Thesis for the Degree of Doctor of Philosophy

Human Activity Recognition Using A Single Tri-axial Accelerometer

Adil Mehmood Khan

Department of Computer Engineering
Graduate School
Kyung Hee University
Seoul, Korea

February, 2011

Thesis for the Degree of Doctor of Philosophy

Human Activity Recognition Using A Single Tri-axial Accelerometer

by

Adil Mehmood Khan

Supervised by

Prof. Young-Koo Lee, Ph.D.

Department of Computer Engineering
Graduate School
Kyung Hee University
Seoul, Korea

February, 2011

Human Activity Recognition Using A Single Tri-axial Accelerometer

by

Adil Mehmood Khan

Supervised by

Prof. Young-Koo Lee, Ph.D.

Submitted to the Department of Computer Engineering and the Faculty of the Graduate School of Kyung Hee University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Dissertation Committee:

Prof. Ok-Sam Chae, Ph.D. (Chairman).. Ok Sam Chae

Prof. Sungyoung Lee, Ph.D. ....

Prof. Tae-Seong Kim, Ph.D. ....

Prof. Brian J. d'Auriol, Ph.D. ....

Prof. Young-Koo Lee, Ph.D. .... Young-Koo Lee

Abstract

With the advent of miniaturized sensing technology, which can be body-worn, it is now possible to collect and store data on different aspects of human activities under the conditions of free living. This technology has the potential to be used in automated activity profiling systems which produce a continuous record of bodily activity patterns over extended periods of time. Such activity profiling systems are dependent on recognition algorithms which can effectively interpret body-worn sensor data and identify different activities.

The automated recognition of bodily activities using body-worn accelerometer data is a challenging area of work. Existing activity recognition systems suffer from several obvious practical limitations such as the number, location and nature of sensors that people will tolerate. Other issues include ease of use, discretion, cost, and the ability to perform daily activities unimpeded. Variations can result in the sensor's output for the same activity across different subjects and for the same individual. Errors can also arise due to variability in sensor signals caused by differences in sensor positioning and from environmental factors such as sensor temperature sensitivity and very little work has been done to validate the idea under the unsupervised real-world circumstances.

This dissertation presents an accurate and robust tri-axial accelerometer-based bodily activity recognition framework. The novelty of the system compared to the previous accelerometer-based bodily activity recognition systems lies in: 1) Unlike previous systems, this system employs a better mathematical model, developed using using stochastic time series analysis, to describe activity acceleration-data. It is shown that such a model is more appropriate as it fits the data well and can be computed in real-time. (2) The system uses a novel state-activity based classification scheme that employs the proposed model for recognizing a diverse set of physical activities with a high accuracy. This scheme is capable of distinguishing the activities for which the previous systems showed difficulty, such as sitting and standing, with a good accuracy. (3) The system also implements a multi-stage classification scheme that employs the proposed model for accelerometer's positions and attachment free activity recognition, offering better convenience for long-term recognition in free-living conditions. It allows users to carry the sensor in any pocket without attaching it firmly to any body part. 4) Finally, the system also implements a light-weight classification scheme that uses the proposed model for recognizing activities in real-time using an accelerometer equipped smartphone independent of phone's position on the human body. It is believed that such technology will turn future smartphones into really clever handsets which would be capable of understanding what people are doing at any moment of time, anticipating what they would do next, and providing services automatically and accordingly.

Acknowledgement

First and foremost, I render my humble and sincere thanks to the Almighty for showering His blessings in every possible form upon me. He gave me strength, courage, patience and introduced to me all those people who made my studies and stay in Korea a pleasant experience.

This dissertation represents a great deal of time and effort not only on my part, but also on part of my advisor, Prof. Young-Koo Lee and co-advisor Prof. Tae-Seong Kim. I am also obliged and thankful to Prof. Sungyoung Lee. I am grateful for the time and advice that you all provided me over the past five years. This thesis owes much of its contents to your ideas and guidance. You helped me in shaping up my research from day one, pushed me to get through the inevitable research setbacks, and encouraged me to achieve the best of my abilities.

I am also thankful to my other thesis committee members for providing insightful and constructive comments to improve the quality of this dissertation, especially to Prof. Brian J. d'Auriol. His constructive criticism on my work and insightful suggestions helped me in improving this dissertation a lot.

I would also like to thank all the current and former members of my lab for their kind support and for providing a pleasant working environment. Every lab member is worthy to be praised and I appreciate them all.

I have no words to express my gratitude to my family for their endless support, love and prayers. I would like to acknowledge the sacrifices made by my parents for my better education and upbringing.

I am very much thankful to my friends Dr. Uzair Ahmad, Dr. Syed M. K. Raazi, Dr. Tahir Rasheed, Dr. Syed Obaid Amin, Muhammad Shoaib Siddiqui, Faraz Idrees Khan, Dr. Sheikh Riaz Ahmed, Ozair Idrees Khan, Asad Masood Khattak, Zafar Khan, Zeeshan Pervez and Hidayath

Mirza, for making my time at Kyung Hee University as good as it was. Finally, I would like to thank my best friend, Ivonne Gutierrez Villa, for her love, support and encouragement that reinforced my spirits at some crucial junctures.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Human activity recognition has emerged as an active area of research over the past few years. It is an important and challenging field which can support many novel ubiquitous applications. These applications range from smart homes, just-in-time information systems for office workers, surveillance and interactive game interfaces to home healthcare. Activity recognition is a multi-disciplinary research area which shares connection with machine learning, artificial intelligence, machine perception, ubiquitous computing, human computer interaction, as well as psychology and sociology. Thus, it has been drawing increasing interest from researchers in a variety of fields.

The aim of an activity recognition system is to recognize the actions or activities of its users by unobtrusively observing the behavior of people and characteristics of their environments and take necessary actions in response. For example, by means of recognizing activities in real time, such systems could allow the development of just-in-time learning environments that educate and inform people by presenting information at the right time as they move through the environment. Knowing what a person is doing will help determine the best time to interrupt the occupant to present them with useful information or messages. Someone preparing dinner represents a good opportunity for a teaching system to show words in a foreign language related to cooking.

In a home environment, activity recognition systems can monitor users' activities over long periods of time in order to remind them to perform forgotten activities or complete actions such as taking medicine, help them recall information, or encourage them to act more safely [1]. In a hospital environment, such systems can remind a doctor or nurse to perform certain tests before operating. In a surveillance system, behavior model can be developed by means of recognized activities which can enable the system to predict the intent and motive of people as they interact with the environment. Moreover, in a production environment, such systems can ensure the quality of the product by monitoring the set of actions. Finally, these systems can also play a vital role in encouraging a healthy life-style among their users by suggesting small behavior modifications. For example, people can be encouraged to use stairs instead of an elevator or stand after a long period of sitting.

Humans are capable of understanding and interpreting what activities the people around them are performing. The ability of recognizing activities appears to be so simple and natural for ordinary people but in fact involves complicated task of sensing, learning and inference. Imagine the following scenario. It is 4:00 P.M in the afternoon on an ordinary day. A girl sees her father standing in his room right beside his desk with a glass of water in his hand. Through her past experience and knowledge on her father's medical history, she can immediately infer that her father is taking daily dose of medicine. However, recognizing this activity would be a great challenge for an automated system and a large number of other sensory evidences will be needed. Humans learn from their past experiences. However, all these functions of sensing the environments, learning from past experience, and applying knowledge for inference are still a great challenge for machines. Therefore, the goal of activity recognition research is to enable computers to have similar capabilities as humans for recognizing people's activities.

1.2 Approaches to Human Activity Recognition

The first step towards achieving the goal of recognizing activities of daily living is to equip activity recognitions systems with sensing capabilities. Three approaches have been mainly employed for this purpose: video based, environmental sensor based and wearable sensor based, as shown in Figure 1.1.

Video based systems: These systems employ video camera for tracking and physical activity recognition. This approach often works fine in laboratory but fails in achieving the same accuracy under a real home settings due to clutter, variable lighting, and highly varied activities that take place in natural environments [2]. Complexity of dealing with changes in the scene, such as lighting, multiple people, and clutter offers additional challenges. Moreover, sensors such as microphones and cameras are mostly expensive. Finally, since these devices commonly serve as

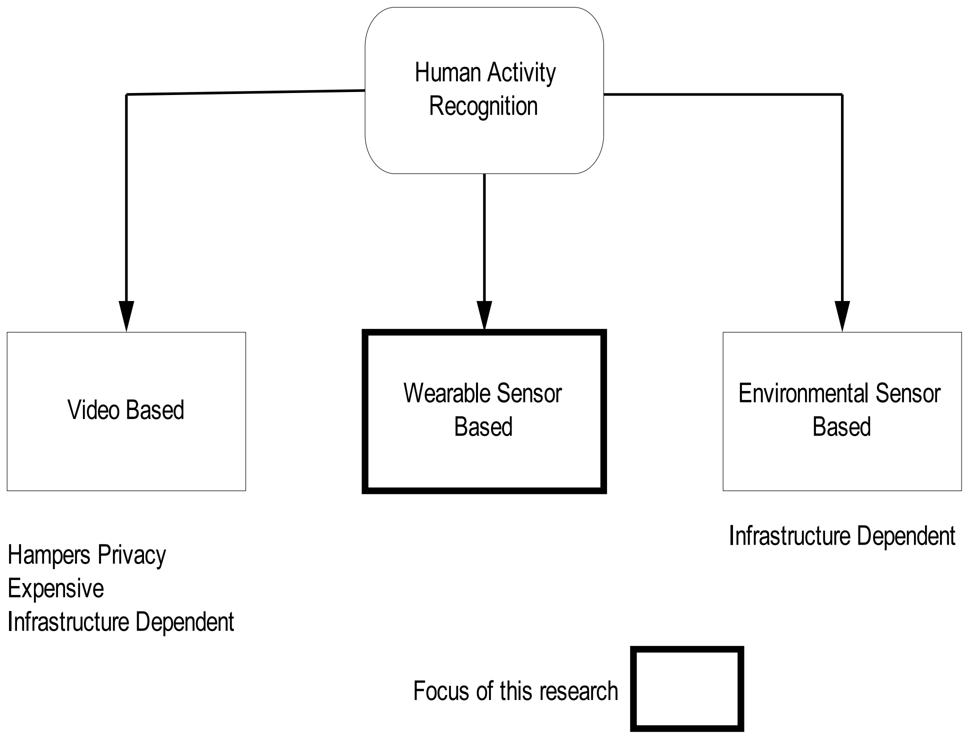

Figure 1.1: Three approaches employed for physical activity recognition

recording devices, they can also be perceived as a threat to privacy by some people.

Environmental sensor based systems: Such systems are developed to monitor the interaction between users and their home environment [2, 3]. This goal is achieved by distributing a number of ambient sensors, especially binary on-off state sensors, throughout the subject's living environment. The data gathered by these environmental sensors can be used to intelligently adapt the environment in the home for its inhabitants. Environmental sensor based systems passively monitor their occupants all day, every day, thus requiring no action on the part of the user to operate. A large number of parameters can be monitored in such systems, by employing a variety of sensors and the processing capabilities of a local PC. Ambient sensors, placed throughout the house, have fewer restrictions (size, weight, and power) than other types of sensors thus simplifying the overall system design. However, such systems are infrastructure dependent and cannot monitor a subject outside of the home setting. Also, they exhibit difficulties distinguishing between the monitored subject and other people in the home.

Wearable sensor based systems: Such systems are designed to be worn during normal daily activity to continually measure biomechanical and physiological data regardless of subject location and thus are an appropriate alternative for the recognition of daily human activities, especially bodily or physical activities [4]. Bodily activities require repetitive motion of the human body and are constrained, to a large extent, by the structure of the body. Examples are walking, running, scrubbing, and exercising. Wearable sensors are well suited to collecting data on daily physical activity patterns over an extended period of time as they can be integrated into clothing [5,6], jewelry [7,8], or worn as wearable devices. Since they are attached to the subjects they are monitoring and are independent of the infrastructure, wearable sensors can therefore measure physiological parameters which may not be measurable using environmental or video sensors. Moreover, such sensors are low-priced and unlike video sensors they are not considered as a threat to people's privacy.

A range of body-attached sensors including electromechanical switches, goniometers , accelerometers, gyroscopes, pedometers, and actometers, have been used to capture and analyze human movement in free-living subjects, as shown in Figure 1.2. Of these, accelerometers are becoming widely accepted as a useful tool for the assessment of human motion in clinical settings and free-living environments [9]. Accelerometers offer a number of advantages in monitoring of human movement. Their response to both frequency and intensity of movement makes them superior to actometers or pedometers, which are attenuated by impact or tilt. Some types of accelerometers can measure both tilt and body movement, and thus are superior to motion sensors that are incapable of measuring static characteristics. Lately, enhancements in microelectromechanical systems (MEMS) technology resulted in miniaturized and low cost accelerometers. These features have made possible the development of small, lightweight, portable systems that can be worn by a free-living subject without hindering movement. Thus accelerometery is emerging as a practical, inexpensive, and reliable method for capturing and analyzing daily physical activities [10]. In this thesis, a human activity recognition system is developed for the recognition of daily physical activities using a single wearable tri-axial accelerometer sensor.

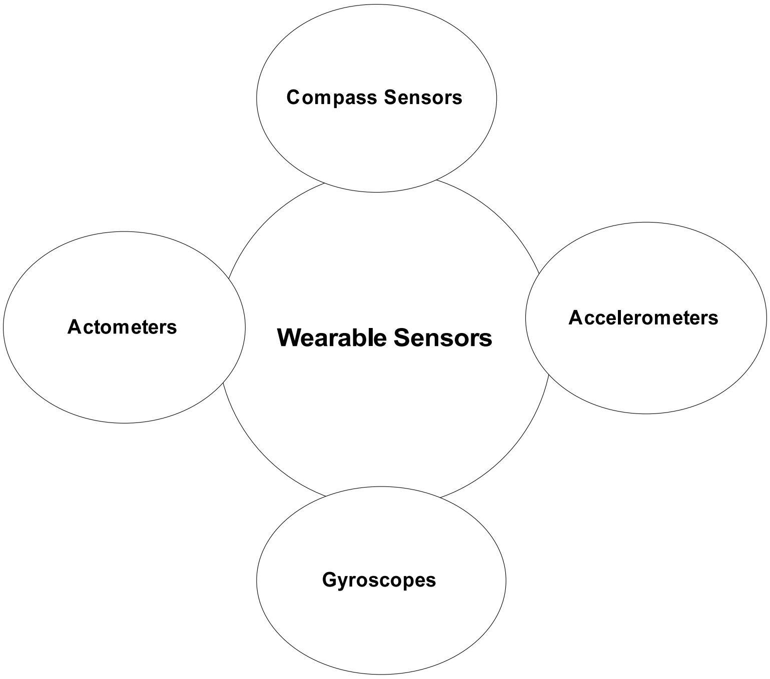

Figure 1.2: Different wearable sensors used to capture and analyze human movement

1.3 Challenges in Bodily Activity Recognition using Wearable Sensors

The automated recognition of daily physical activities using body-worn sensor (such as accelerometers) data is a challenging area of work. There exist several practical limitations such as the number, location and nature of sensors that people will tolerate. Apart from these obvious limitations, there are several other issues that directly impact the success of any given physical activity recognition systems. Factors which contribute to the complexity of the recognition task can categorized into following types [11].

1.3.1 Complexity of the Activities

In the field of wearable sensor based recognition of bodily activities, recognition algorithms can be evaluated on the basis of the complexity of the activities they recognize. The complexity of the activities can vary and depends on different factors including the number of activities, the types of activities and the complexity of the training data collected for those activities.

Number of activities: People perform a large number of different activities in daily life. Therefore, a human activity recognition system should be able to recognize a diverse and large set of activities. However, recognizing a small set of activities is usually easier than recognizing a large set of activities. The reason for this can be attributed to the fact that as the number of activities increases, the classifier has to discriminate among a larger set of activities, which is usually harder.

Types of activities: Activities which are static in nature including postures, such as lying and standing, are easier to recognize than the activities which are periodic in nature, such as running and walking. However, postures that are highly similar, such as sitting and standing, are also very hard to discriminate as they overlap significantly in the feature space. Furthermore, activities with high motion similarities, such as walking along the corridor, walking upstairs and walking downstairs, are also very hard to discriminate as such activities share high similarity in the feature space because of their similar movement patterns.

Moreover, recognizing a large number of activities having both highly different and similar characteristics at the same time makes the recognition problem even harder. In such cases, high similarity among activities is not uniform throughout the whole set of activities. In other words, a subset of activities shares high similarity among its activities but is very different from another subset. For example, sitting and standing are very similar (hard to distinguish), however, they are very different from walking (easily distinguishable).

Data collected for the activities: Training data for the activities to be recognized can be collected either in the laboratory or free-living conditions. Laboratory data are usually collected using a strict protocol. In other words, the activities are performed at the same speed and for the same duration by the participating subjects in constrained ways, whereas during the free-living conditions subjects might behave differently and in less constrained ways. Long-term out-of-lab monitoring means unsupervised, less-controlled and user-annotated data collection which brings along several challenges. The most important of these challenges include:

- Under such settings, subjects tend to annotate the data themselves without researcher's supervision. This results in unreliable annotations which can cause difficulty in classifier training and eventually degrade the classifier's recognition accuracy.

- There is no standard way to perform an activity. For example: 1) a person may lie down on a sofa in a manner that cannot be categorized to be either sitting or lying. 2) A person may perform dynamic activities, such as walking, at different pace at different times. In short, people perform activities in different ways which are hard to categorize. Therefore, activities for which the training and test data are collected in laboratory settings are usually easier to recognize than the ones for which the data have been collected in free-living condition.

1.3.2 Training Data Requirements

Recognition algorithms can be evaluated based on the type and the amount of the the training data that they require.

Subject independent recognition: In an ideal scenario, any activity recognition algorithm should be trained on a given subject population and then should recognize activities for unseen subjects, without requiring any training data from the new subjects. However, some previous works, such as [12], strongly suggest that subject independent recognition of activities is hard to achieve especially in the case of a diverse set of activities due to a high variability in the way people perform those activities.

Amount of training data for subject-dependent recognition: Previous work on bodily activity recognition using wearable accelerometers strongly suggests that the recognition algorithms perform better when trained with more person-specific training data. However, in case of large number of activities, providing this data can be time consuming and burdensome, so ideally training data requirements should be kept to minimum.

1.3.3 Sensor Requirements

The number of sensors employed for the recognition of activities, the types of sensors used, and their location on which they are placed on the human body can significantly impact the complexity of the recognition algorithm.

Number of sensors: Activity recognition systems that employ a small set of sensors to recognize activities are easier and convenient to use in real-world applications. Since a small number of sensors are used, fewer sensor signals are needed to be analyzed than systems that make use of large number of sensors. Consequently, systems with fewer sensors have lower computational requirements. However, the recognition accuracy of such systems is lower than the systems with large set of sensors as less information is available.

Location of sensors: Sensors are usually attached to different parts of the human body for collecting data on activities. Such configurations might be acceptable for short-term activity monitoring, however, they are infeasible for long-term activity monitoring as they take away the ability to perform daily activities unimpeded [13]. Any system which impedes subjects' daily physical activities or forces them into a fixed life pattern due to its size, communication methods or location is most likely to be rejected [11].

Thus an ideal system should allow its users to carry sensors freely in different pockets and should still be able to recognize activities with a high accuracy. In general, the output of any body-worn accelerometer depends on the position at which it is placed and can vary for different positions on a subject's body, even for the same activity. The output patterns for walking, for example, vary at three different positions as shown in Fig. 1.3. The high within-class variance caused by changes in orientation, magnitude, and frequency thus makes accelerometer's position free human activity recognition very challenging.

1.3.4 Real-time Constraints

Activity recognition algorithms, especially those running on hand-held devices, should be fast-enough and light-enough to be able to perform the recognition task in real-time, using as limited resources (such as memory and computational power) as possible. In other words, these systems should employ a small number of sensors, preferably a single sensor, to perform the recognition task. Systems that use multiple sensors need to analyze multiple data streams which increases the processing time and the complexity of such systems significantly.

Moreover, most approaches to activity classification using body-worn sensors involve a multi-stage process. Firstly, the sensor signal is divided into a number of small time segments, referred to as windows, each of which is considered sequentially. For each window, one or more features are derived to characterize the signal. These features are then used as input to a classification algorithm which associates each window with an activity. These mechanism of learning and inference (feature extraction and classification) should also be light-enough to be performed in real-time.

1.4 Limitations of Previous Systems

Majority of the accelerometer-based physical activity recognition systems developed in the past investigated the use of plurality of sensors attached at different sites on subject's body [4, 12, 14–23]. As mentioned earlier, this approach though capable of providing higher recognition rate is not feasible for long-term activity monitoring because of two or more different sites of attachments to the body and cable connections. Comparatively, a very small number of studies have investigated the use of a single accelerometer mounted at waist, sternum or back [24–33]. Such systems provided good recognition results for the basic activities including lying, standing, walking and running. However, they failed to exhibit the same accuracy for static activities such as standing and sitting, transitional activities such as lie-stand, sit-stand and stand-sit, and dynamic activities such as walking-downstairs, and walking-upstairs.

Most of the previous systems developed for the sake of bodily activity recognition, limited their scope to a small activity set. Few systems did try to recognize a large number of activities, however, their accuracy was low due to the problem of high similarity among activities,as mentioned above. A large number of features, both from frequency and time-domain, have been investigated in the previous systems with varying success rates. Frequency domain features require a large number of components to distinguish activities and thus require high computational power and time [4, 16, 19, 22, 27]. Time domain features, on the other hand, can be easily extracted in real-time and thus require less computational power [12, 15, 20, 24, 25, 29, 31, 34]. However, the recognition results using these features have not had high success rates. Finally, these features are calculated over long time-windows which reduce their ability to detect the short-duration movements, e.g., the transitions between sitting and standing or taking a couple of steps.

Moreover, an ideal activity classification scheme should work off-the-shelf. In other words, it should be able to use the data from a range of previous subjects to identify activities from an unseen individual. However, most of the times this is not possible and an intra-subject classification scheme is currently all that can be achieved for some problems. With this approach, sample training data are required for a given individual before classification can be performed.

Although the literature supports the fact that accelerometery has emerged as an effective and inexpensive mean to recognize physical activities, little work has been done to validate the idea under the unsupervised real-world circumstances. Majority of the prior work on physical activity recognition using acceleration signals relies on the data collected in supervised controlled laboratory settings. The researchers investigated a limited number of activities and collected data from a small number of subjects and often these subjects included the researchers themselves. The studies have shown very high success in recognizing the most prevalent everyday physical activities, such as sitting, lying, walking and running. However, when tested for long-term out-of-lab monitoring the recognition accuracy of these systems decreased significantly.

Almost all previous works require accelerometers to be firmly attached to subjects' bodies. Most studies employed multiple accelerometers attached at different sites [4, 12, 16, 17, 19–23], whereas others investigated the use of a single tri-axial accelerometer mounted at waist, chest, thigh, wrist, or sternum [24–33, 35, 36]. Such configurations would force subjects into a fixed life pattern and hinder their daily physical activities and thus make these systems impractical for long-term activity monitoring during unsupervised free living.

1.5 Study Goal and Methodology

In conclusion, physical activity recognition using body-worn accelerometers pose five main requirements. (1) The recognition system should recognize activities in real-time. This demands

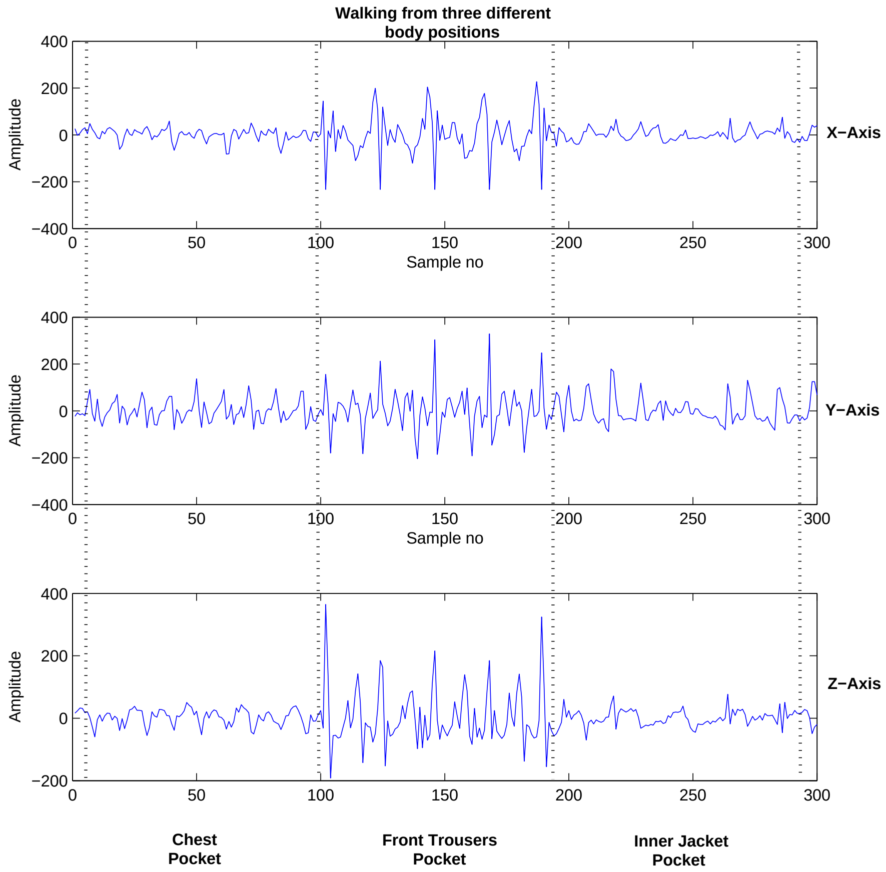

Figure 1.3: Sample acceleration signals for walking from three different positions.

that the features used for classification should be those that can be extracted in real-time. Moreover, short window lengths must be employed to avoid delayed response. Finally, the classification schemes should be simple, light-weight and computationally inexpensive to be able to run on hand-held devices. (2) The classifiers need to be able to discriminate the activities that exhibit significant similarities in their characteristics. This needs increasing the low between-class variance that results due to these similarities (3) The system should employ less sensors, preferably one, and recognize activities independent of sensor's position and its firm attachment on the human body. This requires that the high within-class variance that results due to placing the sensor on different positions must be decreased. (4) The recognition system should work off-the-shelf. In other words, it should recognize activities of the new subjects without going through the training phase again. This is very challenging as people perform the same activities differently, in terms of speed and intensity, and thus huge amount of variations could exist in their activity patterns. (5) Lastly, recognition system's accuracy should not get effected by the variations in the activity patterns for the same subject. This is also very hard to achieve as humans can perform the same activities in infinite different ways and it is difficult to collect enough training data to cater for this need.

The aim of this study was to implement a single triaxial accelerometer-based physical activity recognition system that fulfills only the first three requirements. It provides real-time information on physical activity by employing features that are well-suited to describe activity acceleration-signals and can also be computed in real-time, independent of sensor's position on the human body. Since the last two requirements are not targeted due to the level of difficulty involved, the activity data for training and validation were therefore collected from different subjects using a standard protocol in order to keep the variations in activity patterns, for different subjects and even for the same subject, to minimum. The system employs sliding window protocol, fixed window length for each activity, due to its simplicity and feasibility for real-time applications. The appropriate window length was chosen through careful analysis of the training data. The chosen window length provided good estimates of the features and was also short enough to not result in a delayed response. One limitation of this approach is that problems can arise if an activity lasts for shorter or longer time periods than the window length. However, it provided good approximation for the study objectives.

1.6 Contribution

As mentioned above, this works implemented an accurate and robust a single accelerometer based physical activity recognition system. Unlike previously developed bodily activity recognition systems which considered a small number of activities for recognition, fifteen bodily activities were considered for recognition in this study which are listed in Table 1.1. These activities include postures (such as sitting and standing), short-duration movements (such as sit-stand and stand-sit) and long-duration movements. Thus the chosen activity set was large and diverse with high similarity in posture and movements patterns among different activities. Therefore, achieving an effective discrimination among activities was harder.

The proposed system employed a single triaxial accelerometer for activity recognition. The use of a single accelerometer offered two advantages. Firstly, acceleration signals from a single sensor were needed to be analyzed and thus the computational requirements were very low. Secondly, carrying or attaching a single accelerometer on the human body was easier than attaching multiple sensors on different body-parts. The chances of hindering daily activities of people were, therefore, very low which made the system easier to use and more comfortable.

Unlike previous single accelerometer based bodily activity recognition systems, this system was capable of recognizing a set of large number physical activities with a high accuracy. The reason for system's high accuracy was due to the use of a novel augmented feature model for representing the activities. It is shown that activity-acceleration signals are in fact random signals generated by an autoregressive (AR) process and thus an AR-model is well-suited to represent the activities in the feature space due to the intensity and frequency characteristics of the these signals. The calculation of these features was performed using a fixed window length that not only provided good estimates of these features for both long and short-duration activities such sit-stand and stand-sit but was also feasible for real-time recognition. The feature model was then used for activity recognition in three different case-studies.

Firstly, the feature model was used to implement a two-stage classification scheme to resolve the problem of high similarity (similar postures and movements) among activities which makes the recognition of a diverse set of 15 bodily activities very hard. The accelerometer was firmly attached to a subject's chest in this case. The proposed classification scheme first separated groups of similar activities using their statistical characteristics. These groups of similar activities were then represented by the augmented feature model. The generated feature space for each activity group was then projected to a new feature space using projections that decreased the overlap between similar activities. The features from this projected feature space were then used for final classification.

Secondly, the feature model was used to implement accelerometer's position and attachment free activity classification scheme that was capable of recognizing about seven bodily activities independent of accelerometer's position on the human body. The accelerometer was freely placed into five different pockets instead of a firm attachment to any body-part. The aim was to recognize daily physical activities without posing any preconditions on accelerometer's position and orientation relative to a subject's body but maintaining the same high standards in terms of accuracy. The proposed classification method thus allows more flexibility and convenience in implementing a system for long-term activity monitoring in free-living conditions as it provides people with the freedom of carrying sensor freely in five different pockets.

Lastly, the proposed augmented feature space model was used for the case of physical activity recognition using accelerometer-equipped smartphones. In this case, special attention was paid to keep the computational requirements and the complexity of the classification scheme as minimum as possible. This was important as phones usually have lower computational power unlike normal desktop computers. The proposed classification scheme was validated using activity data collected from five body positions using a phone with a built-in tri-axial accelerometer.

In first case-study, the performance of the proposed activity recognition system was evaluated using datasets collected in both laboratory and free-living conditions. However, in second and third case-studies only data collected in free-living conditions were used for performance evaluation. In every evaluation study, the performance of the recognition algorithms is evaluated using both subject-dependent and subject independent training. The amount of training data required for the subject-dependent case is evaluated using different amounts of training data to determine the minimum amount required to get good recognition results.

Table 1.1: The Classified states and activities recognized in this study

State	Activity
Static	Lying Sitting Standing
Transitions	Lie-Stand Stand-Lie Lie-Sit Sit-Lie Sit-Stand Stand-Sit Walk-Stand Stand-Stand
Dynamic	Walking Walking-upstairs Walking-downstairs Running

1.7 Structure of the Dissertation

The thesis has been organized into seven chapters, as shown in Figure 1.4.

- Chapter 1 has presented a brief introduction of the concepts of HAR. It discussed the importance of HAR, its applications in different fields, its requirements and the factors that make it challenging. The problems associated with the existing wearable accelerometer-based physical activity recognition systems were summarized and finally, an overview of my contributions was given.

- Chapter 2 discusses the related work in the area of wearable accelerometer-based physical activity recognition in detail. Firstly, it describes different types of wearable-sensor based HAR systems, types of wearable-sensors and the reasons behind the high choice of accelerometers for bodily activity recognition. Secondly, this chapter discusses different kinds of physical phenomenas investigated, such as gait analysis and movement classification. Lastly, it discusses different parameters of pattern recognition techniques, including different types of features and the classification algorithms employed so far for the recognition of physical activities in both supervised laboratory and unsupervised real-home settings.

- Chapter 3 provides an overview of the research approach and methodology followed in this work. It also presents details on the sensor devices used and different data collection studies.

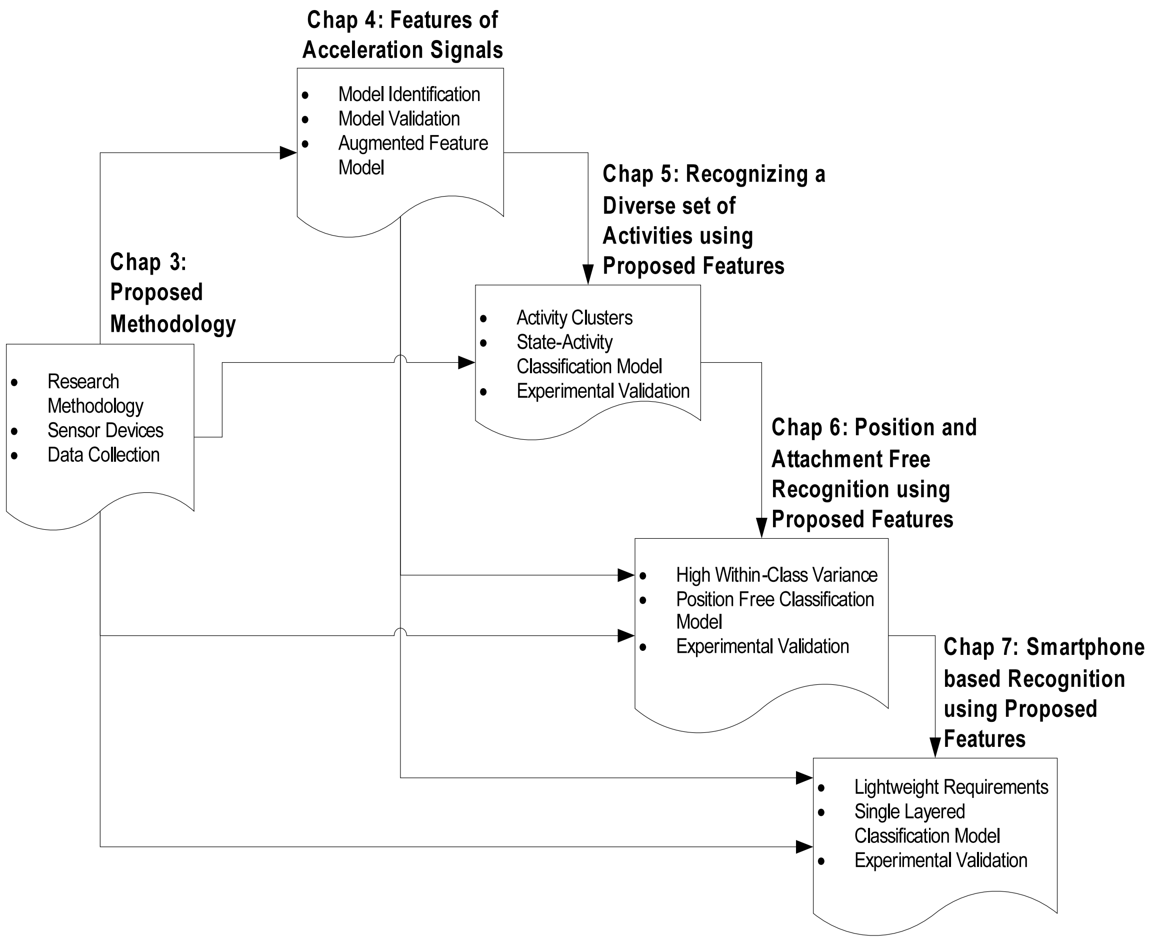

Figure 1.4: The structure of the dissertation

- Chapter 4 describes the proposed augmented feature space model. It explains in detail different stages of the analysis process, including model identification, parameter selection and model validation.

- Chapter 5 describes in detail the procedure employed for evaluating the use of the proposed augmented feature model for the classification of a large and diverse set of physical activities in both controlled and naturalistic settings. It also presents details and the reasoning behind the implementation of a multi-stage classification scheme used for the given classification task.

- Chapter 6 presents details on the implementation of the accelerometer's position-independent physical activity recognition scheme.

- Chapter 7 presents details on the implementation of the accelerometer's position free activity recognition system for accelerometer-equipped smartphones.

- Finally, chapter 8 discusses the application of the proposed system in ubiquitous healthcare. It also discusses the implementation and validation of the real-time personal life log system.

Chapter 2

Related Work

2.1 Types of Wearable Systems

Wearable systems are designed to be worn during normal daily activity to continually measure biomechanical and physiological data regardless of subject's location. Based on their data collection methods, wearable systems can be classified as: data processing, data logging, and data forwarding.

Data Processing Wearable Systems: These systems include a processing element such as a PDA or a microcontroller device. These consume more power than other types of wearable systems but they can provide realtime feedback to a user and do not require large amounts of data storage, as the raw data are typically summarized in real-time before storage or transmission. The use of summarized data also reduces costs by lowering the upload time to the server.

Data Logging Wearable Systems: Data logging are those which simply acquire data from the sensors and log these for offline analysis. They have the advantage of being able to monitor the subject regardless of their location. The disadvantage of data logging systems is that the subjects mobility patterns cannot be analyzed between uploads. If an alarming trend occurs between uploads it will not be discovered until that data is uploaded and analyzed on the pc. This problem will become more significant as improving memory technology increases the time between uploads.

Data Forwarding Wearable Systems: Data forwarding systems are those which simply acquire data from the sensors and forward these directly to a local computer for further analysis. These are used when the weight of the wearable system is a key factor, as a data storage or a data processing unit can be replaced by aminiature transmitter. However, data forwarding wearables, which typically use RF, Bluetooth, or WLAN, are range-limited, and therefore the data from the subject is not recorded when the subject is outside the range of the receiver. This makes data forwarding systems suitable for housebound subjects but not necessarily those who are independent and have the ability to move outside of the house.

2.2 Types of Wearables Sensors

A range of wearable sensors, shown in Figure 1.2, have been used to assess daily mobility levels in free-living subjects. Of these, accelerometers have emerged as the most useful tool for mobility assessment in both clinical and home environments. The reasons for such a wide acceptance of accelerometers are: Firstly, they can respond to both frequency and intensity of movement. This fact makes them superior to actometers or pedometers which are attenuated by impact or tilt [37]. Secondly, most of the widely available accelerometers can measure both the movement and the tilt which makes them superior to motion sensors that lack the capabilities of measuring these characteristics. Thirdly, due to enhancements in microelectromechanical systems (MEMS) technology, today's accelerometers are not only coming in small size and at a low-price but are also capable of demonstrating a high degree of reliability in measurement.

Accelerometers are devices which are capable of measuring the applied acceleration acting along a sensitive axis. Accelerometers use transducers for measuring acceleration. These come in different varieties, such as piezoelectric crystals, piezoresistive sensors, servo force balance transducers, electronic piezoelectric sensors and variable capacitance accelerometers. Some accelerometers require an external power supply whereas others do not. Moreover, some accelerometers are capable of responding to static accelerations (such as the acceleration due to gravity) whereas others do not.

Most physical activity recognition systems have used accelerometers which are capable of responding to acceleration due to gravity as well as acceleration due to movement. At any point in time, the output of such accelerometers is a linear combination of these two components, the acceleration component due to gravity (GA) and the acceleration component due to bodily motion (BA) [25]. Since these two components are linearly combined and overlap both in time and frequency, they cannot be easily separated. However, low pass filtering can be used to make approximation to the two components. Low pass filtering, when applied to an acceleration signal, separates the GA from the actual signal. GA can then be subtracted from the original signal to obtain the BA. Since most human movements occur between 0.3 and 3.5 Hz [38], most investigators have used a filter with a cut off frequency between 0.1 and 0.5 Hz to separate the two components.

2.3 Recognizing Bodily Activities using Accelerometers in the Past

2.3.1 Recognition Problems Investigated in Previous Works

The position at which the accelerometer is placed on the body is important in the measurement of bodily activity [9]. Normally, accelerometers are attached to the part of the body whose movement is being studied. For example, accelerometers attached to the thigh or ankle are used to study leg movement during walking [17, 39, 40], accelerometers attached to the wrist have been used in the measure of Parkinsonian bradykinesia [41].

However, in many cases, the intention is to study whole body movements. In such cases, some investigators have achieved this by using multiple instruments placed across the body [4, 12, 15, 17, 19, 42–44], while others have used a single instrument placed close to the centre of mass, which is located within the pelvis [24–26, 28–30, 45].

The accelerations generated during human movement vary across the body and depend on the activity being performed [9]. Accelerations increase in magnitude from the head to the ankle, and are generally greatest in the vertical direction, although the accelerations in the other two directions cannot be neglected [37].

The major energy band for daily activities is 0.33.5 Hz [38]. Although foot acceleration at heel strike can reach frequencies of up to 60 Hz, 98% of the acceleration power during bare foot walking is contained below 10 Hz and 99% is contained below 15 Hz [46]. Slightly higher frequencies occur during running, but most acceleration is below 18 Hz at the ankle. The maximum frequencies obtained decrease from the ankle to the head, and are greater in the vertical direction than in the transverse plane . In the light of such findings, it was concluded that in order to assess daily physical activity, accelerometers must be able to measure accelerations up to 12g in general, and up to 6g if they are attached at waist level, and that they must also be able to measure frequencies between 0 and 20 Hz [9, 34].

There are design trade-offs between the number of instruments that are used, the cost, the usability and the transferability of an ambulatory monitoring system [9]. The design of the recognition or monitoring systems is usually determined to a large extent by the purpose and duration of the monitoring. In short-term, supervised monitoring situations, large numbers of body-fixed sensors can be used to allow the collection of greater quantities of information, leading to very accurate assessments of movement, however, in long-term, unsupervised monitoring environments, subject compliance is essential if the system is to be used [9]. In this situation, the wearable instrumentation needs to be easy-to-use, comfortable and as unobtrusive as possible. One approach is to embed multiple sensors into an item of clothing [47]. The subject then has only to wear the item of clothing, and all of the sensors are attached in the correct locations. However, increasing the number of sensors increases the complexity and cost of the system. Additionally, items of clothing must be designed in a range of sizes in order to ensure a proper fit on all subjects. A simpler approach is to use only one instrument that is attached at a single location on the body. This greatly simplifies the design and use of the system, but it also reduces the quantity of information that is obtained about the movements. A review of the literature demonstrates that, despite this limitation, useful information can in fact be obtained from a single device attached near the centre of mass of the subject (see, for example, [27, 34]).

2.3.1.1 Gait Analysis

In addition to being an important skill for independent living, parameters of gait can provide indication of deteriorating functional ability and increasing falls risk. Walking speed is related to functional status [48] and is a predictor of falls [49].

It has been shown that simple parameters such as step and cycle time and stride symmetry can be determined during normal gait from waist, thigh or heel accelerations [43, 50]. Accelerometers attached to the legs have been used to enable automated extraction of temporal gait patterns including left and right heelstrikes and toe-offs [40]. Aminian et al [46] used two neural networks to estimate incline and walking speed during unconstrained walking using a triaxial accelerometer attached to the back and a uniaxial accelerometer attached to the top of the right heel. The standard deviation of the estimated incline was less than 2.6%, and the maximum of the coefficient of variation between speed estimation was 6%. However, after applying a similar approach, [51] reported that their system allowed accurate prediction of speed but not of incline during running.

Outdoor walking speed has been accurately measured using a combination of accelerometry and altimetry [52]. Studies have demonstrated thatwalking on level ground and walking on a stairway can be distinguished in the signals of a waist-mounted triaxial accelerometer [27].

The vertical acceleration component of the trunk- or back-mounted TA is the most important in the assessment of gait [34, 42, 43]. This is the component that is most sensitive to the presence of gait disorders and from which elements of the gait cycle can most easily be identified .

2.3.1.2 Sit-Stand and Stand-Sit transfers

The ability to rise from a chair is of fundamental importance for functional independence. Rising from a chair is regarded as the most mechanically demanding functional task undertaken during daily activities and is a prerequisite for gait. An inability to rise from a chair can prevent an otherwise functionally independent subject from independent living [53]. The ability to sit down in a controlled manner is of equal importance.

Little work has been reported using accelerometers for assessment of the sitstandsit movement. Sit-to-stand and stand-to-sit transitions can be automatically identified as periods of activity [24], and they can be classified by identifying the preceding and succeeding postures as sitting and standing [40, 42]. A preliminary study found a moderate correlation ($r = 0.537$) between the accelerometry characteristics of the sit-to-stand transfer measured at the waist and falls risk in 37 elderly subjects [54]. Other useful clinical information may be able to be obtained from the accelerometry signals of the sitstandsit movement, but this remains to be investigated.

2.3.1.3 Fall Detection

One of the biggest risks to the health and well being of the elderly is the risk of morbidity from injury, leading to functional dependence. Falls are a very serious risk for the elderly, particularly for those living in the community. In those aged over 65 years, two thirds of accidents are falls and, for example, in the general Australian community, accidents are the fifth leading cause of death, and one quarter of them are falls.

Accelerometry has been proposed as being suitable for falls detection in free-living subjects but there has been relatively little work done in this field to validate the method. The basic approach was first published in [55]. In this approach, a change in orientation from upright to lying that occurs immediately after an abrupt, large negative acceleration (due to impact) is indicative of a fall. Both of these conditions can be detected using an accelerometer that has a dc response, and have been incorporated into fall detection algorithms using an accelerometer.

However, little real data are available on the ability of an accelerometry-based system to detect falls in a community setting. This remains an area requiring further work.

2.3.1.4 Movement Classification

Accelerometry systems have been used to identify and classify sets of postures and activities. Most of these systems have used multiple sensors, some systems have used only accelerometers, while other systems have used accelerometers together with another type of sensor. The most common placement locations are the chest or waist and the thigh [15, 40, 42–44].

Algorithms for the detection of posture and motion patterns remain a crucial aspect of accelerometry, and the ability to achieve an adequate data reduction while still being able to differentiate between a variety of dynamic activities is still under investigation [43].

Systems have been developed to identify the postural orientation of a subject. Other systems have used accelerometers placed on the chest orwaist and the thigh to discriminate between postures and activities sitting, lying, standing, walking, stair climbing and cycling with a high degree of accuracy [15, 18, 40, 44] by first discriminating between activity and rest, and then between different resting postures, and different activities. Accelerometry systems using multiple instruments placed across the body have been also used to achieve classification of multiple activities and postures [17, 42, 43]. Accelerometry has also been used in conjunction with heart rate, GPS or gyroscopes to classify postures and activities.

The majority of movement classification systems have been custom designed for a specific domain of postures and activities. Although many of these systems have produced excellent results in classification of specific movements, there is still scope for the development of systems that are able to automatically identify and classify arbitrary movements performed in free-living conditions.

2.3.2 Components of the Recognition Algorithms

Most approaches to activity classification, using body-worn accelerometers, involve a multi-stage process. Firstly, the sensor signal is divided into a number of small time segments, referred to as windows, each of which is considered sequentially. For each window, one or more features are derived to characterize the signal. These features are then used as input to a classification algorithm which associates each window with an activity.

2.3.2.1 Kinds of Features Investigated

Previous physical activity recognition schemes have used a large variety of techniques to generate features in order to characterize windows of body-fixed acceleration data. Once generated, these features are then employed as inputs to classification schemes. In this section, we present a brief overview of different feature generation techniques.

Heuristic Features: Output of a body-worn accelerometer comprises two components. The first is the static acceleration. It results due to the effect of gravity and provides a measure of the inclination of the sensor to the vertical. The second is the dynamic acceleration. It is due to the acceleration of the body segment to which the accelerometer is attached. When the subject is at rest, the measured acceleration is equal to the cosine of the sensor orientation angle relative to the vertical. This angle, often known as tilt angle, is often used as an input to classification algorithms, particularly those designed to distinguish static postures [40] and identify postural transitions [4].

All movement patterns result in time varying segmental accelerations. Different methods have been used to derive certain heuristic features to quantify the amplitude of these accelerations. Before these features are derived, a high pass filter is applied to the signal to remove any baseline offset. These features includes the signal magnitude area [24], peak-to-peak acceleration [56], mean rectified value [17] and root mean square [15]. This type of feature is often used to differentiate between static and dynamic activity [24]

Time-domain Features: Some studies derived time-domain features directly from a window of acceleration data and are typically of statistical nature. Examples include the mean, median, variance, skewness, kurtosis [12, 20, 29]. Other studies employed high and low pass filters to separate accelerometer signals on a frequency basis. Means are calculated separately for the low frequency and rectified high frequency components which are then used as inputs to the classification schemes. Cross-correlation coefficients have also been used to quantify the similarity between acceleration signals from different axes on the same body segment and across different segments [12].

Frequency-domain Features: In order to derive frequency-domain features, the window of sensor data must first be transformed into the frequency domain, normally using a fast Fourier transform (FFT). The output of a FFT typically gives a set of basis coefficients which represent the amplitudes of the frequency components of the signal and the distribution of the signal energy. Different methods can then be used to characterize the spectral distribution from these coefficients. For example, median frequency [14] or a subset of the different FFT coefficients can be used [10]. Alternatively, information from a number of coefficients can be combined to give a single feature. Examples include spectral energy, which is the sum of the squared FFT coefficients [57], and frequency-domain entropy, which is the normalized information entropy of the FFT components [12]. This latter feature allows for differentiation between activities which have simple acceleration patterns and those with more complex patterns. For example, as cycling involves a uniform movement of the legs, a frequency-domain analysis of thigh acceleration shows a single dominant frequency. In contrast, running may result in more complex acceleration pattern and often displays many major FFT components. This difference leads to a much higher frequency-domain entropy for running in comparison to cycling [12].

Wavelet Analysis: Unlike Fourier analysis which can only be used to extract information on the frequency content of a signal, wavelet analysis can be used to investigate both time and frequency characteristics. Like Fourier analysis, wavelet analysis can be formulated via a continuous or discrete wavelet transform. Previous work on activity monitoring has employed the discrete wavelet transform (DWT). The discrete wavelet transform is normally implemented using the filter bank interpretation. In this approach, the original signal is successively decomposed into separate low and high pass filtered signals, referred to as approximation and detail coefficients respectively.

Wavelet analysis allows a body-worn sensor signal to be decomposed into a number of individual coefficients, each of which contains data on a specific frequency band. As these coefficients characterize the original signal along its entire length, they contain information on temporal changes in frequency content. Thus, unlike Fourier analysis, wavelet techniques can be used to analyse and characterize non-stationary signals (those in which frequency context changes over time).Wavelet analysis has been applied to three different types of problem within activity monitoring. These are signal enhancement [4], identification of activity transition points [58] and generation of timefrequency features subsequently used for classification [45, 58].

2.3.2.2 Feature Selection and Dimensionality Reduction Methods

People tend to perform the same movement in a variety of different ways which can lead to substantial variability in the features derived from body-fixed sensor data. Therefore, to achieve effective classification, identifying features with high discriminative ability is of high importance. A good feature set should show little variation between repetitions of the same movements and across different subjects but should vary considerably between different activities.

A number of different techniques, of varying complexity, have been used to select appropriate features for activity classification. These include visual and statistical analysis to assess the distribution of a given feature for different activities [59] and correlation-based feature selection [20]. Another method for feature selection is a forwardbackward search in which features are sequentially added and removed from a larger set. Optimal features are identified depending on the resulting classification accuracies for each feature subset [60].

As an alternative to selecting a subset of the existing features, it is often possible to combine the original features to define a new set of variables. There are two benefits associated with such a procedure. Firstly, the often unnecessarily large numbers of features, resulting from many sensors, can be reduced. Secondly, the new reduced set of variables frequently has better discriminative ability for classification problems. Principal component analysis (PCA) and Independent component analysis (ICA) are the two most commonly used dimensionality reduction techniques used in the field of activity monitoring using body-worn accelerometers [16].

2.3.2.3 Classifiers

Once features have been derived to characterize a window of sensor data, they are used as input to a classification algorithm. The degree of complexity of these different classification schemes varies from simple threshold-based schemes to more advanced algorithms, such as artificial neural networks or hidden Markov models. With these advanced classification algorithms, appropriately implemented software learns to recognize and associate patterns in the input features with each activity. As such, this field of study is often referred to as machine learning. Machine learning techniques are generally considered to fall within one of two categories, either supervised or unsupervised.

With supervised learning, a significant amount of fully labelled activity data is required in order to train the classification algorithm. Once the training phase is complete, the classifier is able to assign an activity label to an unknown window of sensor data. With unsupervised approaches no activity labels are required for the training dataset. Instead, all the sensor data are passed to the algorithm which automatically identifies a number of states or data clusters, each of which may correspond to a particular activity.

Within the field of activity classification, the classical cross-validation (CV) can be adapted to evaluate the accuracy of the system in two ways: between-subject and within-subject evaluation. In the former case, the classifier is first trained with data from all subjects except a few and then tested with data from the excluded subjects. The accuracy is then calculated as the proportion of correctly classified windows of data across all activities. The process of excluding some subjects and performing a traintest cycle is repeated until all subjects have participated in the testing datasets. The finally overall accuracy is then calculated as the average accuracy across all traintest cycles. When one subject is used for the testing, for a number of cycles equal to the number of subjects, this is called leave-one-subject-out CV. For within-subject evaluation, training is performed using a portion of windows for a specific subject, while testing takes place with the remaining samples of the same subject. This process is then repeated, each time using a different portion of the subject samples for testing. The overall accuracy is determined from the average of all the cycles for all available subjects.

Although an overall accuracy is often provided, more detailed views of the classifiers performance can be given through sensitivity and specificity. These are calculated separately for each activity by determining whether each data window in the test dataset has been identified as the correct activity or not. Sensitivity represents the ability of the classifier to select instances of a certain activity class, whereas specificity represents the true negative rates of an activity. These measures are based on the analysis of the confusion matrix, which summarizes the predicted and actual instances for each class.

Threshold-based Classification: With threshold-based classification, a derived feature is simply compared to a predetermined threshold to determine whether a particular activity is being performed. This approach has been used successfully to differentiate between static postures, such as standing, sitting and lying, using angles derived from accelerometers placed on combinations of the pelvis/trunk and chest [4, 17, 40, 56]. Moreover, threshold-based classification have also been applied on SMA to differentiate between static postures and dynamic activity [15, 24]

Hierarchical Methods: Several studies employed hierarchical classification methods to classify activities using body-worn sensors [19, 24–26, 42, 59]. To implement a hierarchical classification scheme, a binary decision structure is constructed which consists of a number of consecutive nodes. At each node, a binary decision is made depending on the input features. This decision results in either a definite classification being made or in a transition to another node, where further differentiation between activities is performed. The exact nature and parameters of the decision made at each node is obtained via manual inspection and analysis of the training data, which means that this approach is very time consuming.

Decision Trees: The decision tree approach is similar to hierarchical classification. However, rather than the decision structure being constructed manually by the user, rigorous algorithms exist to automate the process and create a compact set of rules. These algorithms work by examining the discriminatory ability of the features one at a time to create a set of rules which ultimately leads to a complete classification system.

Decision trees have been applied to a wide range of classification problems [19, 29, 59]. One of the most comprehensive studies was carried out by [12] who used both time and frequency features to differentiate between 20 activities. Using five sensors, they obtained an accuracy of 86%. However, additional analysis showed an accuracy reduction of only 3% if only data from a thigh and wrist sensor was used.

***k*-nearest Neighbors:** With a k-nearest neighbor (kNN) classification scheme, a multi-dimensional feature space is constructed, in which each dimension corresponds to a different feature. The feature space is first populated with all training data points, each of which corresponds to a particular activity. Unknown windows of sensor data are represented in the feature space and the k-nearest points (or neighbors) of training data identified. Classification is then determined by the majority of the k-nearest neighbors which correspond to a given activity. The value of k typically varies from 1 to a small percentage of the training data and is selected using trial and error, or ideally using cross-validation procedures.

Foerster et al [14] were the first one to use KNN in activity classification to differentiate between nine activities. Later they extended their original approach, combining a kNN classifier with a hierarchical decision structure and including a frequency-domain feature. At each node of their hierarchical decision structure, they constructed an appropriate feature space using a subset of features. With this approach they were able to accurately classify a wider range of activities than in their previous work.

A similar approach has been used in [17]. However, rather than applying the standard kNN approach, they used training data for each activity to specify a maximum and minimum value along each axis. This effectively defined a volume corresponding to each activity within the feature space. For an unknown window of activity data, classification was determined by the closest activity volume within the feature space. With this approach, they were able to identify a wide range of movements and postures with good levels of accuracy.

Artificial Neural Networks: An artificial neural network (ANN) can be likened to a flexible mathematical function configured to represent complex relationships between its inputs (independent variables) and outputs (dependent variables). The ANN is initially presented with a set of training data and some form of optimization process is employed to enable known outputs to be predicted for a given set of inputs. Once trained, the ANN can then be used to obtain the outputs for any set of inputs. In the field of activity classification, the inputs are normally features derived from sensor data with the outputs being the different classes of activities [16, 18, 19, 23].

One of the most common ANNs is referred to as amulti-layer feedforward neural network or multilayer perceptron. This consists of inputs and outputs which are interconnected via special nodes, distributed in so-called hidden layers. The flow of information through the network is controlled by the weighting of the links between the nodes and the transfer function within each node. This type of network is trained by iteratively optimizing the weights in order to accurately produce the desired training outputs from the corresponding inputs. Several studies have employed such ANNs for the task of activity classification with high success rates [16, 18, 19, 23].

An alterative to the feedforward ANN is the probabilistic neural network. Unlike most ANNs which require an extensive training period, this type of network enables classification to be rapidly performed using example patterns stored in memory. This approach has been used in [18] where ANN was trained using template waveform patterns for each activity, rather than using features derived from sensor signals. Although their classification schemewas straightforward to implement, an individually designed network was required for each subject.

Support Vector Machines: Support vector machines (SVMs) constitute a popular machine learning method which is based on finding optimal separating decision hyperplanes between classes with the maximum margin between patterns of each class. Additionally, by using the so-called kernel functions, they can project the data from the original feature space they lie in, to another higher dimensional space. In this way, a linear separation in the new space becomes equivalent to a non-linear classification in the original space. An optimization technique is used to find the optimal separating hyperplanes that perform the required classifications. SVMs have only been applied in a small number of activity classification studies [29, 61].

Naive Bayes and Gaussian Mixture Models: The Bayesian classifier is based on the estimated conditional probabilities or likelihoods of the signal patterns available from each activity class. Given such likelihoods, the probability of a new unknown pattern having been generated by a specific activity can be estimated directly. With a naive Bayes classifier, the input features are assumed to be independent of each other. With this assumption, it is possible to express the likelihood function for each activity as the product of n simple probability density functions, where n is the number of features. These functions are typically expressed as one-dimensional normal distributions. Although the assumption of feature independence is often violated, the Bayesian approach is popular due to its simplicity and ease of implementation. A more general version of the naive Bayesian is discriminant analysis, where cross-correlations between features are taken into account.

Mixed results have been reported when the Bayesian approach to activity classification has been compared to other methods. For example, [20, 29] found this approach to either outperform or match the classification accuracy of other methods, whereas [12] found low levels of classification accuracy. They suggested that the reason for this poor performance may have been the questionable assumptions that acceleration features can be considered conditionally independent and modelled by a normal distribution. Other studies which have used the Bayesian approach [21, 61].

A Gaussian mixture model (GMM) operates along similar principles to a Bayesian classifier. However, the likelihood function is not assumed to be a single Gaussian probability density. Instead, it is assumed to be of unknown shape and functional form and thus approximated by a weighted mixture of Gaussian functions. The weights and the parameters (centres and covariances) of the mixture components are calculated using the expectation-maximization (EM) algorithm. Allen et al [30] employed this approach using time-domain features to construct separate GMMs for a number of movements/postures. To train the GMMs and calculate the parameters, they used an approach similar to EM but which employed a statistical estimate proposed in the field speech recognition. Classification of test data was achieved by selecting the GMM (activity) with the highest probability of having produced that particular set of features. They showed that, provided subject-specific training was used, the GMM outperformed a hierarchical classifier.

Markov Chains and hidden Markov Models: For certain classification problems, some transitions between activities are more likely to occur than others. For example, it is highly unlikely that an individual would sit down directly after descending stairs, but would be likely to start walking. A Markov chain is a discrete time stochastic process in which each activity is represented as a different state. Markov chains can be used to represent the likelihood of transitions between different activities.

An HMM is similar to the Markov chain, but the state of the model at any given time is unknown (or hidden) and can only be determined from observable parameters which depend on the state. In contrast to the Markov chain, the HMM can be used directly for activity classification problems. The observable parameters are the features derived from body-worn sensor data, with the states corresponding to the different activities. Unlike a Markov chain, states in an HMM can correspond to more than one activity. As with previous classification techniques, an HMM is first trained using example data. Once trained, it can then be used to determine the most likely sequence of state transitions (and thus activities) which could have resulted from an observed sequence of features. HMMs are trained by determining state transitions along with the probabilities that each possible set of observations (features) will be observed for a given state. In activity classification studies, HMMs have been used with varying success rates [21, 22, 28].

Fuzzy Logic: Fuzzy logic is based on fuzzy set theory. The idea is to use reasoning which is approximate rather than specifically defined. The advantage of using fuzzy logic is that it provides the freedom to map from a set of inputs to one or more outputs using a set of simple if-then statements, which are called rules. In case of physical activity classification problem, features extracted from body-worn sensor signals make the inputs, whereas the outputs are the fuzzy truths which correspond to each class of activity. Flow of information through a fuzzy system happens via a number of steps. Firstly, the inputs, or features in this case, are assigned membership to fuzzy sets. This assignment is carried out using appropriate membership functions.

In classical set theory, data points or members are either part of a set or not, in other words, partial membership among multiple sets is not allowed. However, the case of fuzzy set theory, by allowing the membership function to range between 0 and 1, permits partial membership in multiple sets. Once each input has been assigned membership of a fuzzy class, the rules can be applied to produce a corresponding output. In the case of activity classification problem, the output is a membership value, or fuzzy truth, which ranges from 0 to 1 for each class of activity. The classification result is then normally taken to be the activity with the maximum fuzzy truth.

Using fuzzy logic, it is possible to reason with imprecise concepts. As such, fuzzy logic is sometimes better suited for dealing with real-world problems than conventional logic which is normally used in hierarchical or decision tree classification schemes. Despite this, fuzzy logic has only been applied to a limited number of activity classification problems. Lee and Mase [62] applied this approach, first using simple heuristic features to identify different static postures, and then using the fuzzy classifier to differentiate between different movements. They defined membership functions in terms of the standard deviations of the sensor signals and the short-term changes in orientations, calculated from the gyroscope signal. By using a set of rules based around the min operation (the fuzzy equivalent of AND), they were able to distinguish between different gaits with good accuracy.

The Mamdani fuzzy inference method is one of the most common techniques for developing a fuzzy logic classifier. With this approach, it is possible to specify certain membership functions and then to develop a set of rules which allow the training inputs (features) to be mapped to the training outputs (activity classes). Salarian [63] used this method as part of a three-stage activity classification scheme. This scheme first used a statistical classifier to identify sit-to-stand and stand-to-sit transitions, and then employed a threshold-based approach to identify periods of walking and lying. Finally, a fuzzy classifier was used to identify periods of sitting and standing. This classifier was developed using membership functions constructed from a knowledge of activity states before and after the period of interest. Classification accuracies obtained using this approach were shown to be better than those obtained using simple threshold rules [4].

Boissy [64] used Mamdani's fuzzy inference to identify falls. Data from a tri-axial accelerometer were used as input to a fuzzy classifier and the amplitude of each acceleration component was used to determine membership values for the classes: low, medium and high. A total of 27 rules were used to produce the output, which was expressed in terms of a three-class membership function (no, maybe and yes) representing the occurrence of a fall. The value of this output function was then combined with the knowledge of body orientation using conventional Boolean logic to determine whether a fall had occurred. By collecting a large dataset of fall and non-fall events from 10 subjects, they were able to demonstrate average fall detection accuracies ranging between 86 and 93%, depending on sensor location.

Combined Classifiers: The popularity of Meta-level classification schemes, within the biomedical community, has increased recently . They are known for improving the performance of individual classifiers by combining their output. The combination of outputs is achieved using different techniques. These include majority voting (where the majority class is accepted), stacked generalization (which trains the base classifiers and then uses their predictions as data to a new learning stage) or boosting (which assigns weights to the training patterns to combine the performance of weak classifiers) [65]. Ravi [29] used a meta-level classification scheme in a pilot study with two subjects who performed eight common activities. Five base-level classifiers were used in their study, including SVMs, decision trees, kNN and naive Bayes. In general, when an inter-subject design was used, the boosted SVM was shown to outperform other meta-level classification schemes.

AdaBoost is a type of adaptive boosting that incrementally trains classifiers by suitably increasing the pattern weights to favour the misclassified data. Thus, it combines multiple weak classifiers to create a single more powerful one and has been used by [66]. They studied ten common daily activities deriving a large number of statistical and frequency-domain features from a range of sensors. They then constructed a set of weak binary classifiers, each of which accepted only a single feature as input and obtained a classification result from a weighted combination of the weak classifiers. They compared the performance of two different weak classifiers: a discriminative decision-stump (a binary decision tree classifier constrained to the use of a single feature) and a generative naive Bayes model and found the Bayesian approach to perform best. Classification accuracy was then improved by using the output from the weak classifiers as input to a HMM.

2.4 Conclusion

This chapter has presented an overview of the different techniques which have been used for activity classification from body-worn sensor data. Information has been organized into two principal sections, the first dealing with feature generation and simple threshold-based classification and the second dealing with more advanced classification techniques.

Within this framework, features were categorized as heuristic, time-domain, frequency-domain or timefrequency (wavelet). Heuristic features are derived from a fundamental understanding of how a specific movement or posture will produce a characteristic body-worn sensor signal. By using such features in simple threshold-based classification schemes, it is possible to accurately differentiate between static postures and dynamic activity and to identify falls with high levels of accuracy.

In order to differentiate between large numbers of dynamic movements and postures, it is necessary to use advanced classification schemes which accept one or more features as input.

Although a small number of studies, comparing the performance of different classifiers, suggest that either decision trees or artificial neural networks may give the highest classification accuracy, differences are often small. Furthermore, there are many other methods such as support vector machines and hidden Markov models which have shown promise in small pilot studies but have yet to be tested in larger-scale studies. Therefore considerable further work is required to establish the suitability of the different techniques for a range of classification problems. Most previously published activity monitoring studies vary considerably in the choice of sensor placements and in the range of activities analyzed.

Chapter 3

Proposed HAR Methodology

This chapter presents overview of the design of the activity recognition system presented in this work. It also describes the research approach followed to collect the necessary data to develop the data-model and evaluate the model and the classification algorithms.

3.1 Overview of Research Approach

The method used in this work for the development of the activity recognition system consisted of four main steps. (1) Firstly, activity data were collected on different physical activities from different subjects. (2) Once these data had been collected, a set of systematic analysis was performed on the collected dataset to determine some important parameters of the recognition algorithm that enable real-time performance. These parameters include the mathematical model that can best describe the data, the number of parameters of the model, the sliding window length and the final feature set to use. (3) Four more activity-datasets were then collected from different subjects under different settings. (4) Finally, appropriate classification schemes were developed to evaluate the parameters determined in the third step for physical activity recognition using the four datasets.

3.2 Sensor Devices

In this work, two sensor devices were employed for collecting data on physical activities.

3.2.1 WiTilt

Figure 3.1 shows the accelerometer called Witilt v2.5. It's a 2.4GHz Wireless 3-axis Tilt Sensor from Sparkfun. It employs a FreeScale MMA7260Q triple-axis or tri-axial accelerometer and a class 1 Bluetooth link from BlueRadios. MMA7260Q is a surface mounted integrated circuit accelerometer that runs on low voltage $(2.2V - 3.6V)$ and is capable of measuring acceleration along the X, Y, Z axes. This accelerometer also features a 4-level sensitivity scale (1.5g, 2g, 4g, 6g), with 1.5g setting being the most sensitive. It measures both (+) and (-) acceleration.

3.2.2 TOmnia

TOmnia is a smartphone from Samsung, also called SCH-M490. It supports a triaxial accelerometer which can measure acceleration in the range of $\pm 2g$. TOmnia accelerometer's resolution is 0.004g and its axis directions are shown in Figure 3.2.

3.3 Speech Annotation System

In this work, annotations are performed using a bluetooth headset combined with speech recognition software. During each data collection study, the starting and ending points of each activity were marked by using a predefined set of commands. The headset used in this study is called the Jabra BT250v bluetooth headset, shown in Figure 3.3. It offers a range of 10 meters and a battery power for 300 hours standby and 10 hours active talking. The software for storing the annotations was developed following the idea presented in [3]. It is written in C and combines elements of the bluetooth API with the Microsoft Speech API.

This method proved very accurate and efficient as annotations are provided by the user on the spot. It also resulted in very little interference while performing activities. To minimize any mislabeling, data within 5 seconds of the start and stop times are discarded. Since the subject is probably standing still or sitting while he records the start and stop times, the data immediately around these times may not correspond to the activity label.

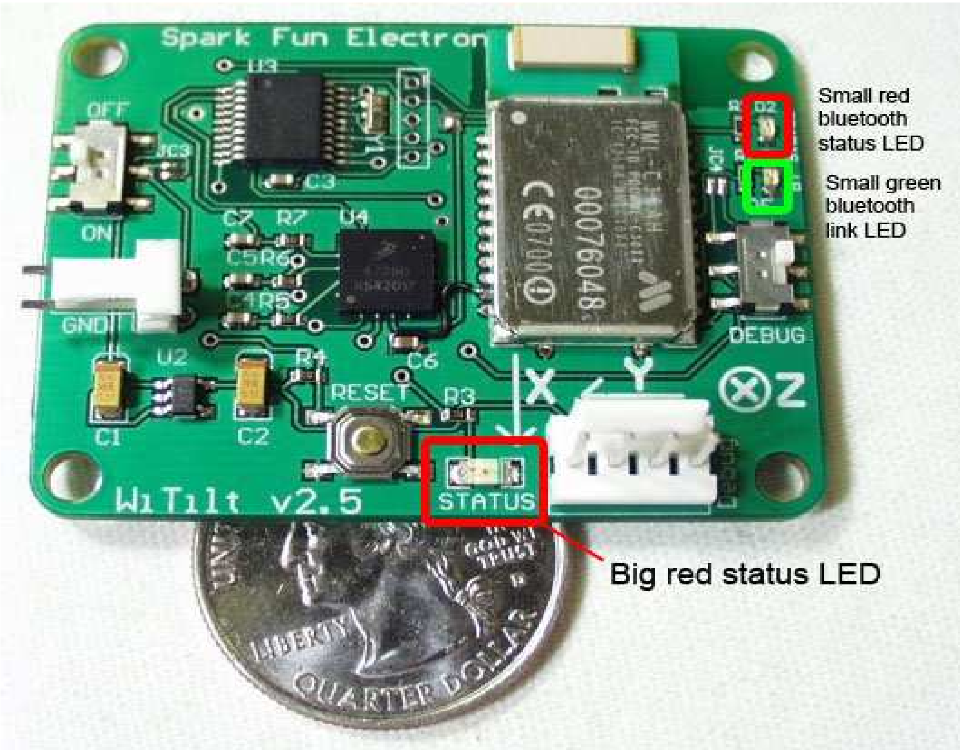

Figure 3.1: WiTilt v2.5: A 2.5 GHz Wireless 3-axis Tilt sensor from Sparkfun

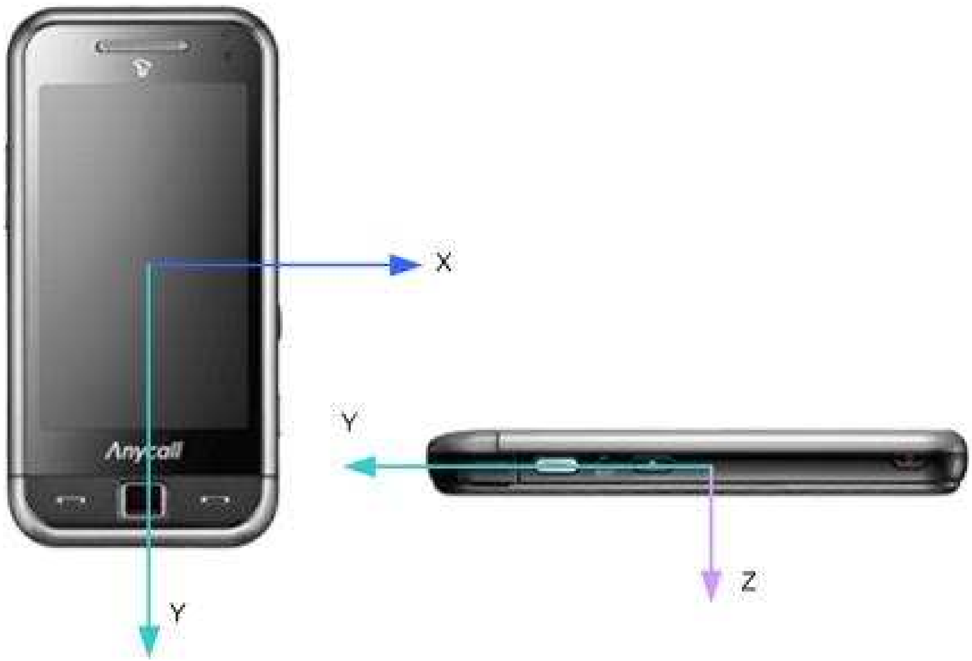

Figure 3.2: TOmnia(SCH-M490), a smartphone from SamSung with a built-in triaxial accelerometer. The X axis is along the width of the device, and positive on the right direction. The Y axis is along the length of the device, and positive on the down direction. The Z axis is along the depth of the device, going into the screen

Figure 3.3: Jabra BT250v bluetooth headset

3.4 Data Collection

3.4.1 Dataset for Model Identification

This dataset contains data on four physical activities collected from six participants (three male, three female, mean age = 28 years old). The four activities include lying, standing, walking and running. In general, the output of any body-worn accelerometer depends on the position at which it is placed [9]. Accelerometers are normally attached to the part of the body whose movement is being analyzed, such as arm, wrist, thigh, etc. However, since the aim was to study the whole body movements, the WiTilt sensor, with a sampling frequency of 45Hz, was placed at a position closer to the center of mass, i.e., the subject's chest as shown in Figure. 3.4.

These data were used to determine important parameters of the recognition algorithm. The most important parameter includes the model for describing the data that was determined using the stochastic time series analysis of the activity-data. The main difference between this dataset and the others is that during this study each participant performed the same activity for longer periods of time, i.e., about 30min per activity per participant thus each file contained data on a single activity. However, in other studies each participant performed different activities in fixed or random sequences where a single activity lasted for not more than 3min and each file contained data on multiple activities. The reason for this approach was to have enough continuous data on each activity for developing a better understanding of the frequency and intensity characteristics of the acceleration-signals for model identification. This dataset was named Data-for-Model-Identification (DMI).

3.4.2 Controlled Laboratory Dataset for Model Evaluation

Ten healthy subjects, i.e., four females and six males with the mean age of 27, participated in this study. The WiTilt was attached to their chests to to collect the 15 physical activities that are listed in Table 1.1. The sampling frequency was 45Hz.

The subjects performed the activities under researcher's supervision. They were told where and how to perform these activities. Annotations were performed using the developed speech annotation system. About 35 hours of activity data were collected. A sample sequence of the

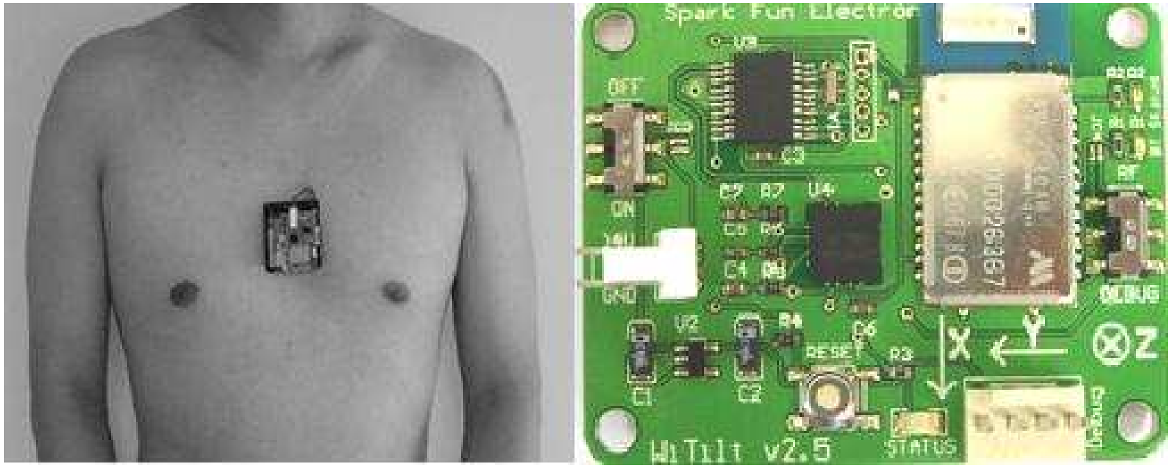

Figure 3.4: WiTilt (a tri-axial accelerometer) being attached to a subject's chest in order to collect data on 13 bodily activities

activities performed by each subject at a time is: sitting (2min) → sit-stand → standing (2min) → stand-lie → lying (2min) → lie-stand → standing (40sec) → walking (2min) → standing (40sec) → walking-upstairs → standing (40sec) → walking-downstairs → standing (40sec) → stand-sit → sitting (40sec) → sit-lie → lying (40sec) → lie-sit.

It is obvious that these data were collected using a strict protocol. In other words, each subject followed a standard activity-sequence, performed each activity with the same speed each time for a fixed timed-interval. Thus there were very less variations in activity patterns across different subjects. The purpose was to have a dataset where the problems that arise due to the variations in the sensor's output for the same activity across different subjects do not exist. In other words, this dataset was for evaluating system's performance for best case scenario. This dataset was named Controlled-Laboratory (CL) dataset.

3.4.3 Naturalistic Dataset for Model Evaluation

The same ten subjects as the previous study participated in this study and wore the same sensor device at their chest with a sampling frequency of 45Hz. This time the participants were not told about how and where to perform the activities. They were just provided with approximate time duration for each activity except for walking-upstairs and downstairs. The time duration of these activities depended on the length of stairs at each subject's home and thus varied among the subjects. The participants were trained on the use of data collection and annotation applications. Each person then collected the activity data at home without the researchers' supervision. They made the annotations themselves throughout the data collection. About 24 hours of data were collected in this study.

Thus these data were collected in less-constrained free-living settings. The purpose of this dataset was to evaluate system's performance for real-life cases where each subject could perform activities differently from the other subjects and the activity acceleration-data thus could vary significantly across subjects. Moreover, whenever system was evaluated using this dataset, activity data from only eight subjects were used as training data whereas the activity data from the last two subjects were used as testing data (subject independent evaluation) to make the classification task more difficult. This dataset was named Naturalistic (NL) dataset.

3.4.4 Sensor's Position Free Dataset

About 24 hours of activity data were collected using the WiTilt sensor, outside the laboratory, in naturalistic less-controlled home settings. The sensor, with the sampling frequency of 45Hz, was placed on eight elderly subjects (six males, two females, age: mean = 65, SD = 3 years old)) on 5 different positions, i.e., chest pocket, front left trousers pocket, front right trousers pocket, rear trousers pocket, and inner jacket pocket. The activities to be recognized were resting (lying/sitting/standing), walking (along the corridor), walking upstairs, walking downstairs, running, cycling, and vacuuming. The cycling activity was recorded in a gym.

An approximate distribution of the data for each subject regarding the five body sites and the seven activities in our study is: resting (40 minutes, 8 minutes per site), walking (40 minutes, 8 minutes per site), running (25 minutes, 5 minutes per site), cycling (25 minutes, 5 minutes per site), vacuuming (25 minutes, 5 minutes per site), walking-upstairs (10 minutes, 2 minutes per site), and walking-downstairs (10 minutes, 2 minutes per site).

The subjects were trained on the use of data collection and annotation applications in the laboratory where they were given short definitions of the seven activity labels. Each subject then collected the data at home, without researcher's supervision, where he/she performed random sequences of the seven activities at their own pace and labeled the start and end points of each activity. Approximately 24 hours of the activity data, i.e., 3 hours per subject were collected. This dataset was named Position-Free (PF) dataset.

3.4.5 Smartphone based Dataset

In this study, TOmnia was used to collect activity-data. Samsung Windows Mobile SDK and Windows Mobile 6 SDK were used to obtain the accelerometer's data and store it on phone's storage card. The accelerometer was configured to provide data with a sampling frequency of 90Hz. Activity data were collected by placing the phone on six healthy subjects on five different positions: shirt's top pocket, jeans' front-left pocket, jeans' front-right pocket, jeans' rear pocket, and coat's inner pocket. The five activities to be recognized were resting (sitting), walking, walk-upstairs, walk-downstairs, and running. For realistic recognition, brief movements such as stretching or changing posture were allowed during resting. For a natural setting, walking, walk-upstairs, walk-downstairs, and running were performed outdoor at various speeds. This dataset was called SmartPhone (SP) dataset.

3.5 Noise Reduction

The raw data from an accelerometer might contain some noise that must be taken care of before using these data for further experimentation. In case of activity acceleration-signals this is usually high frequency noise. This section describes the procedure that was used to handle the noise in acceleration-signals before using the collected datasets in different stages of the research process.

In this work, moving average filtering technique of order 3 was employed to reduce or filter-out the random noise. A moving average filter smooths data by replacing each data point with the average of the neighboring data points defined within a given span (order of the filter). This process is equivalent to low-pass filtering with the response of the smoothing given by the difference equation

$$y_s(i) = \frac{1}{2N+1}(y(i+N) + y(i+N-1) + \ldots + y(i-N)) \tag{3.1}$$

where $y_s(i)$ is the smoothed value for the ith data point, N is the number of neighboring data points on either side of $y_s(i)$ and $2N+1$ the span. The choice of such a filtering technique offers two advantages. Firstly, it reduces random high frequency noise while retaining a sharp step response. Secondly, since each data point is replaced by the average of the neighboring data points, this helps in reducing the random measurement errors that may result while measuring an activity.

3.6 Segmentation Technique

Like any other pattern recognition problem, in activity classification the sensor signal is first divided into smaller time segments more commonly knows as windows. Features are computed separately for each window and fed to the classification algorithms. In real-time applications, windows are defined concurrently with data collection and a continuous real-time activity profile is produced. When the sensor data are processed off-line, the windows are defined first and classification algorithms applied sequentially to each window. This information is then combined to give an activity profile along the entire signal. Three different windowing techniques have been used in activity monitoring, sliding windows, event-defined windows and activity-defined windows.

With the sliding window method, the signal is divided into windows of fixed length with no inter-window gaps. A range of window sizes have been used in previous studies from 0.25 s to 6.7 s, with some studies including a degree of overlap between adjacent windows [10, 12]. The sliding window approach does not require pre-processing of the sensor signal and is therefore ideally suited to real-time applications. Due to its implementational simplicity, most activity classification studies have employed this approach.

In order to use event-defined windows, pre-processing is required to locate specific events, such as heel strike or toe-off. These events are then used to define successive windows. Given that such events may not be uniformly spaced in time, the size of these windows is not fixed. A number of different approaches have been proposed for identifying heel strike and toe-off from body-worn sensor signals. For example, it is possible to define search windows from either a low pass filtered version of the original signal [40, 67] or segmental angles [68] within which maxima or minima correspond to gait events. Another approach is to identify the times at which the anterio-posterior component of the trunk acceleration changes sign. Heel strike is then located at a given time offset from these points [69, 70].

The use of activity-defined windows is dependent on determining the times at which the activity changes. These points are then used to define windows of sensor data, each of which correspond to a different activity. A number of methods have been proposed to identify activity-transition points prior to explicitly identifying the specific activities. For example, wavelet analysis can be used to identify localized changes in frequency characteristics [58] which correspond to a change between activities. Once defined, classification is performed for each window, sometimes using only a subset of the data contained within the window.

In conclusion, (1) The longer the length of the window, the better the quality of the features estimated. However, the longer the window-length, the longer the end-user has to wait for the recognition result. Moreover, longer windows result in failure of recognition of activities performed for short periods, such as sit-stand or walking a couple of steps. (2) The optimal window length to use for the feature-computation depends on the activity being recognized [57]. However, utilizing one window length per activity is computationally expensive.

Since the goal of this study was to implement a system which is light-weight to run on a hand-held device and is capable of recognizing activities in real-time, this work therefore utilized only a single window length for all activities. Performance of different window lengths for different activities across multiple subjects was analyzed to select one that gave good estimates of the selected features using the least number of samples in a given window. One limitation of this approach is that the appropriate window-length is training data dependent. However, it provides reasonable approximation for the study objectives.

Chapter 4

Features of Acceleration Signals

In machine learning, pattern classification is the process which employs a specific algorithm and rules to assign an output (which is mostly a class-label) to a given input. The goal is to assign each input a specific class from a given set of classes. It is a step-by-step process where the data about some real phenomena, after pre-processing of some sort such as noise reduction, are used to extract features. These feature vectors have the ability to describe all known characteristics of the any instance. The features are then fed to a module, the classifier, which implements a specific classification algorithm.

In the case of physical activity recognition using wearable accelerometers, the data are activity acceleration-signals which are measured using wearable accelerometers attached to the human body. As mentioned above, each window of the acceleration data is processed to remove noise and calculate representative features which are then fed as an input to a classifier to recognize a particular physical activity.

This chapter presents details on the procedure used for identifying the model that can best describe the activity acceleration-data. The DMI dataset (section 3.4.1) was used for model identification.

4.1 Need for a better Mathematical Model

Features used in the previous studies can be categorized into three groups:

- Frequency domain features, such as FFT.
- Time-Frequency domain features, such as wavelet analysis.
- Time domain features, such as mean and standard deviation.

Both frequency domain and time-frequency-domain features require much high components to discriminate different activities. Moreover, calculation of these features require longer time-windows. Hence they increase computation and are not suitable for real time application.

On the other hand, time domain features can be easily extracted in real time, therefore, they are more popular in many practical acceleration activity recognition systems. Although activity recognition using time domain features was successful to some limit, the recognition results using these features have not had a high success rate because such methods assume that activity acceleration signals are deterministic. However, in fact, the activity acceleration signals are random signals in their nature. Figure 4.1 shows the probability distribution function of vertical acceleration signals of walking activity that follows a normal or Gaussian distribution. Thus it's important to establish a better mathematical model using stochastic time series analysis to describe these data. There are many reasons for wanting to do this. (1) To get a better understanding of physical mechanism generating the signal. (2) To predict signal's future behavior. (3) To improve the quality of the signal, for example, reduction of noise. (4) To achieve data compression for storage and transmission. (5) To generate artificial signals similar to the natural ones. (6) To classify the signal.

4.2 Autocorrelation Analysis for Model Identification

There are a number of approaches to modeling time series. Autocorrelation plot is a commonly used tool for identifying a model that can best describe a given time-series. Autocorrelation is the average of the product of a data sample $x[n]$ with a version of itself advanced by a lag. The autocorrelation function is described by the equation

$$r_{xx}[k] = \frac{1}{N}\sum_{n=1}^{N-k} x[n]x[n+k] \tag{4.1}$$

where $r_{xx}[k]$ is the autocorrelation value of x at sample delay k, and N is the number of data points. For a very small advance, the values of the two signals at any given instant will be very similar. As the lag increases, the difference between the two values becomes larger. If a signal has both a periodic and a random component, the latter gradually disappears as the lag increases . The property is useful for extracting periodic signals from random noise.

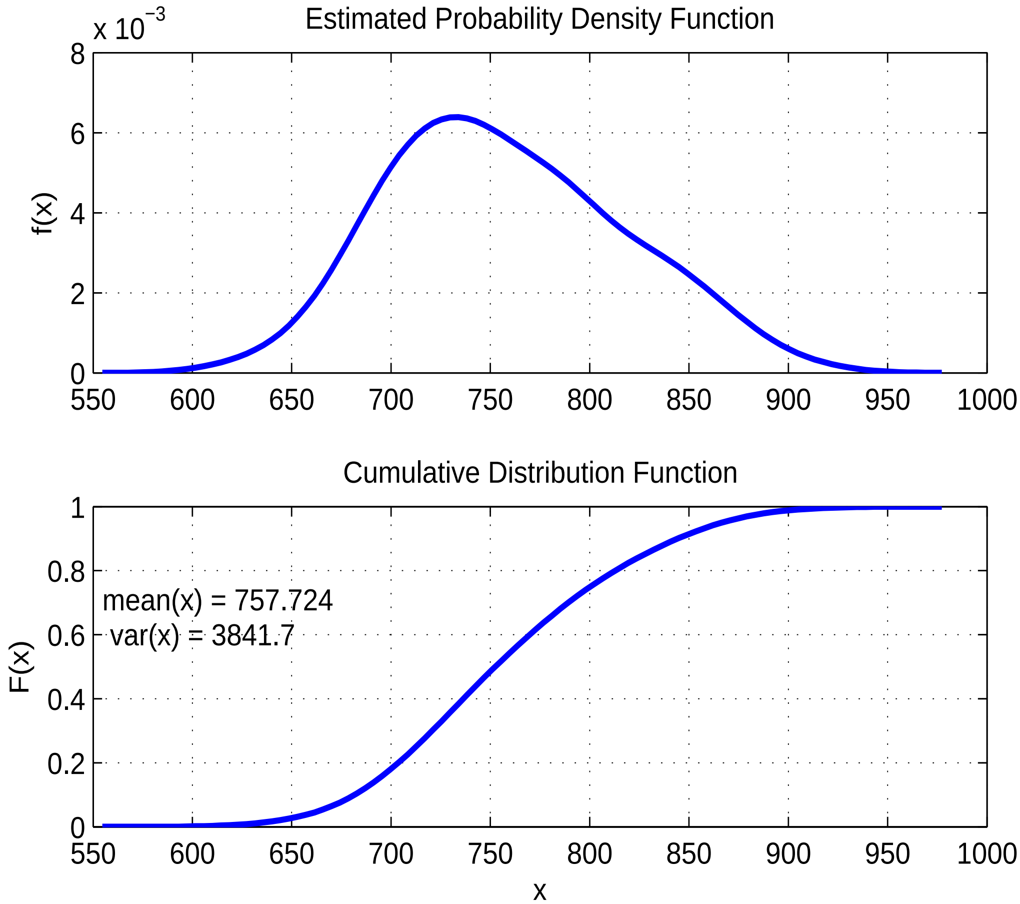

Figure 4.1: Probability Density Function and Cumulative Density Function for of the activity-acceleration data (vertical axis) for walking activity

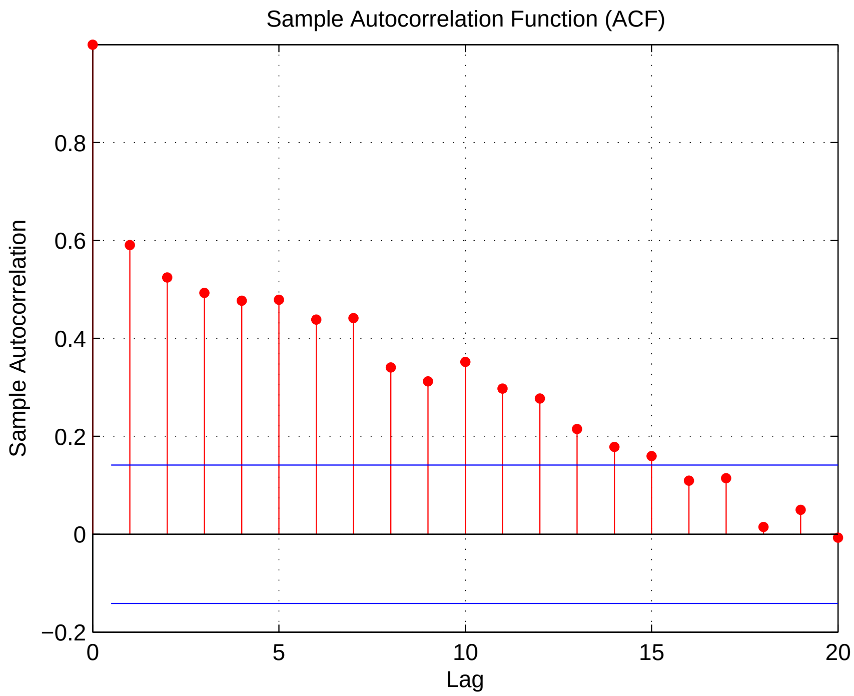

Figure 4.2: Autocorrelation values for 20 lags for the activity-acceleration signals (vertical axis) of standing, showing strong positive autocorrelation suggesting that the data come from an underlying autoregressive process

The autocorrelation plot can provide answers to the following questions. (1) Is an observation related to adjacent observation? (2) Is the observed time series white noise? (3) Is the observed time series autoregressive? Figure 4.2 and 4.3 show autocorrelation plots of a single axis for standing and walking respectively. The plot starts with a high autocorrelation at lag 1 that slowly declines. The conclusion that can be drawn from these plots is that the activity acceleration-data come from an underlying autoregressive model with strong positive autocorrelation [71]. Therefore, autoregressive models are employed for modeling the activity acceleration signals.

4.3 Autoregressive (AR) Models

Autoregressive (AR) modeling utilizes the time history of a signal to extract important information hidden in the signal. It is superior to many other methods, especially in biomedical signal processing as it can take advantage of the noise inherent in a biological system and extract information

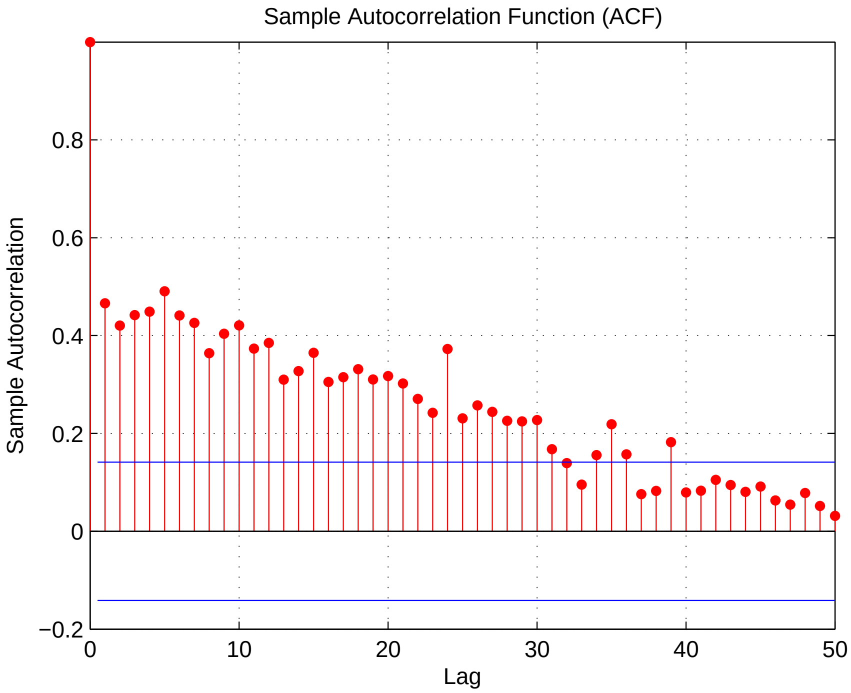

Figure 4.3: Autocorrelation values for 50 lags for the activity-acceleration signals (vertical axis) of walking, showing strong positive autocorrelation suggesting that the data come from an underlying autoregressive process

from propagation of that noise in a signal.

An AR model predicts the current values of a time series from the past values of the same series. Basically, the AR model may be regarded as a set of autocorrelation functions. AR modeling of a time series is based on an assumption that the most recent data points contain more information than the other data points, and that each value of the series can be predicted as a weighted sum of the previous values of the same series plus an error term. The AR model is defined by:

$$x[n] = \sum_{i=1}^{M} a_i x[n-i] + \varepsilon[n] \tag{4.2}$$

where $x[n]$ is the current value of the time series which in our case is the activity acceleration-signal,$a_1 \cdots a_M$ are predictor (weighting) coefficients, M is the model order, indicating the number of the past values used to predict the current value, and $\varepsilon[n]$ represents a one-step prediction error, i.e. the difference between the predicted value and the current value at this point.

The AR model determines an analysis filter, through which the time series is filtered. This produces the prediction error sequence. In the model identification, the AR analysis filter uses the current and past input values to obtain the current output value. By writing equation 3.2 in a form

$$\varepsilon[n] = x[n] - \sum_{i=1}^{M} a_i x[n-i] \tag{4.3}$$

we get the filter with an impulse response $[1, -a_1 \cdots - a_M]$, which produces the prediction error sequence. The predictor coefficients are usually estimated using the least-squares minimization technique so that they produce the minimum error $\varepsilon[n]$. From equation 3.2 we get

$$x[n] = a_1 x[n-1] + a_2 x[n-2] + \cdots a_M x[n-M] + \varepsilon[n] \tag{4.4}$$

If we use equation 4.4 to write the expressions for several estimates of $x[n]$, we get a set of linear equations:

$$\begin{cases} x[M+1] = a_1x[M] + a_2x[M-1] + \cdots a_Mx[1] + \varepsilon[M+1] \\ x[M+2] = a_1x[M+1] + a_2x[M] + \cdots a_Mx[2] + \varepsilon[M+2] \\ \quad \vdots \\ x[N] = a_1x[N-1] + a_2x[N-2] \quad + \cdots a_Mx[N-M] + \varepsilon[N] \end{cases} \tag{4.5}$$

We need M equations to solve the M unknown coefficients $a_i, i = 1 \cdots M$. The least squares solution is easiest to achieve by matrix calculation. The above equation may be rewritten in matrix form:

$$\overline{x} = \begin{pmatrix} x[M] & x[M-1] & \dots & x[1] \\ x[M+1] & x[M] & \dots & x[2] \\ \vdots & & & \\ x[N-1] & x[N-2] & \dots & x[N-M] \end{pmatrix} a + \varepsilon = \overline{X}\overline{a} + \overline{\varepsilon} \tag{4.6}$$

where

$$\overline{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_M \end{bmatrix} \quad and \quad \overline{\varepsilon} = \begin{bmatrix} \varepsilon[M+1] \\ \vdots \\ \varepsilon[N] \end{bmatrix} \tag{4.7}$$

In other words, $\overline{X}$ is a square matrix with M rows and M columns, and $\overline{a}$ and $\overline{\varepsilon}$ are column matrices consisting of M rows and 1 column.

When two vectors form a 90 degree angle, and one vector is projected onto the other, the result is a zero vector. The vectors are then said to be orthogonal, and their inner product equals 0. The inner product of any two column vectors $\overline{a}$ and $\overline{b}$ of the same length is defined as $\overline{a}^T\overline{b}$, where $\overline{a}^T$ is the transpose of $\overline{a}$.

The optimum predictor coefficients $(\overline{a}_{opt})$ can be obtained by applying the orthogonality principle in the least-squares minimization technique. This means that the predictor coefficients are selected so that column vector $\overline{\varepsilon}$ is orthogonal to each explanatory vector $\overline{x_i}, i = 1 \dots M$, i.e. to each column vector in matrix $\overline{X}$. As in normal regression analysis, this minimizes the mean-square error. Then, $\overline{\varepsilon}$ vector is independent of the data $\overline{X}$, i.e. it contains the part of the time series that can not be explained by M previous data points.

Since in this work the activity acceleration-signal is a 3-dimensional, i.e., $x-axis$, $y-axis$, and $z-axis$, we model each axis separately and generate the following feature vector:

$$F = [a_{x1},\dots,a_{xM},a_{y1},\dots,a_{yM},a_{z1},\dots,a_{zM}] \tag{4.8}$$

where $a_{x1},\dots,a_{xM}$ are the AR-coefficients for $x-axis$, $a_{y1},\dots,a_{yM}$ for $y-axis$, and $a_{z1},\dots,a_{zM}$ for $z-axis$ respectively.

4.3.1 Optimum Model-order and Window-length

There is no straightforward way to determine the correct model order for an AR model. A proper order for an AR model would yield a good data fitting effect while retaining a high data compression ratio. In order to determine the optimal AR model order we adapted the following criteria:

Akaike Information Criteria (AIC): Although root mean square error is generally used to achieve a good estimate of an AR model order, it is still not the most appropriate. An information-based criteria which is more appropriate for model order selection is AIC [72]. It is a measure of goodness of fit of an estimated model. Based on the concept of entropy, AIC offers a relative measure of information lost when a given model is used to describe a given time series. Given a dataset, several models are fitted and ranked according to their AIC. The one having the lowest AIC is usually the best model for describing the dataset. AIC is defined by

$$AIC = -2L_m + 2m \tag{4.9}$$

where L_m is the maximized log-likelihood and m is the number of parameters in the model. The index takes into account both the statistical goodness of fit and the number of parameters that have to be estimated to achieve this particular degree of fit, by imposing a penalty for increasing the number of parameters. Lower values of the index indicate the preferred model, that is, the one with the fewest parameters that still provides an adequate fit to the data [73].

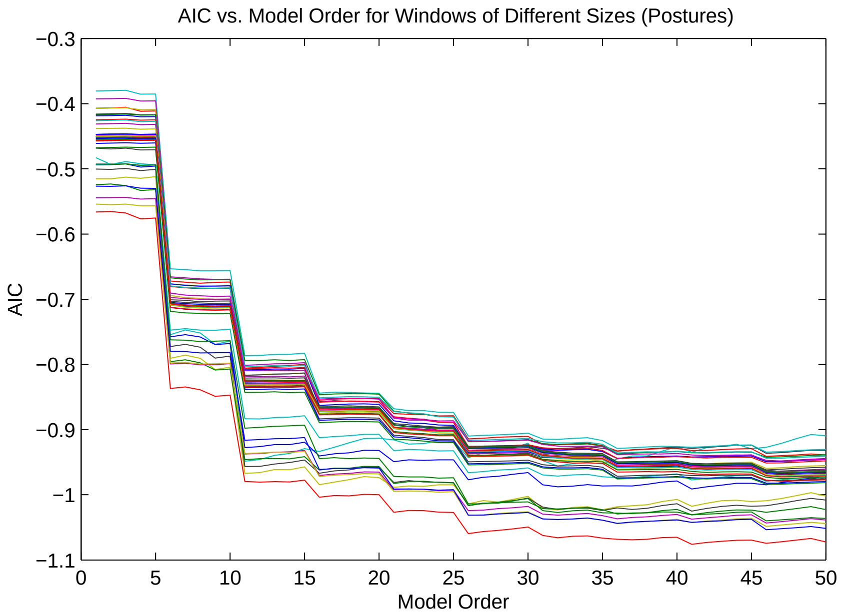

Figure 4.4: Average AIC values for three axes plotted against model order for standing for windows of different length. Different colors represent windows of different lengths (min = 1sec, max = 60sec). AIC-curves for all windows tend to even out near 30 (10 per axis) suggesting that 10 is the appropriate model order in this case

Figures 4.4 and 4.5 show AIC values for different models orders (computed using the aic-function from Matlab) using data-windows of different lengths/sizes for postures standing and walking respectively. The smallest window is 1sec long, i.e., 45 samples (at the sampling frequency of 45Hz) whereas the largest window is 1min long, i.e., 2700 samples. Using these plots, it is difficult to conclude which window-length is the most appropriate. However, the conclusion that can be easily drawn is that although larger windows provided slightly smaller AIC than smaller windows, the AIC-curves for all windows tend to even out near the same model order, i.e., 30 (10 per axis). Each AIC value in these plots represent average of the AIC values for three axes.

In order to determine the suitable window-length, AIC values were calculated for different window-sizes (starting from 1sec to 45sec) for both postures and movements given the model-order 10. Resulting AIC-values are shown in Figure 4.6 and 4.7. It can be easily concluded that the window-size of 3sec, i.e., 135 samples is the most appropriate as it offers the same goodness of fit as larger windows and is not too long to result in a delayed-response, which is desirable

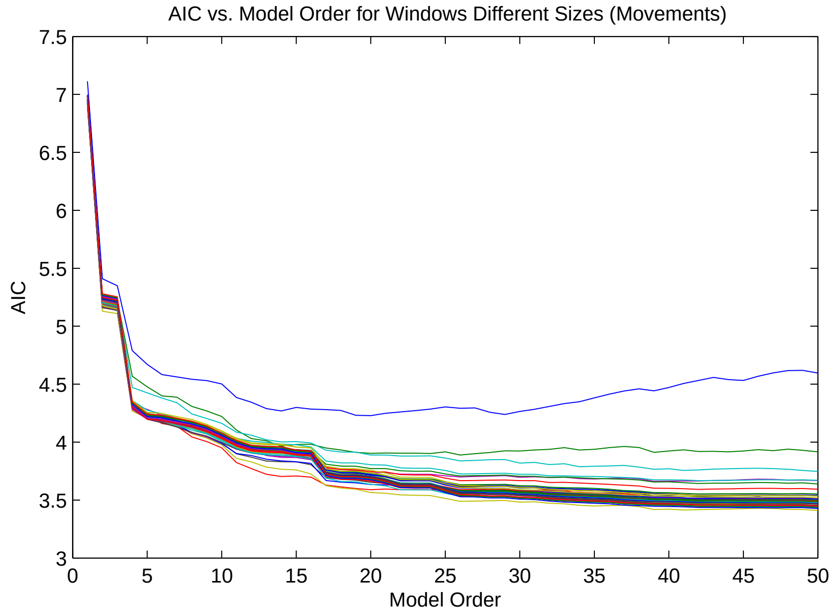

Figure 4.5: Average AIC values for three axes plotted against model order for walking for windows of different length. Different colors represent windows of different lengths (min = 1sec, max = 60sec). AIC-curves for all windows tend to even out near 30 (10 per axis) suggesting that 10 is the appropriate model order in this case

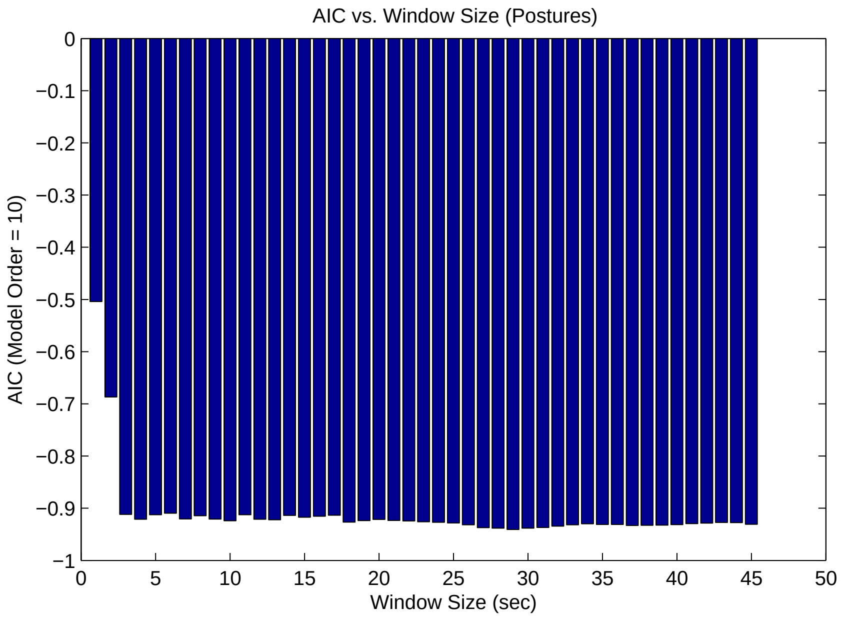

Figure 4.6: Average AIC values for three axes for the chosen model order for standing using windows of different length. No significant decrease can be seen after the window size of 3sec

considering the real-time requirements of human activity recognition systems.

4.3.2 Model validation

After the AR-model has been identified, it's validity must be checked. The primary tool for model diagnostic checking is the analysis of the residual, i.e., the prediction error sequence. If the chosen model is a good model for the data, the residuals should be white noise, drawn from a fixed distribution with a constant mean and variance [71]. Another method to validate the selected model is to treat the AR-model as an all-pole filter and compare its power spectral density with the power spectral density estimate of the modeled signal [71]. To validate whether the selected AR-model of order 10, given a window-length of 3sec (135 samples), is a good model for the activity acceleration-data, both validation methods were employed.

First, the AR-coefficients were estimated using the least square minimization method (discussed above). These parameters were then used to create copies of the modeled signals and residuals were collected. Figure 4.8 shows the estimated probability density function and the cu-

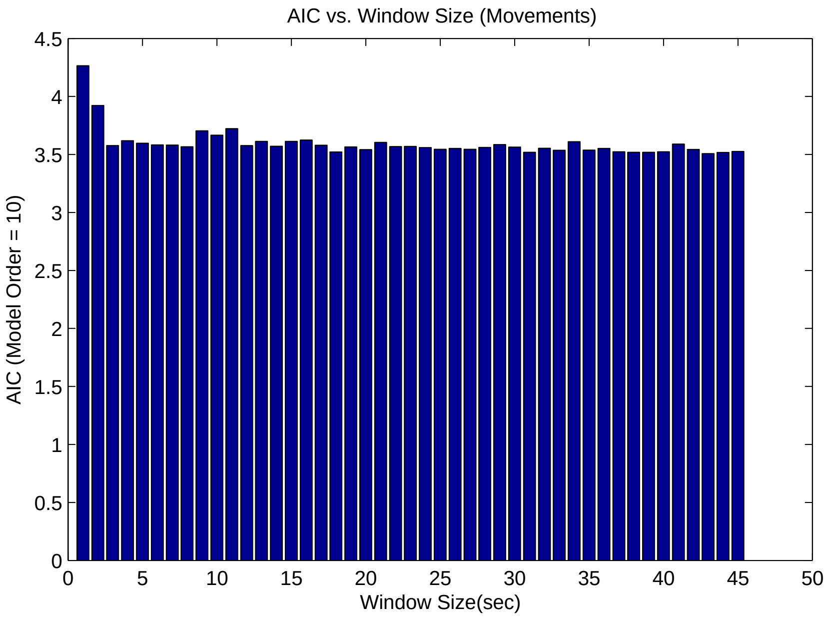

Figure 4.7: Average AIC values for three axes for the chosen model order for walking using windows of different length. Again, no significant decrease can be seen after the window size of 3sec

mulative distribution function of the residuals when fitting the chosen model in case of walking activity and it can be seen that the residuals are in fact white noise with fixed mean and variance.

Second, the power spectral density of the model was also compared with the power spectral density estimate of the modeled activity acceleration-signal. Figure 4.9 and 4.10 show this comparison for standing and walking respectively. These plots show an almost perfect match that indicate the strength of the chosen AR-model in describing the activity acceleration-data. Finally, figure 4.11 shows some exemplary fitting results for walking activity.

4.4 Augmented Feature Vector

Besides AR-coefficients, other time domain features (which have been investigated in previous works) were also calculated from the activity acceleration-data. These features are listed in Appendix . AR-coefficients and these features were combined to create a single large feature-set. The next step was to analyze the classification performance of the different configurations of the front-

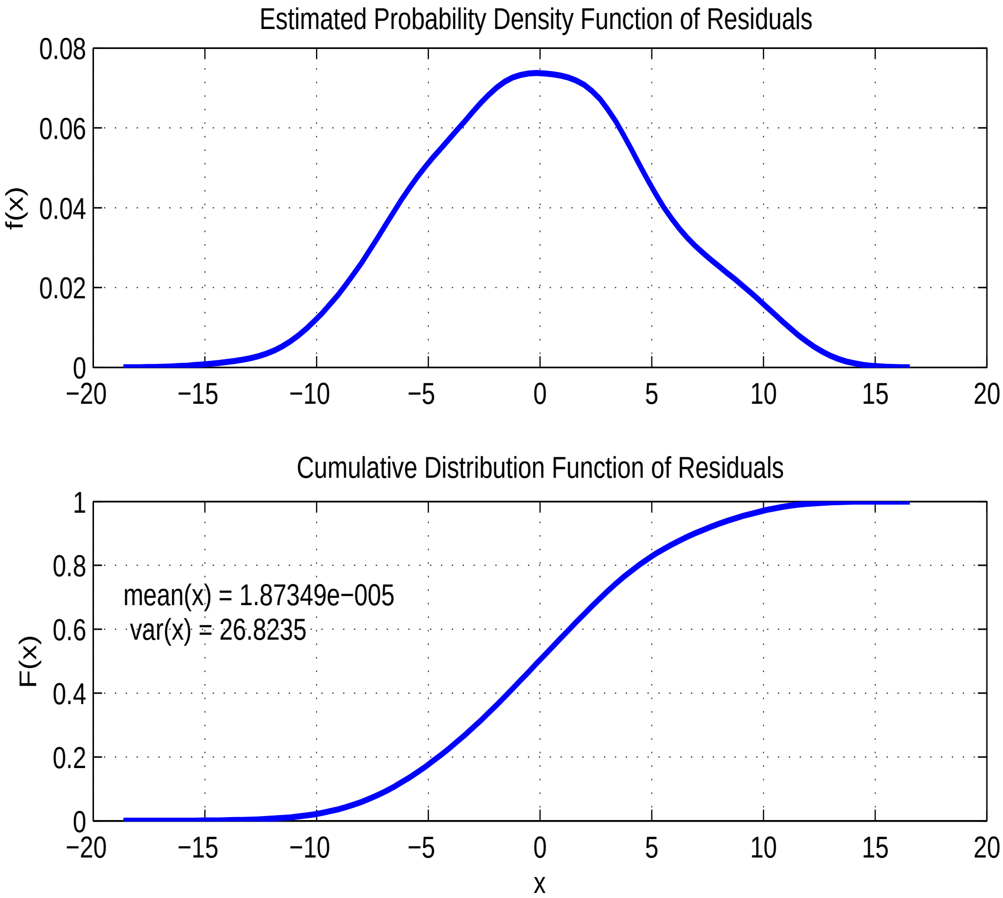

Figure 4.8: Probability Density Function and Cumulative Density Function of the residuals showing that the residuals are random white noise with a fixed mean and variance hence proving the validity of the chosen model

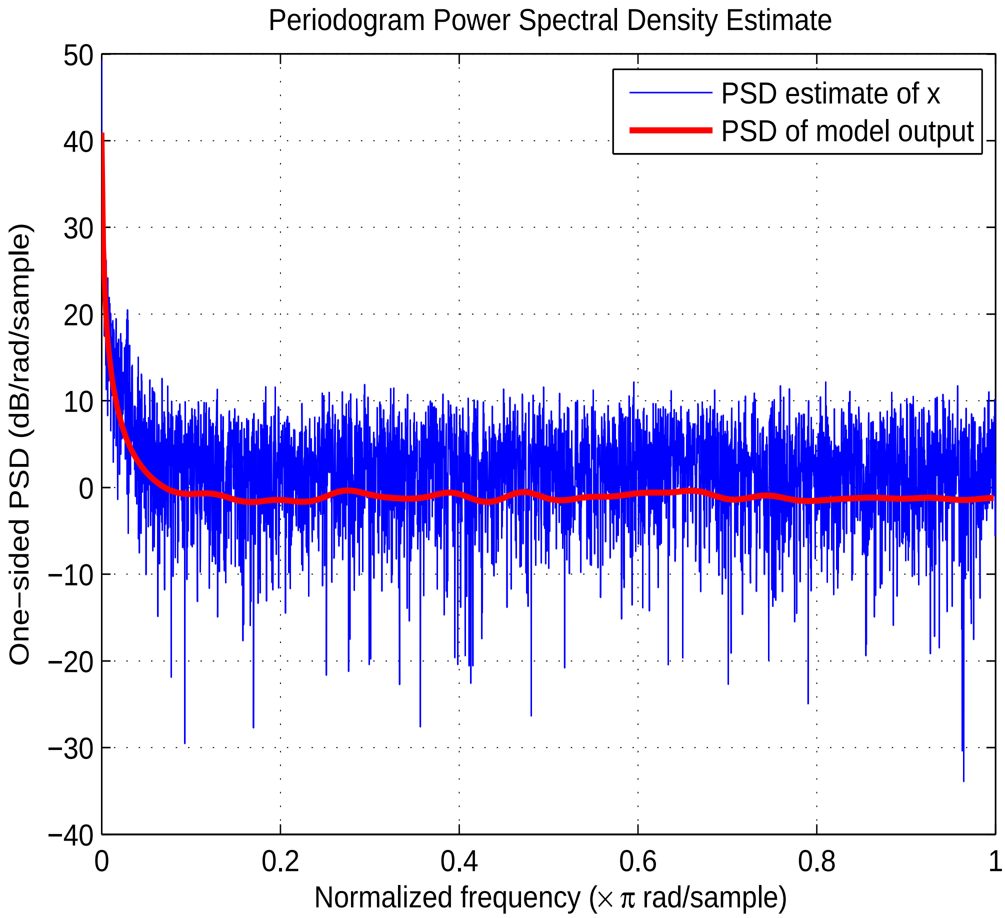

Figure 4.9: Power spectral density of the model vs. the power spectrum of the original data for standing, indicating a perfect match and thus proving the validity of the chosen model in describing the activity-acceleration data

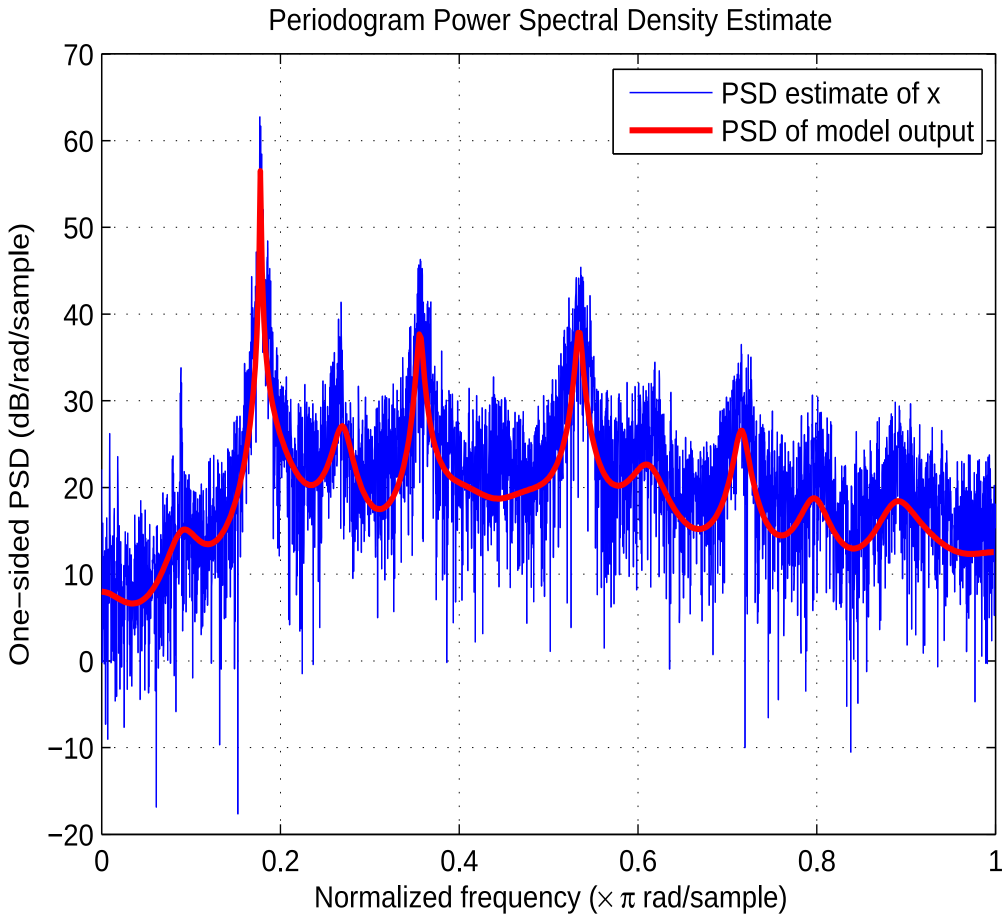

Figure 4.10: Power spectral density of the model vs. the power spectrum of the original data for walking, indicating a perfect match and thus proving the validity of the chosen model in describing the activity-acceleration data

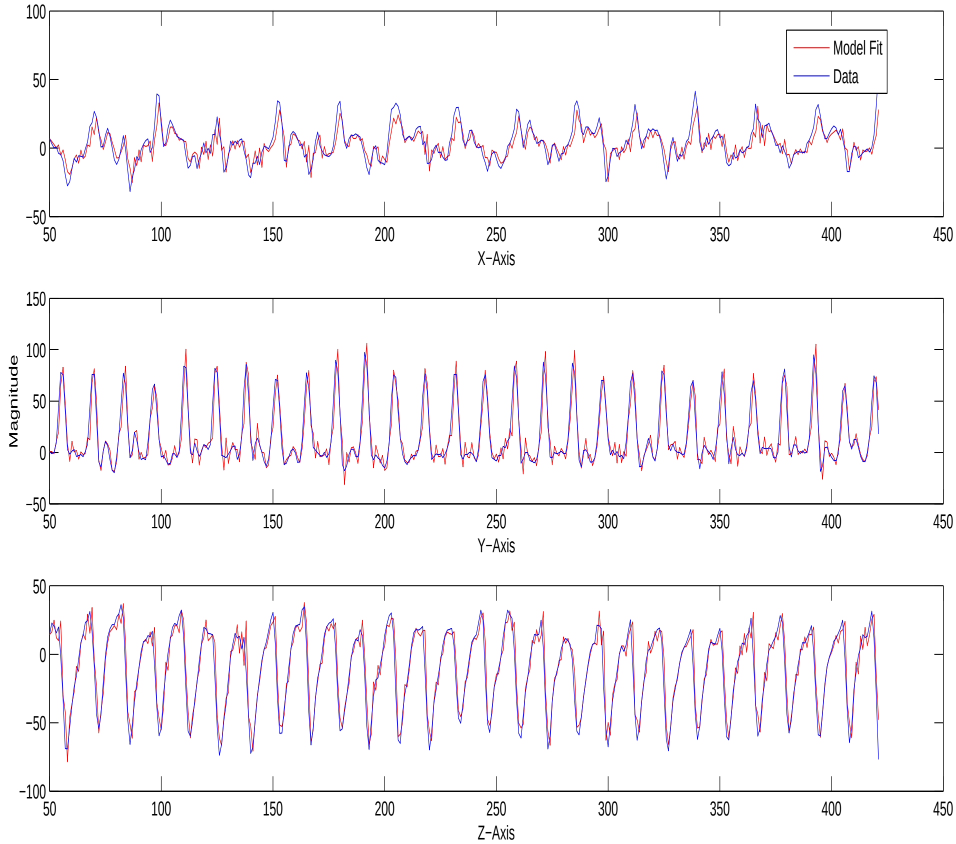

Figure 4.11: Fitting results for activity acceleration-signals of walking for three axes showing a good-fit

end features. The purpose was to identify the feature(s) having the best performance in classifying the activities used for AR-model identification.

All these features were tested with the forward-backward search [60], which is a well-known feature selection algorithm. With this procedure, a subset of best (giving the best classification result) features can be determined for the final analysis. In forward search (FS), every feature is tested for the classification one by one, and the best is selected to a subset of best features. The features that remain are then tested with the selected one, and the best one is selected to the subset and so forth. The procedure starts from one feature. The FS finds the best single features but does not find the best combination subset.

Backward search (BS) starts with classifying all features and removing the one that is lowering the classification result. In forward-backward combination, two features are selected with FS and one is removed with BS. The classification is usually done with a simple classifier. The classification, in this case, was done using artificial neural networks. AR-coefficients along with two other features gave the best classification accuracy for all activities. These two features are:

Signal Magnitude Area (SMA): As mentioned earlier, an acceleration signal is a linear combination of two components: a component due to gravitational acceleration and a component due to bodily motion. These components are separated using the method discussed in chapter 2. The component due to the body movements is then used to calculate SMA. It contains total power of the signal and is calculated as

$$SMA = \sum_{i=1}^{N} (|x(i)|) + (|y(i)|) + (|z(i)|) \tag{4.10}$$

where $x(i)$, $y(i)$ and $z(i)$ indicate the acceleration signal along x-axis , y-axis, and z-axis respectively.

Tilt Angle (TA): It refers to the relative tilt of the body in space and helps in distinguishing postures different in angle such as standing and lying. It can be defined as the angle between the positive z-axis and the gravitational vector g and can be calculated according to

$$\vartheta = \arccos(z) \tag{4.11}$$

Thus the feature vector used in this study for representing activities in the feature space includes the AR-coefficients augmented with SMA and the TA, shown in Figure 4.12. It is named augmented feature vector and can be represented as

$$F = [a_{x1}, \dots, a_{xM}, a_{y1}, \dots, a_{yM}, a_{z1}, \dots, a_{zM}, SMA, TA] \tag{4.12}$$

where $a_{x1}, \dots, a_{xM}$ are the AR-coefficients for X-axis, $a_{y1}, \dots, a_{yM}$ are the AR-coefficients for Y-axis, $4a_{z1}, \dots, a_{zM}$ are the AR-coefficients for Z-axis, SMA is the signal magnitude area and TA is the tilt angle.

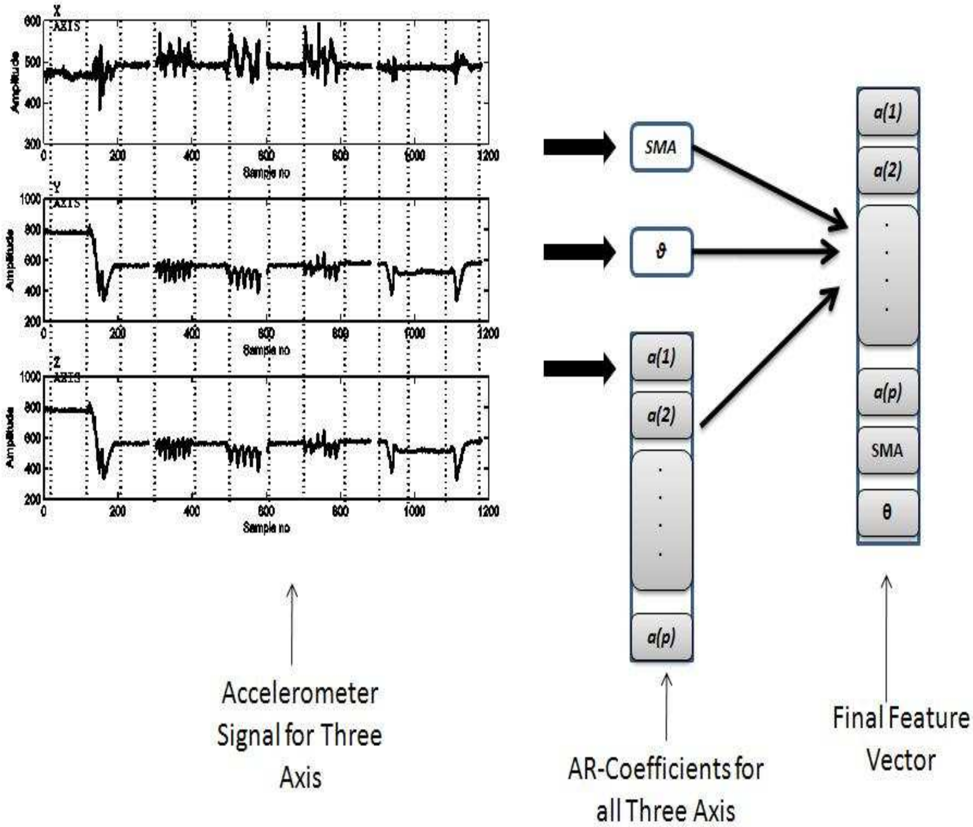

Figure 4.12: Block diagram, showing components of the augmented feature vector

Chapter 5

Recognizing a Diverse Set of Activities using Proposed Features in Controlled and Naturalistic Settings

This chapter describes in detail the procedure employed for evaluating the use of the proposed augmented feature model for the classification of a large and diverse set of physical activities in both controlled and naturalistic settings. It also presents details and the reasoning behind the implementation of a multi-stage classification scheme used for the given classification task.

5.1 Study Goal

The activity-set investigated in this study contained 15 physical activities: Lying, Sitting, Standing, Sit-Stand, Stand-Sit, Lie-Stand, Stand-Lie, Lie-Sit, Sit-Lie, Walking, Walking-upstairs, Walking-downstairs, and Running. Recognizing this activity-set is challenging because: (1) the activities are hard to discriminate as they share highly similar postures-patterns (such as sitting and standing) and movement-patterns (such as walking, walking-upstairs and walking-downstairs). (2) this high similarity among activities is not uniform throughout the whole activity-set. In other words, a subset of activities shares high similarity among its activities but is very different from another subset. For example, sitting and standing are very similar (hard to distinguish), however, they are very different from walking (easily distinguishable). (3) In addition to long-duration movements and postures, short duration movements such as sit-stand and stand-sit are also present which only last for few seconds. These are the most widely performed tasks that represent transition from one physical activity to another. Their recognition plays a vital role if real-time activity recognition is required because if these transitions are not handled properly they could result in a large number of miss classification. Recognizing these short-duration activities with a good accuracy has not be successful previously because the features employed in previous systems needed to be calculated over longer time-windows. Thus the goal of this study was to devise and evaluate a classification scheme that, unlike previous systems, can recognize a large and diverse set of physical activities with a high accuracy in real-time using the proposed augmented feature model.

5.2 Classification using Three Different Neural Network Training Algorithm

As mentioned in chapter 3, the neural network were chosen for initial classification during the model identification phase due to their high accuracy for activity classification in previous works. It was a feed forward backpropogation network with one hidden layer having the same number of neurons as the input layer. However, the scope of this initial classification was limited to only four activities compared to the current classification task that involved classification of 15 physical activities . Therefore, the first step in the overall evaluation approach was to investigate different neural network training algorithms and topologies. This section first provides brief description of the algorithms used and then discusses the classification results of these algorithms for both CL and NL datasets (section 3.4.2 and 3.4.3).

Various algorithms have been proposed proposed in the past to train a multilayer feed-forward network. There exists a theoretical framework that focuses on estimating the generalization ability of a network as a function of architecture and training set considering the region of weight space consistent with the training set; that is, a particular learning rule might favor some regions over others [74]. However, the suitability of a training algorithm in producing good generalization ability, in relation to a particular application, is usually determined by experiments [75]. In this study, three commonly applied neural network learning algorithms were investigated, namely, standard Backpropagation, Scaled Conjugate Gradient Algorithm and Backpropagation with Bayesian Regularization in order to find the best suited algorithm for the given classification problem.

Standard Backpropagation (BP): This algorithm uses gradient descent technique for iteratively updating the weights to map a set of input-output pairs. The input vector is multiplied by weight vectors to produce outputs at the hidden layer. Similarly, hidden layer outputs are multiplied by their respective weights and are propagated to the final output layer. Backpropagation minimizes the sum of squared error. Further details can be found in [74].

Scaled Conjugate Gradient (SCG): The purpose of scaled conjugate gradient technique is to achieve faster convergence in training in multilayer feedforward network. In these methods, a search is performed along conjugate directions [76]. The new search direction is determined by combining the new steepest descent direction with the previous search direction so the current and previous search directions are conjugate. This technique is based on the assumption that the error in the neighborhood of a given point is locally quadratic. Further details can be found in [77].

Bayesian Regularization (BR): The main goal of any classification problem is to develop a classifier that, once trained, should be capable of recognizing not only the training data but also the test data. In other words, the trained network should generalize well on the unseen data.

In order to achieve better generalization in multilayer feed-forward network training, a method has been proposed in [78] which employs Bayesian framework for constraining the size of network parameters by regularization. Regularization technique forces the network to settle to a set of weights and biases having smaller values. This causes the network response to be smoother and less likely to overfit and capture noise. Further details can be found in [78].

Each neural network model had an input layer with 32 neurons (corresponding to the 32-dimensional augmented feature model), one hidden layer (the number of hidden layers was kept limited to one due to the real-time recognition requirements) and an output layer with 15 neurons corresponding to 15 physical activities. Cross-validation was employed to evaluate the performance of the classifiers in case of CL dataset. The data from all the subjects in in this case were divided into six segments of equal length. Data from five segments were used to train the classifier whereas the data from the sixth segment were used as test data. The process was repeated until data from all the subjects'appeared in the test samples. In case of NL dataset, data from eight subjects were used to train the classifiers whereas data from the remaining two subjects were used for validation. Experiments were conducted with increasing number of hidden neurons until no further significant increase in the accuracy was achieved. All the results presented in Table 5.1 are based on the number of hidden neurons that gave the best accuracy in each case.

Table 5.1 summarizes the results for this experiment. It can be noticed that the BR algorithm gave the best performance among the three network training algorithms. However, the best overall performance was just 71.6% in the case of CL dataset and 56.5% in the case of NL dataset for the BR algorithm. The recognition rates for NL dataset are even lower as it is a subject-independent classification case where the classifiers have not seen the data from the test subjects before.

5.3 Need for Dimension Reduction and Discriminating Feature Extraction

All algorithms exhibited better recognition rates for lying and running activities. However, the recognition accuracy for the rest of the activities was low. This is due to the fact that these activities share highly similarity and thus overlap significantly in the feature space. For example, figure 5.1 and 5.3 show the range of possible estimates of SMA and TA, across all the subjects, for sitting whereas figure 5.2 and 5.4 show the range of possible estimates of these parameters for standing respectively. It can be seen that the average values of these parameter for the two activities are almost the same. Moreover, figure 5.5 and 5.6 show the power spectral density estimates for sitting and standing respectively which show the presence of almost the same frequency components in acceleration-signals.

It is due to these similarities in time and frequency-domain parameters which made the discrimination of these activities very difficult. In other words, the presence of very low between-class variance in the activity-data resulted in low recognition accuracy. Moreover, the BR algorithm (which showed the best performance among three algorithms) achieved its best accuracy with 30 neurons in the hidden layer. Such a large number of neurons is not feasible when real-time recognition (especially on hand-held devices) is desirable. One probable reason for this high number of hidden units might be the high number of neurons in the input layer (32 neurons) as some theories suggest that neural networks usually require at least the same number of hidden units as the input neurons in order to converge properly. Therefore, a method was required which not only achieves dimension reduction but also increases the low between-class variance to increase the class separability before the features were fed to the classifier. Dimension reduction by means of extracting discriminating features works on the idea of maximizing total scatter of the data while

Table 5.1: Average recognition results(%) for the three algorithms for both CL and NL datasets

Algorithm	Back Propogation		Scaled Conjugate Gradient		Bayesian Regularization	
Activity	CL Data	NL Data	CL Data	NL Data	CL Data	NL Data
Lying	93	92	94	92	95	95
Sitting	71	52	70	51	74	52
Standing	62	50	70	51	74	52
Lie-Stand	64	51	64	50	64	52
Stand-Lie	87	50	88	51	90	52
Lie-Sit	61	42	60	43	61	44
Sit-Lie	50	40	52	40	54	44
Sit-Stand	62	42	62	41	68	44
Stand-Sit	50	40	50	41	50	45
Walk-Stand	80	52	80	53	81	61
Stand-Walk	71	51	72	51	74	54
Walking	70	60	70	61	74	62
Walking-Upstairs	69	55	70	58	72	59
Walking-Downstairs	68	52	68	52	70	55
Running	81	60	80	61	85	66
Total	69	52.6	69.3	53	71.6	56.5

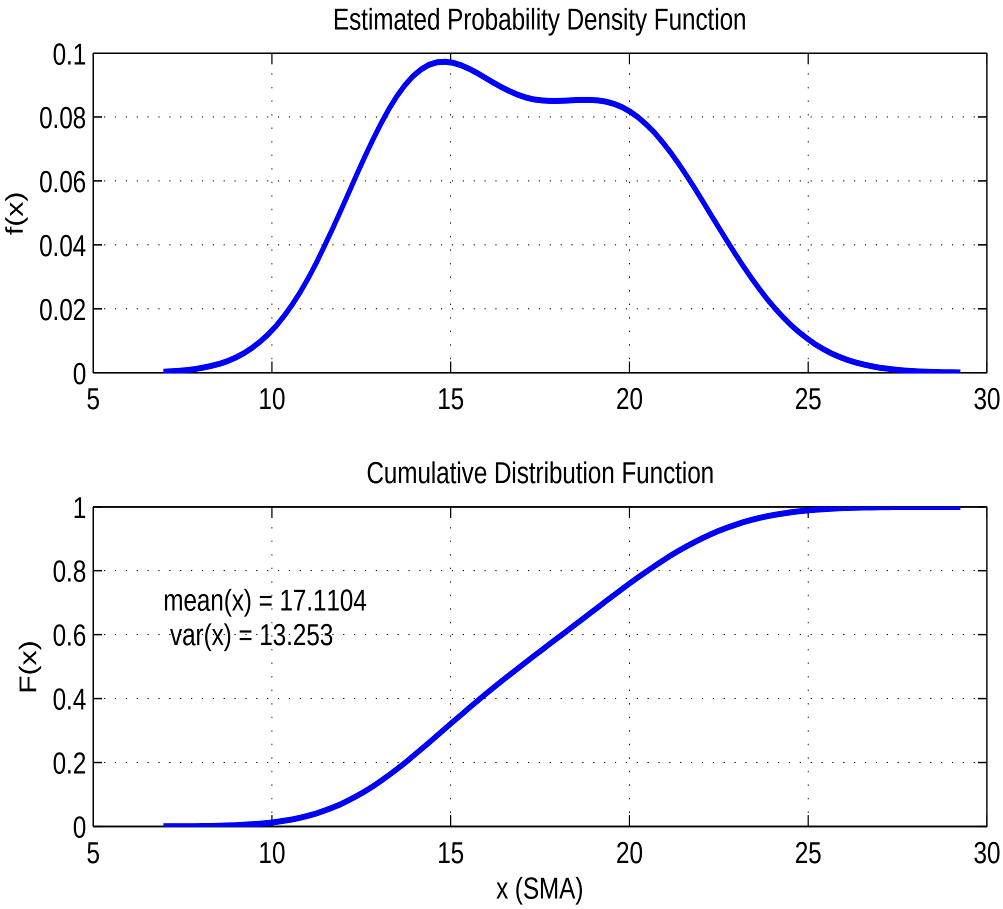

Figure 5.1: Probability density, Cumulative density functions, mean and variance of SMA estimates across all subjects for sitting

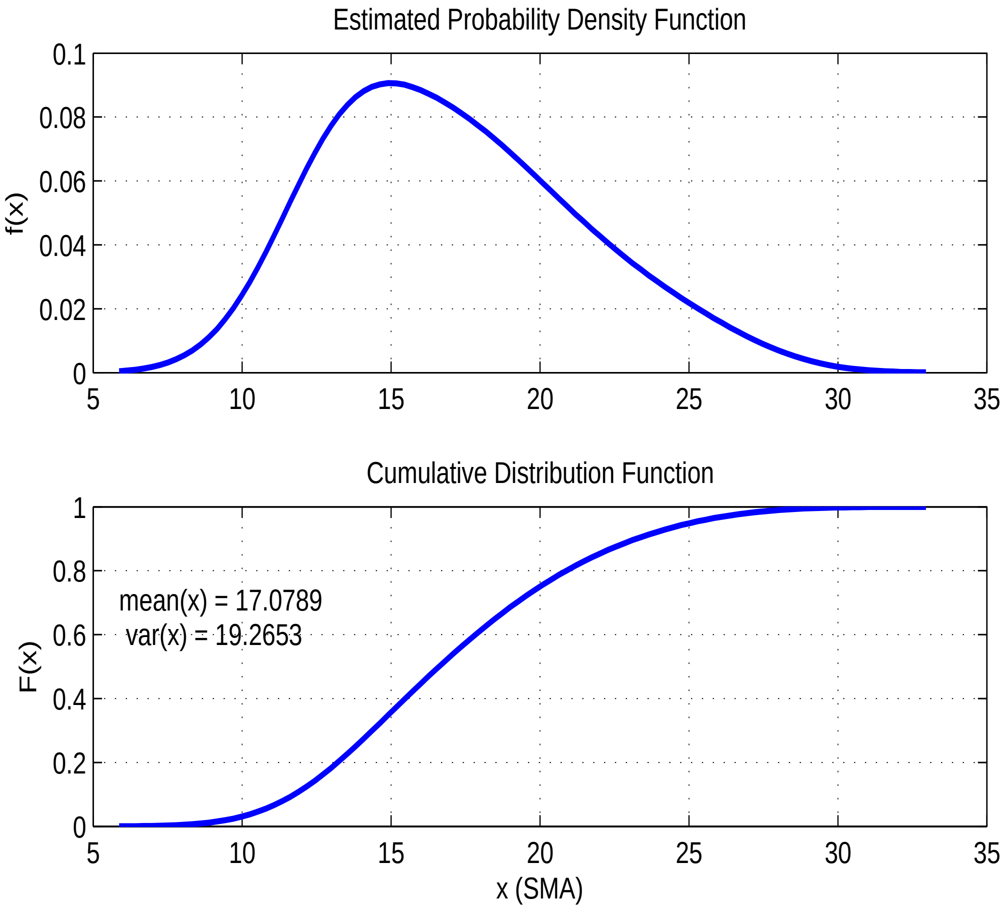

Figure 5.2: Probability density, Cumulative density functions, mean and variance of SMA estimates across all subjects for standing

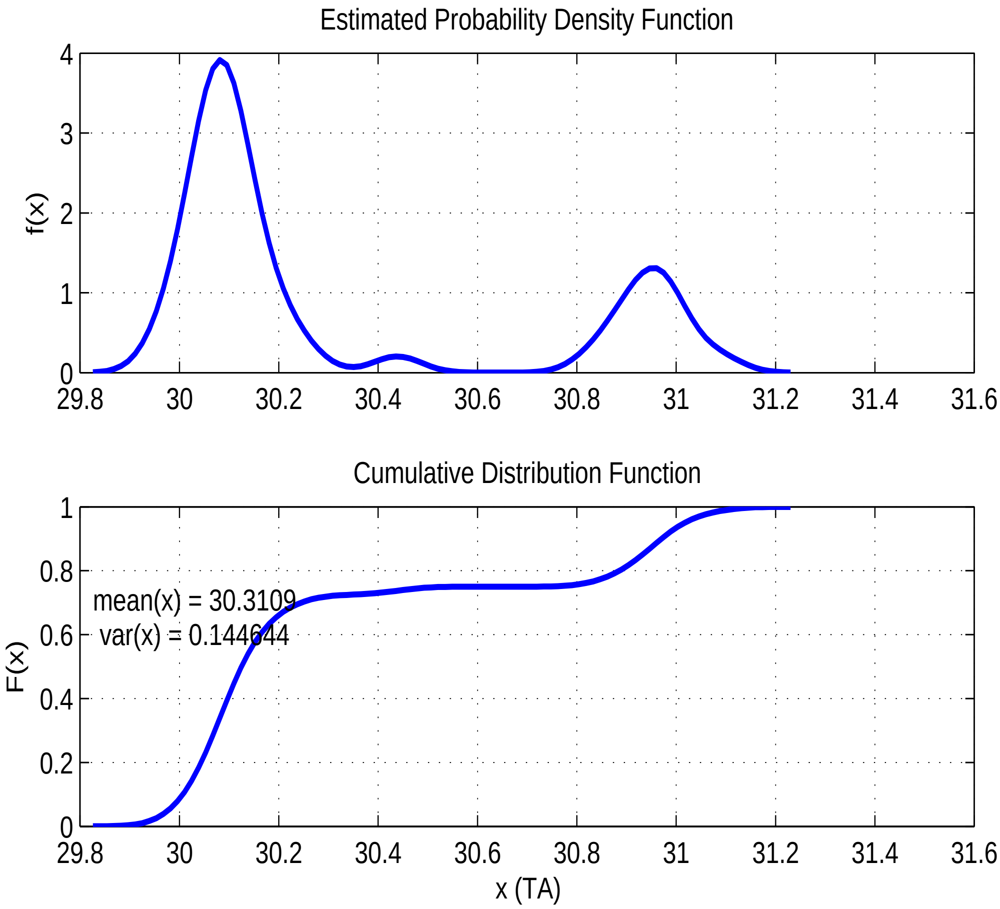

Figure 5.3: Probability density, Cumulative density functions, mean and variance of TA estimates across all subjects for sitting

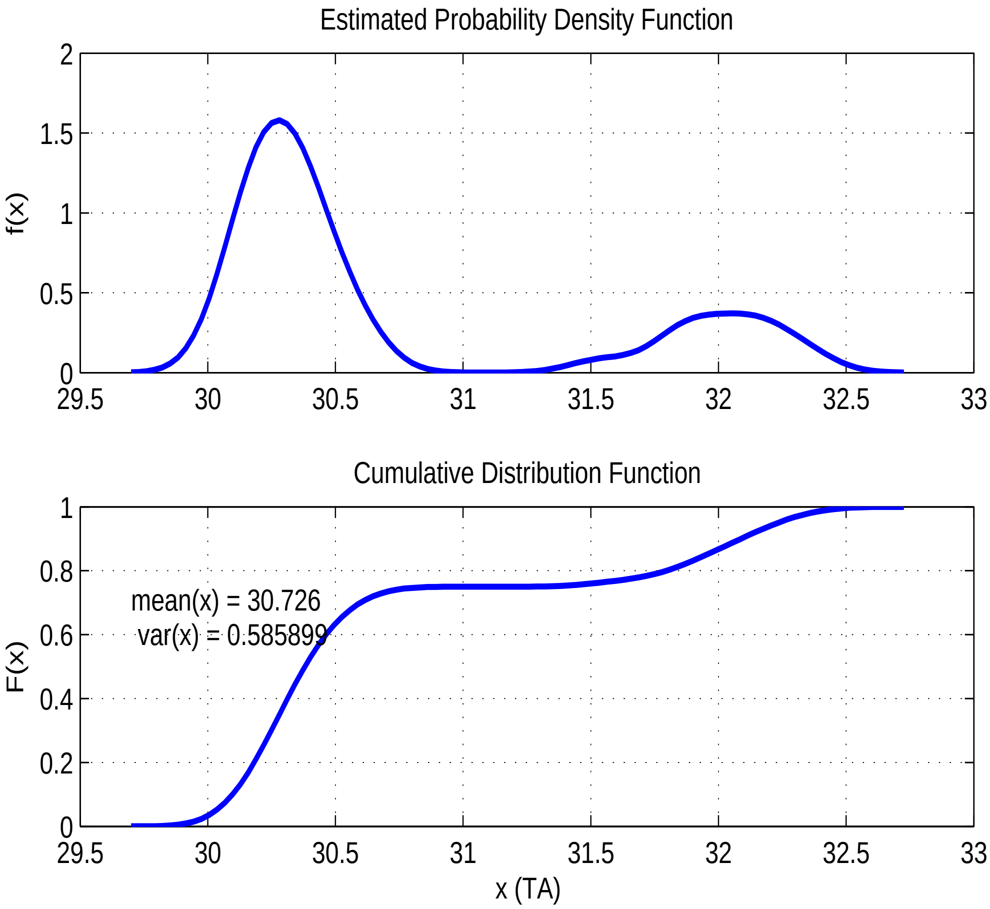

Figure 5.4: Probability density, Cumulative density functions, mean and variance of TA estimates across all subjects for standing

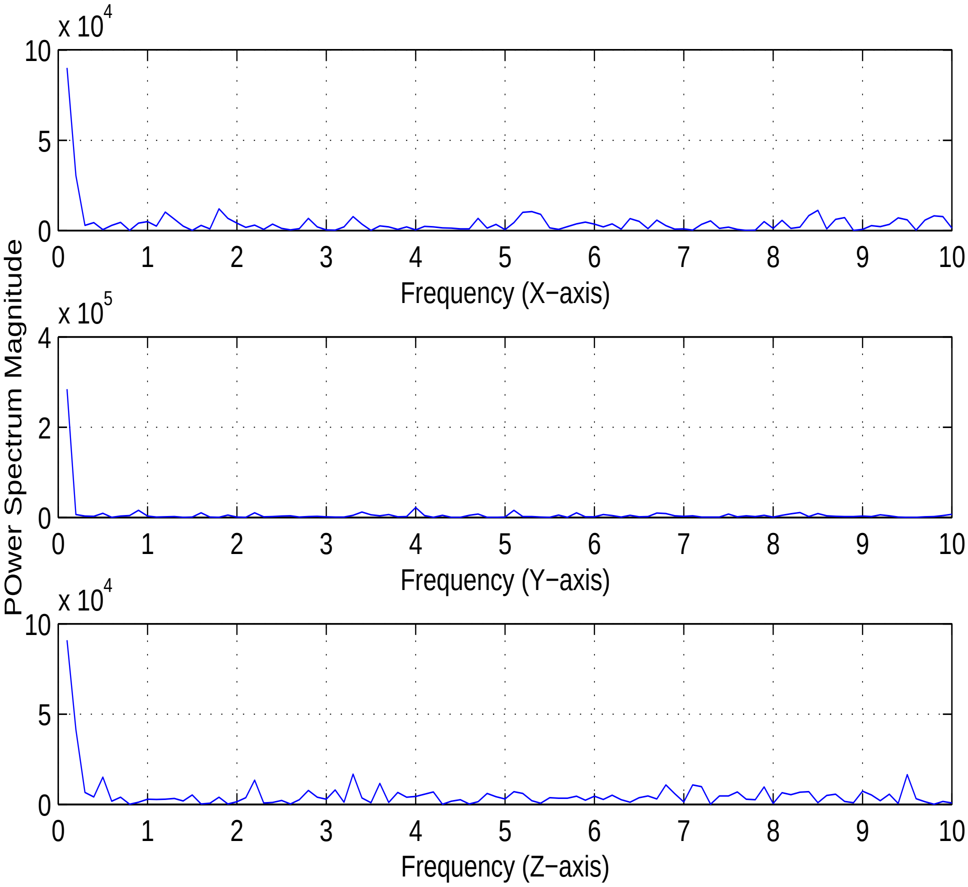

Figure 5.5: Power Spectral Density estimates for sitting, showing the presence of low frequency components

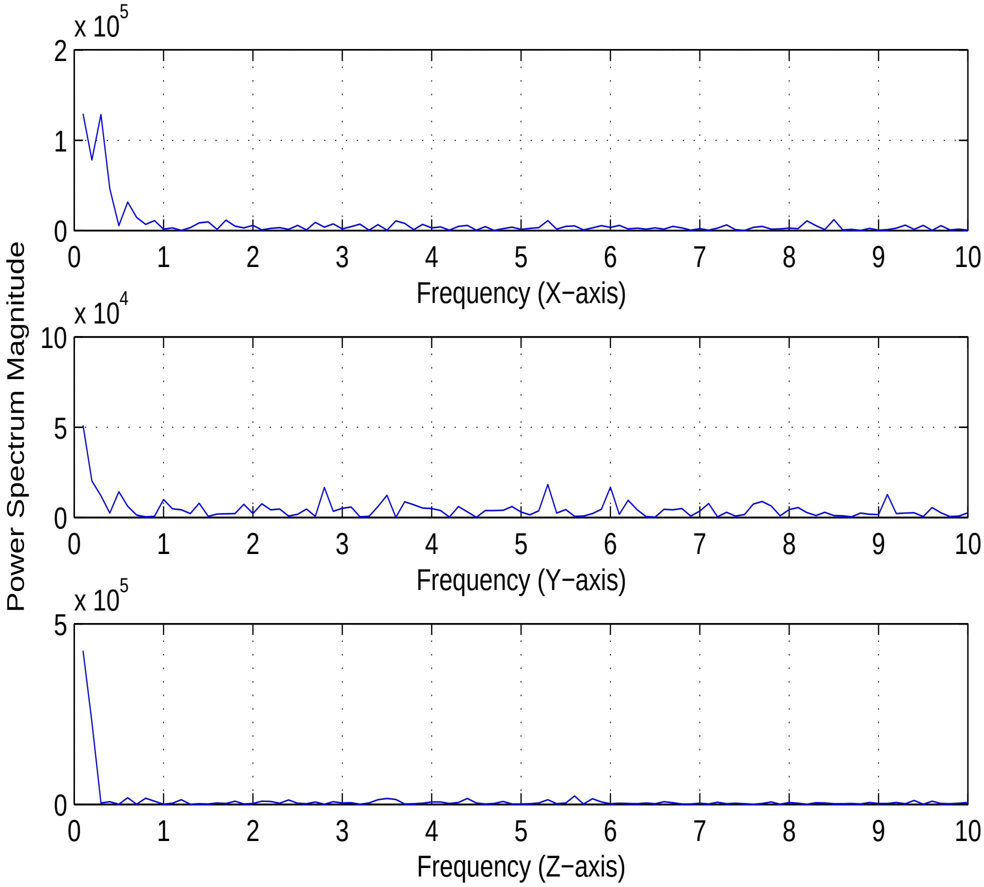

Figure 5.6: Power Spectral Density estimates for standing, also showing the presence of low frequency components

minimizing the variance within classes. One of the best techniques used for such purpose is Linear Discriminant Analysis (LDA).

Linear Discriminant Analysis (LDA): LDA easily handles the case where the within-class frequencies are unequal and their performances has been examined on randomly generated test data. This method maximizes the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability. The use of LDA for data classification is applied to classification problem in speech recognition. LDA produces an optimal linear discriminant function which maps the input into the classification space on which the class identification of the samples is decided. The within S_W and between S_B class comparison is done by following equations.

$$S_B = \sum_{i=1}^{c} J_i(\overline{m}_i - \overline{\overline{m}})(\overline{m}_i - \overline{\overline{m}})^T \tag{5.1}$$

$$S_W = \sum_{i=1}^{c} \sum_{m_k \in C_i} (m_k - \overline{m}_i)(m_k - \overline{m}_i)^T \tag{5.2}$$

where J_i is the number of vectors in ith class C_i. c is the number of classes and in our case, it represents the number of activities within each state. $\overline{\overline{m}}$ represents the mean of all vectors, $\overline{m}$ the mean of the class C_i and m_k the vector of a specific class. The optimal discrimination projection matrix D_{opt} is chosen from the maximization of ratio of the determinant of the between and within class scatter matrix as

$$D_{opt} = \arg\max_{D} \frac{|D^T S_B D|}{|D^T S_W D|} = [d_1, d_2, ..., d_t]^T \tag{5.3}$$

where D_{opt} is the set of discriminant vectors of S_W and S_B corresponding to the $c-1$ largest generalized eigenvalues λ and can be obtained via solving (10). The size of D_{opt} is $t \times r$ where $t \leq r$ and r is the number of elements in a vector.

$$S_B d_i = \lambda_i S_W d_i \;\; i = 1, 2, ..., c-1 \tag{5.4}$$

where the rank of S_B is $c-1$ or less and hence the upper bound value of t is $c-1$.

Thus, LDA maximizes the total scatter of the data while minimizing the within scatter of the classes. Augmented feature vectors were calculated for each window of the acceleration data for both CL and NL datasets and then to acquire a better feature space, the LDA was applied to the extracted augmented feature vectors. The new feature vectors using LDA on the augmented feature space can be represented as

$$F_i = A_i D_{opt}^T \tag{5.5}$$

where F_i and A_i represent the LDA-feature vector and augmented feature vector for the $i-th$ data window respectively. These LDA features were then fed to the BR classifier (only BR was used as its performance was better than the other two algorithms) for activity classification using the same cross-validation procedure adopted previously. The results are summarized in Table 5.2. It can be seen that there was a slight improvement in the accuracy in the case of CL dataset, however, no improvement in the overall recognition rate for NL dataset was observed. The recognition rates for lying and running improved a little but the accuracy for the other activities still remained low. One benefit of the use of LDA was the decrease in the number of hidden units. In this experiment, the BR neural network achieved its best accuracy with only 10 neurons in the hidden layer which was much better than the previous case (30 neurons). Thus the use of LDA succeeded in providing an effective data reduction for achieving the same accuracy as before with a less complex neural network however it failed in resolving the overlap or low-between class variance among the activities.

Table 5.2: Average recognition results(%) for the BR neural network for both CL and NL datasets after applying LDA

Activity	CL Data	NL Data
Lying	96	95
Sitting	74	51
Standing	65	54
Lie-Stand	66	52
Stand-Lie	90	51
Lie-Sit	62	45
Sit-Lie	54	44
Sit-Stand	67	44
Stand-Sit	51	46
Walk-Stand	82	60
Stand-Walk	73	54
Walking	74	61
Walking-Upstairs	73	59
Walking-Downstairs	70	55
Running	90	68
Total	72.4	56.6

5.4 Activity-clusters

The reason for failure in resolving the problem of low between-class variance in the previous experiment is due to the fact that activities tend to lie in clusters in the augmented feature space. In other words, a subset of activities share similarities and this subset is different from another subset in both its frequency and intensity characteristics. For example, sitting and standing are very similar and tend to lie in one cluster but they are very different from walking-upstairs and walking-downstairs which lie in a different cluster. Since LDA tried to improve total-scatter of the data, therefore, such a technique when applied to the augmented feature space for the whole dataset (for all activities) failed in extracting effective discriminating features as it worked on global mean instead of local mean (within the cluster). In other words, the extracted discriminating features maximized the separation among activity-clusters further, however, the separation among the classes within these cluster still remained small. Figure 5.7 shows the 3D feature plots of four transitions from the original feature space of the 15 activities. It can be seen that these transitions are clustered together with a very low between-class variance among them. Figure 5.8 shows the

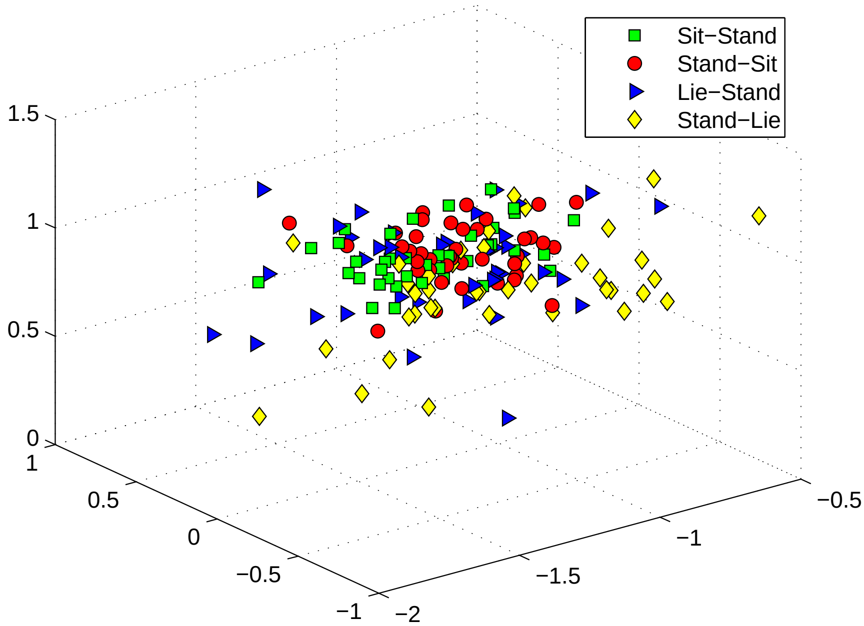

Figure 5.7: 3D feature plot showing just four transitions from the original feature space of 15 activities before applying LDA

3D feature plot of the same four transitions from the LDA-feature space of the 15 activities. In this case, LDA was applied using the global mean (mean for all the activities). It can be seen that the cluster has become more compact, however, the classes still exhibit a strong overlap. Therefore, it was proposed that LDA should be applied to each activity-cluster separately.

5.5 State-Activity-based Classification

5.5.1 Architecture

Based on their characteristics, postures and movement patterns activities were divided into three clusters or groups. These clusters were named 'States': Static, Transitions, and Dynamic. The grouping of 15 physical activities into three states is shown in Table 1.1. In the case of first three activities the human body is at rest and the net acceleration is due to the gravitational acceleration (lower frequency components only and smaller SMA). Therefore, they were grouped as static-

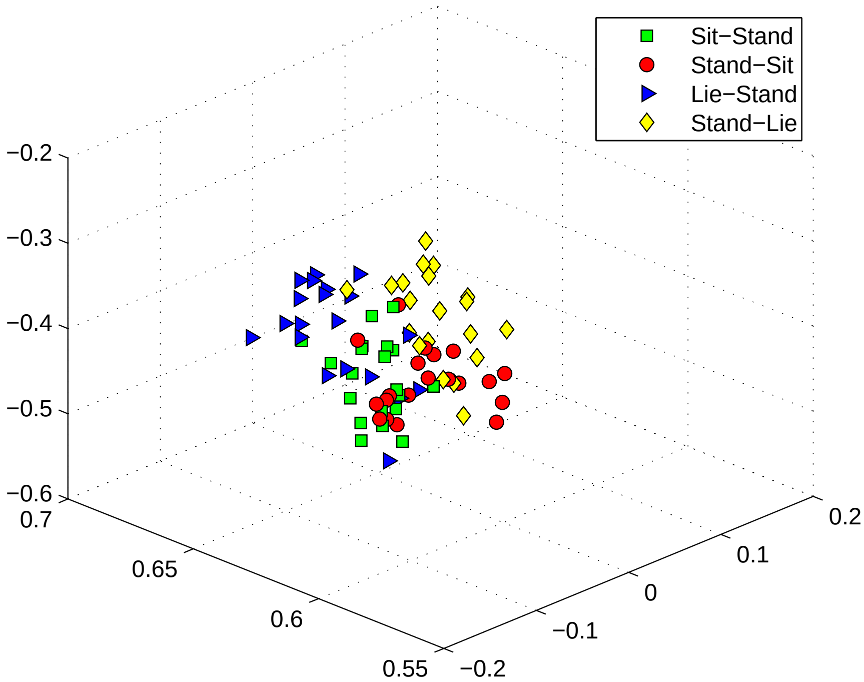

Figure 5.8: 3D feature plot showing just four transitions after applying LDA to the whole feature space of 15 activities

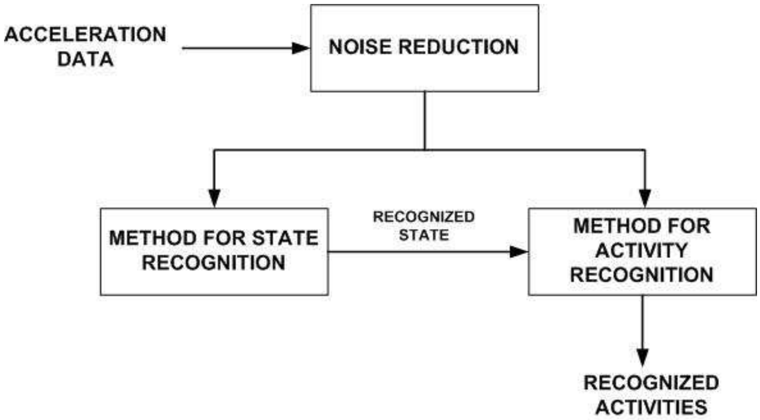

Figure 5.9: Block diagram for the proposed state-activity based recognition technique. AT the lower level, state (static, transitions or dynamic) was recognized by means of statistical signal features followed by activity recognition at the lower-level

activities. The next eight activities in Table 1.1 are short-duration movements from one activity to another activity. Therefore, these activities were assigned to transitions. Finally, in the case of last four activities in Table 1.1, the acceleration is mainly due to the bodily motion (higher frequency components and greater SMA). Therefore, these activities were grouped as dynamic-activities.

A multistage classification scheme was then proposed. At the first stage of the recognition process, state of a given data window was identified. Once the state was known, augmented feature vectors were calculated only for the activities within the recognized state. LDA was then applied to the augmented feature space to achieve dimensionality reduction and a better class separation. These LDA features were then fed to the classifier for the final activity recognition. Thus two classifiers were used for classifying a data window as a particular activity.

The proposed State-Activity-based classification scheme's architecture is illustrated in Figure 5.9. As mentioned above, it is a two-staged architecture which incorporates a different set of features at each stage. The first layer is called the state-layer where the second layer is called the activity-layer.

5.5.2 Results for State Recognition

The purpose of the state recognition is to determine the state to which an activity belongs. Since the three states differ significantly in their physical characteristics, as shown in figures, simple time-domain parameters such as mean and standard deviation were therefore employed for state recognition. A brief description of these features is provided in Appendix A and they were called the ‘State-features’. A neural network based on BR algorithm was trained using these features for both CL and NL datasets. The network converged to give a high accuracy in both cases using only 2 and 3 hidden units in the case of CL and NL datasets respectively. The recognition accuracy for in the case of CL dataset was almost 100% because (1) states were easily distinguishable (2) dataset was collected in the controlled settings with almost no variations in activity patterns across differnt subjects and (3) data from all the subjects were used for both training and testing the system. The recognition accuracy for NL dataset was lower than the CL dataset, as shown in Table 5.3. Nevertheless, 97.1% is a very good accuracy considering that the NL dataset was collected under naturalistic settings and test subjects were not part of the training process. These results confirm that states, clusters of highly similar activites, are easily distinguishable from each other even in free-living scenarios where the way of performing activties may vary significantly among the subjects.

Table 5.3: Average recognition results(%) for for state recognition for both CL and NL datasets

State	CL Data	NL Data
Static	100	99
Transition	100	95
Dynamic	100	97
Total	100	97

5.5.3 Final Results for Activity Recognition

Once the state for a given data-window was recognized, the tri-axial activity acceleration-signals were used to calculate augmented feature vectors. LDA was then applied to the extracted augmented feature space to extract discriminating features which maximize the between-class variance and minimize the within-class variance for the recognized state. The LDA features were then used as inputs to a neural network for final classification, as shown in figure 5.10. Figure 5.11 shows the three-dimensional feature plots for the four transitions after the application of LDA to the transition-cluster only. These plots prove the success of employing LDA algorithm on each state separately. The activities which were otherwise very hard to discriminate due to a significant overlap in the feature space have clearly been separated from each other.

Three separate networks were trained using the BR algorithm for three separate states: the neural network to recognize static activities (SNN), the neural network to recognize transitions (TNN), and the neural network to recognize dynamic activities (DNN). The input to each of these NNs was the output of the LDA module as shown in figure 5.10.

Different number of layers and neurons were tested in order to optimize the performance. The maximal value of the weights in the neuron connections was normalized to the modulus of 1. Different steps of the increment for the weights were also investigated. The training of ANN was also repeated several times by changing the input order in a random fashion.

SNN gave its best performance with two hidden units for the CL dataset and three hidden units for the NL dataset. Further increase in the number of hidden units did not result in any significant increase in SNN's accuracy for both datasets. TNN started giving its best results when the number of hidden neurons was increased to seven for both CL and NL datasets. No significant increase in accuracy was achieved beyond this number. One probable reason for slightly higher number of neurons in the case of TNN was the fact that it was required to classify larger number of activities as compared to static and dynamic cases. DNN started giving good recognition rate when the number of hidden neurons was increased to two in the case of CL and four in the case of NL datasets. The overall recognition results of the state-activity based classification scheme for all activities for both CL and NL datasets are summarized in Table 5.4 which show an average recognition rate of 97.9% and 85% for the CL and NL datasets respectively. Once again the recognition rate for the CL dataset was high because (1) There were very little variations in activity-data across different subjects and (2) Data from all the subjects were used to train the system and thus it was subject-dependent classification.

Recognition accuracy of 85% in the case of NL dataset is in fact very promising considering an activity set of 15 activities which include both short and long-duration activities. Moreover, the NL

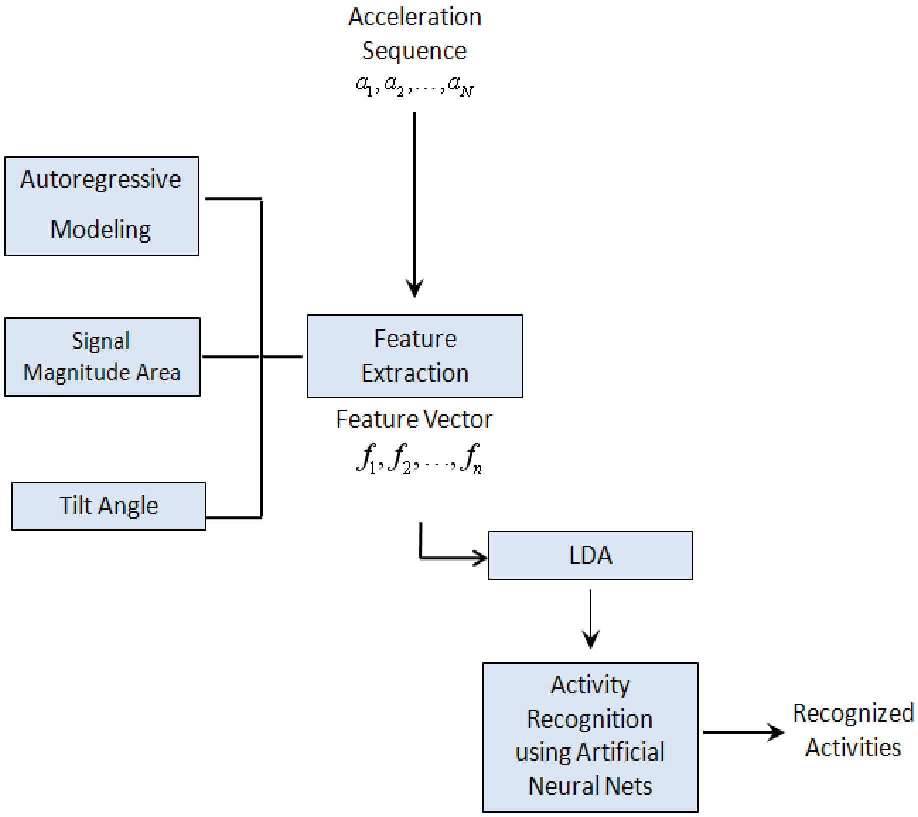

Figure 5.10: Block diagram for the activity recognition method: Once the state is recognized, activity acceleration-signals are used to calculate the augmented feature vectors, LDA is applied to increase the class separation and the LDA-features are then fed to the classifier to recognize the activities

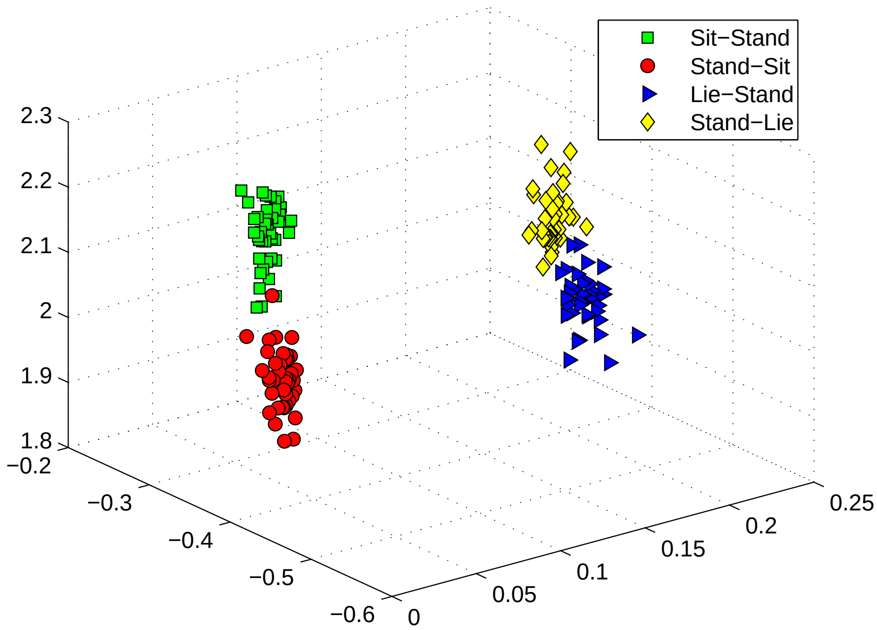

Figure 5.11: 3D feature plot for four transitions after LDA showing a much better class separation. Thus the application of LDA to the augmented feature space within the transition-cluster improved the class separation significantly

dataset represents the subject-independent classification case. In other words, the activity patterns from the test subjects were not seen by the system before. Therefore, an average recognition accuracy of 85%, especially high recognition accuracies in the case of dynamic activities, proved the success of the proposed framework for bodily actvities recognition under conditions close to those found in real-world settings.

Table 5.4: Average recognition results(%) for the complete state-activity classification scheme for both CL and NL datasets after applying LDA

Activity	CL Data	NL Data
Lying	99	99
Sitting	95	74.7
Standing	99	78.6
Lie-Stand	94	82.3
Stand-Lie	96	78
Lie-Sit	92	81
Sit-Lie	94	80
Sit-Stand	99	80.1
Stand-Sit	99	79.2
Walk-Stand	99	91
Stand-Walk	99	90
Walking	99	92.2
Walking-Upstairs	99	87.7
Walking-Downstairs	99	86.3
Running	99	96.2
Total	97.9	85

5.6 Conclusion

This study aimed to develop an accurate and robust classification scheme using the proposed augmented feature model of human activities for recognizing an activity-set of 15 physical activities in both lab and free-living conditions. The proposed classification scheme is effective in a sense that it was capable of recognizing a broad set of daily physical activities with an average accuracy of 97.9% in the lab-settings and 85% in naturalistic free-living settings. It was able to distinguish between the activities with high accuracy that exhibited difficulty in discrimination in the previous works. Examples include sitting and standing postures, sit-stand and stand-sit transitions, and walking-upstairs and walking-downstairs movements.

There are two main reasons behind the high recognition accuracy of the system. Firstly, the augmented feature model used in this study employs AR-coefficients obtained by AR-modeling of the activity acceleration signals. Since these signals are generated by an autoregressive process (shown in chapter 4), the AR-coefficients therefore provide very reliable estimate of the frequency spectrum of these signals and are an appropriate choice to be used as features for the their classification. Moreover, unlike previous features, they can be computed in real-time and are independent of the length of the data-sequence. This makes them an ideal choice to be used as features for the classification of short-duration activities (transitions).

Secondly, the low between-class variance in the activity-data or overlap between different activities is resolved by applying LDA to the augmented feature space of separate groups of activities (states) to extract discriminating features that correspond to a single state only. This not only increased the class separation within a particular state but also provided effective dimensionality reduction which helped reducing the complexity of the neural network as fewer hidden units were required to perform the classification task.

Lastly, one important advantage of the proposed multistage classification scheme is that it is simple and makes it easy to focus on states. For example, in cases where only dynamic activities are of interest, the rest of the two states can be ignored.

Chapter 6

Accelerometer's Position and Attachment Free HAR using Proposed Features

Previous chapter presented details on the state-activity-based classification scheme which used the proposed augmented feature model to recognize a variety of daily physical activities using only a single accelerometer with a high accuracy both in controlled and uncontrolled environments. This chapter presents details and the reasoning behind the implementation of the accelerometer's position-independent physical activity recognition scheme. In this study, the PF dataset (section 3.4.4) was used for analysis and evaluation.

6.1 Study Goal

Long-term activity recognition in free-living conditions brings along several technical requirements which must be addressed. These include instrument usability, ease-of-use, energy consumption, reliable wireless communications and secure transfer of information. Many of these issues are being resolved with the development of home wireless network technologies and very low power instruments that are designed to be used in wearable monitoring systems. However, the requirement of developing easy-to-use cost-effective recognition algorithms that can function robustly in free-living conditions without forcing subjects into a fixed life pattern or hindering their daily activities still needs to be addressed.

Ensuring the ease-of-use requires addressing several factors such as the number of the sensors used, their comfort and their location on the human body. In general, the output of any body-worn accelerometer depends on the position at which it is placed. The output of an accelerometer, when positioned at a lower-body position such as legs, registers higher frequency components and greater magnitude compared to the scenarios where the sensor is positioned at an upper-body position, such as chest. Moreover, in order to have an accurate estimate of certain parameters such as tilt-angle, an accelerometer needs to be firmly attached to the human body. This requirement ensures that the sensor's orientation will not change while users perform any activity which involves bodily motion, such as walking and running.

The activity acceleration-signals, therefore, can vary significantly for different positions on the human body, even for the same activity. The problem gets further compounded if the placement of the sensor is not firm, in other words, if the sensor is placed freely in any pocket without a firm attachment to any specific human body-part. Such changes in orientation, magnitude, and frequency thus make accelerometer's position and attachment free physical activity recognition very challenging. Almost all previous works thus require accelerometers to be firmly attached to subjects' bodies. Most studies employed multiple accelerometers attached at different sites [4, 12, 16, 17, 19–23, 79–83], whereas others investigated the use of a single tri-axial accelerometer mounted at waist, chest, thigh, wrist, or sternum [24–33, 35, 36]. Such configurations would force subjects into a fixed life pattern and hinder their daily physical activities and thus make these systems impractical for long-term activity monitoring during unsupervised free living.

The aim of this study was therefore to recognize physical activities without posing any preconditions on accelerometer's position and orientation relative to a subject's body yet maintaining the same high standards in terms of accuracy.

6.2 Exclusion of Tilt Angle from the Feature Model

Table 6.1 summarizes the features used in the state-activity-based classification scheme (section 5.5). The first column lists the features used for the state-recognition, whereas, the second column lists the features employed for the activity-recognition task. Among these features, tilt-angle played a vital role in distinguishing static postures: lying, sitting and standing, as well as, the transitions between these static postures. However, in order to have an accurate estimate for the tilt-angle from the tri-axial activity acceleration-signals, accelerometer needs to be firmly attached to the human body. Loosely placing the sensor on the human body could result in changes in sensor's orientation while subjects perform a dynamic activity and, thus, makes it very difficult to land

Table 6.1: Features employed in the State-Activity-based classification scheme

State-Recognition	Activity Recognition
Mean	Autoregressive-coefficients
Variance	Signal Magnitude Area
Standard Deviation	Tilt-Angle

on a correct estimate for the tilt-angle. Therefore, tilt-angle was excluded from the feature-list in the case of accelerometer's position and attachment free recognition. Consequently, the three static postures: sitting, standing and lying, were combined into a single class called the resting-activity. Moreover, transitions between different activities were also not included in this study.

6.3 Feature Analysis

The feature extraction phase was proceeded by the feature analysis phase, which was performed in the following two steps.

- Firstly, the classification performance of the different configurations of the front-end features for a single sensor position was analyzed. The purpose was to identify the feature(s) having the best performance in classifying activities from a single sensor-site (one of the five pockets/sites used for data collection). The features mentioned above were tested with the forward-backward search (section 4.4), which is a well-known feature selection algorithm. The AR-coefficients augmented with the SMA, i.e., the AR-SMA proved to be the best discriminating features for all activity classes for all sensor positions, considering one at a time.

- Secondly, the activity-data from all 5 sensor positions were combined into a single dataset to evaluate the classification performance of the AR-SMA. A significant decrease in the performance was witnessed due to high within-class variance resulting from positioning the sensor on 5 different sites. The output patterns for walking, for example, vary at three different positions as shown in Figure 6.1. To minimize this variance, a two-level classification scheme was proposed, i.e., classifying the acceleration-signal to be either from upper-body (chest and inner jacket pocket) or from lower-body positions (front and rear trousers pocket) before classifying the activity itself. Thus the idea was to perform position classification before the activity classification to achieve position-independent activity recognition. During the analysis it was revealed that the activity acceleration-signals for all dynamic activities registered higher frequency components for the lower-body sensor positions, i.e., front and rear trousers pocket, and lower frequency components for the upper-body sensor positions, i.e., chest and inner jacket pocket. Since during the resting activity the body is at rest, same frequency components (very low) were therefore seen for all sensor positions. Therefore, Spectral Entropy (SE) was employed for the initial position classification.

6.4 Position-Free Classification Scheme

6.4.1 Architecture

Based on our findings, a two-level classification approach was finalized. Its architecture is illustrated in Figure 6.2. At the lower level, the SE was employed to recognize 3 classes, i.e., the resting activity, dynamic-activity (upper-body), and dynamic-activity (lower-body). Such a division helped reducing the high within-class variance for dynamic activities resulting from the upper and lower-body sensor positions and avoiding the cost of computing the AR-SMA feature when the subject is at rest.

If the resting activity was not recognized at the lower-level, the system output the sensor position as upper or lower-body for the case of dynamic activities. The AR-coefficients and SMA were then calculated from the noise reduced acceleration signal to form an augmented feature vector. However, a high within-class variance and low between-class variance due to different sensor positions, i.e., front and rear trousers pockets in the case of lower-body whereas chest and inner jacket pockets in the case of upper-body, could still exist in this new augmented feature space.

As mentioned in section 5.3, LDA produces an optimal linear discriminant function which maps the input into the classification space on which the class identification of the samples is decided. Thus to acquire a better feature space, the LDA algorithm was applied to the extracted

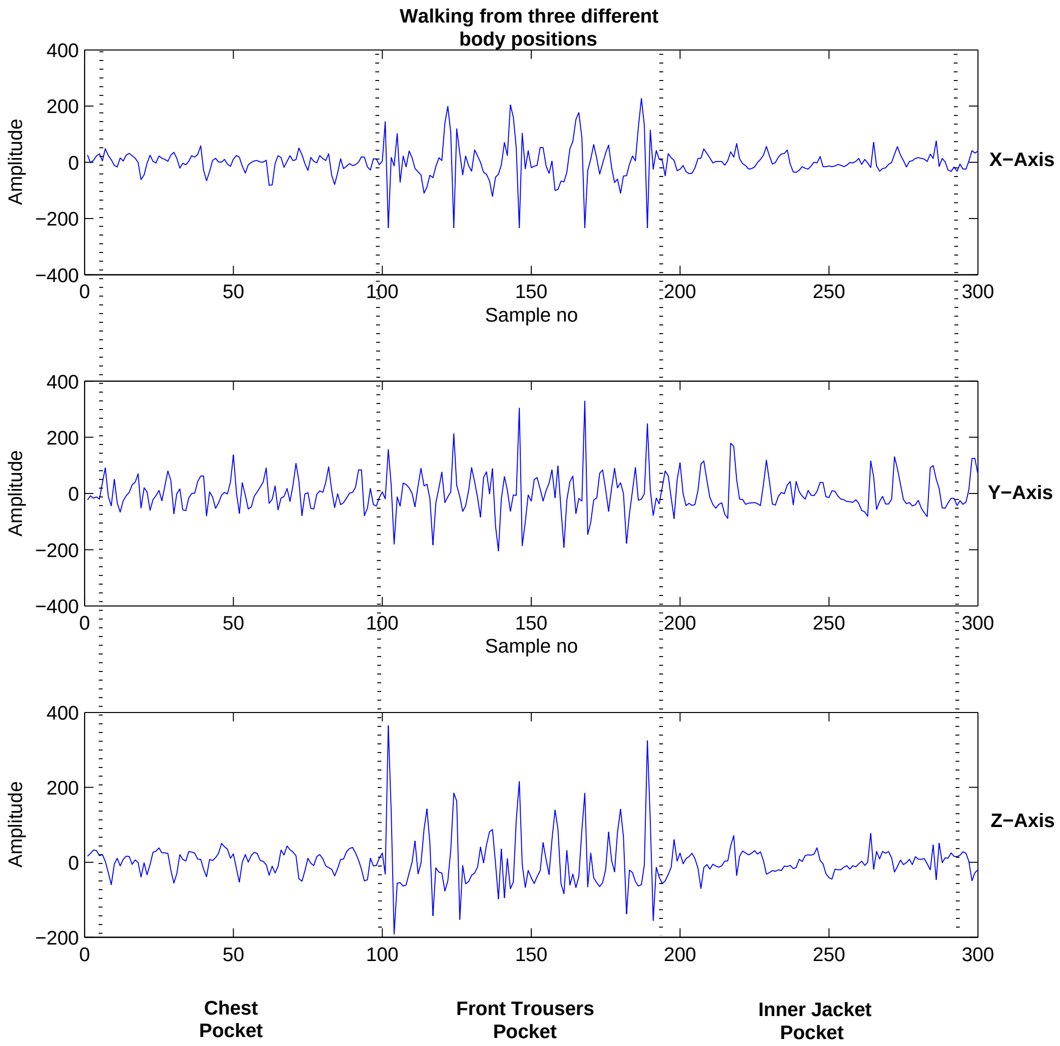

Figure 6.1: Sample acceleration signals for walking from three different positions.

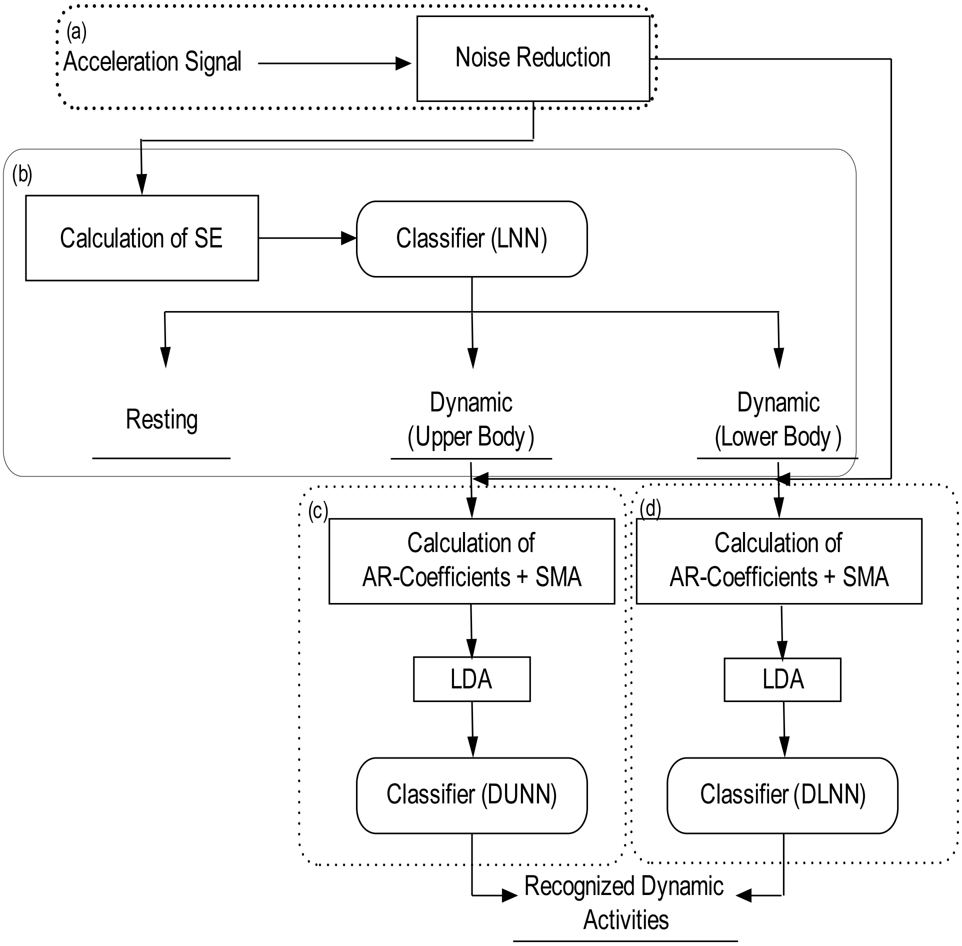

Figure 6.2: Block diagram of the proposed recognition scheme: (a) A moving average filter of order 3 was used to filter out the random noise from the acceleration signal. (b) At the lower-level, the SE and the neural net (LNN) was employed to recognize three classes. The sensor position was outputted as lower or upper-body in the case of dynamic activities (absence of resting state). (c) Augmented features (AR-coefficients + SMA) were calculated, LDA (see the text) was applied and the neural net (DUNN) was employed to recognize dynamic activities in the case of upper-body. (d) DLNN was used to recognize dynamic activities in the case of lower-body.

augmented feature vectors of different dynamic activities. The new feature vectors using LDA on the augmented feature space can be represented as

$$F_i = A_i D_{opt}^T \tag{6.1}$$

where F_i and A_i represent the LDA-feature vector and augmented feature vector for the $i-th$ dynamic activity sample respectively. Each neural network was trained using the BR algorithm (chapter 5). The training of each network was also repeated several times by changing the input order in a random fashion. The training and the testing datasets were composed of mixture of activity data collected from the five sensor positions.

For the lower-level recognition, only one network (LNN) was trained. The inputs to LNN were the SE-features. It consisted of one hidden layer with three neurons (chosen after experimenting with different number of neurons) and an output layer with three neurons corresponding to three classification outputs, i.e., the resting activity, dynamic activity (lower-body), and dynamic activity (upper-body). For the upper-level recognition, two networks were used, i.e., a neural net to recognize the dynamic activities from the lower-body positions (DLNN) and a neural net to recognize the dynamic activities from the upper-body positions (DUNN). The inputs to each of these networks were the LDA-features. Each of these networks had one hidden layer with five neurons (again chosen after experimenting with different number of neurons) and an output layer with six neurons corresponding to six dynamic activities.

The classical cross-validation [84] was adopted to evaluate the between-subject accuracy of the system. In other words, the networks were first trained with data from all subjects except few and then tested with data from the excluded subjects. The accuracy was then calculated as the proportion of correctly classified windows of data across all activities. This process was repeated until all subjects had participated in the testing datasets. The final overall accuracy was then calculated as the average accuracy across all train-test cycles.

6.4.2 Experimental Results

Performance of the proposed hierarchical recognition system was then validated in the following three studies.

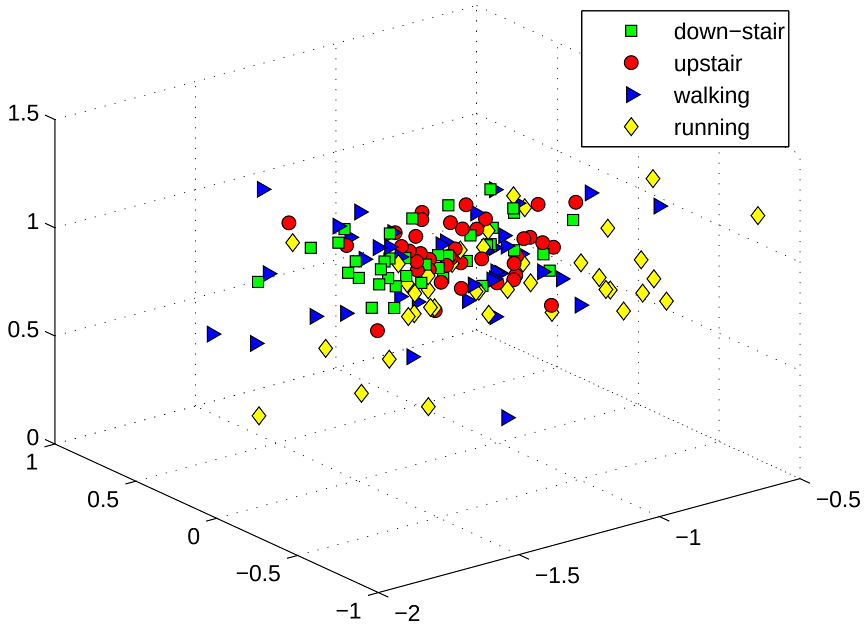

Figure 6.3: 3D-feature plot for four dynamic activities recorded from five different body positions, showing a high within-class variance.

Single-Level Recognition without LDA: In this study, a single BR-based neural network was used to recognize all seven activities without employing the proposed hierarchical recognition scheme. Features including the AR-coefficients, SMA, and SE were calculated to form a single feature vector. The 3D-representation of the feature space is shown in Figure 7.2, only four classes are shown for the sake of visualization. Severe non-linearity and a high within-class variance could be observed. These features were used to train the network. During testing, each test activity was modeled in a similar fashion and the ANN was used for recognition. The network had one hidden layer and it gave its best performance with 15 hidden neurons. No significant improvement in the accuracy was achieved beyond this number. The recognition results are summarized in Table 6.2, showing an average recognition of only 47% only.

Single-Level Recognition with LDA: In this study, after calculating the AR-coefficients, SMA, and SE, the LDA was applied to the extracted feature space. The LDA-features were then used to train a single BR-based neural network. The LDA-features for the four activities are shown

Table 6.2: Average recognition results(%) for the first experiment

Activity	Single-Level (S-L)
Resting (Lying/Sitting/Standing)	72
Walking downstairs	42
Walking upstairs	39
Walking	44
Running	52
Cycling	44
Vacuuming	36
Total	47

Table 6.3: A comparison of average recognition results(%) for the first and the second experiment

Activity	Single-Level (S-L)	S-L with LDA
Resting (Lying/Sitting/Standing)	72	89
Walking downstairs	42	53
Walking upstairs	39	51
Walking	44	56
Running	52	68
Cycling	44	50
Vacuuming	36	44
Total	47	58.7

in Figure 7.3. They show improved class separability. However, a high within-class variance could still be observed. During testing, each test activity was modeled in a similar fashion and the network was used for recognition. The network had one hidden layer and it gave its best performance for nine hidden neurons. No significant increase in accuracy was achieved beyond this number. The recognition results are summarized in Table 6.3, showing an average recognition rate of 58.7% only.

Proposed Hierarchical Recognition: In this study, the proposed hierarchical recognition scheme was used to achieve accelerometer's position-independent activity recognition. Figure 7.4 demonstrates the LDA-features for the four dynamic activities collected from the lower-body sensor positions, i.e., front and rear trouser pockets.. A significant improvement in class separabil-

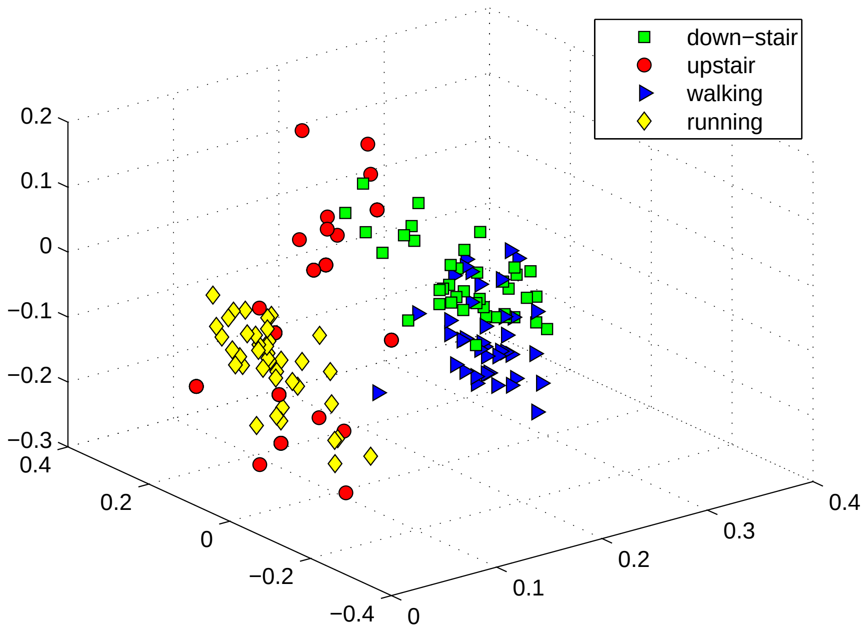

Figure 6.4: LDA feature space for four dynamic activities, recorded from five different body positions, after applying the single-level recognition system.

ity and a very low within-class variance could be observed. The recognition results for this study are summarized in Table 6.4, showing an average recognition rate of 94.4% which is a significant improvement over the recognition rates of the two previous studies.

6.5 Conclusion

In general, the output of any body-worn accelerometer depends on the its location on the human body and can vary significantly for the same for different locations. These variations result in high within-class variance which reduces the recognition accuracy significantly. Therefore, most of these systems require accelerometers to be firmly attached to specific body-parts, thereby forcing subjects to live into a fixed life pattern which can be burdensome especially during long-term recognition.

Though the system showed high accuracy in distinguishing 15 activities, it still required users to attach the accelerometer firmly to their chests. As mentioned earlier, this requirement is not feasible for real-life scenarios. The accuracy of the proposed system, when tested by freely placing the sensor in different pockets, went down to 47%.

The aim of this study was to implement a single tr-iaxial-accelerometer-based human activity recognition system without posing any preconditions on accelerometer's position and orientation relative to a subject's body. About 24 hours of activity data were collected on 7 bodily activities of the daily living from 8 elderly subjects at home, outside the laboratory. Activities were recognized from the data by loosely placing a tri-axial accelerometer in 5 different pockets, without attaching it firmly to the subjects' bodies. Annotations were performed on the spot by the subjects using a bluetooth headset together with speech recognition software which resulted in very little interference while performing the activities.

In the state-activity based classification framework , the tilt angle (TA) was used as a part of an augmented feature vector to recognize three static activities, including lying, sitting, and standing, with an above 90% average recognition accuracy. The TA refers to the relative tilt of the body in space and its computation requires accelerometer's firm attachment to the body. In cases where the sensor is placed freely in different pockets, the sensor's orientation can undergo arbitrary changes while performing an activity, it is therefore very hard to compute a reliable estimate of

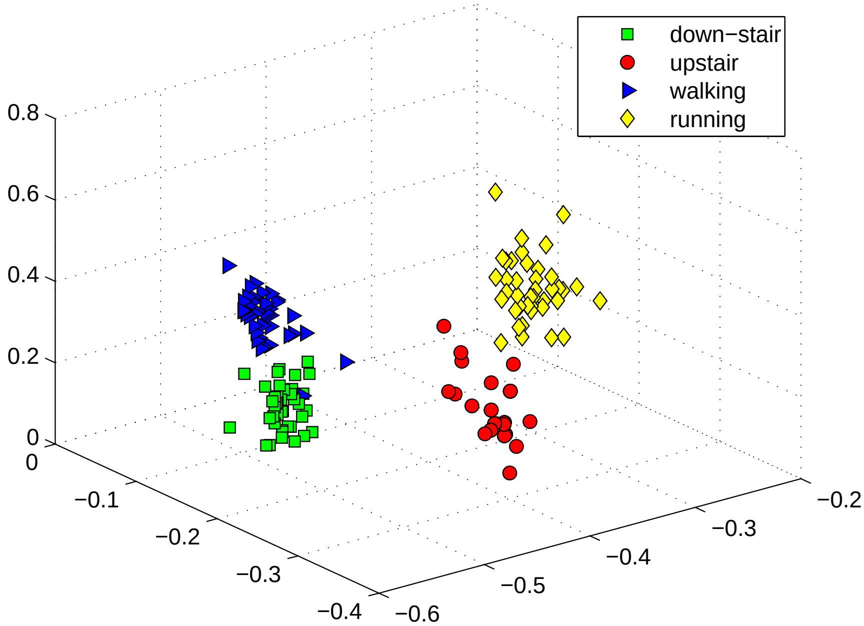

Figure 6.5: LDA feature space for four dynamic activities, from lower-body, i.e., front and rear trousers pockets, after applying the proposed hierarchical recognition system.

Table 6.4: A comparison of average recognition results(%) for all three experiments

Activity	Single-Level (S-L)	S-L with LDA	Hierarchical Scheme
Resting (Lying/Sitting/Standing)	72	89	98
Walking downstairs	42	53	96
Walking upstairs	39	51	94
Walking	44	56	96
Running	52	68	96
Cycling	44	50	94
Vacuuming	36	44	87
Total	47	58.7	94.4

the TA. Moreover, since the body is at rest, the three postures register almost the same frequency components and the intensity. Therefore, these activities were combined into a single group, i.e., the resting activity.

However, a higher level analysis can be employed to achieve further classification of the resting activity as lying, sitting, and standing. For instance, by employing an improved knowledge of the transitional movements, such as lie-to-sit, sit-to-lie, sit-to-stand, and stand-to-sit, these activities could be distinguished from each other with a greater accuracy. In other words, if the system recognizes that the subject is currently resting and he/she has just undergone a sit-to-stand transition, then it can infer that the subject is now standing.

A high within-class and a low between-class variance, caused by change in sensor orientation, magnitude, and frequency, makes accelerometer's position free human activity recognition very challenging. Therefore, extracting discriminating features, which minimize and maximize these variances respectively, was crucial. Linear discriminant analysis was employed this purpose. However, it is a linear technique in nature and does not perform well when severe non-linearity is involved. The experimental results of our the second study, i.e., single level recognition with LDA, support this fact.

To improve the recognition accuracy, a hierarchical recognition approach was employed to separate the dynamic activities from the upper and lower-body sensor positions using their SE-values. Consequently, the high within-class variance resulting from these positions was removed, which reduced the complexity of the classification task. A better feature space was then created by applying the LDA on the augmented feature space, i.e. the AR-coefficients augmented with the SMA.

There are clear limits on what can be achieved in a free-living activity monitoring environment using a single accelerometer, without posing any pre-conditions on its position and orientation. A greater number of sensors, attached firmly to different body-parts allow more accurate activity classification. However, the proposed system is more practical for continuous long-term activity monitoring in free-living subjects because of its simplicity, ease-of-use, compliance, lower cost, and the ability to recognize some important dynamic activities of daily living with a 94.4% average recognition accuracy.

Chapter 7

Smartphone based HAR using Proposed Features

This chapter presents details on the implementation of light-weight classification scheme that employs the augmented feature model for sensor's position-free physical activity recognition using an accelerometer-enabled smartphone.

These days, more and more people are using hand-held computers such as mobile phones with advanced features like Internet, touch screens, built-in-cameras, accelerometers for user interface control, and so on. As the popularity of such devices increases and their cost decreases, opportunities for the novel healthcare applications arise. More importantly, mobile phones are part of people's daily life. People carry these devices with them nearly everywhere they go. Also, they mostly tend to keep their phones functioning and charged. Consequently, such devices can be employed to collect healthcare related information and thus deliver new tailored health-related services continuously over long periods of time during free-living conditions.

Though hundreds of applications appear every day which exploit the capabilities of these phones, their usage in creating smart, low-cost and timely healthcare oriented services is yet to be explored. One important area where mobile phones, especially accelerometer-enabled smartphones can be applied is in creating valid and reliable measures of physical activity and energy expenditure. As explained in chapter 1, automatic detection of physical activity would enable new types of health assessment and intervention tools that help people maintain their energy balance and stay physically fit and healthy. For example, mobile phones could be used to run algorithms that automatically recognize physical activities and estimate energy expenditure from body worn accelerometers and display this information as behavioral feedback in real-time.

Another powerful extension of mobile technology is to use it to deliver just-in-time interventions at the point of decision, for example, to encourage a positive behavior change. In this scenario, accelerometer-based mobile phones can be used to detect activities of interest (e.g. walking slowly) and encourage increases in intensity levels (e.g. brisk walking over walking slowly). Obviously, for these types of applications to be possible, activity recognition algorithms running on mobile phones have to be capable of recognizing the intensity of physical activity. A new area of research where such accelerometer-enabled mobile phones might also be applied is non-exercise activity thermogenesis. Recent results suggest that small changes to daily routine such as walking upstairs vs. riding the elevator, sitting fidgeting feet vs. sitting and brisk walking vs. walking can accumulate over the course of a day to meaningful amounts.

Therefore, the R&D labs at major cell phone/OS vendors plan to turn accelerometer-enabled future smartphones into really clever handsets capable of understanding what people are doing at any moment of time, anticipating what they would do next, and providing services automatically and accordingly. Recently, [85] described systems and techniques for automatically activating applications on a mobile device based on a comparison of current real-time acceleration data measured by the mobile device and acceleration profiles that are stored in the mobile device. Each stored acceleration profile can be associated with an activity that the user may engage in while using the corresponding mobile device application.

A user profile is a collection of personal data, such as favorite applications, associated to a specific user. Today's smartphones are capable of storing a large number of such profiles, each of which can be further associated with a specific activity and the corresponding application that the user prefers to engage in while performing that particular activity. The presence of accelerometer in these phones makes them capable of recognizing these activities in many cases. Therefore, the aim of activity-aware smartphones is to recognize these activities by means of built-in accelerometers, match it with stored user-activity profiles and then provide the services and applications associated with the target activity to the user automatically and accordingly.

Figure 7.1 illustrates an example in order to provide a better understanding of user-activity profiles and accelerometer-enabled activity-aware smartphones. It is 5 A.M Monday morning. I wake up and get ready for jogging. Before leaving my house, I put my smartphone and the headset in my trouser's pocket. Once inside the park, I activate the music player on my smartphone to play an already complied list of my favorite jogging-songs. I enjoy jogging listening to the jogging-

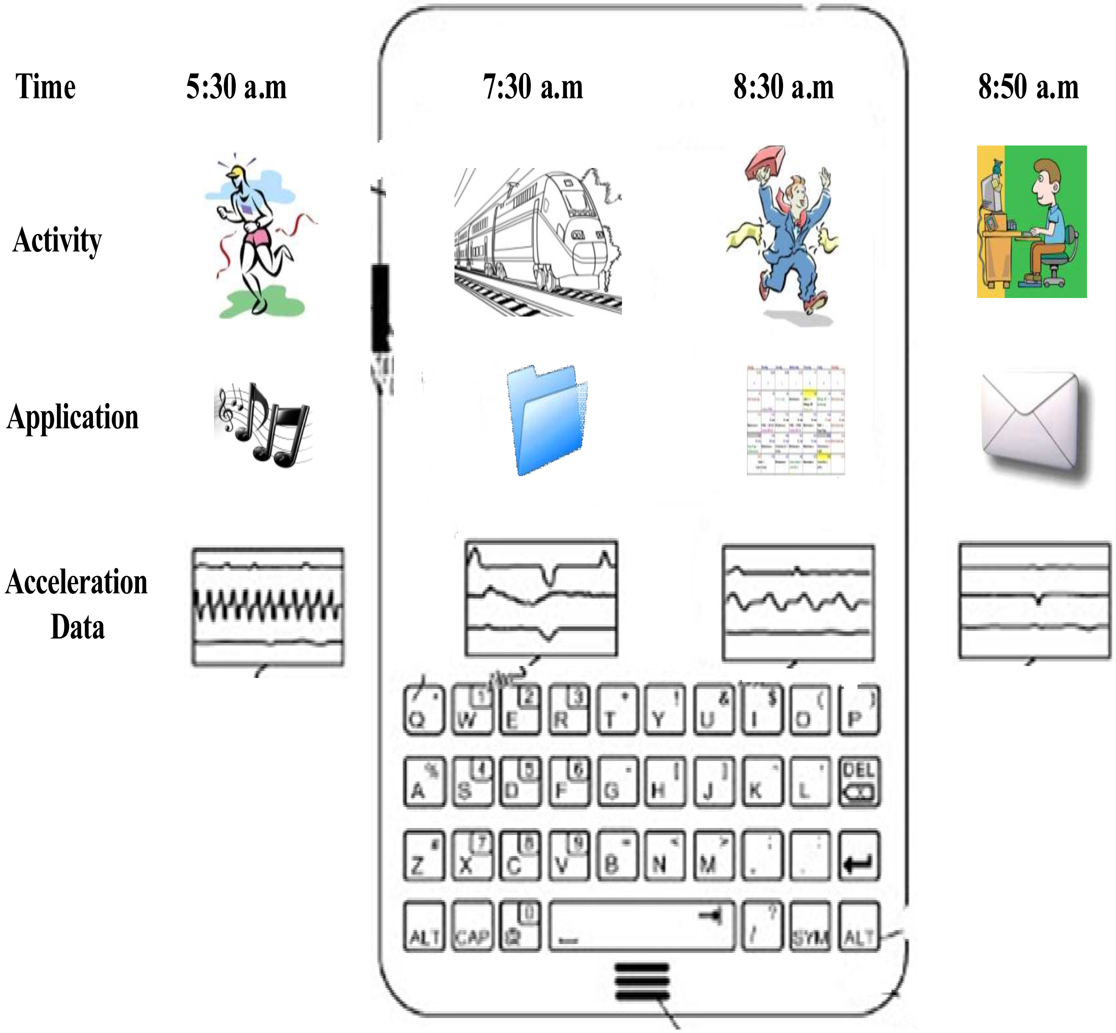

Figure 7.1: An example of an activity-aware smartphone: Recognizing activities by means of acceleration-signals and activating applications automatically

songs for about 30 minutes and then walk for another 15 minutes. Since I prefer listening to a different set of songs while walking, I once again reach the music player in my smartphone and choose the walking-songs.

It is 7:30 A.M now. I am in a train on my way to work. During the commute, I once again reach my phone and log on to my favorite newspaper's website to check the latest news. It is 8:30 A.M and I am walking towards my office-building. During the walk, I check my day's schedule on my phone's schedule application. Finally, it is 8:50 A.M and I am sitting in my desk. I once again reach my phone and go online to check my emails.

In the above example, I used my phone at different times, every day, for different purposes while performing different activities. Let us consider that the phone that I carry is an activity-aware smart phone. During the first week, the training period, I store my activity profiles on the phone. I perform and label activities along with the activated application information, which I want my phone to store for automatic activation in future. The phone collects the acceleration data on labeled activities in order to train itself for recognizing these activities automatically in future. After the training period, the phone now has information on what applications do I activate at different times while I perform a specific activity.

Its 5:40 A.M on the first Monday after the training-week. I just started jogging. The phone recognizes the running activity and activates the music player to play the jogging-songs automatically. After 30 minutes I stop jogging and start walking. The phone recognizes the change in activity and switches to walking-songs automatically. At 7:30 A.M, the activity recognition module reactivates to see if I am in the train or not. Upon successful recognition, it logs on to the newspaper website and displays the latest news. The other two applications, scheduler and email, are also activated automatically by recognizing the walking and sitting activities using built-in accelerometer.

7.1 Study Goal

The owners of smartphones are more likely to carry their handsets freely in their pockets, hands or even bags rather than attaching them firmly to a specific body part. The acceleration data thus could vary significantly for the same activity, leading to poor recognition results. Translating the idea of activity-aware smartphones into an actual product thus requires an activity recognition method that can function independent of phone's position along subject's body and is capable of providing high recognition results even in the absence of adequate amount of training data from different positions. For such a recognition system extracting discriminating features, which maximize the between-class variance and minimize the within-class variance, is crucial.

The sensor's position and attachment free activity recognition scheme, presented in the previous chapter, exhibited a high accuracy in distinguishing seven physical activities. However, its hierarchical structure and the use of spectral entropy as a feature makes it infeasible for smartphone considering their limited processing and memory resources.

Therefore, the aim of this study was to implement a light-weight system which uses the proposed augmented feature model (time domain features only, spectral entropy was not used anymore) and employs only one neural network for the classification task. The goal was to find a method that can resolve the high within-class variance in the feature space, that results due to freely placing the phone in different pockets, before feeding the features to the classifier. In this study, the SP dataset (section 3.4.5) was used for analysis and evaluation.

7.2 System Design

Figure presents the overall design of the system. After computing the augmented feature model from the activity acceleration signals, some method would be applied to decrease the high within-class variance. Several techniques exist in the literature for this purpose. Principal Component Analysis (PCA), also known as eigenface method, has been widely adopted for solving such problems. However, it is worth noticing that the features extracted by PCA are actually global features for all face classes, thus they are not necessarily representative for discriminating one face class from others [86]. LDA (section 5.3) seeks to find a linear transformation by maximizing the between-class variance and minimizing the within-class variance, proved to be a more suitable technique for classification [86]. Although LDA can provide a significant discriminating improvement to the task of recognition, it is still a linear technique in nature. When severe non-linearity is involved, this method is intrinsically poor. Kernel Discriminant Analysis (KDA), a kernel based technique, has been developed to compute the non-linear discriminating basis vectors which has shown good performance in cases where LDA failed [86]. In this study, both LDA and KDA were used to solve the problem of high within class variance and there results were compared. A detailed description of LDA is provided in section 5.3 whereas details of KDA algorithm are given below.

Kernel Discriminant Analysis **(KDA):** KDA is a non-linear discriminating approach based on kernel techniques to find non-linear discriminating features. Suppose we have a set of m augmented feature vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m \in \mathrm{R}^{3p+1}$ belonging to C activity classes where p is the AR-model order. Let

$$\mathbf{x}_i = [a_{x1}, a_{x2}, \cdots, a_{xp}, a_{y1}, a_{y2}, \cdots, a_{yp}, a_{z1}, a_{z2}, \cdots, a_{zp}, s]^T$$

where a_{xi}, a_{yi}, and a_{zi} are the AR coefficients for three axes and s is the SMA. We considered the problem in a feature space F induced by some nonlinear mapping $\varphi : \mathrm{R}^{3p+1} \to F$. Our choice of φ was the radial basis function. For a properly chosen φ an inner product $\langle , \rangle$ can be defined in F which makes for so called reproducing the kernel Hilbert space. More specifically, $\left\langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}_j) \right\rangle = \mathrm{K}(\mathbf{x}_i, \mathbf{x}_j)$ holds where $K(.,.)$ is a positive semi-definite kernel function. To find the linear discriminant in F, we needed to maximize

$$J(\omega) = \frac{\omega^T S_b^{\varphi} \omega}{\omega^T S_w^{\varphi} \omega} \tag{7.1}$$

where

$$S_b^{\varphi} = \sum_{k=1}^{C} m_k (\mu_{\varphi}^k - \mu_{\varphi})(\mu_{\varphi}^k - \mu_{\varphi})^T \tag{7.2}$$

$$S_w^{\varphi} = \sum_{k=1}^{C} \left(\sum_{i=1}^{m_k} \left(\varphi\left(x_i^k\right) - \mu_{\varphi}^k \right) \left(\varphi\left(x_i^k\right) - \mu_{\varphi}^k \right)^T \right) \tag{7.3}$$

are the between-class and within-class scatter matrices respectively in F and ω is the KDA basis vector. μ_{φ}^k and μ_{φ} are the mean of the k-th class and the global mean respectively. m_k is the number of samples in the k-th class. The solution to equation (3) is a linear combination of $\varphi(\mathbf{x}_i)$ [4] with coefficients α_i such that

$$\omega = \sum_{i=1}^{m} \alpha_i \varphi(\mathbf{x}_i) \tag{7.4}$$

Let $\alpha = [\alpha_1, \cdots, \alpha_m]^T$ it can be proved [4] that equation (3) is equivalent to

$$J(\alpha) = \frac{\alpha^T \mathbf{KWK}\alpha}{\alpha^T \mathbf{KK}\alpha} \tag{7.5}$$

and the optimal αs are given by the eigen vectors with respect to the maximum eigen values of

$$\mathbf{KWK}\alpha = \lambda \mathbf{KK}\alpha \tag{7.6}$$

where $\mathbf{K}$ is the kernel matrix ($\mathbf{K}_{ij} = \mathrm{K}(\mathbf{x}_i, \mathbf{x}_j)$) and $\mathbf{W}$ is defined as

$$\mathbf{W}_{ij} = \begin{cases} 1/m_k, & \text{if } \mathbf{x}_i \text{ and } \mathbf{x}_j \text{ belong to } k\text{−th class} \\ 0, & \text{otherwise} \end{cases} \tag{7.7}$$

For a new pattern $\mathbf{x}$ its projection onto a KDA basis vector ω in F is calculated as

$$(\omega, \varphi(\mathbf{x})) = \alpha^T \mathrm{K}(:, \mathbf{x}) \tag{7.8}$$

where

$$\mathrm{K}(:, \mathbf{x}) = [\mathrm{K}(\mathbf{x}_1, \mathbf{x}), \cdots, \mathrm{K}(\mathbf{x}_m, \mathbf{x})]^T \tag{7.9}$$

More details are available in [87].

7.3 Experimental Results

The performance of the proposed recognition system was validated in the following three studies:

Recognition using Original Features: In this study, the augmented feature vectors i.e., AR-

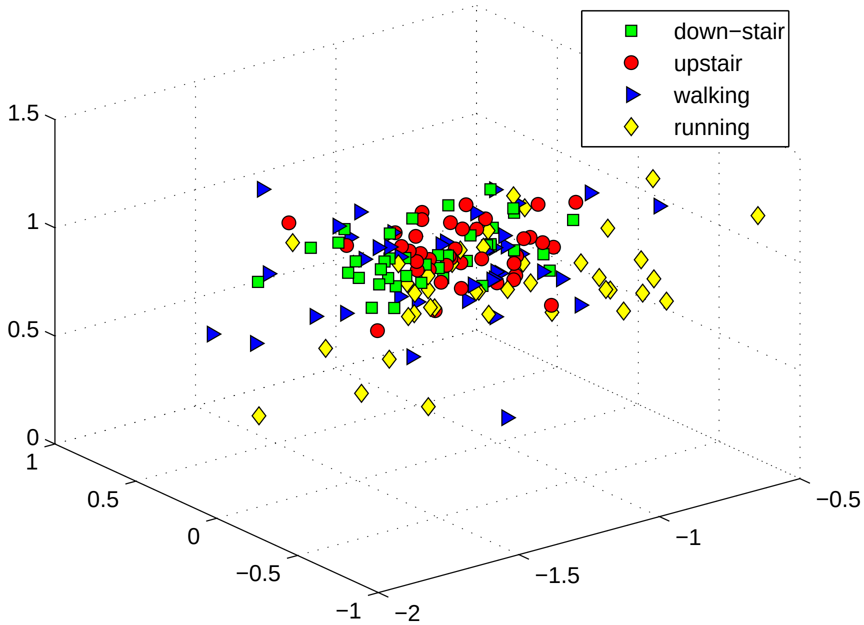

Figure 7.2: Feature plot for four activities before LDA and KDA showing high with-in and low between-class variances.

coefficients and SMA, were calculated from the acceleration data and used to train a neural network using BR (chapter 5) algorithm. The network started giving its best performance when the number of hidden layers was reached 24. No significant increase in the accuracy was seen after this number. During testing, each test activity was modeled in a similar fashion and the resulting augmented feature vector was fed to the network for final recognition. Freely placing the sensor at four different positions resulted in high within-class and low between-class variances in the input feature space as shown in Figure 7.2. Only four activities are shown for the sake of visualization. The average recognition rate was only 46%. Results are summarized in Table I.

Recognition using LDA Features: The purpose of this study was to evaluate the effectiveness of LDA in minimizing the within-class and maximizing the between-class variances. By taking LDA on the original features, one can improve the feature set as shown in Figure 7.3. However, being a linear technique in nature, it was not effective enough and the average recognition rate was 60%. Results are summarized in Table I.

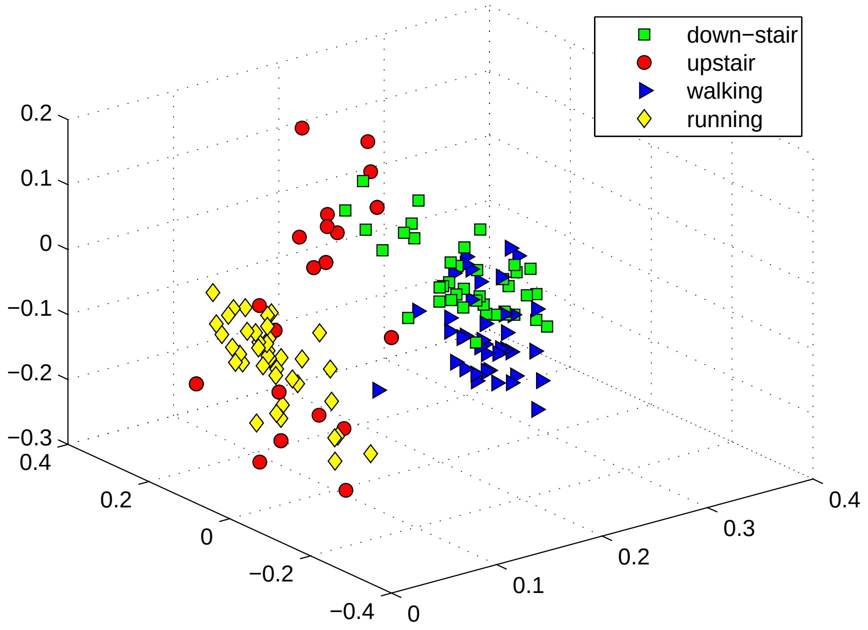

Figure 7.3: Feature plot for four activities after LDA.

Table 7.1: Average recognition results(%) for three studies

Activity	Original Features	LDA Features	KDA Features
Resting	61	74	99
Walking	41	52	95
Walk upstairs	41	56	95
Walk down-stairs	37	49	92
Running	50	69	99
Total	46	60	96

Recognition using KDA Features: In this study, we applied KDA to the same feature set used in the previous study. The distribution of KDA patterns for four classes is shown in Figure 7.4. Compare to that of LDA patterns, the improvement on class separability is significant. The average recognition rate for five activity classes was 96%, in this case. Results are summarized in Table I.

7.4 Conclusion

All existing accelerometer-based physical activity recognition systems and the proposed position and attachment free recognition framework (though it allows users to carry the sensor in any pocket) require users to carry an extra device with them all the time. The data collected by these devices is either sent to a computer in real-time or stored on portable devices carried by users and later read off-line. This is acceptable during a short-term monitoring. However, carrying these extra devices could be considered as a burden when long-term monitoring is the goal.

Todays mobile phones, called smartphones, come equipped with built-in accelerometers and better computational power. Moreover, people don't consider mobile phones as a burden and are used to carrying them all day long in pockets or handbags. These characteristics make these devices an ideal mean for recognizing physical activities for longer durations. Therefore, a prototype of the position and attachment free activity recognition scheme was implemented for smartphones. Considering their low computation power compared to traditional personal computers, the hierarchical scheme was altered to use only a single ANN for the recognition task, without calculating any frequency-domain features. The high within class variance which results due to carrying the phone in different pockets is reduced by employing KDA. The technique was validated using the activity data collected from five body positions using a smartphone. Thus the proposed system increases the applicability of activity classification systems. By using an accelerometer enabled smartphone, which could be placed in any pocket without firm attachment to a specific body part, activities could be monitored throughout a longer period of time.

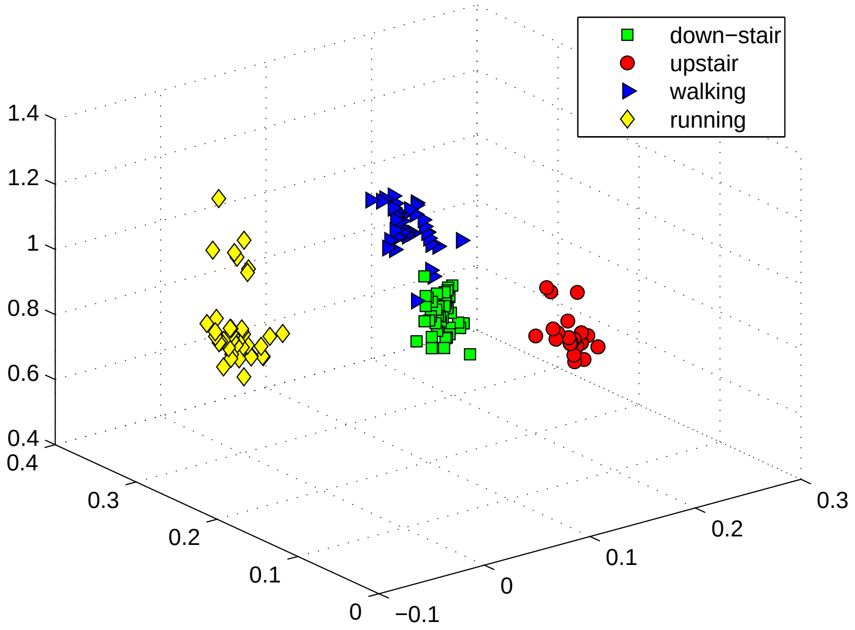

Figure 7.4: Feature plot for four activities after KDA.

Chapter 8

Application in Healthcare

8.1 Overview

I pose a question for us to ponder upon. What is our most valuable possession? A thing which if we possess makes us feel as if we have everything. Some of us would say it is their family, others would describe it as their wealth, profession, or even religion. For me, it is my health. When we have health, we have everything. But I also wonder that what does it take to make a person healthy? Perhaps being healthy means being free from disease and infirmity and moreover, includes a state of complete mental and social well-being. I would describe myself healthy if I look and feel great and have abundance of energy, free of any sickness.

Maintaining a healthy state and preventing sickness requires a healthcare infrastructure which includes a healthcare system. The motivation behind such a system is to prevent, treat and manage sickness and preserve physical and mental well-being of a person through the services offered by medical, nursing and allied health professions [88]. However, traditional healthcare delivery system failed in providing consistent, timely and high quality medical care to all people [88]. Such systems are established to help people receive the medical care that is tailored to meet their needs and is based on the best scientific knowledge, yet evidence suggests that this frequently is not the case. In fact, between the healthcare systems we have today and the healthcare systems we could have lies a huge gap [88].

There are many problems with today's healthcare infrastructure which contribute to this huge gap. The biggest and the most important one of which is the approach it takes. It focuses almost entirely on treating diseases and health problems and very little effort is spent on preventing them [89]. This leads to many problems including the high cost of treatment which is massively larger than the cost of prevention [89]. Numerous programs have shown that spending just an hour on preventive care with patients would cut down the annual medical cost significantly [89].

Another problem that can be attributed to reactive healthcare infrastructure is the use of antibiotics for treating diseases [89]. Once sick, people are bound to have these antibiotics and thus all good bacteria are also being wiped out of their bodies. This may seem unimportant, but when another virus or bacteria enters the human body, it will be much less able to fight off the intruder and build up immunity to it. Once the good bacteria are gone from the human body, it will take a great deal of time to regain the immune function.

Moreover, there exist certain factors, more commonly known as lifestyle-diseases, which can further lead to certain chronic diseases such as diabetes, stroke, high blood cholesterol, hypertension and cardiac failure [90]. One such lifestyle-disease, which has seen rapid increase over the past decades, is obesity. In fact, obesity is now regarded as a global epidemic that may dramatically impact health, especially in the industrialized world [91]. The prevalence of obesity from 1960 to 1994 in the U.S alone increased approximately 50% from 13.4% to 22.3%. Nowadays, about 65% of adult Americans aged 20 years or more are considered overweight and about 30% are considered obese. Furthermore, 16% of children and teenagers aged between 6 and 9 years are considered overweight and the numbers are increasing [90]. If obesity continues to increase at such a rate and no action is taken to halt its growth, the majority of the adult population could be overweight within in few generations [92].

Obesity can be termed as a complex condition which results from the interaction of many factors, including genetic makeup, neuroendocrine disorders, emotions and even secondary effects from medical treatments [90]. However, the recent rapid increase in its widespread, throughout the world, is generally believed to result from a caloric imbalance [92]. Nowadays, most people have high caloric intake due to easy access to foods and beverages with high caloric content and extremely low levels of physical activity to relative to that caloric intake [93]. Physical activities such as sports and outdoor activities, which used to be a part of average daily life, are now being increasingly replaced by sedentary behaviors such as television viewing, videogame playing and internet surfing [94]. In fact, a study conducted in 2003 revealed that Americans aged 13 years and older spend on average eight hours a day sitting and four hours a day watching TV, playing video games or surfing the web [95].

Such a lifestyle-disease apparently poses no immediate threat to people's health or life but is in fact a time-bomb in itself, ready to explode in years to come. It is this very characteristic of these diseases which leaves them unnoticed or unattended due to the reactive nature of the current healthcare system [89].

Moreover, the world is experiencing a so-called grey population. In other words, the ratio of the number of persons aged between 16 and 65 to those aged 65 or over, also called the care-ratio, is in decline. According to a recent study [96], the number of Americans aged 65 years or older (the elderly group) in the year 2008 was about 38.9 million, which is about 12.8% of the U.S population. In terms of gender, there were about 22.4 million elderly women and 16.5 million elderly men. The study suggested that the percentage of the elderly group has tripled, i.e., from 4.1% in 1900 to 12.8% in 2008. Furthermore, the older population itself is getting increasingly older. According to the statistics [96], the number of people aged between 65-74 years was 20.1 million in 2008 which was over 9 times larger than in 1900. In contrast, the number of people aged between 75-84 years was 13 million which was 17 times larger and the number of Americans aged 85 or older was 5.7 million which was 47 times larger than in 1900.

Futhermore, the statistics also suggest that persons reaching age of 65 in 2007 had an average life expectancy of additional 18.6 years [96]. In other words, a child born in 2007 was expected to live 77.9 years, which is about 30 years more than a child born in 1900. This increase can be attributed to the reduced death rates for children and young adults. The number of people who celebrated their 65th birthday in 2008 was about 2.7 million and about 1.8 million persons aged 65 or older died. Thus, the final figures revealed an annual increase of 927,305 in the number of persons aged 65 years or older [96]. Such a trend suggests that there will be less people to take care of the elderly in the coming years. This problem is compounded further by the fact that the many of the older people are living on a low income, suffering a disability and living either by themselves or prefer to live at home, being cared by their friends and family rather than being hospitalized.

Faced with such circumstances the health care delivery system, therefore, needs reinvention to meet the challenges at the hand [88]. In order to combat the increasing prevalence of the lifestyle diseases and the grey-population, the healthcare system must focus on not only treating people but also advising and guiding them about how to deal with and prevent chronic medical conditions [89]. It is generally expected that to create an efficient, high-quality but low-cost health care delivery system the use of information and communication technologies will be required [97]. One such technology is telemedicine. It involves the use of communication methods and information technology in order to deliver efficient, timely, cost-effective and high-quality medical care to people, especially the elderly [88].

An application of telemedicine is telemonitoring [13]. It involves remotely monitoring the patients who are living at their homes, with their own community away from the health care service providers. It works on the idea of bringing medical care to our homes which are the best and the most natural place to implement modern telecommunications technology for delivering healthcare to all people [13]. Telemedicine and telemonitoring offer huge reduction in healthcare costs by providing nursing services to the home [98]. Several companies are using this idea to provide home healthcare services at a very low-price then an on-site visit [98].

A study conducted in [99], investigated the use of telemonitoring technology in the home care settings. Their preliminary findings indicated that the technology is dependable and that average telehealth video visits are cost-effective and are about 60% shorter (18 minutes vs. 45 minutes) than on-site visits, with no decrease in patient satisfaction. In the past, positive effects have been reported on diabetes, asthma, and hypertension patients when treated by means of telemonitoring [99]. These positive results could mainly be associated with the fact that telemonitoring, by means of for more frequent follow-up of patients, may provide earlier detection of warning signs that a patient's state of health is deteriorating.

Let us use an example to get a better understanding of how cost-effective and time-efficient telemonitoring technology can be. It is 9:00 A.M Monday and a nurse is preparing to see her first patient at a home healthcare center in a hospital. However, her patient lives 50 miles away. Rather than drive there, the nurse steps into the center's video room as it is time to become a video nurse. Through a video interface, the nurse asks her patient and the patient's family a series of routine questions, which are indeed the same questions the nurse would have asked if it were a face-to-face visit. At the same time, the nurse accesses the patient's life log, medical readings such as heart rate captured by the devices installed at the patient's home and transmitted to the center via internet on continuous bases. Using the patient's life log and their answers, the nurse concludes her visit and leaves a voice message for the doctor summarizing the health-status of her patient. Such a technology not only saves extra-cost and time but also gives patients and their family a sense of confidence as they know that help is only a phone call away and they can see the nurse.

In general, the telemonitoring based health care delivery systems require patients to have monitoring devices at their homes [13]. The result of these devices is transmitted via telephone or internet to the health care service provider, as shown in Figure 8.1. This information holds clinically useful trends that can allow physicians to make informed decisions, to monitor deterioration in chronic conditions, or to assess the response of a patient to a particular treatment [13]. Telemonitoring, therefore, has the potential to provide safe, effective, patient-centered, timely, efficient, and location independent monitoring; thus, fulfilling the six key aims for improvement of healthcare as proposed by the Institute of Medicine, Washington, DC [88].

Some of the more common things that telemoitoring devices keep track of include heart rate, blood pressure, blood glucose and mobility. Mobility refers to the amount of time we spend in the dynamic activities, such as walking or running, as well as the static activities, such as sitting, standing, and lying [13].

It is mostly believed that being engaged in vigorous physical activities such as high-intensity exercise programs is associated with reductions in physical decline. However, such perceptions exclude the elderly patients, especially the oldest ones who suffer from arthritis, cardiovascular, or neurodegenerative diseases which result in limitation of mobility and physical activity of the affected persons. Therefore, such patients are unable to take part even in low-intensity exercise programs.

It is in such scenarios where assessment of daily mobility levels, time spent in performing simple daily physical activities such as walking, standing straight, standing up from a chair etc., play a vital role in determining physical independence and functional ability. Increased mobility, especially in case of elderly patients, improves stamina and muscle strength. It promotes their psychological well-being and quality of life by increasing their ability to perform a greater range of activities of daily living [100]. Thus, objective mobility data can be used to monitor health to assess the relevance of certain medical treatments and to determine the quality of life of an elderly

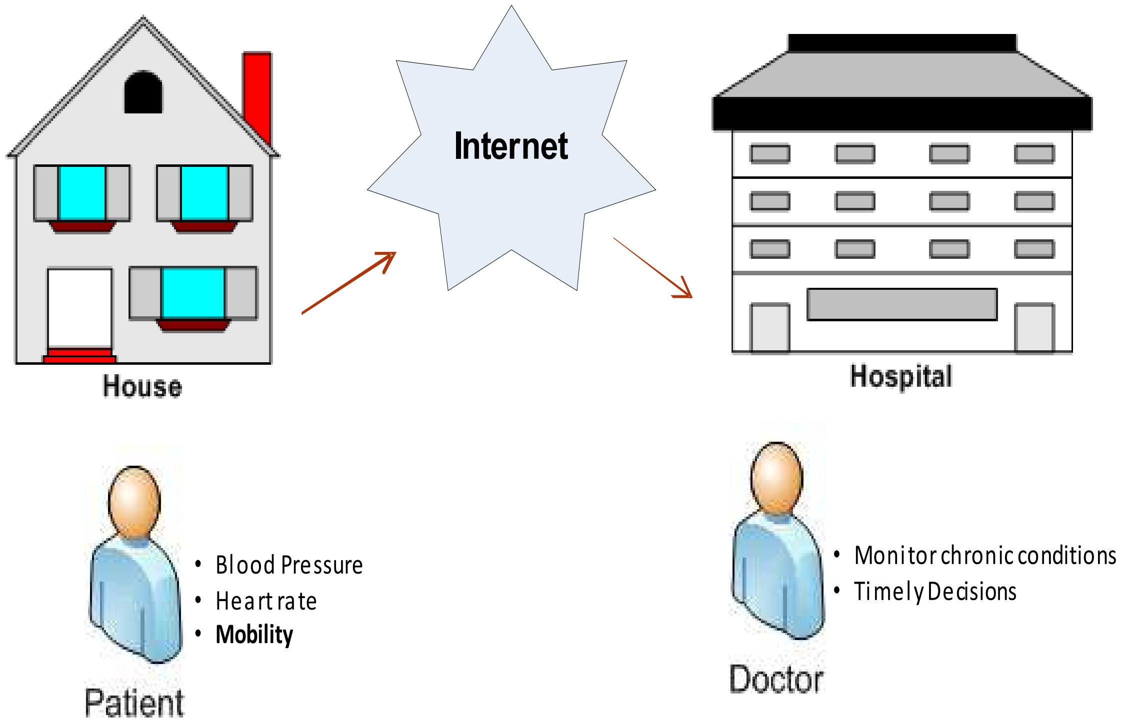

Figure 8.1: Telemonitoring: Patients are monitored using monitoring devices at their homes and the result is transmitted via Internet to the health care service provider

patient.

Moreover, the thermodynamic expression of the principle of the conservation of energy states that when energy is added to a system, it is either stored or used to perform work. When we apply the same physical law to living things, such as animals and humans, we can easily reach the conclusion that when total energy intake, by means of food, is greater than energy expenditure, excess energy would be stored in the body as fat. In other words, such state of disequilibrium between the amount of energy taken into the body and the amount of energy consumed or expended results in obesity, which is a preventable cause many chronic diseases including type II diabetes, hypertension, stroke, degenerative arthritis, sleep apnea, and cancer [90].

There are two ways humans can spend the calories. One is to perform vigorous exercises such as weight training and rowing. The other is through all the activities of daily living, also known as Non-Exercise Activity Thermogenesis (NEAT). NEAT has got a lot of attention in the research community over the past decade. The theory behind NEAT is based on the fact that minor behavior modifications to a person's daily routine, such as sitting fidgeting legs vs. sitting, standing vs. sitting, brisk walking vs. walking, and using stairs vs. elevator, can sum up over the course of day and boost overall energy expenditure and thus provide a protective effect against lifestyle diseases like obesity [101, 102]. The reason for this can be attributed to the fact that most of the energy expended everyday comes from non-exercise activity. Therefore, if a handheld device like mobile phone could recognize non-exercise activities and the energy expenditure associated with them, it could then suggest people minor changes in their daily routine that might impact their daily energy expenditure positively.

8.2 Personal Life Log

Personal life log (PLL) is a set of data containing an individuals daily activities collected in one or multiple media forms. These life log data is able to help provide personalized support for various real world applications such as health monitoring, activity level checking, etc.

This section explains the real-time personal life log (PLL) that was developed in this work to evaluate the real-time performance of the proposed system. It employed the proposed state-activity based classification scheme to recognize physical activities of a person. From the recognized activities, medically useful information, such as step counts from walking, going up-stairs, and going down-stairs, walking distance and duration, energy expenditure, etc., are extracted. Upon the computation of activity classification and exercise information, all information is stored in a database as a personal life log for future reference, as shown in Figure 8.2 .

8.2.1 Generation of Exercise Information

Once each activity is classified, exercise information is computed based on the activity recognition result. In this study, exercise information includes stride length, step count, walking distance, walking speed, and energy expenditure. A rough estimate of a stride length of each user is obtained based on a subjects height. Table 8.1 is used for assistance.

Table 8.1: The Ratio between Stride Length and Height in General Walking Phase of 10-60 Aged Men and Women

Age	Gender	The Ratio between Stride Length and Height(%)
10-30 Age	Male	42.36
Group	Female	43.56
40-60 Age	Male	41.17
Group	Female	40.55

Step counting is performed based on a zerocrossing detector which is activated only for walking, going up-stairs, or going down-stairs. To reduce the influence of noise, a threshold of three times the standard deviation of the static activities was used. The number of steps get computed by equation 8.1, and an example of detected zero crossings is shown in Figure 8.3.

$$\mathrm{Number\,of\,Steps} = \mathrm{Number\,of\,Zero\,Crossings}/\,2 \tag{8.1}$$

The total walking distance and the walking speed are computed by equation 8.2 and 8.2, respectively.

$$\mathrm{Distance} = \mathrm{StrideLength} \times \mathrm{StepCounts} \tag{8.2}$$

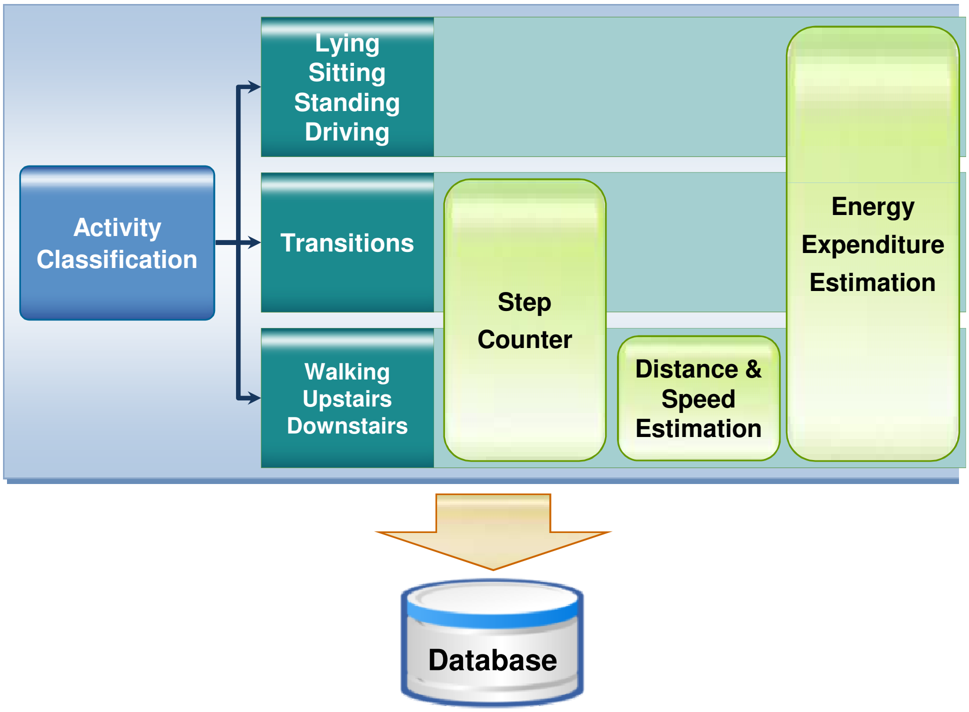

Figure 8.2: An overview of the architecture of the implemented Personal Life Log (PLL)

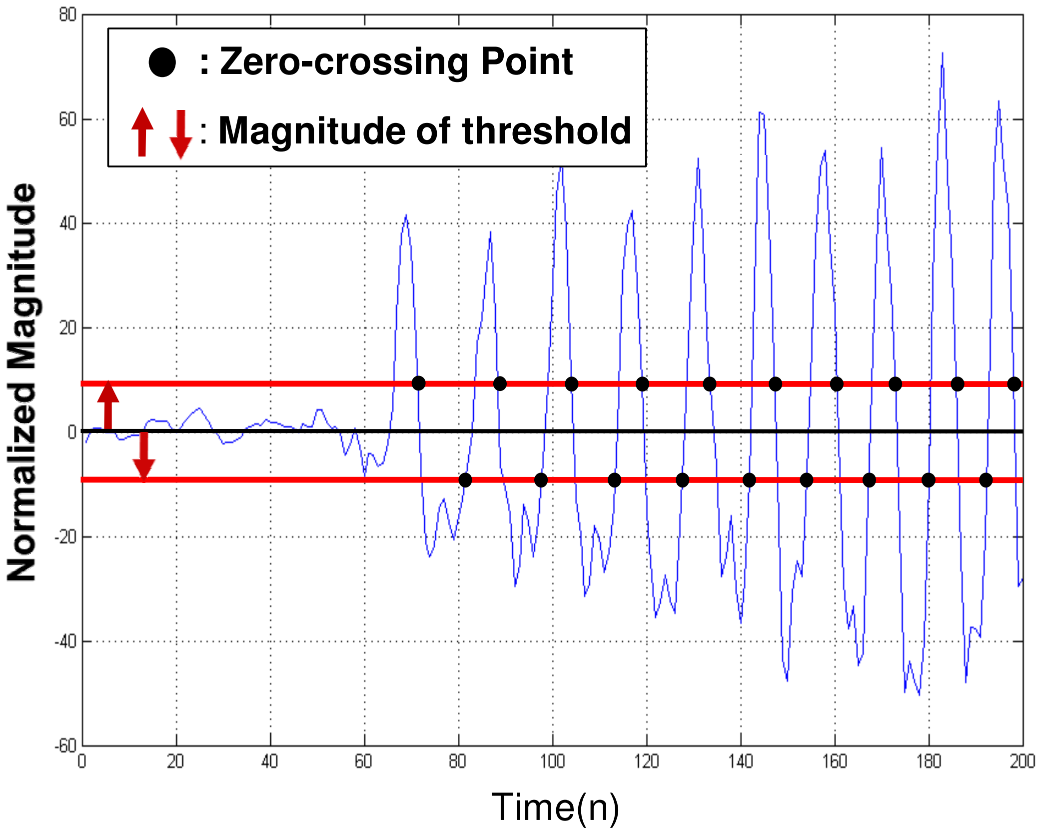

Figure 8.3: AN example of zero-crossings to get the step counts

$$\text{Speed} = \text{Distance}/\text{Duration}(\text{Walking}) \tag{8.3}$$

The standard reference for the measurement of physical activity is the metabolic energy expended due to that physical activity. Unfortunately, estimating an accurate measurement of energy expenditure is challenging. At present there is no technology that allows people to measure these variables comfortably, accurately and continuously over the course of a day and obtain real-time feedback. Therefore, true total energy expenditure is very difficult to measure, and nearly all techniques use approximations.

In order to calculate the metabolic energy expanded, the Metabolic Equivalents (METS) values [8] were used which are most frequently used for the calorie count, to compute energy consumed during each activity. MET is defined as the ratio of a per- sons working metabolic rate relative to the resting metabolic rate. METS values correlate with oxygen requirements. Starting with 1, which is the least amount of activity (such as resting), the values increase with the amount of activity. For example, running at 9.7 km/h has a METS value of 10. Standard tables exist that provide METS values for a wide range of exercises and activities. A persons calorie consumption can be easily calculated using this METS values given by equation:

$$\text{EnergyExpenditure}(\text{kcal}) = 1.05 \times \text{METS} \times \text{Duration}(\text{hour}) \times \text{Weight}(\text{kg}) \tag{8.4}$$

The METS values for six activities (lying = 1.0, sitting = 2.0, standing = 2.3, up-stairs = 8.0, down-stairs = 3.0, driving = 2.0) are obtained from [2]. Since the METS value for walking can be very different depending on speed, we computed the METS value for walking according to the following equation [6]:

$$\text{METS}(\text{walking}) = 0.0272 \times \text{Speed}(\text{m}/\text{min}) + 1.2 \tag{8.5}$$

8.2.2 PLL Database

A PLL database is created using Microsoft Office Access, which is composed of four tables such as Activity-Definition (AD), State-Definition (SD), User-List (UL), and Result- Recorder (RR). The AD and SD tables predefine the human states and the human activities respectively, which the PLL system is able to recognize. The UL contains a list of users and user-associated physical information (e.g., height, weight, etc.), and the RR is used to record the recognized activities and the estimated exercise information.

8.2.3 Experimental Validation

A new data collection study was conducted in order to validate the real-time PLL. Ten subjects participated in the training. These subjects were given a brief introduction of each activity, however, they were not provided with any protocol. Each subject was allowed to perform activities in random, with varying speed and postures. To train the hierarchical ANNs (Section 5.5), activity acceleration-signals were collected for several hours per subject, performing activities randomly with varying speeds and postures. After training the ANNs with the data from ten subjects, the system was tested with two new subjects, which did not take part in data collection. During the experiment, the test subjects activities were recorded by a camera for the evaluation of accuracy rate.

The activity classification performances were evaluated after the experiment. As summarized in Table II showing the average classification accuracy of activities for the two subjects, the overall accuracy rate was 84.8%. An example of extracted exercise information is shown in Figure 8.4: all the exercise information was recorded and delivered to the database in real-time. In Figure 8.5, the error rates for the step counter are shown which are 6.5%, 13.3%, and 10.1% for walking, going up-stairs, and going down-stairs, respectively. The errors were mainly due to the misclassified activities. One should note that not all exercise information can be validated due to some dependcy on each individuals physical conditions.

```
Stride length: 75cm
The number of steps
  - Walking: 466
  - Upstairs: 212
  - Downstairs: 275
Walking Distance: 349.5m
Average Speed(walking): 1.34m/s
Energy Expenditure
  - Lying     :   2.21kcal
  - Sitting   :   5.51kcal
  - Standing  :   3.68kcal
  - Walking   :  23.72kcal
  - Upstairs  :  31.58kcal
  - Downstairs:  13.08kcal
  - Driving   :  20.83kcal
  - Total     : 100.61kcal
```

Figure 8.4: Exercise information extracted from acceleration signals of a subject

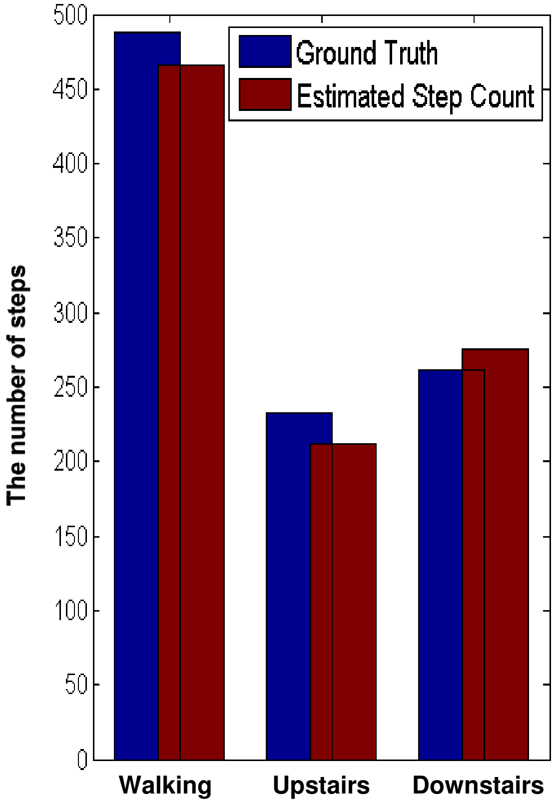

Figure 8.5: The result of step counting: The error rates for the step counter for walking, going upstairs and downstairs

References

[1] J. Yang, B. N. Schilit, and D. W. McDonald, “Activity recognition for the digital home,” *Computer*, vol. 41, pp. 102–104, 2008.

[2] E. M. Tapia, S. S. Intille, and K. Larson, “Activity recognition in the home using simple and ubiquitous sensors,” in *Pervasive Computing*. Springer Berlin / Heidelberg, 2004, pp. 158–175.

[3] T. van Kasteren, A. Noulas, G. Englebienne, and B. Kröse, “Accurate activity recognition in a home setting,” in *UbiComp ’08: Proceedings of the 10th international conference on Ubiquitous computing*. New York, NY, USA: ACM, 2008, pp. 1–9.

[4] B. Najafi, K. Aminian, A. Paraschiv-Ionescu, F. Loew, C. J. Bla, and P. Robert, “Ambulatory system for human motion analysis using a kinematic sensor: Monitoring of daily physical activity in the elderly,” *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 6, 2003.

[5] N. Noury, A. Dittmar, C. Corroy, R. Baghai, D. B. J. L. Weber, F. Klefstat, A. Blinovska, S. Vaysse, and B. Comet., “Vtamna smart clothe for ambulatory remote monitoring of physiological parameters and activity,” in *Proc. of 26th Annual IEEE International Conference on Engineering in Medicine and Biology Society*, 2004.

[6] S. Park and S. Jayaraman, “Enhancing the quality of life through wearable technology,” *Engineering in Medicine and Biology Magazine, IEEE*, vol. 22, pp. 41–48, 2003.

[7] H. Asada, P. Shaltis, A. Reisner, S. Rhee, and R. Hutchinson, “Mobile monitoring with wearable photoplethysmographic biosensors,” *Engineering in Medicine and Biology Magazine, IEEE*, vol. 22, no. 3, pp. 28 – 40, may-june 2003.

[8] A. Sarela, I. Korhonen, J. Lotjonen, M. Sola, and M. Myllymaki, "Ist vivago reg; - an intelligent social and remote wellness monitoring system for the elderly," april 2003, pp. 362 – 365.

[9] M. J. Mathie, A. C. F. Coster, N. H. Lovell, and B. G. Celler, "Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement," *Physiol. Meas*, vol. 25, p. R1R20, 2004.

[10] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, D. Howard, K. Meijer, and R. Crompton, "Activity identification using body-mounted sensors–a review of classification techniques," *Physiological Measurement*, vol. 30, pp. R1–R33, 2009.

[11] P. Peeters, "Design criteria for an automatic safety-alarm system for elderly," *Technology and Health Care*, vol. 8, pp. 81–91, 2000.

[12] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing*. Springer Berlin / Heidelberg, 2004, pp. 158–175.

[13] C. N. Scanaill, S. Carew, P. Barralon, N. Noury, D. Lyons, and G. M. Lyons, "A review of approaches to mobility telemonitoring of the elderly in their living environment," *Annals of Biomedical Engineering*, vol. 34, pp. 547–563, 2006.

[14] F. Foerster, M. Smeja, and J. Fahrenberg, "Detection of posture and motion by accelerometry: A validation study in ambulatory monitoring," *Comput. Hum. Beh*, vol. 15, pp. 571–583, 1999.

[15] P. H. Veltink, H. B. J. Bussmann, W. de Vries, W. L. J. Martens, , and R. C. van Lummel, "Detection of static and dynamic activities using uniaxial accelerometers," *IEEE Transactions on Rehabilitation Engineering*, vol. 4, pp. 375–85, 1996.

[16] J. Mantyjarvi, J. Himberg, and T. Seppanen, "Recognizing human motion with multiple acceleration sensors," in *Proc. of 2001 IEEE International Conference on Systems, Man, and Cybernetics*, vol. 2, 2001, pp. 747–752.

[17] J. B. J. Bussmann, W. L. Martens, J. H. M. Tulen, F. C. Schasfoort, H. J. van den Berg-Emons, and H. J. S. ., “Measuring daily behaviour using ambulatory accelerometry: the activity monitor,” *Behavior Research Methods, Instruments, and Computers*, vol. 33, p. 34956, 2001.

[18] K. Kiani, C. J. Snijders, and E. S. Gelsema, “Computerized analysis of daily life motor activity for ambulatory monitoring,” *Technol. HealthCare*, vol. 5, p. 30718, 1997.

[19] M. Ermes, J. Parkka, J. Mantyjarvi, and I. Korhonen, “Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, p. 2026, 2008.

[20] U. Maurer, A. Smailagic, D. Siewiorek, and M. Deisher, “Activity recognition and monitoring using multiple sensors on different body positions,” in *Proc. Of International Workshop on Wearable and Implantable Body Sensor Networks*, 2006, pp. 113–116.

[21] N. Kern, B. Schiele, H. Junker, P. Lukowicz, and G. Troster, “Wearable sensing to annotate meeting recordings,” *Personal and Ubiquitous Computing*, vol. 7, pp. 263–274, 2003.

[22] D. Minnen, T. Starner, J. Ward, P. Lukowicz, and G. Troester, “Recognizing and discovering human actions from on-body sensor data,” in *Proc. of IEEE International Conference on Multimedia and Expo (ICME)*, 2005, pp. 1545–1548.

[23] D. Giansanti, V. Macellari, and G. Maccioni, “New neural network classifier of fall-risk based on the mahalanobis distance and kinematic parameters assessed by a wearable device,” *Physiol. Meas*, vol. 29, pp. N11–N19, 2008.

[24] M. J. Mathie, A. C. F. Coster, N. H. Lovell, and B. G. Celler, “A pilot study of long term monitoring of human movements in the home using accelerometry,” *J. Telemed. Telecare*, vol. 10, pp. 144–151, 2004.

[25] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, “Implementation of a real-time human movement classifier using a tri-axial accelerometer for ambulatory monitoring,” *IEEE Transactions on Inf Technol Biomed*, vol. 10(1), pp. 156–67, 2006.

[26] S. H. Lee, H. D. Park, S. Hong, K. J. Lee, and Y. H. Kim, “A study on the activity classification using a tri-axial accelerometer,” in *Proc. of 25th Annual IEEE International Conference on Engineering in Medicine and Biology Society*, vol. 3, 2003.

[27] M. Sekine, T. Tamura, M. Akay, T. Fujimoto, T. Togawa, and Y. Fukui, “Discrimination of walking patterns using wavelet-based fractal analysis,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 10, pp. 188–196, 2002.

[28] M. Sung, C. Marci, and A. Pentland, “Wearable feedback systems for rehabilitation,” *J. Neuroeng. Rehabil*, vol. 2:17, 2005.

[29] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, “Activity recognition from accelerometer data,” in *Proc. 20th National Conf. on Artificial Intelligence*, 2005, p. 15416.

[30] F. Allen, E. Ambikairajah, N. Lovell, and B. Celler, “Classification of a known sequence of motions and postures from accelerometry data using adapted gaussian mixture models,” *Physiol. Meas*, vol. 27, p. 935951, 2006.

[31] M. Mathie, B. Celler, N. Lovell, and A. Coster, “Classification of basic daily movements using a triaxial accelerometer,” *Medical and Biological Engineering and Computing*, vol. 42, p. 679687, 2004.

[32] M. R. Narayanan, M. E. Scalzi, S. J. Redmond, S. R. Lord, B. G. Celler, and N. H. Lovell, “A wearable triaxial accelerometry system for longitudinal assessment of falls risk,” in *Proc. of 30th Annual IEEE International Conference on Engineering in Medicine and Biology Society*, 2008, pp. 2840–2843.

[33] M. Marschollek, K. Wolf, M. Gietzelt, G. Nemitz, H. M. Z. Schwabedissen, and R. Haux, “Assessing elderly persons’ fall risk using spectral analysis on accelerometric data - a clinical evaluation study,” in *Proc. of 30th Annual IEEE International Conference on Engineering in Medicine and Biology Society*, 2008, pp. 3682–3685.

[34] C. V. Bouten, K. T. Koekkoek, M. Verduin, R. Kodde, and J. D. Janssen, “A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity,” *IEEE Transactions on Biomedical Engineering*, vol. 44, p. 13647, 1997.

[35] A. M. Khan, Y. K. Lee, and T.-S. Kim, "Accelerometer signal-based human activity recognition using augmented autoregressive model coefficients and artificial neural nets," in *Proc. of 30th Annual IEEE International Conference on Engineering in Medicine and Biology Society*, 2008, pp. 5172–5175.

[36] A. M. Khan, P. T. H. Truc, Y. K. Lee, and T.-S. Kim, "A tri-axial accelerometer sensor-based human activity recognition via augmented signal features and hierarchical recognizer," in *Proc. of 5th International Conference on Ubiquitous Healthcare*, 2008.

[37] M. J. Mathie1, A. C. F. Coster, N. H. Lovell, and B. G. Celler, "Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement," *Physiol. Meas*, vol. 25, p. R1R20, 2004.

[38] M. Sun and J. O. Hill, "A method for measuring mechanical work and work efficiency during human activities," *J. Biomech*, vol. 26, p. 22941, 1993.

[39] M. A. Lafortune, "Three-dimensional acceleration of the tibia during walking and running," *J. Biomech*, vol. 24, p. 87786, 1991.

[40] K. Aminian, P. Robert, E. Buchser, B. Rutschmann, D. Hayoz, and M. Depairon, "Physical activity monitoring based on accelerometry: validation and comparison with video observation," *Medical and Biological Engineering and Computing*, vol. 37, pp. 304–308, 1999, 10.1007/BF02513304. [Online]. Available: http://dx.doi.org/10.1007/BF02513304

[41] P. Veltink, E. Engberink, B. Van Hilten, R. Dunnewold, and C. Jacobi, "Towards a new method for kinematic quantification of bradykinesia in patients with parkinsons disease using triaxial accelometry," 1995.

[42] J. Fahrenberg, F. Foerster, M. Smeja, and W. Muller, "Assessment of posture and motion by multichannel piezoresistive accelerometer recordings," *Psychophysiology*, vol. 34, p. 60712, 1997.

[43] F. Foerster and J. Fahrenberg, "Motion pattern and posture: correctly assessed by calibrated accelerometers," *Behav. Res. Methods Instrum. Comput.*, vol. 32, p. 4507, 2000.

[44] M. Uiterwaal, E. B. Glerum, H. J. Busser, and R. C. van Lummel, “Ambulatory monitoring of physical activity in working situations, a validation study,” *J. Med. Eng. Technol.*, vol. 22, p. 16872, 1998.

[45] M. Sekine, T. Tamura, T. Togawa, and Y. Fukui, “Classification of waist-acceleration signals in a continuous walking record,” *Med. Eng. Phys.*, vol. 22, p. 28591, 2000.

[46] K. Aminian, P. Robert, E. Jequier, and Y. Schutz, “Aincline, speed, and distance assessment during unconstrained walking,” *Med. Sci. Sports Exerc.*, vol. 27, p. 22634, 1995.

[47] S. Park and S. Jayaraman, “Enhancing the quality of life through wearable technology,” *IEEE Eng. Medl. Biol. Mag.*, vol. 22, p. 418, 2003.

[48] J. M. Potter, A. L. Evans, and G. Duncan, “Gait speed and activities of daily living function in geriatric patients,” *Arch. Phys. Med. Rehabil.*, vol. 76, p. 9979, 1995.

[49] H. Luukinen, K. Koski, P. Laippala, and S. L. Kivela, “Predictors for recurrent falls among the home-dwelling elderly,” *Scand. J. Prim. Health Care*, vol. 13, p. 2949, 1995.

[50] P. Terrier, Q. Ladetto, B. Merminod, and Y. Schutz, “High-precision satellite positioning system as a new tool to study the biomechanics of human locomotion,” *J. Biomech.*, vol. 33, p. 171722, 2000.

[51] R. Herren, A. Sparti, K. Aminian, and Y. Schutz, “The prediction of speed and incline in outdoor running in humans using accelerometry,” *Med. Sci. Sports Exerc.*, vol. 31, p. 10539, 1999.

[52] O. Perrin, P. Terrier, Q. Ladetto, B. Merminod, and Y. Schutz, “Improvement of walking speed prediction by accelerometry and altimetry, validated by satellite positioning,” *Med. Biol. Eng. Comput.*, vol. 38, p. 1648, 2000.

[53] B. J. Munro, J. R. Steele, G. M. Bashford, and N. RyanMand Britten, “Kinematic and kinetic analysis of the sit-to-stand transfer using an ejector chair: implications for elderly rheumatoid arthritic patients,” *A J. Biomech.*, vol. 31, p. 26371, 1998.

[54] B. S. Troy, D. E. Kenney, and E. E. Sabelman, “Sit-to-stand as an evaluation tool for balance,” *GSA 52nd Annual Scientific Meeting (San Francisco, CA, 1923 Nov. 1999)*.

[55] G. Williams, K. Doughty, K. Cameron, and D. A. Bradley, “A smart fall and activity monitor for telecare applications,” *Proc. 20th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, 1998.

[56] M. Makikawa and Iizumi, “Development of an ambulatory physical activity memory device and its application for the categorization of actions in daily life,” *Medinfo*, vol. 8, p. 74750, 1995.

[57] T. Huynh and B. Schiele, “Analyzing features for activity recognition,” in *Proc. Conf. Smart Objects and Ambient Intelligence: Innovative Context-Aware Services: Usages and Technologies (Grenoble*, 2005.

[58] M. N. Nyan, F. E. Tay, K. H. Seah, and Y. Y. Sitoh, “Classification of gait patterns in the time-frequency domain,” *J. Biomech*, vol. 39, p. 264756, 2006.

[59] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, “Activity classification using realistic data from wearable sensors,” *IEEE Trans. Inf. Technol. Biomed*, vol. 10, p. 11928, 2006.

[60] S. Pirttikangas, K. Fujinami, and T. Nakajima, “Feature selection and activity recognition from wearable sensors,” *Ubiquitous Computing Systems*, vol. 4239, pp. 516–27, 2006.

[61] T. Huynh and B. Schiele, “Towards less supervision in activity recognition from wearable sensors,” in *Proc. of the 10th IEEE International Symposium on Wearable Computers (Montreaux)*, 2006.

[62] S. Lee and K. Mase, “Activity and location recognition using wearable sensors,” *IEEE Perv. Comp.*, p. 2532, 2002.

[63] A. Salarian, H. Russmann, F. J. G. Vingerhoets, P. R. Burkhard, and K. Aminian, “Ambulatory monitoring of physical activities in patients with parkinsons disease,” *IEEE Trans. Biomed. Eng*, vol. 54, p. 22969, 2007.

[64] P. Boissy, S. Choquette, M. Hamel, and N. Noury, "User-based motion sensing and fuzzy logic for automated fall detection in older adults," *Telemed J. E. Health*, vol. 13, p. 68393, 2007.

[65] S. Theodoridis and K. Koutroumbas, "Pattern recognition," *3rd edn (San Diego: Academic)*, 2006.

[66] J. Lester, T. Choudhury, and G. Borriello, "A practical approach to recognizing physical activities pervasive computing," *(Lecture Notes in Computer Science vol 3968) (Berlin: Springer)*, p. 116, 2006.

[67] R. W. Selles, M. A. G. Formanoy, J. B. J. Bussmann, P. J. Janssens, and H. J. Stam, "Automated estimation of initial and terminal contact timing using accelerometers; development and validation in transtibial amputees and controls," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, p. 818, 2005.

[68] J. M. Jasiewicz, J. H. Allum, J. W. Middleton, A. Barriskill, P. Condie, B. Purcell, and R. C. Li, "Gait event detection using linear accelerometers or angular velocity transducers in able-bodied and spinal-cord injured individuals," *Gait Posture*, vol. 24, p. 5029, 2006.

[69] A. Mansfield and G. M. Lyons, "The use of accelerometry to detect heel contact events for use as a sensor in fes assisted walking," *Med. Eng. Phys.*, vol. 25, p. 87985, 2003.

[70] W. Zijlstra, "Assessment of spatio-temporal parameters during unconstrained walking," *Eur. J. Appl. Physiol.*, vol. 92, p. 3944, 2004.

[71] R. Takalo, H. Hytti, and H. Ihalainen, "Tutorial on Univariate Autoregressive Spectral Analysis," *Journal of Clinical Monitoring and Computing*, vol. 19, no. 6, pp. 401–410, December 2005.

[72] Akaike and Hirotugu, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19 (6), 1974.

[73] B. Everitt, *The Cambridge dictionary of statistics / B.S. Everitt*, 2nd ed. Cambridge University Press, Cambridge :, 2002.

[74] J. Hertz, R. G. Palmer, and A. S. Krogh, *Introduction to the Theory of Neural Computation*, 1st ed. Perseus Publishing, 1991.

[75] R. Begg and J. Kamruzzaman, “Neural networks for detection and classification of walking pattern changes due to ageing,” *Australasian Physical and amp; Engineering Science in Medicine*, vol. 29, pp. 188–195, 2006.

[76] M. T. Hagan, H. B. Demuth, and M. H. Beale, *I Neural Network Design*. PWS Publishing, 1996.

[77] M. F. Moller, “A scaled conjugate gradient algorithm for fast supervised learning,” *NEURAL NETWORKS*, vol. 6, no. 4, pp. 525–533, 1993.

[78] D. J. C. Mackay, “A practical Bayesian framework for backpropagation networks,” *Neural Computation*, vol. 4, pp. 448–472, 1992.

[79] J. Zheng, Z. Zhang, T. Wu, and Y. Zhang, “A wearable mobihealth care system supporting real-time diagnosis and alarm,” *Medical and Biological Engineering and Computing*, vol. 45, pp. 877–885, 2007, 10.1007/s11517-007-0221-y. [Online]. Available: http://dx.doi.org/10.1007/s11517-007-0221-y

[80] H.-Y. Lau, K.-Y. Tong, and H. Zhu, “Support vector machine for classification of walking conditions using miniature kinematic sensors,” *Medical and Biological Engineering and Computing*, vol. 46, pp. 563–573, 2008, 10.1007/s11517-008-0327-x. [Online]. Available: http://dx.doi.org/10.1007/s11517-008-0327-x

[81] J. Han, H. Jeon, W. Yi, B. Jeon, and K. Park, “Adaptive windowing for gait phase discrimination in parkinsonian gait using 3-axis acceleration signals,” *Medical and Biological Engineering and Computing*, vol. 47, pp. 1155–1164, 2009, 10.1007/s11517-009-0521-5. [Online]. Available: http://dx.doi.org/10.1007/s11517-009-0521-5

[82] B. Coley, B. Jolles, A. Farron, and K. Aminian, “Detection of the movement of the humerus during daily activity,” *Medical and Biological Engineering and Computing*, vol. 47, pp. 467–474, 2009, 10.1007/s11517-009-0464-x. [Online]. Available: http://dx.doi.org/10.1007/s11517-009-0464-x

[83] F. Heinze, K. Hesels, N. Breitbach-Faller, T. Schmitz-Rode, and C. Disselhorst-Klug, “Movement analysis by accelerometry of newborns and infants for the early detection of movement disorders due to infantile cerebral palsy,” *Medical and Biological Engineering and Computing*, vol. 48, pp. 765–772, 2010, 10.1007/s11517-010-0624-z. [Online]. Available: http://dx.doi.org/10.1007/s11517-010-0624-z

[84] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. Wiley-Interscience Publication, 2000.

[85] E. Andrew Rubin and P. David Conway, “Activating applications based on accelerometer data,” Patent Application US 2009/0 132 197 A1, 05 21, 2009.

[86] Y. Li, S. Gong, and H. M. Liddell, “Recognising trajectories of facial identities using kernel discriminant analysis,” *Image Vision Comput.*, vol. 21, no. 13-14, pp. 1077–1086, 2003.

[87] G. Baudat and F. Anour, “Generalized discriminant analysis using a kernel approach,” *Neural Computation*, vol. 12, p. 23852404, 2000.

[88] CommitteeOnQualityOfHealthCareInAmerica, “Crossing the Quality Chasm: A New Health System for the 21st Century,” July 2001.

[89] A. Braun. (2010, February) Proactive vs. reactive: Shifting paradigms in health care provision. [Online]. Available: http://blogs.ubc.ca/phar330/2010/02/13/proactive-vs-reactive-shifting-paradigms-in-health-care-provision/

[90] CentersforDiseaseControlandPrevention, “Overweight and obesity,” *http://www.cdc.gov/obesity/index.html*, 2010.

[91] WorldHealthOrganization, “Obesity: Preventing and managing the global epidemic,” *Geneva: Stationery Office Books*, 2001.

[92] J. P. Foreyt and G. K. Goodrick, “The ultimate triumph of obesity,” *Lancet)*, vol. 346, 8968, pp. 134–135, 1995.

[93] K. H. Schmitz, D. R. Jacobs, A. S. Leon, P. J. Schreiner, and B. Sternfeld, "Physical activity and body weight: Associations over ten years in the cardia study," *International Journal of Obesity*, vol. 24, pp. 1475–1487, 2001.

[94] F. B. Hu, M. F. Leitzman, M. J. Stampfer, G. A. Colditz, W. C. Willett, and E. B. Rimm, "Physical activity and television watching in relation to rik for type 2 diabetes mellitus in men," *Archives of Internal Medicine*, vol. 161, pp. 1542–1548, 2001.

[95] D. Bush, "America on the move lanunches national grassroots initiative will inspire healthy eating and active lying and fight obesity epidemic," *America On the Move HHS Press Office. Boston, MA*, 2003.

[96] AdministrationOnAging. (2009) A profile of older americans. [Online]. Available: http://www.aoa.gov/aoaroot/aging_statistics/Profile/2009/3.aspx(LastaccessedonOctober25,2010)

[97] HeadlinesIndia. (2010) Strategy needed to tackle lifestyle diseases. [Online]. Available: http://headlinesindia.mapsofindia.com/health-and-science-news/lifestyle-diseases/strategy-needed-to-tackle-lifestyle-diseases-president-54945.html(LastaccessedonOctober25,2010)

[98] D. C. Willems, M. A. Joore, J. J. Hendriks, E. F. Wouters, and J. L. Severens, "Cost-effectiveness of a nurse-led telemonitoring intervention based on peak expiratory flow measurements in asthmatics: results of a randomised controlled trial," *Cost Eff Resour Alloc*, vol. 5, p. 10, 2007.

[99] B. Johnston, L. Wheeler, and J. Deuser Kaiser, "Permanente medical center's pilot telehome health project," *Telemed Today*, vol. 5:(4), p. 167, 1997.

[100] "Physical activity and health: A report of the surgeon general," *Washington D.C., U.S. Department of Health and Human Services, Public Health Service*, 1996.

[101] J. A. Levine and Edberhardt, "Role of non-exercise activity thermogenesis in resistance to fat gain in humans," *Science*, pp. 212–214, 1999.

[102] J. A. Levine, M. W. V. Weg, J. O. Hill, and R. C. Klesges, “Non-exercise activity thermogenesis: The crouching tiger hidden dragon of societal weight gain,” *Journal of American Heart Association*, vol. 26, pp. 729–736, 2006.

Appendix A

Different Features investigated in This Study

Mean vector: Each sample from the sensor device at any given time is a three dimensional point and can be represented as

$$A(t) = \begin{bmatrix} a_x(t) \\ a_y(t) \\ a_z(t) \end{bmatrix} \in R^3 \tag{A.1}$$

where x, y and z represent the three axes of an accelerometer. A sequence of acceleration of length T with N samples can be represented as

$$A(T) = \begin{bmatrix} a_x(1)\ a_x(2)\ ,\ldots,\ a_x(N) \\ a_y(1)\ a_y(2)\ ,\ldots,\ a_y(N) \\ a_z(1)\ a_z(2)\ ,\ldots,\ a_z(N) \end{bmatrix} \tag{A.2}$$

where N is the total number of samples for each axis. The mean vector for the above sequence of samples can be written as

$$\bar{A}(T) = \begin{bmatrix} \bar{a}_x \\ \bar{a}_y \\ \bar{a}_z \end{bmatrix} \in R^3 \tag{A.3}$$

where

$$\bar{a} = \frac{1}{N}\sum_{i=1}^{N} a(i) \tag{A.4}$$

Standard deviation: The standard deviation for an acceleration sequence represented in Equation (3) is a three dimensional vector written as

$$S(T) = \begin{bmatrix} s_x \\ s_y \\ s_z \end{bmatrix} \in R^3 \tag{A.5}$$

where

$$s = \left[\frac{1}{N-1}\sum_{i=1}^{N}[a(i) - \bar{a}]^2\right]^{\frac{1}{2}} \tag{A.6}$$

where $\bar{a}$ is the mean for a given axis given by Equation (5).

Spectral entropy: Spectral entropy S_N of the acceleration signal for the frequency band $f_1 - f_2$ was calculated as

$$S_N(f_1, f_2) = \frac{-\sum_{f_i=f_1}^{f_2} P(f_i)\log(P(f_i))}{\log(N[f_1, f_2])} \tag{A.7}$$

where $P(f_i)$ represents the power spectral density (PSD) value of the frequency f_i. The PSD values are normalized so that their sum in the band $[f_1 - f_2]$ is one. $N[f_1 - f_2]$ is the number of frequency components in the corresponding band in PSD.

Correlation: The aim of including this feature was to find out the relationship among three axes. Correlation indicates the strength and the direction of a linear relationship between two random variables. A sample from the sensor device at any given time t is given by Equation (2) and a sequence of such samples for a time segment of length T is given by Equation (3). The correlation among three axes for the aforementioned acceleration sequence was represented in a matrix form as

$$R(T) = \begin{bmatrix} r_{xx}\ r_{xy}\ r_{xz} \\ r_{yx}\ r_{yy}\ r_{yz} \\ r_{zx}\ r_{zy}\ r_{zz} \end{bmatrix} \tag{A.8}$$

where

$$r_{12} = \frac{1}{N-1} \sum_{i=1}^{N} \left(\frac{a_1(i) - \overline{a}_1}{s_1} \right) \left(\frac{a_2(i) - \overline{a}_2}{s_2} \right) \tag{A.9}$$

where r_{12} represents the correlation between two axes of an accelerometer, $a(i)$ the value of the ith-sample for a given axis, $\overline{a}$ the mean for a given axis, while s_1 and s_2 the standard deviation for both axes respectively.

Appendix B

List of Publications

International Journal Papers:

[1] **Adil Mehmood Khan**, Young-Koo Lee, Sungyoung Lee, Tae-Seoung Kim. Accelerometer's Position Independent Physical Activity Recognition System for Long-term Activity Monitoring in the Elderly (Accepted: October 12, 2010) (To be published in) Medical and Biological Engineering and Computing (MBEC), SCI, IF: 1.78.

[2] **Adil Mehmood Khan**, Young-Koo Lee, Sungyoung Lee, Tae-Seoung Kim. A Triaxial Accelerometer-based Physical Activity Recognition via Augmented Features and a Hierarchical Recognizer. IEEE Transactions on Information Technology in Biomedicine. Sep 2010. SSN: 1089-7771, vol: 14, Issue: 5, pp: 1166 - 1172.

[3] Asad Masood Khattak,**A.M. Khan**, Young-Koo Lee, Sungyoung Lee. Analyzing Association Rule Mining and Clustering on Sales Day Data with XLMiner and Weka. International Journal of Database Theory and Application. June, 2010

International Conference Papers:

[4] **A. M. Khan**, Y. K. Lee, S. Y. Lee. Accelerometer's Position Free Human Activity Recognition Using A Hierarchical Recognition Model. 12th International Conference on e-Health Networking, Application Services (IEEE HealthCom 2010). July 1-3 2010. Lyon, France.

[5] **A. M. Khan**, Y. K. Lee, S. Y. Lee, T.-S. Kim. Human Activity Recognition via An Accelerometer-Enabled-Smartphone Using Kernel Discriminant Analysis. The 5th International Conference on Future Information Technology(FutureTech2010). May 20-24 2010. Busan, Korea.

[6] **A. M. Khan**, P. T. H. Truc, Y. K. Lee, T.-S. Kim. A Tri-axial Accelerometer Sensor-based Human Activity Recognition via Augmented Signal Features and Hierarchical Recognizer. Proceedings of 5th International Conference on Ubiquitous Healthcare. November 2008. Busan, South-Korea.

[7] **A. M. Khan**, Y. K. Lee, T.-S. Kim. Accelerometer signal-based human activity recognition using augmented autoregressive model coefficients and artificial neural nets. in Proc. of 30th Annual IEEE International Conference on Engineering in Medicine and Biology Society. 2008. pp. 5172-5175.

[8] **Adil Mehmood Khan**, Faraz Idrees Khan, Yong Koo Lee, Sungyoung Lee. Prospects Identification Scheme for Supermarkets by Classification of Customer Behavior Using Time Based Analysis of Transactional Data. Proceedings of the 9th International Conference on Advance Communication Technology. February 12-14, 2007. Phoenix Park, South Korea.

[9] Myong-Woo Lee, **Adil Mehmood Khan**, Ji-Hwan Kim, Young-Sun Cho, Tae-Seong Kim. A Single Tri-Axial Accelerometer-Based Real-Time Personal Life Log System Capable of Activity Classification and Exercise Information Generation. in Proc. of 32nd Annual IEEE International Conference on Engineering in Medicine and Biology Society. 2010.

[10] Asad Masood Khattak, **Adil Mehmood Khan**, Tahir Rasheed, Young-Koo Lee, Sungyoung Lee. Comparative Analysis of XLMiner and Weka for Association Rule Mining and Clustering. The International Conference on Database Theory and Application. Jeju Island, Korea.

[11] Khan Ozair Idris, Kim Sang-Hyuk, Rasheed Tahir, **Khan Adil**, Kim Tae-Seong. Extraction of P300 using constrained independent component analysis. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 2009;1():4031-4.