best overall ranking of accelerometers. The rankings of sensors that are combined result from the accelerometer selection – instance selection methods (AS-ISMs) used in this thesis Discrete Empirical Interpolation Method (DEIM), an alternate formulation of DEIM: Q-DEIM, and Projection Coefficients (PCs).

	1	2	3	4	5
1	12	12	12	4	12
2	9	4	8	7	9
3	7	7	3	11	7
4	2	9	9	8	2
5	3	3	5	9	4
6	14	2	13	3	3
7	4	11	2	6	5
8	8	13	4	13	11
9	11	8	7	12	13
10	13	6	6	5	6
11	6	5	11	1	8
12	5	14	14	14	14
13	1	1	1	2	1
14	10	10	10	10	10

The matrix (A – 14 x 60) represents the output from only of the three reduction methods (i.e. DEIM, dime, Projection coefficients). The columns represent walking trials (there are 60 of them) and the rows represent the rank of any sensor (from 1st independent to 14th most redundant). The entries are the sensors used in the test. For instance, sensor 12 was the first ranked sensor in trials 1-3 and 5.

	1	2	3	4	5
1	0	0	0	2	0
2	0	6	2	11	0
3	0	1	9	7	0
4	0	3	11	8	4
5	2	4	13	10	2
6	3	3	12	7	0
7	0	2	5	5	10
8	0	8	5	3	5
9	3	4	0	2	4
10	1	9	0	3	7
11	10	6	2	0	6
12	7	6	0	2	16
13	34	8	1	0	6
14	0	0	0	0	0

The count matrix (Count – 14 x 14) is a count of number of time a specific sensor appears in a specific row. The column is now the sensors used in the test and the row remains the ranking from 1st (most independent) to 14th (most redundant). For instance, sensor 3 was the fifth ranked sensor 13 times.

	1
1	1
2	0.2639
3	0.1530
4	0.0930
5	0.0786
6	0.0676
7	0.0566
8	0.0555
9	0.0464
10	0.0402
11	0.0387
12	0.0347
13	0.0313
14	0.0104

The average singular values for all walking trials (a – 14 x 1).

	1	2	3	4	5
1	0	0	0	2	0
2	0	1.5835	0.5278	2.9030	0
3	0	0.1530	1.3774	1.0713	0
4	0	0.2791	1.0235	0.7443	0.3722
5	0.1571	0.3143	1.0213	0.7856	0.1571
6	0.2029	0.2029	0.8115	0.4734	0
7	0	0.1132	0.2829	0.2829	0.5659
8	0	0.4442	0.2776	0.1666	0.2776
9	0.1391	0.1854	0	0.0927	0.1854
10	0.0402	0.3616	0	0.1205	0.2812
11	0.3872	0.2323	0.0774	0	0.2323
12	0.2428	0.2081	0	0.0694	0.5550
13	1.0645	0.2505	0.0313	0	0.1879
14	0	0	0	0	0

The Weighted_Count matrix (14 x 14) is the count matrix with each row scaled by the corresponding singular value (in that row). This method handicaps the lowest ranked rows (i.e. 14) as it has the lowest singular value score.

The matrix R (not pictured) is the summation along the columns.

I =

Columns 1 through 5

12 7 4 9 8

Columns 6 through 10

3 2 13 11 5

Columns 11 through 14

6 14 1 10

The Overall_Ranking (I – 1 x 14) is the best overall ranking of the 14 sensors as solved for by this method. It results from sorting the from largest to smallest of the R matrix.

APPENDIX C: ANALYSIS FOR STATISTICAL SIGNIFICANCE

There are many data treatments included in this thesis work. A data treatment is defined as a method that altered the original walking database to produce a characteristic classification such as machine learning algorithm, feature type, instance reduction method, etc. These data treatments must be fairly compared using statistics to offer suggestions as to which treatments were most effective in producing a quality characteristic classifier. To complete this statistical determination, an analysis of variance (ANOVA) is completed followed by a multiple comparison test, known as the Tukey test. During this analysis, all other variables are considered. That is, using the statistical methods to determine which machine learning algorithm produces the most accurate classifier considering all other types of data treatment (e.g., feature type, instance reduction method, etc.,).

An ANOVA is used to analyze the differences in the mean error for each data treatment and ultimately if a significant difference exists between treatment types [100]. The analyzed mean errors for each data treatment are from the testing stage. In an ANOVA, a null hypothesis is tested and is either accepted or rejected. The null hypothesis is that all treatments will result in the same average classification error. That is, there is no difference in classification error between using any particular treatment to classify. An ANOVA rejects the null hypothesis if the recorded mean error is statistically improbable under the assumption that all the means are the same (null hypothesis). The threshold to determine what is statistically significant is chosen to be 95 %, which is customary in an ANOVA. In summary, the ANOVA allows for understanding if there is a difference in performance of specific treatments of data.

Following an ANOVA, which established a difference between groups, a multiple comparison test allowed for the determination of which treatment was better for characteristic classification. A Tukey multiple comparison test is chosen for the analysis as it has been proven to be robust to recognizing existing differences between groups when the number of groups small number and is a commonly used comparison test [101]. The Tukey test compares all pairs of mean errors for similarity within some confidence interval. This confidence level was again chosen as 95 % in this analysis.

Certain assumption must be met to use both an ANOVA and a Tukey test. The assumptions of the ANOVA and Tukey test are given in Eisenhart [102] and addressed below:

- The observations are independent. This work meets this assumption because the gait of each walker is independent of another.
- The distribution of the group means is normal. A reasonable assumption as the mean errors produced in each analysis are at least 25 in number. This 25 sample requirement is the rule of thumb to assume a normal distribution by invoking the Central Limit Theorem (CLT). The CLT essentially states that sampling from any distribution behaves effectively normally when normalized appropriately [99].

- There is equal within-treatment variance across the treatments. There is no reason to believe that, for example, the bagged decision trees would have more variance than the neural network model meaning that this work meets this assumption.

The statistical analysis was completed using the MATLAB ‘anova1’ and ‘multcompare’ commands. Appendix A contains sample coding. These techniques are used in Chapter 5 to determine the best performing data treatments for gender classification considering:

- (a) machine learning algorithm,
- (b) feature type,
- (c) domain type (i.e. time or frequency), and
- (d) walking direction.

These techniques are used in Chapter 6 to conclude the best performing data treatments for gender and weight classification considering:

- (i) instance selection method,
- (ii) walking direction,
- (iii) group type,
- (iv) machine learning algorithm type, and
- (v) feature type.

Chapter 6.1.1.1 will formally introduce the reduction and grouping types.

The statistical analysis allows for the determination of the most appropriate data treatments for classification. While this chapter is short in length, the methods discussed within it are integral for answering the question of, “how to developed the best classifier?” This determination of best data treatments finishes the required analysis discussion required before the presentation of the results, discussion, and conclusion of the technical work in Chapter 5 and Chapter 6. Chapter 5 examines the ability to classify walker gender through a machine learning approach. Chapter 6 examines gender and weight classification using instance reduction methods.

APPENDIX D: FULL GENDER CLASSIFICATION RESULTS TABLE

This appendix houses the full gender classification results that accompany the work in Chapter 5. The gender classification error for each data treatment (machine learning algorithm, feature type, domain type, and walking direction) is shown in Table 9 and Table 10. Table 9 shows the classification error for all of the time domain features and Table 10 shows the error for the frequency domain. Chapter 5.1 contains all of this data in a visual manner in Figure 27.

Table 9. Algorithm and Feature Set Error Summary in Time Domain. The entries of the table represent the classification error of a particular feature type (column) and machine learning algorithm (rows). The Westward and Eastward trials are both shown. The bolded entries are the best performers.

Time											
Machine Learning Algorithm	Sensor Averaged Step 10 Time		Sensor Averaged Step 5 Time		Sensor Averaged Step 3 Time		Multiple Sensor Step Time		Single Sensor Step Time		Average
	Westward	Eastward	Westward	Eastward	Westward	Eastward	Westward	Eastward	Westward	Eastward	
Bagged Decision Trees	23%	5%	17%	12%	22%	25%	45%	33%	33%	27%	24%
Boosted Decision Trees	20%	7%	10%	13%	30%	18%	45%	27%	27%	18%	22%
SVM - 24th Order Poly	23%	20%	15%	13%	17%	23%	55%	32%	30%	27%	26%
NN - 1 Neuron	50%	48%	50%	53%	50%	50%	37%	53%	53%	45%	49%
Average – Eastward/ Westward	29%	20%	23%	23%	30%	29%	45%	36%	36%	29%	30%
Average - Total	25%		23%		29%		41%		32%		30%

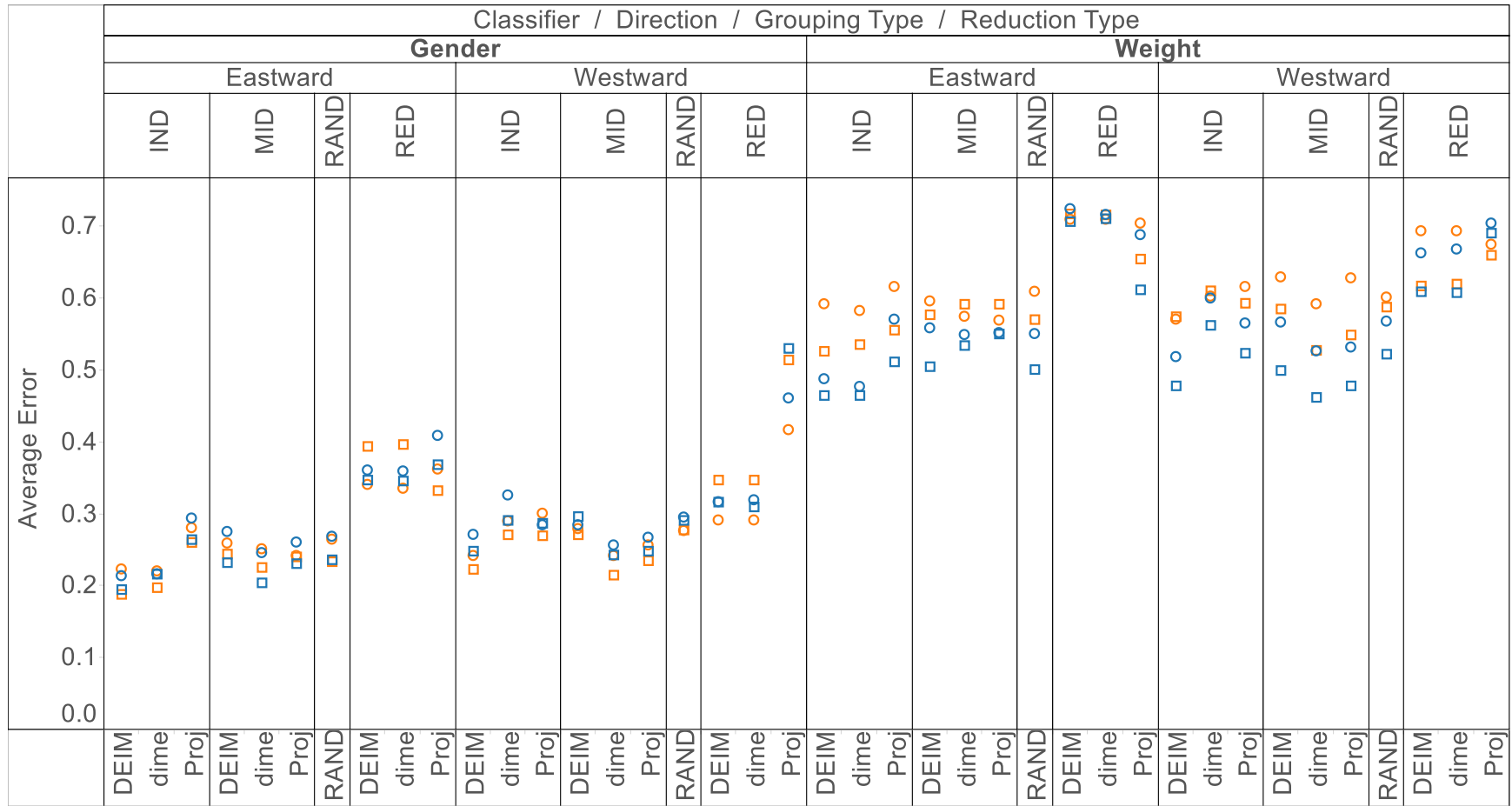
Table 10. Algorithm and Feature Set Error Summary in Frequency Domain. The entries of the table represent the classification error of a particular feature type (column) and machine learning algorithm (rows). The Westward and Eastward trials are both shown. The bolded entries are the best performers.

Frequency											
Machine Learning Algorithm	Sensor Averaged Step 10 Frequency		Sensor Averaged Step 5 Frequency		Sensor Averaged Step 3 Frequency		Multiple Sensor Step Frequency		Single Sensor Step Frequency		Average
	Westward	Eastward	Westward	Eastward	Westward	Eastward	Westward	Eastward	Westward	Eastward	
Bagged Decision Trees	25%	27%	27%	20%	30%	33%	53%	23%	35%	25%	30%
Boosted Decision Trees	27%	25%	30%	20%	32%	28%	35%	23%	33%	25%	28%
SVM - 24th Order Poly	32%	32%	33%	32%	28%	35%	38%	45%	45%	33%	35%
NN - 1 Neuron	50%	53%	53%	55%	48%	50%	55%	52%	48%	58%	52%
Average – Eastward/ Westward	33%	34%	36%	32%	35%	37%	45%	36%	40%	35%	36%
Average - Total	34%		34%		36%		41%		38%		36%

APPENDIX E: FULL INSTANCE SELECTION RESULTS

This appendix houses the results of the accelerometer selection – instance selection method (AS-ISM). The results are presented in both a figure and table. These results are vital to the understanding of the Algorithm Design Confirmation (ADC) methodology, but are rather large to include in the thesis chapter. These results expand upon the results show in Figure 32.

Figure 31 shows the full results of the AS-ISM study completed in Chapter 6. Chapter 6 in the thesis only shows the DEIM AS-ISM as this method produces the best results and the figure is difficult to digest without the explanation of all of the results of the statistical testing first. This figure further iterates the results discussed in Chapter 6. Namely, the poorer performance of the weight classifier in comparison to the gender classifier. Also, the poorer performance of the RED grouping type. This figure shows that the Projection Coefficients (PCs) perform worse than the DEIM and Q-DEIM AS-ISMs. These results are shown in the statistical testing of Figure 31. The same information in Figure 43 is tabulated in Table 11.



Data Type **Machine Learning Type**
 ○ SAS5 ■ Bagged ■ Boosted
 □ SAS10

Figure 43. Classification Error for all Data Treatments. This figure is organized into many columns that examine first the gender or weight classifiers, then walking direction, then grouping type, and finally the AS-ISM. Color includes detail.

Table 11. The tabulated form of the classification error for all data treatments. The ID is a unique identifier of the combination of the data treatments and the remaining columns have previously been discussed.

ID	Classifier	Machine Learning Type	Feature Type	Grouping Type	Direction	AS-ISM	Average Error
1	Gender	Bagged	SAS10	RAND	Westward	RAND	0.291
2	Gender	Bagged	SAS5	RAND	Westward	RAND	0.295
3	Gender	Boosted	SAS10	RAND	Westward	RAND	0.277
4	Gender	Boosted	SAS5	RAND	Westward	RAND	0.276
5	Weight	Bagged	SAS10	RAND	Westward	RAND	0.522
6	Weight	Bagged	SAS5	RAND	Westward	RAND	0.567
7	Weight	Boosted	SAS10	RAND	Westward	RAND	0.587
8	Weight	Boosted	SAS5	RAND	Westward	RAND	0.601
9	Gender	Bagged	SAS10	RAND	Eastward	RAND	0.236
10	Gender	Bagged	SAS5	RAND	Eastward	RAND	0.267
11	Gender	Boosted	SAS10	RAND	Eastward	RAND	0.233
12	Gender	Boosted	SAS5	RAND	Eastward	RAND	0.263
13	Weight	Bagged	SAS10	RAND	Eastward	RAND	0.500
14	Weight	Bagged	SAS5	RAND	Eastward	RAND	0.550
15	Weight	Boosted	SAS10	RAND	Eastward	RAND	0.570
16	Weight	Boosted	SAS5	RAND	Eastward	RAND	0.609
17	Gender	Bagged	SAS10	IND	Westward	DEIM	0.248
18	Gender	Bagged	SAS5	IND	Westward	DEIM	0.270
19	Gender	Boosted	SAS10	IND	Westward	DEIM	0.222
20	Gender	Boosted	SAS5	IND	Westward	DEIM	0.241
21	Weight	Bagged	SAS10	IND	Westward	DEIM	0.478
22	Weight	Bagged	SAS5	IND	Westward	DEIM	0.517
23	Weight	Boosted	SAS10	IND	Westward	DEIM	0.574
24	Weight	Boosted	SAS5	IND	Westward	DEIM	0.570
25	Gender	Bagged	SAS10	IND	Eastward	DEIM	0.195
26	Gender	Bagged	SAS5	IND	Eastward	DEIM	0.213

27	Gender	Boosted	SAS10	IND	Eastward	DEIM	0.188
28	Gender	Boosted	SAS5	IND	Eastward	DEIM	0.223
29	Weight	Bagged	SAS10	IND	Eastward	DEIM	0.464
30	Weight	Bagged	SAS5	IND	Eastward	DEIM	0.487
31	Weight	Boosted	SAS10	IND	Eastward	DEIM	0.526
32	Weight	Boosted	SAS5	IND	Eastward	DEIM	0.591
33	Gender	Bagged	SAS10	RED	Westward	DEIM	0.316
34	Gender	Bagged	SAS5	RED	Westward	DEIM	0.316
35	Gender	Boosted	SAS10	RED	Westward	DEIM	0.347
36	Gender	Boosted	SAS5	RED	Westward	DEIM	0.290
37	Weight	Bagged	SAS10	RED	Westward	DEIM	0.609
38	Weight	Bagged	SAS5	RED	Westward	DEIM	0.662
39	Weight	Boosted	SAS10	RED	Westward	DEIM	0.617
40	Weight	Boosted	SAS5	RED	Westward	DEIM	0.693
41	Gender	Bagged	SAS10	RED	Eastward	DEIM	0.346
42	Gender	Bagged	SAS5	RED	Eastward	DEIM	0.360
43	Gender	Boosted	SAS10	RED	Eastward	DEIM	0.393
44	Gender	Boosted	SAS5	RED	Eastward	DEIM	0.340
45	Weight	Bagged	SAS10	RED	Eastward	DEIM	0.707
46	Weight	Bagged	SAS5	RED	Eastward	DEIM	0.724
47	Weight	Boosted	SAS10	RED	Eastward	DEIM	0.717
48	Weight	Boosted	SAS5	RED	Eastward	DEIM	0.709
49	Gender	Bagged	SAS10	IND	Westward	Q-DEIM	0.291
50	Gender	Bagged	SAS5	IND	Westward	Q-DEIM	0.326
51	Gender	Boosted	SAS10	IND	Westward	Q-DEIM	0.270
52	Gender	Boosted	SAS5	IND	Westward	Q-DEIM	0.289
53	Weight	Bagged	SAS10	IND	Westward	Q-DEIM	0.562
54	Weight	Bagged	SAS5	IND	Westward	Q-DEIM	0.600
55	Weight	Boosted	SAS10	IND	Westward	Q-DEIM	0.610

56	Weight	Boosted	SAS5	IND	Westward	Q-DEIM	0.602
57	Gender	Bagged	SAS10	IND	Eastward	Q-DEIM	0.216
58	Gender	Bagged	SAS5	IND	Eastward	Q-DEIM	0.216
59	Gender	Boosted	SAS10	IND	Eastward	Q-DEIM	0.196
60	Gender	Boosted	SAS5	IND	Eastward	Q-DEIM	0.220
61	Weight	Bagged	SAS10	IND	Eastward	Q-DEIM	0.465
62	Weight	Bagged	SAS5	IND	Eastward	Q-DEIM	0.476
63	Weight	Boosted	SAS10	IND	Eastward	Q-DEIM	0.535
64	Weight	Boosted	SAS5	IND	Eastward	Q-DEIM	0.582
65	Gender	Bagged	SAS10	RED	Westward	Q-DEIM	0.309
66	Gender	Bagged	SAS5	RED	Westward	Q-DEIM	0.318
67	Gender	Boosted	SAS10	RED	Westward	Q-DEIM	0.347
68	Gender	Boosted	SAS5	RED	Westward	Q-DEIM	0.290
69	Weight	Bagged	SAS10	RED	Westward	Q-DEIM	0.607
70	Weight	Bagged	SAS5	RED	Westward	Q-DEIM	0.667
71	Weight	Boosted	SAS10	RED	Westward	Q-DEIM	0.620
72	Weight	Boosted	SAS5	RED	Westward	Q-DEIM	0.693
73	Gender	Bagged	SAS10	RED	Eastward	Q-DEIM	0.345
74	Gender	Bagged	SAS5	RED	Eastward	Q-DEIM	0.359
75	Gender	Boosted	SAS10	RED	Eastward	Q-DEIM	0.396
76	Gender	Boosted	SAS5	RED	Eastward	Q-DEIM	0.335
77	Weight	Bagged	SAS10	RED	Eastward	Q-DEIM	0.711
78	Weight	Bagged	SAS5	RED	Eastward	Q-DEIM	0.716
79	Weight	Boosted	SAS10	RED	Eastward	Q-DEIM	0.715
80	Weight	Boosted	SAS5	RED	Eastward	Q-DEIM	0.709
81	Gender	Bagged	SAS10	IND	Westward	PC	0.287
82	Gender	Bagged	SAS5	IND	Westward	PC	0.283
83	Gender	Boosted	SAS10	IND	Westward	PC	0.269
84	Gender	Boosted	SAS5	IND	Westward	PC	0.299

85	Weight	Bagged	SAS10	IND	Westward	PC	0.523
86	Weight	Bagged	SAS5	IND	Westward	PC	0.564
87	Weight	Boosted	SAS10	IND	Westward	PC	0.593
88	Weight	Boosted	SAS5	IND	Westward	PC	0.615
89	Gender	Bagged	SAS10	IND	Eastward	PC	0.263
90	Gender	Bagged	SAS5	IND	Eastward	PC	0.294
91	Gender	Boosted	SAS10	IND	Eastward	PC	0.260
92	Gender	Boosted	SAS5	IND	Eastward	PC	0.281
93	Weight	Bagged	SAS10	IND	Eastward	PC	0.511
94	Weight	Bagged	SAS5	IND	Eastward	PC	0.569
95	Weight	Boosted	SAS10	IND	Eastward	PC	0.555
96	Weight	Boosted	SAS5	IND	Eastward	PC	0.615
97	Gender	Bagged	SAS10	RED	Westward	PC	0.530
98	Gender	Bagged	SAS5	RED	Westward	PC	0.460
99	Gender	Boosted	SAS10	RED	Westward	PC	0.514
100	Gender	Boosted	SAS5	RED	Westward	PC	0.416
101	Weight	Bagged	SAS10	RED	Westward	PC	0.690
102	Weight	Bagged	SAS5	RED	Westward	PC	0.703
103	Weight	Boosted	SAS10	RED	Westward	PC	0.660
104	Weight	Boosted	SAS5	RED	Westward	PC	0.674
105	Gender	Bagged	SAS10	RED	Eastward	PC	0.368
106	Gender	Bagged	SAS5	RED	Eastward	PC	0.408
107	Gender	Boosted	SAS10	RED	Eastward	PC	0.332
108	Gender	Boosted	SAS5	RED	Eastward	PC	0.361
109	Weight	Bagged	SAS10	RED	Eastward	PC	0.612
110	Weight	Bagged	SAS5	RED	Eastward	PC	0.688
111	Weight	Boosted	SAS10	RED	Eastward	PC	0.655
112	Weight	Boosted	SAS5	RED	Eastward	PC	0.703
113	Gender	Bagged	SAS10	MID	Westward	DEIM	0.296

114	Gender	Bagged	SAS5	MID	Westward	DEIM	0.283
115	Gender	Boosted	SAS10	MID	Westward	DEIM	0.270
116	Gender	Boosted	SAS5	MID	Westward	DEIM	0.278
117	Weight	Bagged	SAS10	MID	Westward	DEIM	0.499
118	Weight	Bagged	SAS5	MID	Westward	DEIM	0.566
119	Weight	Boosted	SAS10	MID	Westward	DEIM	0.584
120	Weight	Boosted	SAS5	MID	Westward	DEIM	0.629
121	Gender	Bagged	SAS10	MID	Eastward	DEIM	0.232
122	Gender	Bagged	SAS5	MID	Eastward	DEIM	0.274
123	Gender	Boosted	SAS10	MID	Eastward	DEIM	0.244
124	Gender	Boosted	SAS5	MID	Eastward	DEIM	0.259
125	Weight	Bagged	SAS10	MID	Eastward	DEIM	0.504
126	Weight	Bagged	SAS5	MID	Eastward	DEIM	0.558
127	Weight	Boosted	SAS10	MID	Eastward	DEIM	0.577
128	Weight	Boosted	SAS5	MID	Eastward	DEIM	0.595
129	Gender	Bagged	SAS10	MID	Westward	Q-DEIM	0.243
130	Gender	Bagged	SAS5	MID	Westward	Q-DEIM	0.256
131	Gender	Boosted	SAS10	MID	Westward	Q-DEIM	0.215
132	Gender	Boosted	SAS5	MID	Westward	Q-DEIM	0.242
133	Weight	Bagged	SAS10	MID	Westward	Q-DEIM	0.462
134	Weight	Bagged	SAS5	MID	Westward	Q-DEIM	0.526
135	Weight	Boosted	SAS10	MID	Westward	Q-DEIM	0.527
136	Weight	Boosted	SAS5	MID	Westward	Q-DEIM	0.592
137	Gender	Bagged	SAS10	MID	Eastward	Q-DEIM	0.203
138	Gender	Bagged	SAS5	MID	Eastward	Q-DEIM	0.245
139	Gender	Boosted	SAS10	MID	Eastward	Q-DEIM	0.226
140	Gender	Boosted	SAS5	MID	Eastward	Q-DEIM	0.250
141	Weight	Bagged	SAS10	MID	Eastward	Q-DEIM	0.533
142	Weight	Bagged	SAS5	MID	Eastward	Q-DEIM	0.549

143	Weight	Boosted	SAS10	MID	Eastward	Q-DEIM	0.591
144	Weight	Boosted	SAS5	MID	Eastward	Q-DEIM	0.573
145	Gender	Bagged	SAS10	MID	Westward	PC	0.248
146	Gender	Bagged	SAS5	MID	Westward	PC	0.267
147	Gender	Boosted	SAS10	MID	Westward	PC	0.234
148	Gender	Boosted	SAS5	MID	Westward	PC	0.256
149	Weight	Bagged	SAS10	MID	Westward	PC	0.477
150	Weight	Bagged	SAS5	MID	Westward	PC	0.531
151	Weight	Boosted	SAS10	MID	Westward	PC	0.549
152	Weight	Boosted	SAS5	MID	Westward	PC	0.627
153	Gender	Bagged	SAS10	MID	Eastward	PC	0.231
154	Gender	Bagged	SAS5	MID	Eastward	PC	0.260
155	Gender	Boosted	SAS10	MID	Eastward	PC	0.240
156	Gender	Boosted	SAS5	MID	Eastward	PC	0.241
157	Weight	Bagged	SAS10	MID	Eastward	PC	0.550
158	Weight	Bagged	SAS5	MID	Eastward	PC	0.551
159	Weight	Boosted	SAS10	MID	Eastward	PC	0.591
160	Weight	Boosted	SAS5	MID	Eastward	PC	0.568

APPENDIX F: FIGURE AND IRB FIGURE PERMISSIONS

The majority of the figures in this work were created for the purpose of this work by the author. Figures 3 and 4 of the thesis were reproduced from publications and the permissions for those figures are given below. Figure 1, 2, 5, 6, and 7 were developed by the Virginia Tech Smart Infrastructure Laboratory and have been reproduced with the permission of the VTSIL director and co-chair for this thesis Dr. Pablo Tarazaga. The remainder of the figures were developed by the author.

IRB # 15-681 was used to advise and oversee the collection of all of the human gait information collected in this work. Attached there is the approval letter for this IRB.

**ELSEVIER LICENSE
TERMS AND CONDITIONS**

Apr 14, 2016

This is a License Agreement between Dustin Bales ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard, Langford Lane Kidlington, Oxford, OX5 1GB, UK
Registered Company Number	1982084
Customer name	Dustin Bales
Customer address	1917 Tabor Village Drive BLACKSBURG, VA 24060
License number	3847661383083
License date	Apr 14, 2016
Licensed content publisher	Elsevier
Licensed content publication	Journal of Sound and Vibration
Licensed content title	Experimental identification and analytical modelling of human walking forces: Literature review
Licensed content author	V. Racic, A. Pavic, J.M.W. Brownjohn
Licensed content date	25 September 2009
Licensed content volume number	326
Licensed content issue number	1-2
Number of pages	49
Start Page	1
End Page	49
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	figure 1

Title of your thesis/dissertation	Characteristic Classification of Walkers via Underfloor Accelerometer Gait Measurements through Machine Learning
Expected completion date	May 2016
Estimated size (number of pages)	150
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.00 USD / 0.00 GBP
Total	0.00 USD

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be

deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation:** This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com>; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier

homepage at <http://www.elsevier.com> . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. For journal authors: the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
 - via their non-commercial person homepage or blog
 - by updating a preprint in arXiv or RePEc with the accepted manuscript
 - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
 - directly by providing copies to their students or to research collaborators for their personal use
 - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- after the embargo period
 - via non-commercial hosting platforms such as their institutional repository
 - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all

value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a [CrossMark logo](#), the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's [posting policy](#) for further information.

18. For book authors the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. Thesis/Dissertation: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our [open access license policy](#) for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the

Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by/4.0>.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-nd/4.0>.

Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. Other Conditions:

v1.8

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.


[Home](#)
[Create Account](#)
[Help](#)


Title: A floor sensor system for gait recognition

Conference Proceedings: Automatic Identification Technologies, 2005. Fourth IEEE Workshop on

Author: L. Middleton; A. A. Buss; A. Bazin; M. S. Nixon

Publisher: IEEE

Date: 17-18 Oct. 2005

Copyright © 2005, IEEE

LOGIN

If you're a **copyright.com** user, you can login to RightsLink using your copyright.com credentials. Already a **RightsLink** user or want to [learn more?](#)

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

- 1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
- 2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
- 3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

- 1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
- 2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
- 3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

[BACK](#)
[CLOSE WINDOW](#)

Copyright © 2016 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#). Comments? We would like to hear from you. E-mail us at customercare@copyright.com

MEMORANDUM

DATE: October 5, 2015

TO: Pablo Alberto Tarazaga, Dustin Bennett Bales, Mary E Kasarda, Jeffrey Poston, Vijaya Venkata Narasimha Sriram Malladi, Mico Woolard

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires July 29, 2020)

PROTOCOL TITLE: Human Subject Gait Measurement

IRB NUMBER: 15-681

Effective October 5, 2015, the Virginia Tech Institution Review Board (IRB) Chair, David M Moore, approved the New Application request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

<http://www.irb.vt.edu/pages/responsibilities.htm>

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As: **Expedited, under 45 CFR 46.110 category(ies) 7**

Protocol Approval Date: **October 5, 2015**

Protocol Expiration Date: **October 4, 2016**

Continuing Review Due Date*: **September 20, 2016**

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.

Date*	OSP Number	Sponsor	Grant Comparison Conducted?

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this IRB protocol is to cover any other grant proposals, please contact the IRB office (irbadmin@vt.edu) immediately.