HUMAN MOTION RECOGNITION USING A WIRELESS WEARABLE SYSTEM

BY JOHN PAUL VARKEY

A thesis submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Electrical and Computer Engineering

Written under the direction of

Prof. Dario Pompili

and approved by

____

____

____

New Brunswick, New Jersey

October, 2010

ABSTRACT OF THE THESIS

Human Motion Recognition Using A Wireless Wearable System

by John Paul Varkey

Thesis Director: Prof. Dario Pompili

The future of human computer interaction systems lies in how intelligently these systems can take into account the user's context, that is, how well the data that they produce characterizes the user's current situation. Context awareness is essential for ubiquitous and wearable computing. Research on recognizing the daily activities of people has progressed steadily, but little focus has been devoted to recognizing activities along with the movements involved in it. For many applications such as rehabilitation, sports medicine, geriatric care, and health/fitness monitoring, the importance of combined recognition of activity and movements within an activity can drive health care outcomes. Motion recognition aims at recognizing the actions of one or more users from a series of observations on the users' actions and environmental conditions.

Sensor-based motion recognition integrates the emerging area of wireless sensor networks with novel machine learning techniques to model a wide range of human motions. A novel algorithm is proposed that can be tuned to recognize on-the-fly either range of activities or fine motor movements within a specific activity using wirelessly connected sensor motes (equipped with accelerometers and gyroscopes) attached to different body sites. This thesis describes a novel algorithm for both situations and also presents a case study on optimal feature set from sensor values and various parameter values for the algorithm to detect the fine motor movements within an activity.

Acknowledgements

I would like to take this opportunity to thank all the wonderful people who have played a significant role in my years as a graduate student. Firstly, I would like to thank my graduate advisor, my role model and guru Prof. Dr. Dario Pompili. I consider him my mentor and am immensely indebted to him for his faith in my abilities. I have always been awed, inspired and motivated by his vision, simplicity and timely words of encouragement and I feel honored to have been associated with him.

I am greatly indebted to my labmates - Eun Kyung Lee, Hariharasudhan Viswanathan, Indraneel Kulkarni, Baozhi Chen, Rahul Suryawanshi and Yang Yuan in Cyber Physical Systems Laboratory supervised by Prof. Dr. Dario Pompili. I would like extend my special thanks to Hariharasudhan Viswanathan and Eun Kyung Lee for their help, feedback and comments whole through out my stint in the lab and also for being subjects of all my experiments for the thesis. I would also like extend special to Yang Yuan for being another subject of my experiments.

I would like to thank my parents, Mrs. Mariamma Varkey and Mr. P.V. Varkey and my brothers, Mr. George Joseph Varkey and Mr. Mathew Varkey, and their family for their constant support and encouragement through out my entire career. I was lucky to have wonderful roommates and friendship in Naveen Atmakuri, Mahesh Lakshminarayanan, Sandilya Bhamidipati, Teja Indrakanti, Mahesh Tadepalli, Lavesh Gupta, and Anudhar Emani. I cannot thank them enough for their show of love and support in all my endeavors.

I would like to thank Prof. Ivan Marsic and Prof. Zoran Gajic for accepting to be the committee members in my thesis defense.

I thank all the faculty and staff of ECE department at Rutgers for their help and support.

Dedication

To my family and to all my friends.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The future of human computer interaction systems lies in how intelligently these systems can take into account the users context, that is, how well the data that it produces characterize the users current situation. Motion recognition is a key feature of many ubiquitous computing applications ranging from office worker tracking to health care. In general, motion recognition systems unobtrusively observe the behavior of people and characteristics of their environments, and, when necessary, take actions in response ideally with little explicit user direction. Motion recognition aims to recognize the actions of one or more users from a series of observations on the users' actions and the environmental conditions.

Sensor-based motion recognition integrates the emerging area of sensor networks (network of spatially distributed autonomous sensors to cooperatively monitor vital signs, temperature, etc, from different body sites) with novel data mining and machine learning techniques to model a wide range of human motions. Sensor-based motion recognition researchers believe that by empowering ubiquitous devices and sensors to monitor the behavior of users (under consent), these devices will be better suited to act on their behalf. Human motion recognition systems composed of wirelessly connected sensor motes (equipped with accelerometers and gyroscopes) attached to different body sites will enable a variety of applications such as rehabilitation, sports science/medicine, geriatric care, and health/fitness monitoring [7]. For example, such a system can be used to measure the effectiveness of active physiotherapy, to perfect techniques of sport persons, to remotely monitor and trigger emergency response for the elderly, and to help people lose weight by providing accurate estimates of their expended calories. In addition, this system will also help facilitate next-generation 3D human-machine interactions. This chapter gives the motivation behind the thesis and our contribution through this work.

1.1 Motivation

To understand the importance of motion recognition better, consider the following scenario. An elderly man wakes up at dawn in his small studio apartment, where he stays alone. He lights the stove to make a pot of tea, switches on the toaster oven, and takes some bread and jelly from the cupboard. After taking his morning medication, a computer-generated voice gently reminds him to turn off the toaster. Later that day, his daughter accesses a secure website where she scans a check-list, which was created by a sensor network in her father's apartment. She finds that her father is eating normally, taking his medicine on schedule, and continuing to manage his daily life on his own. That information puts her mind at ease. Hence, recognizing the users current actions will help study the behavioral pattern of the user by correlating it with the users context.

To understand human motion it is imperative to know the difference between an *activity* and the *movements* that comprise it. A physical movement is a body posture/gesture that typically lasts for several milliseconds or seconds, while an activity lasts several minutes or hours and comprises of different physical movements that may be repeated over time. For example, a "walking" activity would comprise of several short leg movements. There has been some research on recognizing the daily activities of people such as whether a person is walking, jogging, standing, etc. However, most prior research has focused on activity recognition without directly considering the movements involved in that activity. Recognizing specific fine motor movements within activities for individual persons will help provide a clear picture of the intensity of his/her activity. For example, in "walking" by knowing the number of steps (leg movements) taken by the person will help to calculate the pace at which the person is walking. In addition to that , if electrocardiogram (ECG) of the person is monitored over the time he/she is walking his/her heart variability can be measured, which would help in determining whether the person is healthy or not. However, what makes movement recognition more challenging than activity recognition is that we are dealing with much shorter time scales.

There are also many vision-based techniques proposed for human motion recognition but none of those techniques can be applied in situations where privacy is required or ecological validity is paramount. Analysis in [10, 30] has shown that human activity recognition algorithms based on inertial sensors like accelerometers, which provide acceleration values of body motion [9, 27], outperform those based on other non-invasive sensors. To the best of our knowledge, there has not been an holistic approach proposed in the literature that addresses the challenges of using wireless wearable system to recognize activities as well as the movements within a specific activity.

1.2 Our Contribution

In this thesis, we propose a novel window-based algorithm that can be tuned to recognize on the fly either various activities or movements in a specific activity using a supervised learning approach based on Support Vector Machines (SVMs). To recognize various activities and movements, we use acceleration values (collected using accelerometers) and angular rate of motion (collected using gyroscopes) for each activity and movement type considered. However, using all the raw data (acceleration and angular rate) would be inefficient in addition to adding complexity. Hence, meaningful features are extracted from the raw data. We used features such as mean, standard deviation, maximum, peak-to-peak, root-mean-square and correlation between pair of accelerometer and gyroscope axes. We use these extracted features to first train the SVM. We then use our proposed window-based algorithm to recognize all the activity and movement types for which the SVM was trained for.

Our approach consists of three phases - training, tuning and motion recognition. In the training phase, we train the SVM for different activity types or movement types within a specific activity and so for that we take several observations of the activity types(if activity recognition) or movement types (if movement recognition) and extract features from the raw data to be fed into the SVM. Then, we have the tuning phase. Unlike activity recognition, for movement recognition we tune the parameters of the window-based algorithm as movement recognition requires much more fine graining due to shorter time scales. Although tuning can be done for activity recognition, in this thesis we do not focus on the tuning for activity recognition as it does not bring significant improvement in the accuracy when considering the added complexity. Third and final phase is the motion recognition. We call it as motion recognition phase, as the it involves either activity recognition or movement recognition depending what the SVM is trained for (activities or movements within a specific activity). In this thesis, we do a case study on activity recognition accuracy for three different scenarios of SVM training. Finally, we also do a case study on parameter tuning of the algorithm for movement recognition.

1.3 Thesis Organization

The remainder of this thesis is organized as follows.

Chapter 2 introduces some of the existing work in activity and motion recognition and also discuss how our work is different from the existing work. We further discuss about context awareness and how it is related to our work. Finally, we also discuss some of the applications of motion recognition systems.

Chapter 3 explains the window-based algorithm to recognize activities as well as movements within a specific activity. First, we explain the classification method used for motion recognition using wearable sensors such as accelerometers, which provide linear acceleration values of motion and gyroscope, which provide angular rate of motion. Further, we explain the various phases involved in our motion recognition system such as training, tuning and motion recognition. Finally, we introduce the window-based algorithm to recognize activities as well as movements within a specific activity.

Chapter 4 discusses the performance of the algorithm in recognizing both activities and movements within a specific activity. We also presents a case study on parameter tuning of the algorithm for movement recognition given that we are dealing with much shorter time scales compared to those of activity recognition.

Chapter 5 concludes by highlighting the accuracy of the window-based algorithm for activity recognition as well as movement recognition. The chapter further provides suggestions for possible future work such as including contextual information such as vital signs to enable numerous application like health behavior monitoring, geriatric care, etc.

Chapter 2

Related Work and Background

This chapter gives an overview of the state of the art, current topics and challenges in motion recognition using wearable sensors. One of the main promises of wearable computing is to enable personal applications that can adapt and react to the current context of the user. While the term context is usually broadly defined and can in principle encompass any kind of information that relates to the current situation of the user or the objects surrounding him [16], this thesis focuses on the users activity, which is often considered one of the most important ingredients of context besides the users location. Hence, in the following we give an overview of work that uses wearable sensors to infer the current activity of the user. We further discuss about context awareness and other existing activity recognition methods [24]. Finally, we also highlight some of the applications of motion recognition systems.

2.1 Previous Work

Previous work on activity recognition using acceleration values have considered features like mean [9, 21, 25, 36], standard deviation [21, 25, 36], maximum [9, 36], peak-to-peak [18], root-mean-square [21] of acceleration values and correlation of acceleration values between pair of axes of the accelerometer [9, 36]. However, in addition to the features collected from the accelerometer values we also extract and use features from angular rate values (gyroscope value) as knowing the orientation of various points of the body would help in differentiating similar activities or movements. Almost all of these works [9, 21, 25, 36] has focused on activity recognition without taking into account the movements involved. However, for some applications like behavioral study of patients and health monitoring, knowing just the activities is not enough. In such applications, knowing the movements involved would help further in deducing the intensity of the activity performed as explained in the introduction. Hence, in our work we focus on both activity and movement recognition and we propose an approach that can be tuned to different timescales to be able to recognize both activities and movements and their starting and ending time instants.

The authors in [28] define a general framework for activity recognition by building upon and extending Relational Markov Networks. However, they had used some constraints like one activity per location which is improbable in real life. In [41], authors introduce a sensor and annotation system for performing activity recognition in a house setting and used probabilistic models to learn the parameters of activities in order to detect them in future sensor readings. In [7], the authors discuss activity recognition results for stereotypical hand flapping and body rocking using data collected from children with Autism in both laboratory and classroom settings and also they present a case study on the various challenges encountered when applying machine learning for recognizing activities. In [18], the authors show that movements have a grammatical framework such as a spoken language and introduce a linguistic framework for symbolic representation of inertial information by constructing primitives across the network of sensors for each movement and then using a decision tree. However, in [18] the approach proposed relies on a perfect segmentation assumption, i.e., it does not identify the starting and finishing instants of movements within an activity.

The computer vision community also has conducted some research on human motion recognition using time frames of a video sequence [33, 17]. However, the downside of this method is that processing video data is very costly and those techniques require external infrastructure - e.g., (infrared) cameras - and may be biased by environmental conditions such as background light or heat. Also, such techniques cannot be directly applied to those scenarios that require privacy, even if the image is blurred. Conversely, inertial sensors like accelerometers and gyroscopes are unbiased by environmental conditions and give a good accuracy for motion analysis. In addition, motion recognitions systems using inertial sensors can be used in application where privacy is an important issue as these systems can be trained to recognize only specific activities or movements.

Some researchers have implemented basic activity recognition using machine learning techniques such as neural networks (Backpropagation, Kohonen Self - Organizing Maps), probabilistic models, and fuzzy logic. In [42], authors used Kohonen maps to perform adaptive recognition, i.e. the system was not pre-trained but trained specifically by the user. The KSOMs and probabilistic models used to train the system also preprocess data using techniques such as standard deviation, variance, derivatives, and FFTs (Fast Fourier

Transforms) before clustering the data to determine the output. The two disadvantages of the KSOM approach described are catastrophic forgetting and the curse of dimensionality. Catastrophic forgetting implies that the system initially has a high learning rate that is exhausted after some time. In order to keep learning previous values are over - written. The curse of dimensionality states that as the number of inputs to the classifier increases, so does system complexity which results in slower learning rates.

In [19], Golding used a probabilistic approach based on Bayes rule to develop context-awareness for an indoor navigation system. Several sensors are placed in a utility belt and are used to measure distinctive patterns in the readings to deduce location. The sensors are sampled every 50 msec (20 samples per second) after which Bayes rule is used to calculate the users location probability. An important result from that paper is that the error rate in predicting user location is dramatically reduced from 50 percent to two percent by preprocessing the data. The data can be preprocessed in a number of ways such as computing the mean, standard deviation, and variance of the last N samples. There is no general rule of thumb for processing data from different sensors. Also, the system has to be trained to recognize the constraints (maxima, minima) of the body. This training helps solve issues with state transitions, eliminates erroneous data, and takes more accurate absolute and relative measurements.

Authors in [35] applied clustering neural network algorithms to determine user context. Their work examined the validity of using a single accelerometer and minimizing power by using a very low sampling rate. The measured accuracy for several motions was roughly 75 %. In contrast to this single accelerometer approach, work by authors in [43] incorporated a relatively large number of sensors, the majority of which were accelerometers, to determine user activity. Van Laerhoven et al. developed a harness-like system that was worn on-top of clothing and examined the performance of classification as a function of the number of sensors. The classification accuracy was also monitored as the number of contexts increased. It was determined that the performance heavily depends on the number of sensors and contexts, as well as the nature of the contexts.

In most machine learning techniques the rate of learning slows down as the number of inputs increases. It is extremely difficult to visualize and work with high-dimensional data. Researchers who work with some variation of neural networks often complain about the opacity in its behavior. Due to the opaqueness of the inner workings of neural networks the selection of the number of inputs, hidden layers, and type of transfer functions is difficult to optimize. In addition to all the choices for neural network design, even the randomization of inputs affects the training model. Non-random selection of inputs can lead to over-training the system for a specific activity [2].

2.2 Context Awareness

Context awareness has been a widely researched topic for ubiquitous computing. Context is defined by [6] as any piece of information used to characterize the situation of a person, place, or object. The context space for what a person is doing can be categorized hierarchically [39]. It contains many attributes such as physiological, environmental, and social conditions. Context-aware applications are developed to intuitively deduce and interpret the context of the current situation and react appropriately. The Active Badge [46] system developed at Olivetti Research Lab is considered to be one of the first context-aware applications. In this system the office personnel wear badges that transmit IR (infra red) signals to a network of sensors placed around the building. The system is used to locate people and forward calls to the closest phone near them. The ParcTab system developed at Xerox PARC is based on palm-sized wireless ParcTab computers with an IR communication system that links them to each other and to desktop computers [45]. Many applications were developed for the ParcTab system, such as determining user location, the presence of other mobile devices, and nearby non-mobile machines. The Forget-Me-Not [37] application used ParcTabs to records where its user is, who they are with, whom they phone, and other autobiographical information, which are stored it in a database for later retrieval. The StartleCam developed by [20] integrated a wearable video camera, computer, and sensing system, which stores images to a remote server whenever it detects events of interest to the wearer [20]. These events are determined by the change in skin conductivity, which can be correlated to arousal level.

This thesis explores a subset of context detection, where everyday activities or movements performed by the user are recognized. Examples of these activities include but are not limited to walking, jogging, standing up, and running. Activity recognition by itself provides low level information, but can be used in conjunction with other on-body physiological sensors to enhance medical monitoring. Many health monitoring applications require that the patient keep a log of activities or movements so that when the physiological data is analyzed off-line, the data can be correlated to what the patient was doing before or during the health events of interest. Patients often do a poor job of self-reporting, so it would be desirable to automatically annotate the physiological data with the users activity. The creation of an activity diary automatically from sensor data eliminates the need for user input, potentially increasing accuracy and detail over a hand-written diary. Also, the log of activities can easily be translated to more useful information such as caloric energy expenditure of the user. The next section discusses the some of the application areas of motion recognition.

2.3 Applications

In the following we outline application areas for motion recognition systems in wearable or mobile settings. We begin with applications for health care and assisted living, which represent an important class of applications. Besides that, there exist a number of other application areas [22], such as industrial applications, applications for entertainment and gaming and various other applications.

Healthcare and Assisted Living. A major goal of current research in motion recognition and context-aware computing in general is to enable new health-related applications and technologies for the aging. Longer life expectancy and declining fertility rates are increasing the proportion of the elderly population in societies worldwide and posing challenges to existing healthcare systems. It is hoped that technology can help in addressing these challenges, for instance by helping elderly people to live more independent lives and thus reducing the burden of care-givers.

One type of system designed for elderly people aims to detect potentially dangerous situations in a persons life in order to call for external help automatically. Such systems can be seen as a complement to traditional emergency systems such as smoke- or fire alarms, by detecting e.g. when a person has fallen [23, 13], or when vital body signs indicate imminent health threats [29, 44, 8]. Preventing age-related diseases or severe medical conditions before they actually happen is the goal of another class of applications, which employ long-term monitoring to detect changes or unusual patterns in a persons daily life that may indicate early symptoms of diseases such as Alzheimers. While automatic detection of subtle behavioral changes is highly challenging and still a long-term goal of current research, applications that accumulate and summarize statistics about daily activities [14] or perform continuous recordings of physiological parameters [26, 8, 29, 34] can already be valuable for physicians and care-givers to estimate the physical well-being of a person.

Industrial Applications. Currently several research groups explore the next generation of industrial applications which, among other improvements, take better advantage of the multi-modal sensing capabilities of wearable platforms by inferring context information such as the users current activity. For instance, [31] investigate the use of wearable computing technology for scenarios in aircraft maintenance, car production, hospital environments and emergency response. In these scenarios, wearable technology and activity recognition are used to provide interactive and hands-free access to information such as electronic manuals or patient records, assist in training of new workers, provide summaries of performed activities, as well as to help in navigation and communication. [40] use information gathered from wearable and environmental sensors for tracking activities of workers in car manufacturing plants, e.g. to provide realtime feedback to the worker about upcoming assembly steps or to issue warnings when procedures are not properly followed. [47] combine data from wearable microphones and accelerometers in order to track wood shop activities such as sawing or hammering.

Entertainment and Games. Wearable systems using motion recognition are appealing for applications in the performing arts, e.g. by allowing dancers to augment their performance with interactive multimedia content that matches their motions. Such systems are described by [11], who employ wearable inertial sensors combined with machine learning techniques in order to record, classify and visualize the motion of dancers. For entertainment and gaming systems in general, the adoption by users may be faster than in other domains, since recognition accuracy is less crucial than e.g. for health care systems, and since they usually raise less privacy concerns. The recent popularity of game controls based on accelerometers, sparked by systems such as Nintendos Wii platform [3] or the Apple iPhone [1], has introduced a wide audience to ideas that originated in the context-aware computing research community and are now being widely adopted by companies and independent developers.

Other Application Areas. There are numerous other possible application areas for wearable computing combined with activity recognition. For example, [38] explore the use of activity-recognition for mobile context-aware advertising. In an educational context, [12] investigate the use of wearable RFID readers combined with tagged objects for casual learning of a foreign language vocabulary. Finally, [32] use a wearable sensing platform to categorize soldier activities, in order to automatically compile action reports or help in recalling situations during missions.

Chapter 3

Proposed Work

This chapter presents the window-based algorithm that can recognize on the fly an activity or the movements involved in a particular activity and also find the starting and finish instants of the current activity or the movements in it. The proposed approach can be used for either activity recognition or movement recognition. According to our approach, we first pursue the activity recognition phase and followed by the movement recognition phase using the same window-based algorithm but separately trained for the two phases. The algorithm is based on machine learning in which it is required to train the system to the type of activities or movements that need to be recognized. For the machine learning approach, we use Support Vector Machines (SVM's). First, we discuss about the classification method used in the SVM; then, we explain the various phases involved in our approach for human motion recognition. Finally, we describe the original window-based algorithm and provide details on its functioning.

In order to recognize on-the-fly human physical activities such as walking, running, driving, dancing, gesturing or the specific movement types composing these activities, the proposed wearable system will 1) collect raw data through distributed sensing and 2) employ supervised learning methods (requiring minimal user inputs) to extract information about the physical activities and movements through localized computation. Interestingly, the system is inherently pervasive (while not interfering with the day-to-day lifestyle of the user) and ensures privacy unlike camera-based solutions (because it is trained to classify only a predefined set of activities and movements). It also generates recording of user' behavior with little or no subject reactivity (user's behavior is unlikely to change based on the monitoring); these features enable research studies with higher internal consistency and support individually tailored treatment regimes especially when used for health monitoring.

3.1 Classification Method

Classifying data is a common task in machine learning. Suppose there are some data points each belonging to either one of two classes, and the goal is to decide which class a new data point will be in. In SVMs [15], a data point is viewed as a p-dimensional vector (i.e., a list of p numbers); the objective is to separate such points using a (p-1)-dimensional hyperplane. This can be done using a linear classifier. A *classifier* is a function that maps input data samples to a defined set of object class after having 'seen' a number of training examples. The classification is achieved by an hyperplane that separates the data samples into different classes, as it is shown in Fig. 3.1. In general, there are many hyperplanes that can classify the data. One reasonable choice as the 'best' hyperplane is the one that represents the largest separation, or margin, between the two classes. Hence, we choose the hyperplane so that the distance from it to the nearest data point on each side is maximized. The samples on the margin are called the *Support Vectors (SVs).* This hyperplane - if it exists - is known as the *maximum-margin hyperplane.* An hyperplane can be defined as shown in the following,

$$F(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b, \tag{3.1}$$

where $\mathbf{w}$ is the normal vector to the hyperplane and can be derived as,

$$\mathbf{w}^* = \sum_{i=1}^{l} \alpha^{i*} \cdot y^i \cdot \mathbf{x}^\mathbf{i}, y^i \in \{-1, 1\}, \tag{3.2}$$

subject to the condition,

$$\alpha^{i*}[y^i \cdot (\mathbf{w}^{*T}\mathbf{x}^\mathbf{i} + b^*) - 1] = 0, \forall \alpha^i \neq 0, \tag{3.3}$$

where l is the number of support vectors (SVs), α^i is the i^{th} Lagrange multiplier.

The function of the hyperplane in (3.1) is not suitable for solving more complicated, linearly non-separable problems and when dealing with more than two classes. In such cases, rather than fitting nonlinear curves to the data, SVM handles this by using a kernel function to map the data into a high-dimensional space where the hyperplane can easily do the separation. There are four basic kernels - linear, polynomial, radial basis function (RBF), and sigmoid. In our system, we train the algorithm using the SVM-based machine learning toolbox available in Matlab (called "Spider" [5]) with RBF as the kernel as it is the most robust according to our experiments.

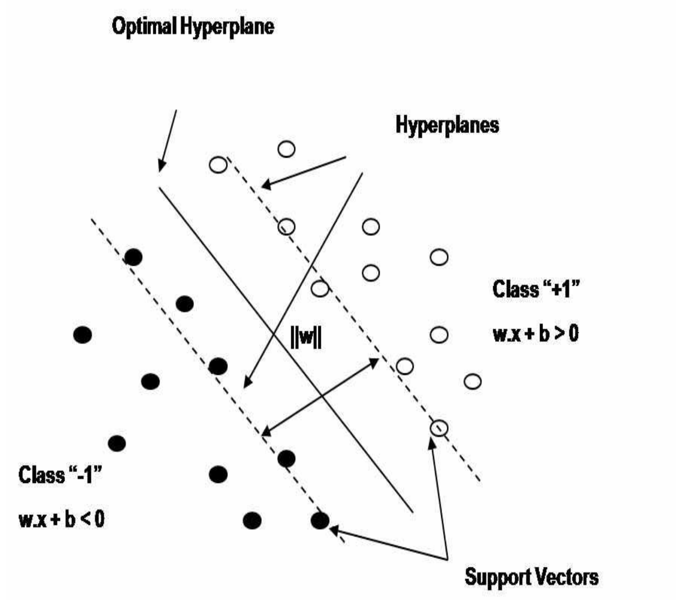

Figure 3.1: Support Vector Machine (SVM).

In order to classify multiple classes, the multiclass problem is reduced into multiple binary classification problems. Each of the problems yields a binary classifier, which is assumed to produce an output function that gives, for example, relatively large values if the vector to classify belongs to the positive class and relatively small values for the negative class. Two common methods to build such binary classifiers are where each classifier distinguishes between (1) one of the labels to the rest ("one-versus-all") or (2) between every pair of classes ("one-versus-one"). Classification of new instances for one-versus-all case is done by a winner-takes-all strategy, in which the classifier with the highest output function assigns the class. Conversely, for the one-versus-one approach, the classification is performed by a max-wins voting strategy, in which every classifier assigns the instance to one of the two classes; then, the vote for the assigned class is increased by one vote, and finally, the class with most votes determines the instance classification. In our approach, we use "one-versus-one" method as it gives the best results.

3.2 Problem Formulation

Our approach for recognizing activities or the movements in a specific activity consists of three phases - *training*, *tuning*, and *motion recognition*, as shown in Fig. 3.2. In the training phase, the SVM is trained with sets of linear acceleration values (from accelerometers)

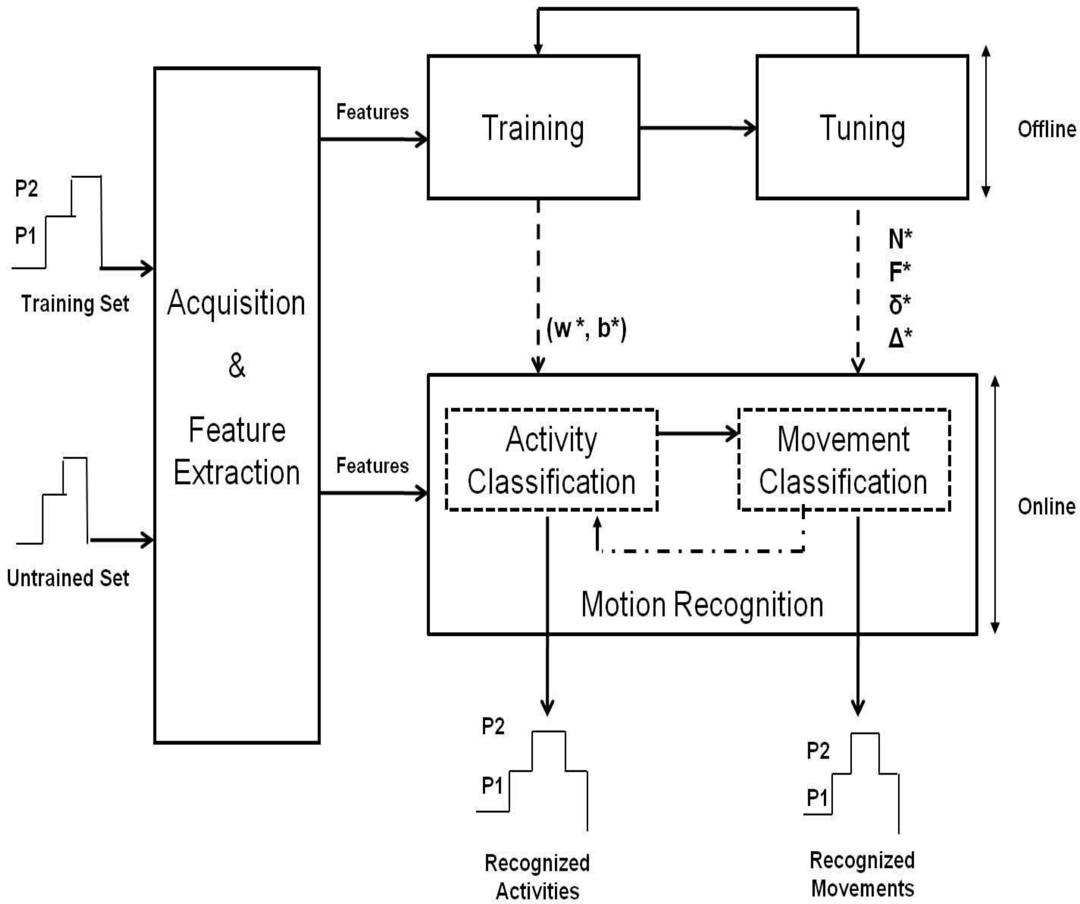

Figure 3.2: Phases involved in motion recognition.

and angular rate values (from gyroscopes) for each activity type or the movement types (depending on whether we want to recognize the activity or the movements in a particular activity). The tuning phase is then used to tune the parameters involved in our algorithm so as to improve the accuracy of recognition. We mainly tune the parameters when it involves recognizing the movements in a particular activity as movements have very shorter time scales compared to activities. In the motion recognition phase, once the SVM is trained with activity types or the movement types, we use a windowing-based algorithm to recognize the activities or the movement types. As shown in Fig. 3.2 we have combined both activity recognition and movement recognition under the phase *motion recognition*. In addition, in the *motion recognition* block we have shown a feedback from the movement recognition phase to the activity recognition phase (although not applied for this thesis), which could be used to improve the accuracy of the activity phase.

Before explaining the windowing algorithm, let us define the parameters used to represent an activity or the movements within a specific activity in the training phase. We represent $\mathcal{D}$ as the training set, which is nothing but a set of activities or movements (depending on whether it is activity or movement recognition), which includes many observations of the same type of activities or movements. Hence, without loss of generality, $\mathcal{D}$ is a set of say P observation of activities (or movements) such that each acitivity/movement $p = 1, 2, \cdots, P$. In addition, let ζ be the set of R types (or classes) of activities/movements considered for training given as,

$$\zeta = \{\mathcal{C}_1, \mathcal{C}_2, ..., \mathcal{C}_r, ..., \mathcal{C}_R\}, \tag{3.4}$$

where $\forall p = 1, 2, ..., P$ and $p \to \mathcal{C}_r$ if the p^{th} activity/movement is of type $\mathcal{C}_r$ (r^{th} activity/movement type) considered for training where $r = 1, 2, \cdots R$. Each activity/movement p can then be associated with *any and only one* of the activity/movement type in ζ such that $R \ll P$.

Further, we represent $\mathcal{S}$ as the set of sensor nodes used to collect both linear acceleration and angular rate values from different body sites such that $s = 1, 2, \cdots, |\mathcal{S}|$. As mentioned before we do not use the collected raw values as such for training but we extract as set of features (list of features mentioned in 2) from the raw values. We represent $\mathcal{F}$ as the set of features extracted from N sub-intervals (not from the entire time interval) from each axis of the accelerometer and gyroscope for each activity/movement p such that each feature $f = 1, 2, \cdots, |\mathcal{F}|$. Using all the parameters, v_p^s represent the p^{th} activity or movement with respect to a sensor node s and is given as,

$$v_p^s = [v_p^{sx}, v_p^{sy}, v_p^{sz}], \tag{3.5}$$

where, v_p^{sx}, as in (3.6), denotes the set of features extracted from the x-axis of the accelerometer and x-axis of the gyroscope attached to the sensor node s.

$$v_p^{sx} = [f_1^{1x}, f_2^{1x}, ..., f_{|\mathcal{F}|}^{1x}; ...; f_1^{nx}, ..., f_{|\mathcal{F}|}^{nx}; f_1^{Nx}, ..., f_{|\mathcal{F}|}^{Nx}]. \tag{3.6}$$

Then using features extracted from readings of all the sensor nodes, p^{th} activity or movement, v_p can be given as,

$$v_p = [v_p^1, v_p^2, ..., v_p^s, ..., v_p^{|\mathcal{S}|}]. \tag{3.7}$$

Finally, the entire training set $\mathcal{D}$ can be represented as,

$$\mathcal{D} = \{(v_1, \mathcal{C}_r^1), ...(v_p, \mathcal{C}_r^p), ..., (v_P, \mathcal{C}_r^P)\}, \mathcal{C}_r^p \in \zeta. \tag{3.8}$$

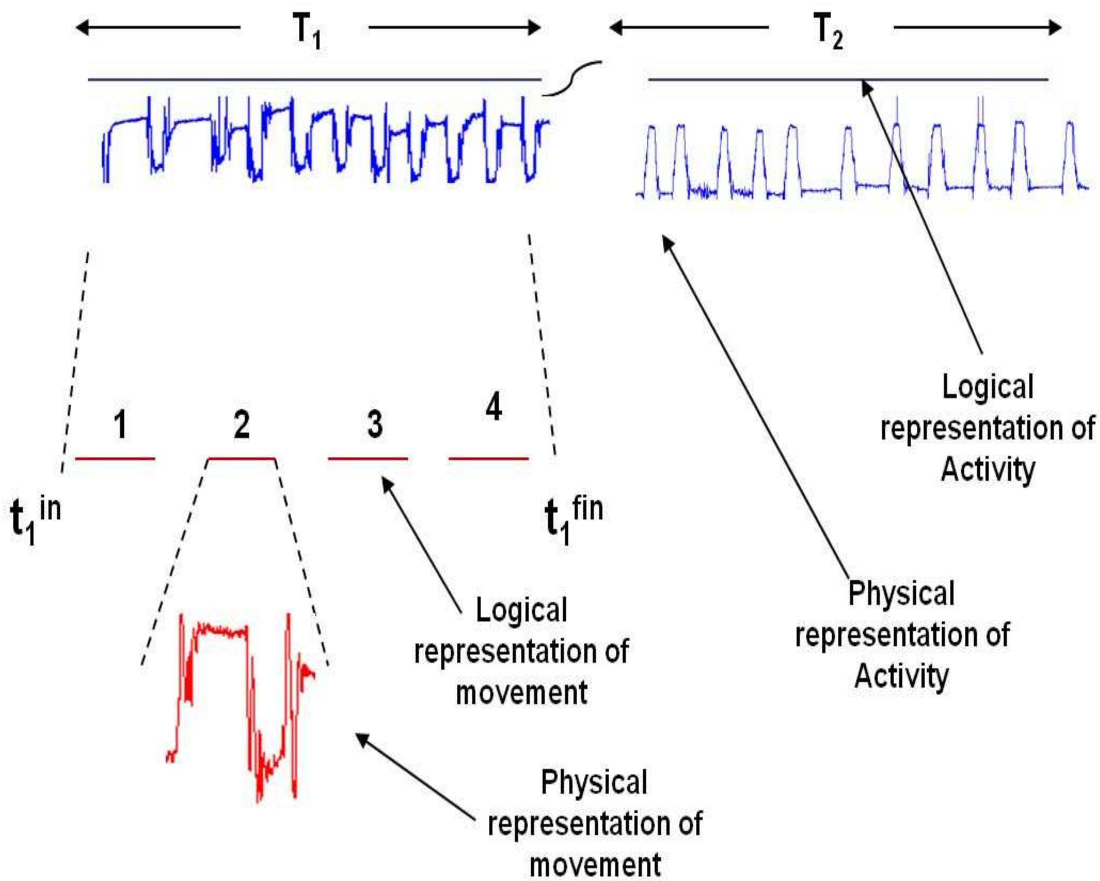

Figure 3.3: Logical and Physical representation of activities and movements.

Once the training phase is complete, we derive the normal vector $\mathbf{w}$ to the hyperplane as shown in (3.3) using the support vectors and the Lagrange multipliers [15] obtained from the SVM, which is then used to recognize the current activity or the movements in a particular activity during the recognition phase using the windowing algorithm. We also define the concept of "confidence", which forms the basis of our recognition algorithm. For a generic vector $\mathbf{x}$, we define its confidence as its distance from the hyperplane ($\mathbf{w}$, b) given as,

$$d(\mathbf{w}, b; \mathbf{x}) = \frac{\mathbf{w} \cdot \mathbf{x} + b}{||\mathbf{w}||}. \tag{3.9}$$

This distance, $d(\mathbf{w}, b; \mathbf{x})$, indicates how confident the classifier is about the type or class of the vector $\mathbf{x}$: the higher the confidence, i.e., larger the d and so more certain is the classifier.

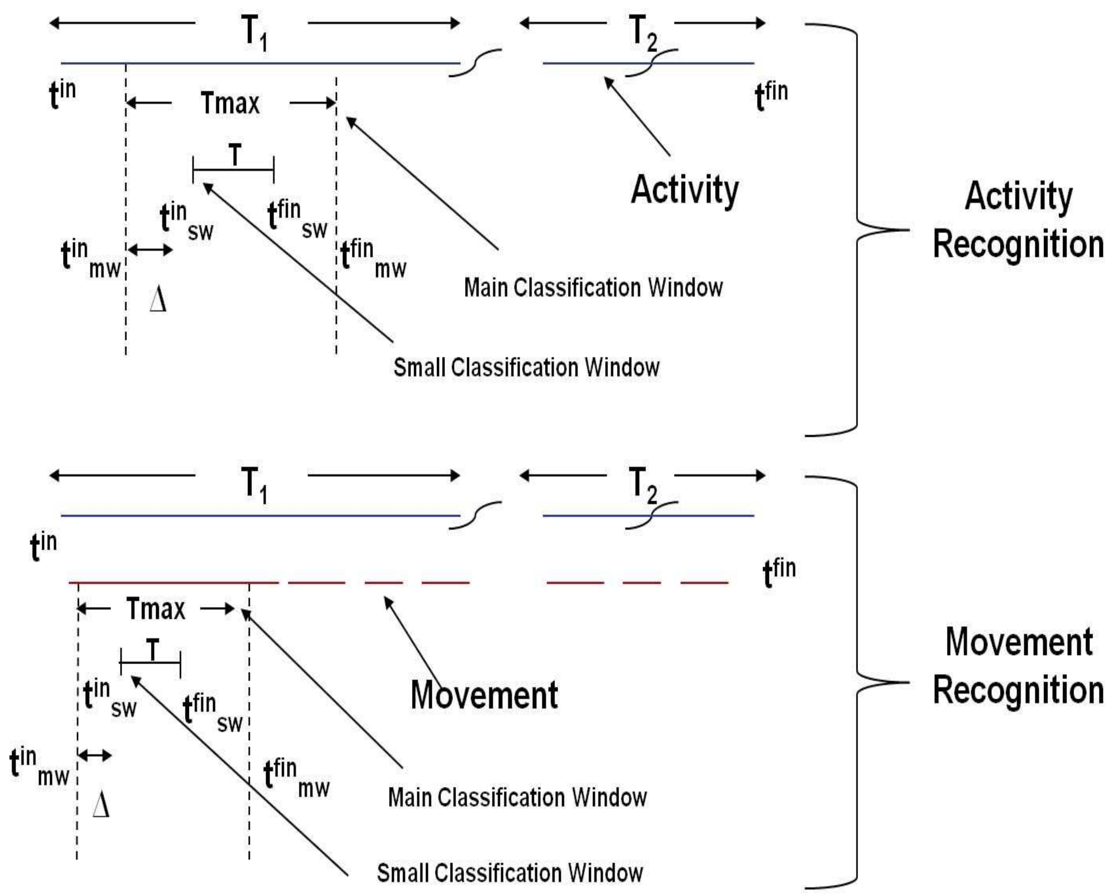

Figure 3.4: Classification windows applied on activities and movements.

3.3 Proposed Solution

We propose to use a window-based algorithm for recognizing the activities or movements (depending whether activity or movement recognition is required). We first pursue the activity recognition phase sequentially followed by movement recognition using the same window-based algorithm, but separately trained for activity and movement recognition. Figure 3.3 shows both the *logical* (a line representing the class) and *physical representation* (the raw signal, i.e., linear acceleration and angular rate values) of an activity and the movements in it. Hereafter we will be using the logical representation for our further discussion as using raw representation of activity and movements would be difficult in our explanation.

The algorithm uses two classification windows, a *main classification window* and a variable size *small classification window* that always moves only within the main classification window as shown in Fig. 3.4. We define the intervals of activity (or movement) and main classification window to be respectively as $[t_p^{in}, t_p^{fin}]$ and $[t_{mw}^{in}, t_{mw}^{fin}]$. In addition, we also consider two other intervals, T^{min} and T^{max}, which we define as minimum and maximum time interval. Based on the demonstrated success for recognizing activities in previous works [9], [36] on accelerometer-based activity recognition using a window size of 32 samples with 16 samples overlapping between consecutive windows, we fixed the value of T^{max} and T^{min} to be $32 \cdot 0.05$, where 0.05 is the sampling time in seconds. The following relationships hold,

$$T^{min} = \begin{cases} \min_{1 \le p \le P} t_p^{fin} - t_p^{in} & \text{if } p \text{ is movement} \\ 32 \cdot 0.05 & \text{if } p \text{ is activity} \end{cases} \tag{3.10}$$

$$T^{max} = \begin{cases} \max_{1 \le p \le P} t_p^{fin} - t_p^{in} & \text{if } p \text{ is movement} \\ 32 \cdot 0.05 & \text{if } p \text{ is activity} \end{cases} \tag{3.11}$$

We vary the size of small classification window interval T based on,

$$T = T^{min} + (h - 1) \cdot \delta, \tag{3.12}$$

where $h = 1, 2, ..., H$ and $H = \lfloor \frac{T^{max} - T^{min}}{\delta} \rfloor + 1$.

The main classification window interval (in Fig. 3.4) is always set to T^{max} and is shifted over the entire time interval that needs to be recognized. For each h value, as given in (3.12), there is a small classification window interval T. The number of small classification windows within the main classification window H depends on T^{min} and the δ value. The small classification window is shifted by Δ within the main classification window for each T value, until it reaches the end of main classification window. The number of such shifts is represented by K. For each small classification window T shift within the main classification window, the confidence is calculated as in (3.9). Once the confidence for all the small classification windows T $\forall h = 1, 2, ..., H$ within the main classification window is computed, we select three intervals with the best confidence out of those initially considered, and check for any overlap in their time intervals. If there is any overlapping windows, we combine those time intervals and then compute the confidence for the combined interval. If there is no overlapping window, then we avoid those window intervals that have confidence less than the average confidence among all the small windows considered within the main classification window. The recognized activity/movement will be the class indicated by the SVM for the window interval that has best confidence. Once the activity/movement is recognized, the main classification window is shifted by the recently recognized time interval. The pseudo-code of the windowing algorithm is given in Algorithm 1.

Algorithm 1: Window-based Recognition Algorithm

```
if Activity Recognition then
|   T^max = T^min = 32 · 0.05;
else
|   T^max = max_{1≤p≤P} |t_p^fin − t_p^in|;
|   T^min = min_{1≤p≤P} |t_p^fin − t_p^in|;
|   δ from tuning phase;
|   Δ from tuning phase;
H = ⌊(T^max − T^min)/δ⌋ + 1;
t_mw^in = t^in;
t_mw^fin = t_mw^in + T^max;
while t_mw^fin <= t^fin do
|   for h = 1 to H do
|   |   T = T^min + (h − 1) · δ;
|   |   K = ⌊(T^max − T)/Δ⌋ + 1;
|   |   for k = 1 to K do
|   |   |   t^in = t_mw^in + (k − 1)Δ;
|   |   |   t^fin = T + t_mw^in + (k − 1)Δ;
|   |   |   extract features;
|   |   |   find confidence;
|   Take 3 recognized small windows with best confidence;
|   if Overlapping intervals then
|   |   find confidence of combined intervals;
|   else
|   |   Avoid intervals with confidence < average;
|   Recognized Activity/ Movement = Class of window with best confidence;
|   t_mw^in = t^fin;
|   t_mw^fin = t_mw^in + T^max;
```

The performance of the algorithm depends on the following parameters - (1) Right feature set $\mathcal{F}$, (2) Number of sub-intervals N for feature extraction, (3) Number of small classification windows H that depends on δ, and (4) Number of shifts of small classification window K that depends on Δ. Hence, fine tuning is required before the algorithm can be used for recognition. The proposed window-based algorithm can be used to recognize activities or movements within a particular activity (once you know what the activity is). Since, activities have much broader time interval compared to movements, we could simplify the algorithm for activity recognition by fixing the values of some of the parameters used.

However, movement recognition involves shorter time scales and so optimal parameter values need to be identified using the tuning phase.

For tuning the system to the right values of $\mathcal{F}$, N, δ, and Δ, we place the main classification window at the exact location of movements in the training set and try to recognize the movements by changing values of $\mathcal{F}$, N, δ, and Δ. Hence, for tuning the system we feed the training data set itself and try to recognize the activities or movements with the training set. As a feedback metric to assess the optimality of these parameters we use *cumulative misclassification ratio*, the ratio of total misclassification in time of recognized activity or movements with the total time interval of the actual activity or movement. Once optimal values for $\mathcal{F}$, N, δ and Δ as indicated by starred values in the Fig. 3.2 are known, we find the optimal support vectors to recognize movements.

Chapter 4

Results and Discussion

This chapter outlines the performance of the window-based algorithm in recognizing activities or movements in a specific activity as well finding their starting and finishing instants. As we mentioned before, the algorithm can be used either for activity recognition or movement recognition, only one at a time or we can do it sequentially (first activity recognition then movement recognition). To recognize activities the SVM needs to be trained for the various types of activities that need to be recognized. However, if the focus is on movement recognition within an activity then the SVM needs to be trained for the various movement types within the concerned activity. We have organized this discussion as follows. First, we discuss how we collected acceleration and gyroscope values for the various activities and movements. Next, we present the performance of the algorithm when dealing with recognizing activities. Finally, we present the accuracy of the algorithm in recognizing movements (which is much more complex than recognizing activities as movements have shorter time interval). In addition, we also present a case study on the optimal values of each parameter in the window-based algorithm to improve the accuracy of movement recognition problem. To assess the performance of the algorithm for both activity and movement recognition over time we used an index called *Cumulative misclassification ratio*, which we define as the ratio of total misclassification (measured in time) at the current time with the current time.

4.1 Data Collection

We used Shimmer [4] motes for collecting linear acceleration (using accelerometers) and angular rate (using gyroscopes) values for the activities and movements we took for our study. Shimmer mote has a triaxial accelerometer MMA7260Q made by Freescale and is capable of sensing accelerations ranging from $\pm 1.5g$, $\pm 2g$ and $\pm 6g$ where $g = 9.8\text{m/s}^2$.

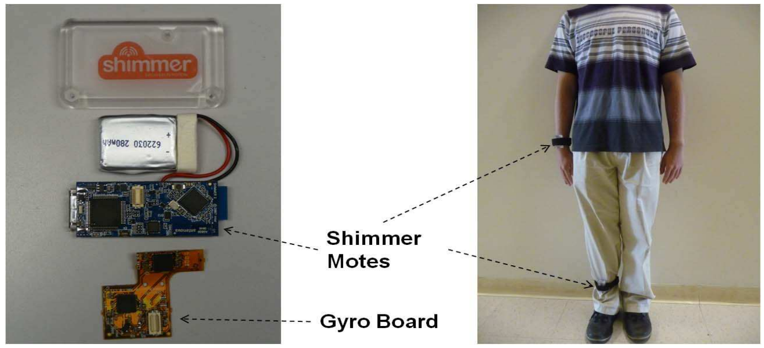

Figure 4.1: Shimmer mote and gyro board; Shimmer motes attached to a subject.

There is also a 3-axis gyroscope board having a full range of $\pm 500°/sec$. The motes were placed on the right arm wrist and right foot of the subject as shown in Fig. 4.1. Each of these motes gathered linear acceleration (from accelerometer) and angular rate(from gyroscope) values at 20 Hz and transmitted it wirelessly to a destination node connected to a desktop that aggregated all the samples and extracted the features. From the raw samples, we extracted features such as mean, standard deviation, maximum, peak-to-peak, root-mean-square and correlation between pair of accelerometer and gyroscope axes and also correlation between the values of the accelerometers. We collected acceleration and gyroscope samples for six different activities - “walking”, “standing”, “writing”, “smoking”, “jacks”and “jogging”. For recognizing movements we consider the movements within a specific activity, smoking. In smoking, there are basically two movements involved- “moving arm upward” for taking a cigarette puff and “moving arm down” after taking the puff.

4.2 Activity Recognition

Activity recognition is the initial step in our approach. To show the performance of the window-based algorithm in recognizing various activities, we considered six different activities. An activity is “static” when there is no movement involved while performing that activity. A “dynamic” activity, on the other hand, requires movement in order to accomplish the activity. The six different activities we considered are - “walking”, “standing”, “writing”, “smoking”, “jacks”and “jogging” out of which ‘standing” is “static” and rest all are “dynamic”. We asked three subjects to perform all the six activities in the order they preferred. To evaluate the performance of activity recognition we did three case studies - 1) Train separately for each subject and recognize activities of each subject, 2) Train for

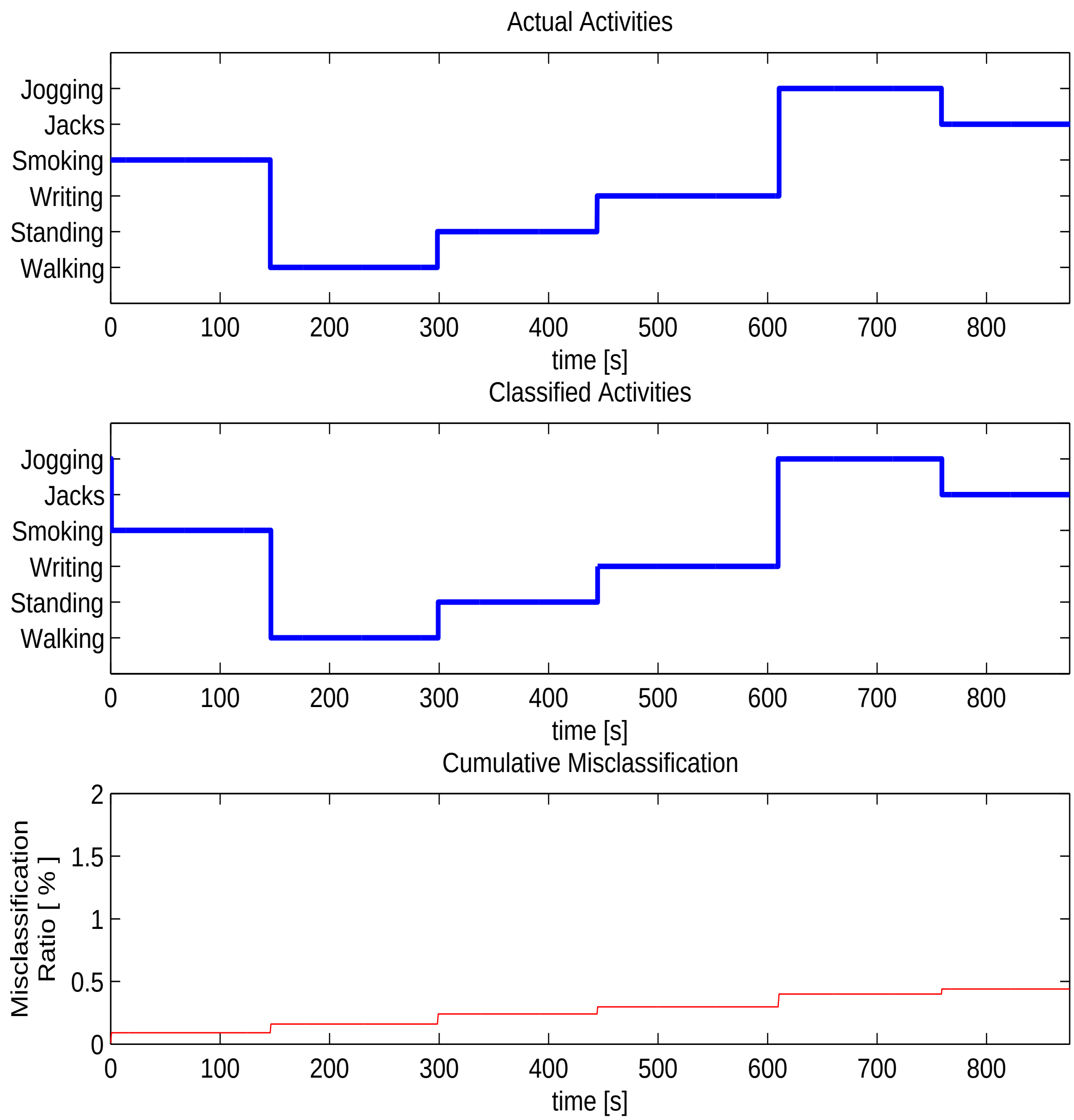

Figure 4.2: Case 1 : Actual and recognized activities of Subject 1.

activities of all subjects at one and recognize the activities of each subject and 3) Train for one subject and recognize the activities of another subject. We asked the three subjects (all male) to perform all the six activities at least 8 times. Each subject was asked to perform the activities in succession in the order he preferred.

Case 1 : Train separately for each subject and recognize activities of each subject - In this case study, we train the SVM separately for each subject for all the activities one

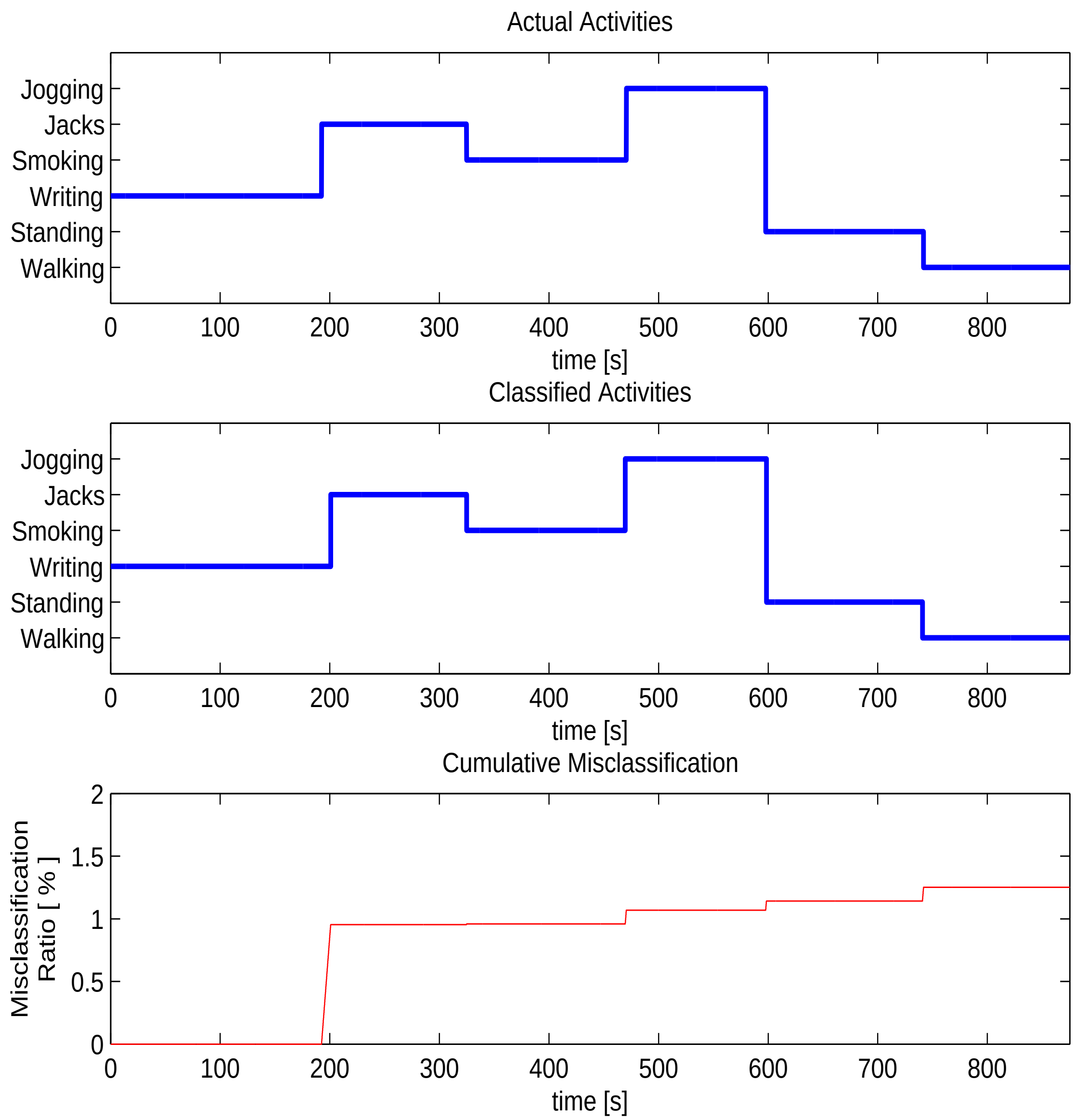

Figure 4.3: Case 1 : Actual and recognized activities of Subject 2.

at a time and then recognize the activities each of the subject performed. Out of the 8 observations for each activity 4 observations were used to train the SVM and the other 4 were used for testing the algorithm. Hence, in this case study, we trained the SVM for say "Subject 1" using first 4 observations and then tried to recognize the activities he performed during the other 4 observations collected. Similarly for the other 2 subjects. Figure 4.2 shows the actual activities and their time instants at which the "Subject 1"

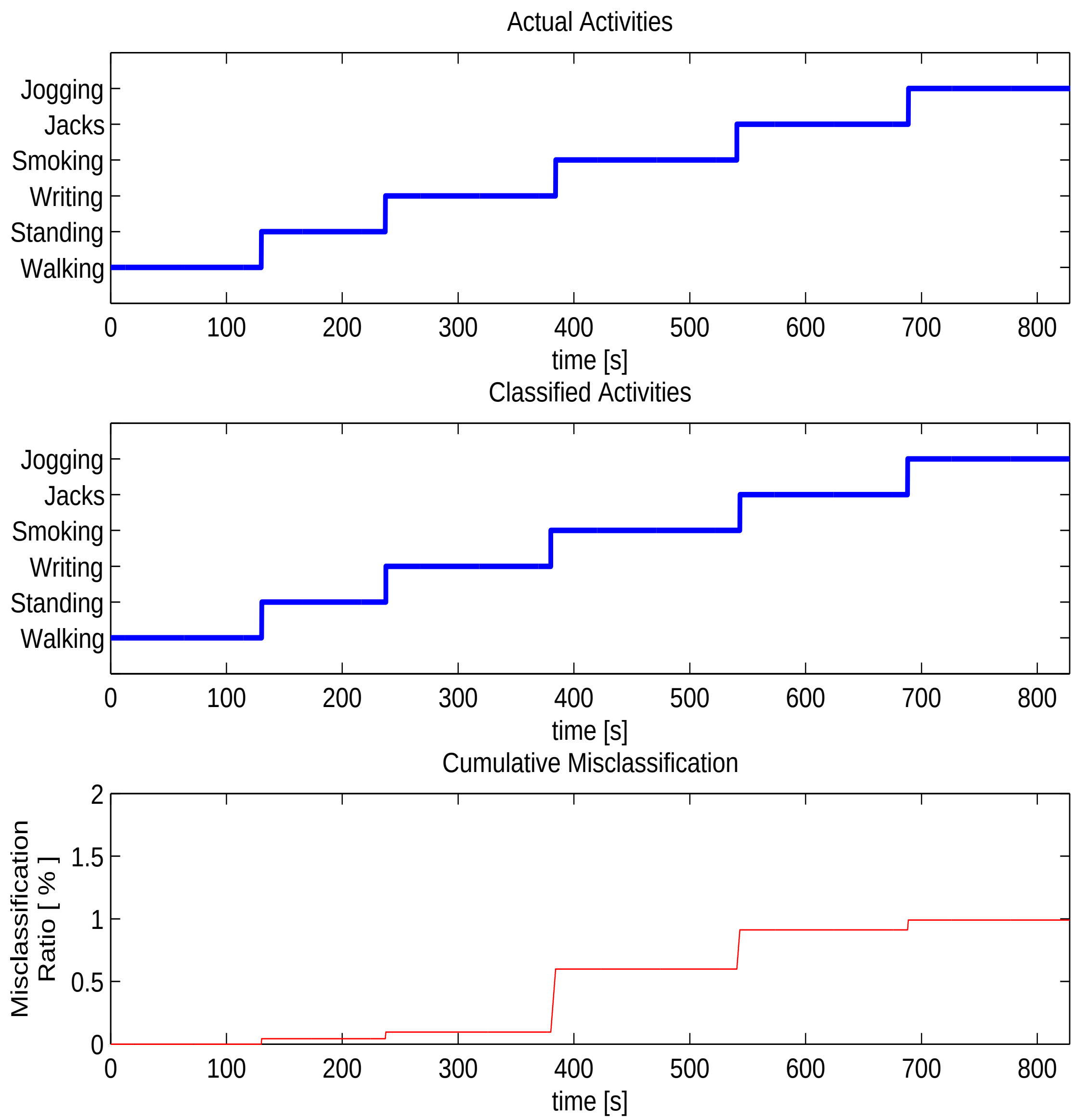

Figure 4.4: Case 1 : Actual and recognized activities of Subject 3.

performed the activities and also shows activities and their time instants recognized by the window-based algorithm when the SVM is trained separately for each subject. "Subject 1" performed six activities in the order - "Smoking", "Walking", "Standing", "Writing", "Jogging" and "Jacks" as shown in Fig. 4.2. The cumulative misclassification ratio for the activities performed by subject 1, is less than 1%. Similarly Fig. 4.3 and Fig. 4.4 shows the actual and recognized time instants of the activity performed by subject 2 and subject

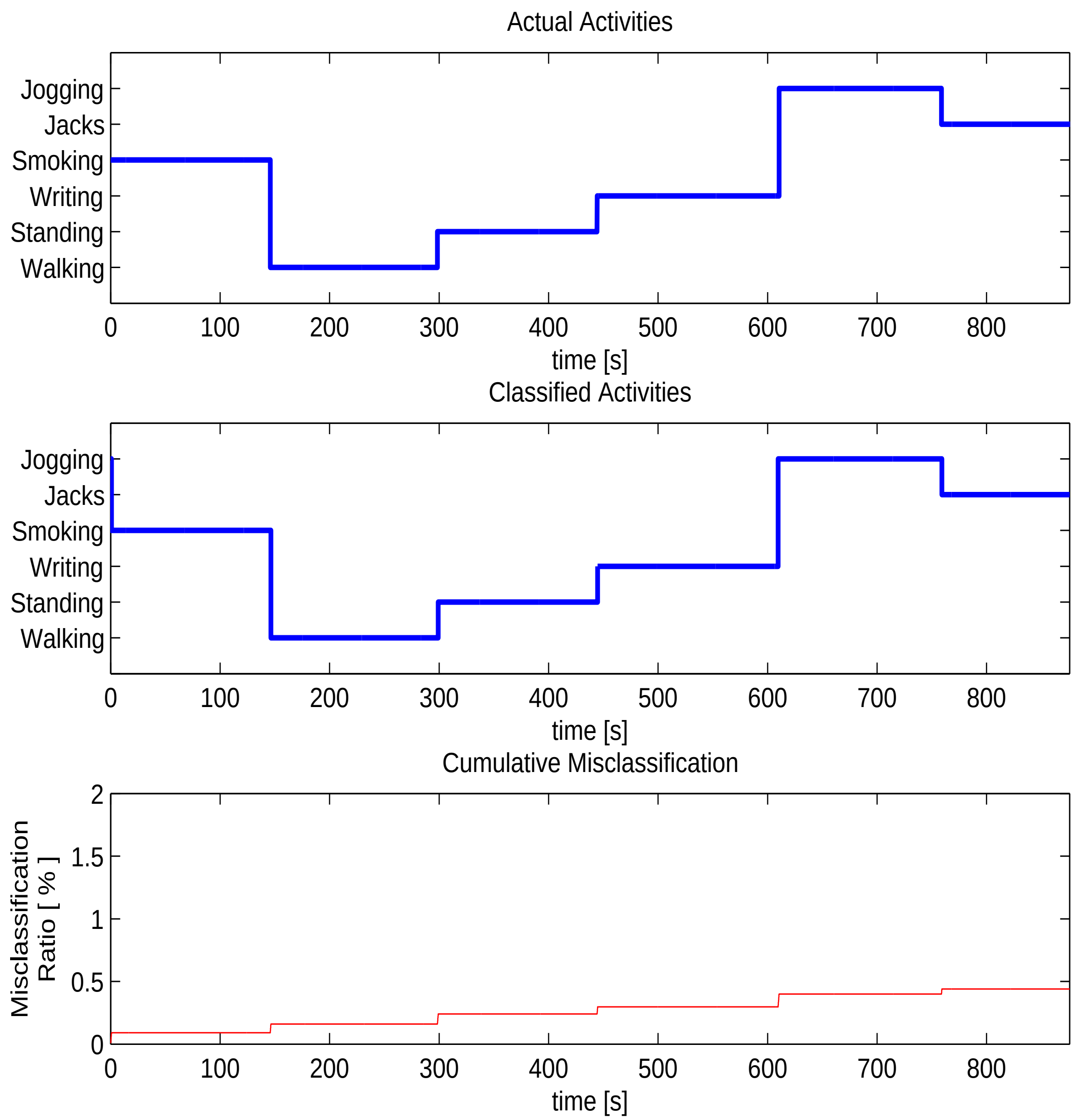

Figure 4.5: Case 2 : Actual and recognized activities of Subject 1.

3 respectively as well as the cumulative misclassification ratio. Subjects 2 and 3 performed the same six activities but in different order. The overall cumulative misclassification ratio is around 1.5%.

Case 2 : Train for activities of all subjects all at once and recognize the activities of each subject -

In this case study, we train the SVM for all subjects all at once and then recognize the

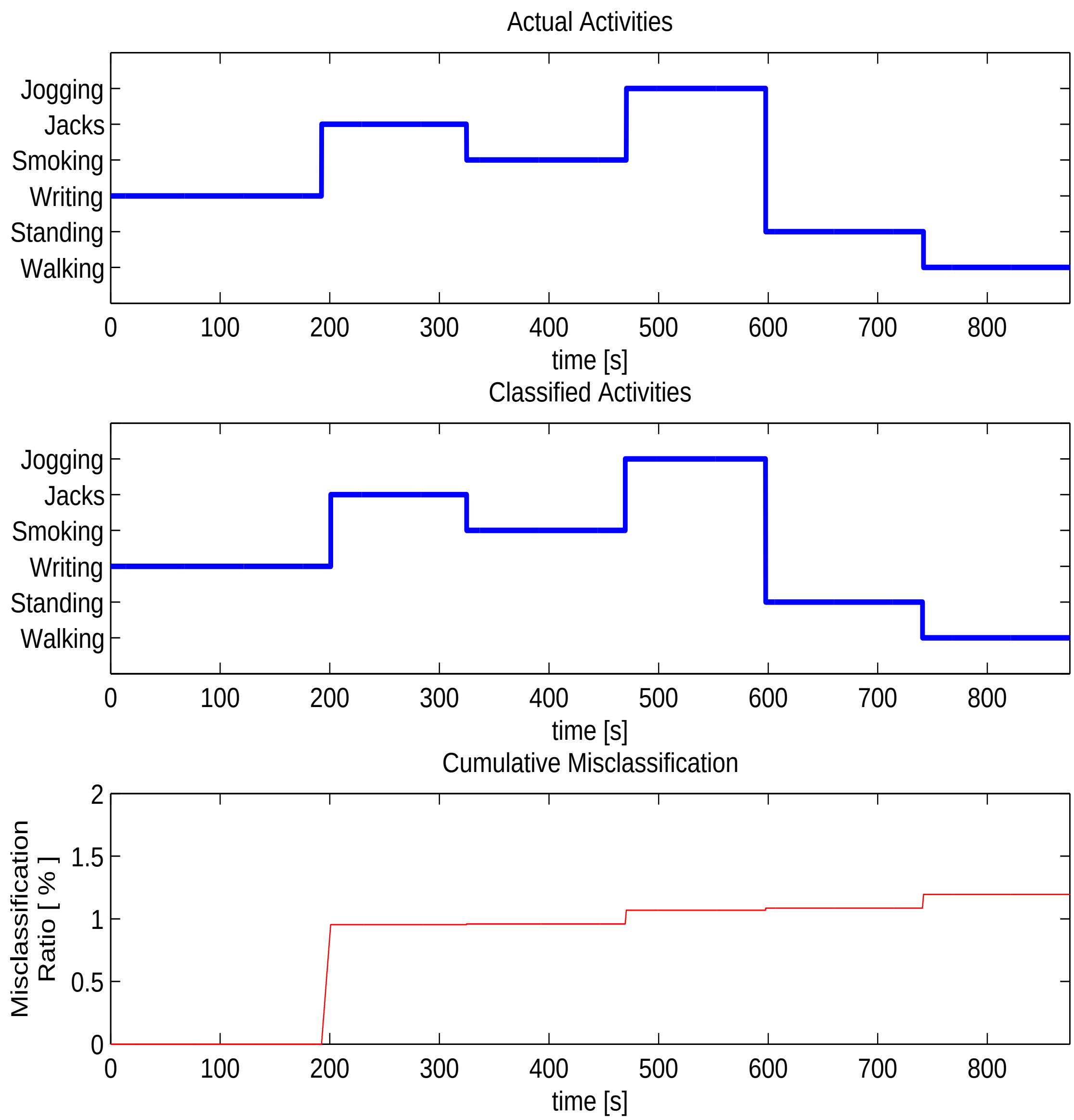

Figure 4.6: Case 2 : Actual and recognized activities of Subject 2.

activities performed by each subject along with the time instants at which those activities were performed. For this case study, we trained the SVM using the first 4 observations collected for each activity from all the subjects together and then we tried to recognize the activities the subjects performed during their next 4 observations. Figure 4.5 shows the actual activities and their time instants at which the "Subject 1" performed those and also the activities and their time instants recognized by the window-based algorithm

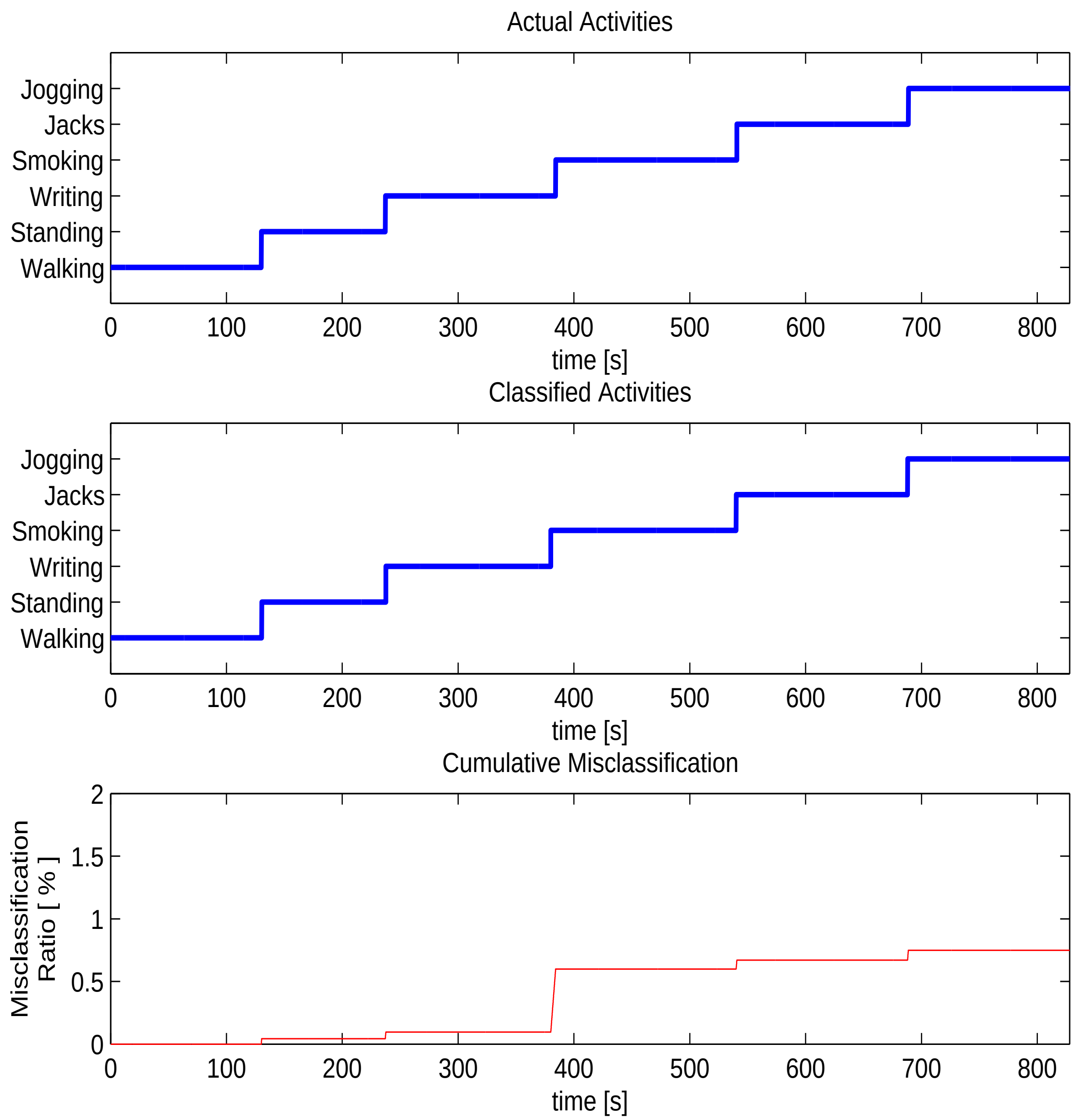

Figure 4.7: Case 2 : Actual and recognized activities of Subject 3.

when the SVM is trained for all the subjects all at once. The cumulative misclassification ratio over time is found to be less than 1%. Similarly Fig. 4.6 and Fig. 4.7 shows the actual and recognized time instants of the activities performed by "Subject 2" and "Subject 3" respectively as well as the cumulative misclassification ratio. The overall cumulative misclassification ratio is less than 1.5%.

Case 3 : Train for one subject and recognize the activities of another subject -

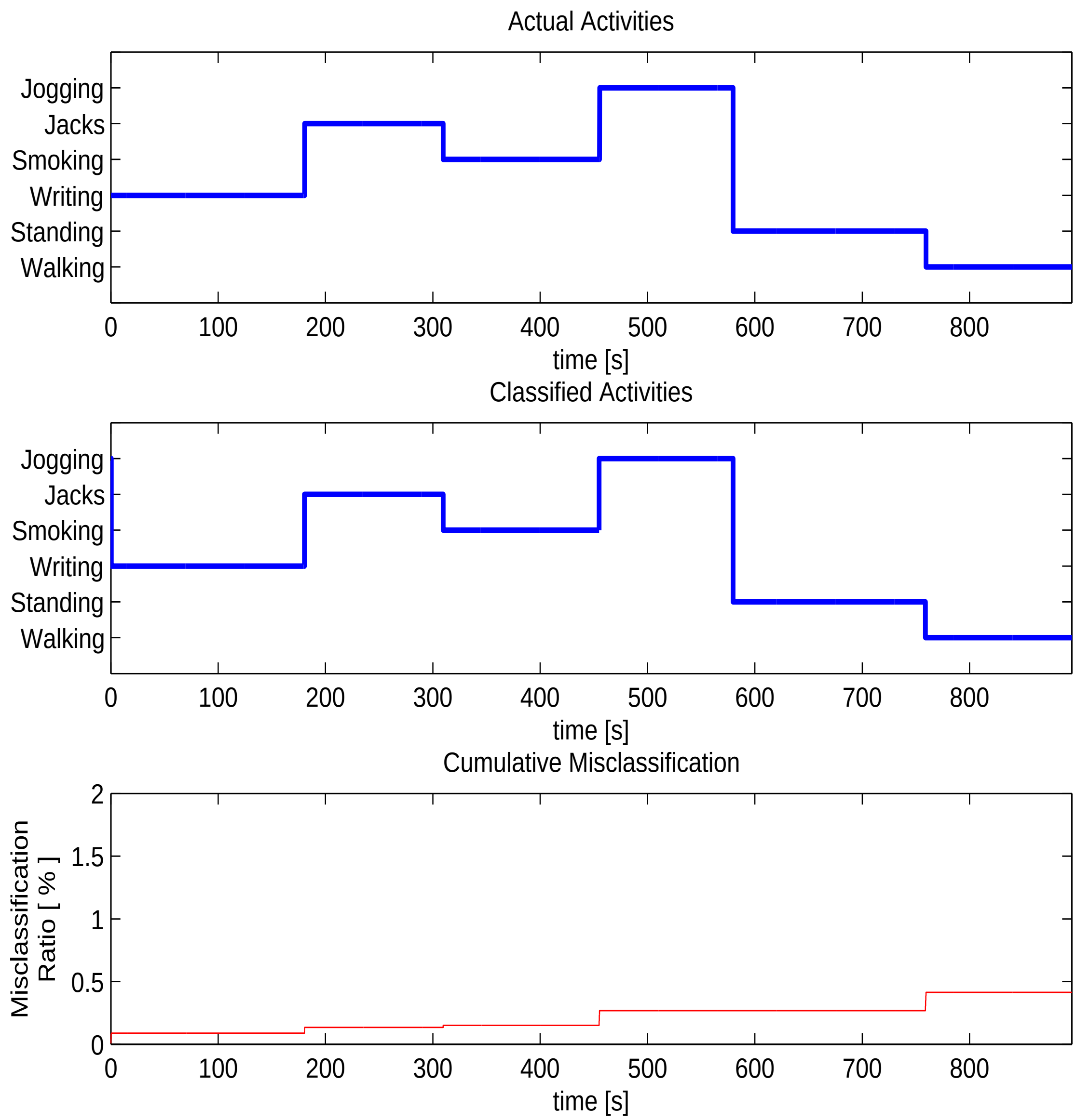

Figure 4.8: Case 3 : Actual and recognized activities of Subject 2 when SVM was trained for Subject 3.

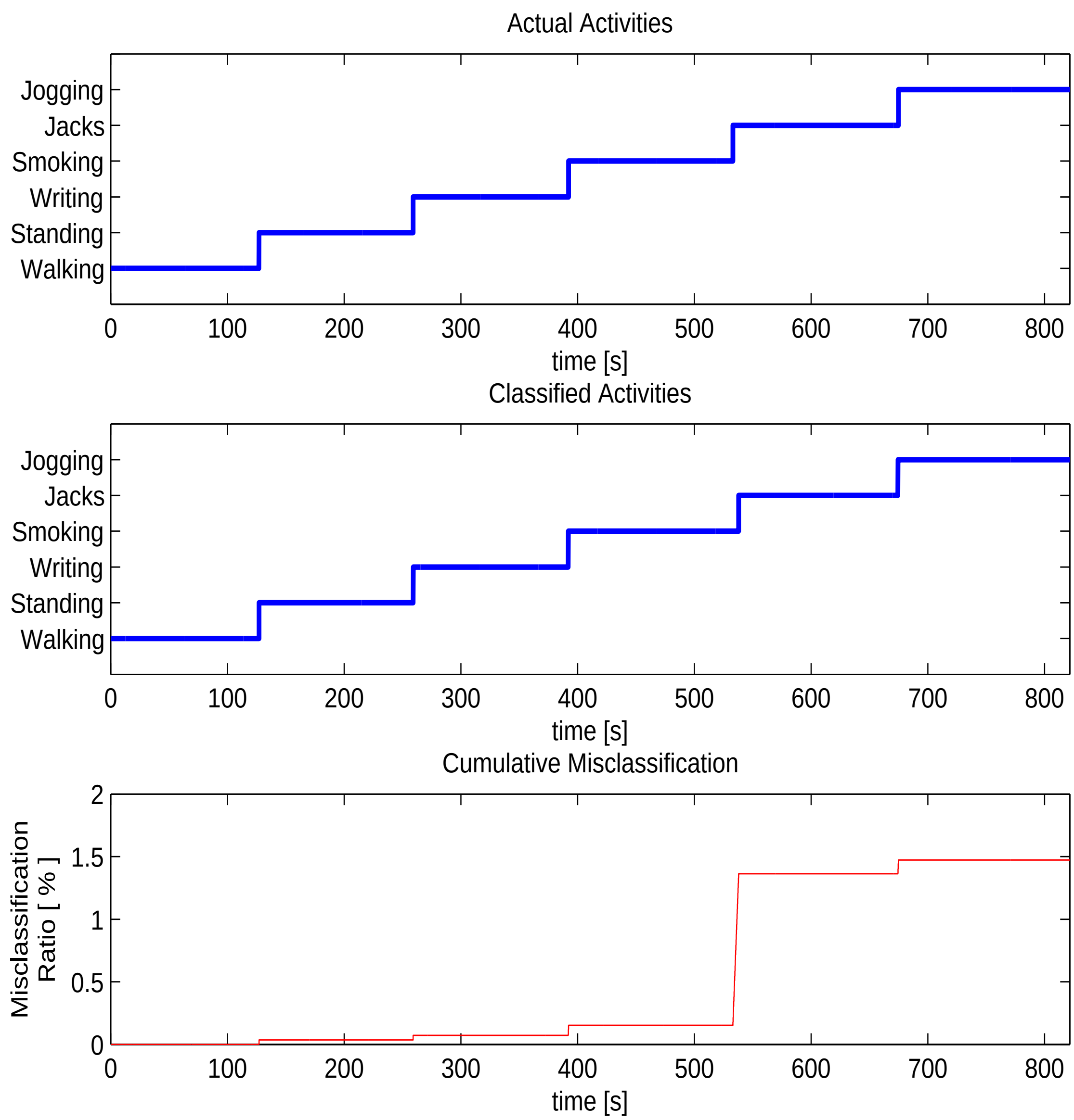

Figure 4.9: Case 3 : Actual and recognized activities of Subject 3 when SVM was trained for Subject 2.

In this case study, we train the SVM for one subject and then recognize the activities another subject performed. For this case study, we trained the SVM using all the 8 observations collected for one subject and then tried to recognize the activities performed for all the 8 observations of another subject. Figure 4.8 shows the actual activities and their time instants at which the “Subject 2” performed the activities and also the recognized activities along with time instants by the window-based algorithm when the SVM is trained for Subject 3. It is evident from the figure that the cumulative misclassification is around than 1.6% slightly higher than other two cases. Similarly Fig. 4.9 shows the actual activities and their time instants at which the “Subject 3” performed the activities and also the recognized activities and their time instants by the window-based algorithm when the SVM is trained for Subject 2. The overall cumulative misclassification is just more than 1.6%, which is just more than for Case 1 and Case 2 study.

Evaluation of the test results involved comparing the subjects actual activities with the recognized activities by the system. If the recognized activity actually occurred during the appropriate time interval, then this outcome was recorded as a correct recognition; if a particular movement produced an unexpected recognition, then this outcome was considered an incorrect recognition. In this way, the accuracy of the system in correctly recognizing the activities was measured. We show the test results for all the case studies done on the three subjects considered in a table called confusion matrix, which basically shows how accurate the window-based algorithm was in recognizing the type of activities and their time instants (see Table 4.1). It can be noticed that for Case 1 when the SVM was trained separately for each subject one by one, all the activities were recognized accurately. However for Case 2 and Case 3 there are some incorrect recognitions. In Case 2 when SVM was trained for all the subjects, there was one misclassified result for activities - “walking”, “smoking” and “jogging”. “Walking” was once incorrectly recognized as “Jogging” and also vice versa. This might have happened because of the change in pace of walking and jogging by the subjects. Similarly “Smoking” was incorrectly recognized as “Standing”. This incorrect classification may be because the subject was standing while he was smoking and there could have been a huge gap between the puffs he took while smoking the cigarette. In Case 3, when SVM is trained for one subject and activities of another subject was used for recognition, there are some incorrect recognitions just like Case 2. The overall accuracy of the system in recognizing the activities out of total 288 tests was found to be 97.2% and

Table 4.1: **Confusion Matrix**

Cases	**Actual Activity**	**No.Tests**	**Activity Recognized**						**No. Incorrect**	**Accuracy** [%]
			Walking	**Standing**	**Writing**	**Smoking**	**Jacks**	**Jogging**		
Case 1	Walking	12	12	0	0	0	0	0	0	100
	Standing	12	0	12	0	0	0	0	0	100
	Writing	12	0	0	12	0	0	0	0	100
	Smoking	12	0	0	0	12	0	0	0	100
	Jacks	12	0	0	0	0	12	0	0	100
	Jogging	12	0	0	0	0	0	12	0	100
Case 2	Walking	12	11	0	0	0	0	1	1	**91.6**
	Standing	12	0	12	0	0	0	0	0	100
	Writing	12	0	0	12	0	0	0	0	100
	Smoking	12	0	1	0	11	0	0	1	**91.6**
	Jacks	12	0	0	0	0	12	0	0	100
	Jogging	12	1	0	0	0	0	11	1	**91.6**
Case 3	Walking	24	22	0	0	0	0	2	2	**91.6**
	Standing	24	0	24	0	0	0	0	0	100
	Writing	24	0	0	24	0	0	0	0	100
	Smoking	24	0	2	0	22	0	0	2	**91.6**
	Jacks	24	0	0	0	0	24	0	0	100
	Jogging	24	1	0	0	0	0	23	1	**95.8**
Total		**288**							**8**	**97.2**

the lowest accuracy among all the case studies was found to be 91.6%.

4.3 Movement Recognition

The second part of our approach involves recognizing movements in an activity. First, we present a case study on the best feature set for classifying movements, and other parameters that impact the classification performance of the algorithm such as the number of small classification windows considered and the number of shifts of a small classification window within the main classification window. Next, using the best feature set and optimal values of the window-based algorithm we discuss the movement recognition accuracy of the algorithm. To recognize each movement type within an activity the SVM needs to be trained for each movement type. For movement recognition, we consider movements within the activity ”smoking” for several reason. Non-technical reason is that the movements in a smoking activity are pretty much evident and can be segregated. A smoking activity comprises of a set of movements such as the “arm moving up” for smoking followed by the “arm moving down” after taking the puff, which are repeated over time until a cigarette gets over.

As mentioned earlier, studying the smoking pattern of people helps in knowing how adversely smoking affects the health of a person over a period of time. Current methods of studying smoking behavior *in vivo*, include diary-based self-reports through electronic devices such as PDAs, cell phones and other devices, in addition to the traditional end of day surveys, interviews and paper and pencil diaries. However, these methods carry substantial deficiencies in validity and reliability of reports due to poor recall of frequent events and their timing, patient discomfort about reporting their actual level of smoking, and even there can be changes in actual smoking behavior of a person over a period of time. In such a case, a movement recognition system would help increase the reliability and validity of reports and provide a more accurate assessment of a persons condition. Beyond focusing on recognizing movements within the smoking activity, this entire approach can be used for recognizing movements within any activity. Movement recognition helps to extract some statistics of movements which could better help in understanding a person’s condition as time progresses or for better understanding of behavioral lifestyle. For example, by recognizing the starting and finish instant of the each of the movement in the smoking activity we can extract statistics like puff length, inter-puff length, number of puffs, etc., from the sequence of movements involved in the smoking activity. Getting the length of puff taken by the person over a duration of time helps to analyze the smoking pattern of a person and also get to know the amount of nicotine s/he would have inhaled during one puff.

Within an activity, we represent a movement logically by a line with length equal to the time interval for which the movement occurs. And to differentiate each movement in the results we put each movement type at a different height within the activity time interval. We consider a movement in an activity is misclassified if either the movement type recognized is wrong or even if the movement type recognized is correct but the recognized interval is less than 20% of the actual interval of the movement, which we call as *jitter*. As mentioned before, we used Cumulative misclassification ratio to evaluate the performance of the algorithm. In addition, for movement recognition, which is much more complex than activity recognition due to shorter time scales we also used another index *Moving Average misclassification ratio*, which we define as the ratio of misclassification of movements over a moving time window. For the Moving average misclassification ratio, we took a time window of interval 10% of the activity time and shifted it over the activity to calculate the misclassification over each window.

Once the activity is recognized as smoking, to recognize the movements within one specific activity we need to train the SVM separately for each movement type within that activity. Hence, we trained the SVM with the two movement types in the smoking activity. For training we took a set of 8 observations of smoking activity and we tried to recognize the movements within another set of 8 smoking activities. We applied the same window-based algorithm but certainly tuned for movement recognition. We consider around 3-4 minutes as the time to smoke one cigarette by a person and for movement recognition we considered features such as mean, maximum, standard deviation, peak-to-peak, root mean square of the acceleration values on each of the three axes as well as the correlation of acceleration values between pair of axes from each sub-intervals of the movements.

4.3.1 Best Feature Set

For identifying the best feature set for recognizing movements within a specific activity, we considered in total six features. The best set of features for recognizing movements, depends on the type of activity. In $\mathcal{F}$ each feature is represented by a bit and features are considered

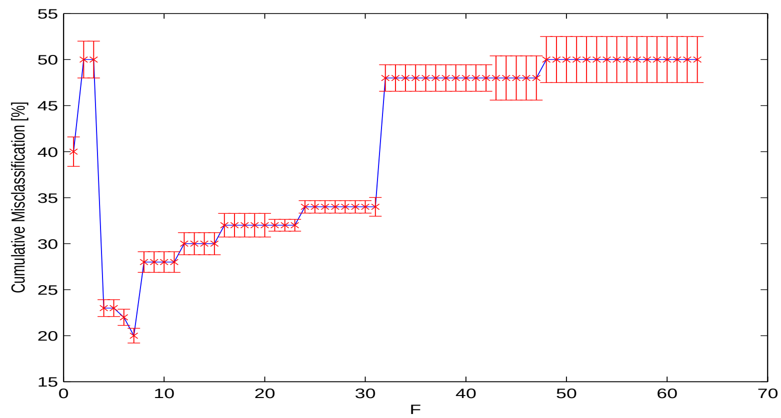

Figure 4.10: Cumulative misclassification Vs $\mathcal{F}$.

in the order - mean, maximum, standard deviation, peak-peak, RMS and then correlation. If a binary element in $\mathcal{F}$ is 0, it means that particular feature is excluded, otherwise it is included. So, if $\mathcal{F} = [000110]$ it means only the features peak-peak, RMS of acceleration and gyroscope values. We used cumulative misclassification ratio for assessing the best feature set and we tried to recongize the movements in the training set itself. We have total 63 combinations for $\mathcal{F}$ as we took six features. We took the values of N, δ and Δ as 8, 3 and 2, respectively, for each value of $\mathcal{F}$. For finding the optimal $\mathcal{F}$, we placed the main classification window at position of occurrence of each of the movements in the training set and classified each of those movements. The optimal $\mathcal{F}$ would be the one that gives less cumulative misclassification. Figure 4.10 shows the cumulative misclassification ratio for various $\mathcal{F}$ values from which we can say that the optimal value of $\mathcal{F}$ is 7. Hence, the best set of features for classifying movements in a smoking activity includes RMS, Peak-Peak and Correlation of acceleration and gyroscope values.

4.3.2 Optimal Parameters

The performance of the classification algorithm also depends on the N sub-intervals considered within a window for extract features, the number of small classification windows, H considered within the main classification window that in turn depends on the δ value taken and the number of shifts of small classification windows, K within the main classification window that in turn depends on the Δ value taken. To find the optimal value of N, we took 16 different N values with $\delta = 3$ and $\Delta = 2$ and $\mathcal{F} = 7$, the optimal value of $\mathcal{F}$. Figure

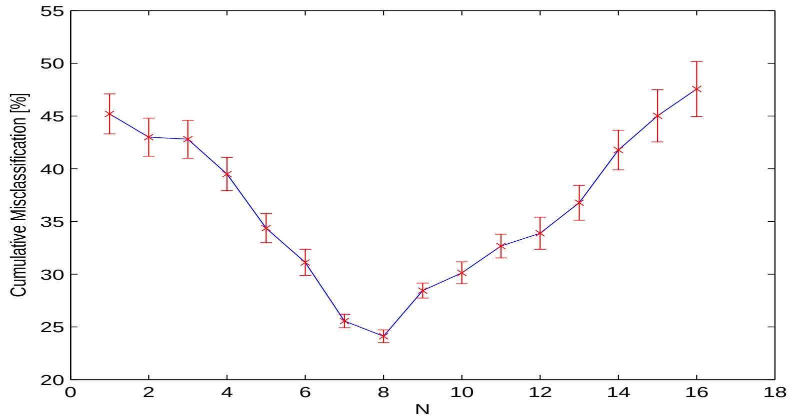

Figure 4.11: Cumulative misclassification Vs N @ $\delta = 3$, $\Delta = 2$.

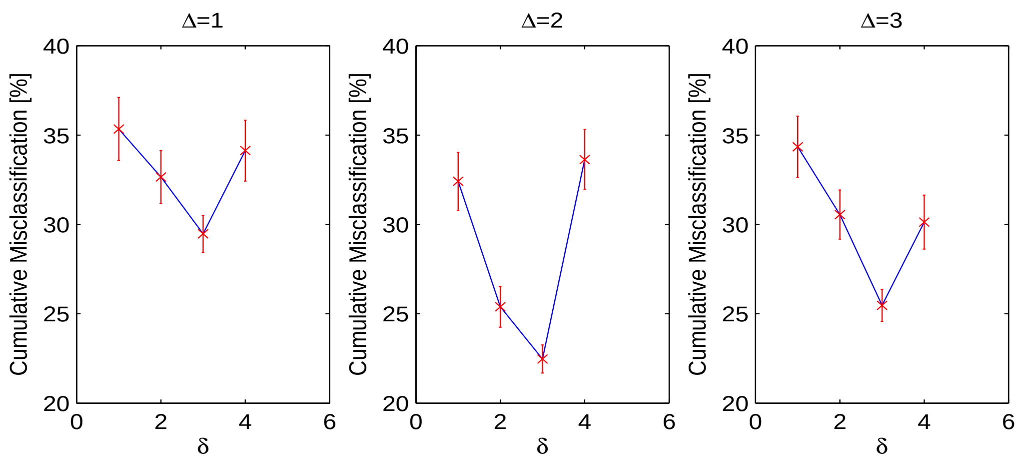

Figure 4.12: Cumulative misclassification Vs δ for variable Δ.

4.11 shows the cumulative misclassification ratio for various values of N from, which we can say that optimal value of N is around 8. Figure 4.11 shows that if N is too low, the misclassification is more and also if N is too high as it would be like taking raw acceleration values of the movements. It is also essential to find the optimal values of δ and Δ. Figure 4.12 shows cumulative misclassification ratio Vs δ for various value of Δ, which indicates that the optimal values of δ and Δ are 3 and 2, respectively, as for these values cumulative misclassification is the lowest.

4.3.3 Movement Recognition Accuracy

Here we show the accuracy of the algorithm in recognizing movements within the smoking activity. For recognizing the starting and finishing instants of the movements we took for

$\mathcal{F} = 7$, $N = 4$ and $\delta = 3$ and $\Delta = 2$, which were identified as optimal values. Figure 4.13 show the actual and classified positions of both the movements involved in a smoking activity where "Movement 1" refers to the arm moving up for taking the puff while, "Movement 2" refers to the arm moving down after smoking. Different movement types are shown at different levels in Fig. 4.13.

In Fig. 4.13 "Actual" refers to the actual position of the movements in the activity, whereas "Classified" refers to the time instants recognized by the algorithm. Figure 4.13 show both the Cumulative misclassification ratio and Moving Average misclassification ratio. Figure 4.14 also shows the recognized movements within another data set of the smoking activity. The results show that cumulative misclassification is around 20%.Considering the fact we did not use any a priori knowledge on the order of occurrence of the movements in the activity the misclassification rate is quite acceptable and also we consider the case of jitter in our misclassification calculation even if the movement is correctly classified. For example in smoking activity, "Movement 2" always follows "Movement 1"and so the order of occurrence of the two type of movements involved in the smoking activity can be used to improve movement recognition. We showed movement recognition results of two smoking activity to indicate the fact that for movement recognition the best and worst cases are almost the same as for movements we are dealing with very shorter time scales.

Misclassification for movement recognition, which is around 20% is much greater than for the activity recognition, which is around $5-9\%$ because of several reasons 1) in movement recognition we are dealing with very shorter time scales and 2) there multiple movements within an activity giving way to multiple transitions and also these transitions are not negligible because of shorter time scales of the individual movements. Although we show the performance of the algorithm for only six different activities and movements within "smoking activity" the approach can be generalized for any number and type of activities and movements. Moreover, even if some of the activities which have considered are similar the algorithm shows considerable accuracy in recognizing all those activities.

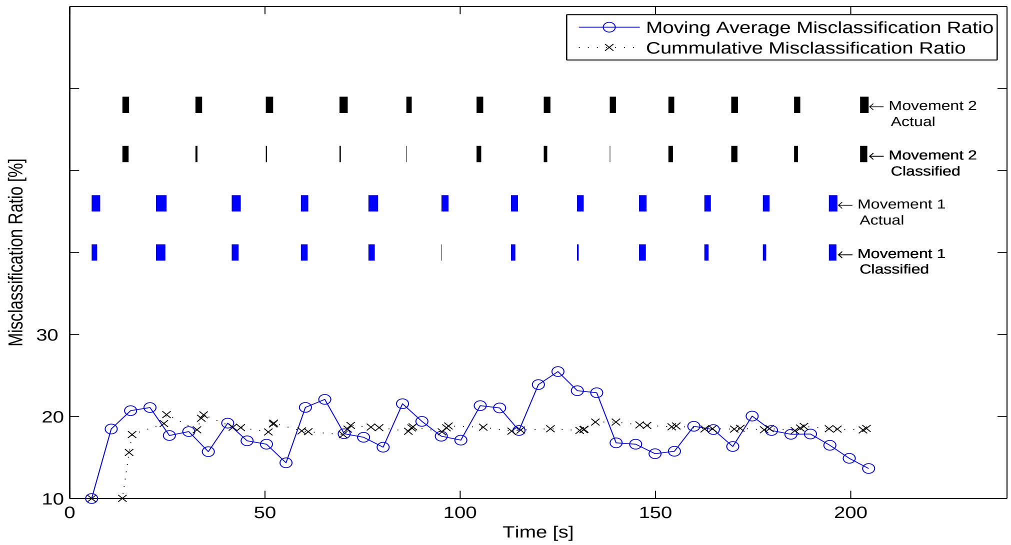

Figure 4.13: Movement recognition in Smoking Activity data set 1.

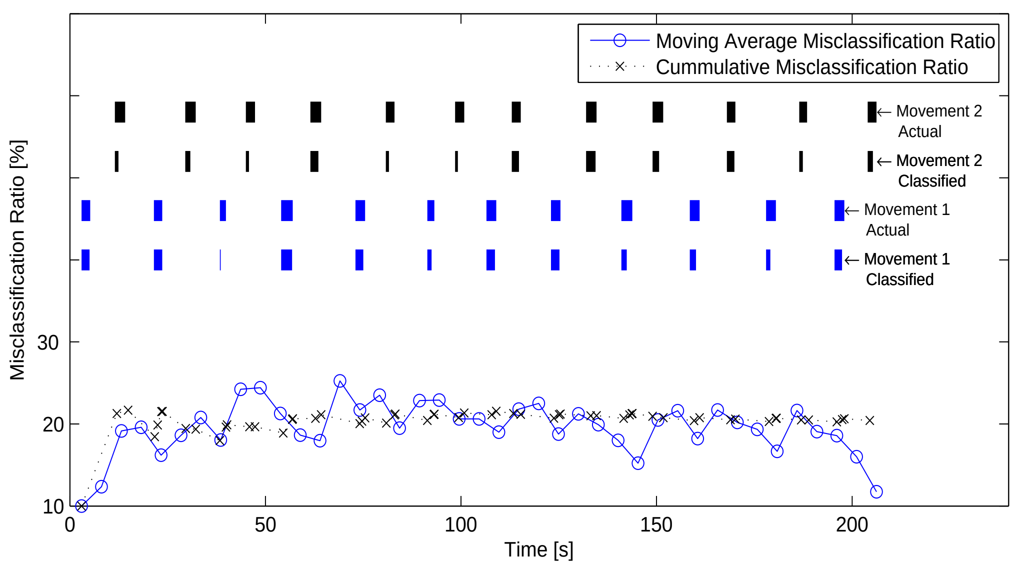

Figure 4.14: Movement recognition in Smoking Activity data set 2.

Chapter 5
Conclusion

In summary, we proposed a window-based algorithm that can be tuned to recognize on the fly either various activities or movements in a specific activity along with identifying their starting and finishing instants. The thesis also identifies the best set of features and optimal parameters values of the algorithm for improving the accuracy of movement recognition. The results show the accuracy of the algorithm for activity recognition is around 91% and around 80% for movement recognition. In addition, this work shows that using a machine learning based approach, accelerometers and gyroscope can be used to recognize a variety of daily-life activities and movements within an activity for context-aware computing.

5.1 Suggestions for Future Work

As future work, we will further optimize the algorithm for movement recognition so as to improve the movement recognition accuracy and also we will train the system for various other daily-life activities. In addition, we will use feedback from the movement recognition phase to the activity recognition phase to further improve the accuracy of activity recognition. Our system can be easily extended to capture context information, e.g., by i) collect vital signs such as ECG of the subject and correlate with recognized physical activity to give a day-to-day report of subject's health for health risk analysis and promoting healthy life style, ii) by sensing the environment (temperature, humidity, etc.), and iii) by performing localization (using GPS or communicating with fixed sensors that play the role of "anchors").

References

[1] Apple iPhone. http://www.apple.com/iphone.

[2] Mathworks, MATLAB Help. http://www.mathworks.com.

[3] Nintendo Wii. http://www.nintendo.com/wii.

[4] Shimmer. http://shimmer-research.com.

[5] Spider. http://www.kyb.mpg.de/bs/people/spider.

[6] Gregory D. Abowd, Anind K. Dey, Peter J. Brown, Nigel Davies, Mark Smith, and Pete Steggles. Towards a better understanding of context and context-awareness. In *Proc. of the International symposium on Handheld and Ubiquitous Computing*, Karlsruhe, Germany, September 1999.

[7] F. Albinali, M. S. Goodwin, and S. S. Intille. Recognizing stereotypical motor movements in the laboratory and classroom: A case study with children on the autism spectrum. In *Proc. of International Conference on Ubiquitous Computing*, Orlando, Florida, October 2009.

[8] U. Anliker, J.A. Ward, P. Lukowicz, G. Trster, F. Dolveck, M. Baer, F. Keita, E.B. Schenker, F. Catarsi, L. Coluccini, A. Belardinelli, D. Shklarski, M. Alon, E. Hirt, R. Schmid, and M. Vuskovic. Amon: a wearable multiparameter medical monitoring and alert system. *IEEE Transactions on Information Technology in Biomedicine*, 8(4):415–427, November 2004.

[9] L. Bao and S. S. Intille. Activity Recognition from User-annotated Acceleration Data. In *Proc. of Pervasive Computing (PERVASIVE)*, Vienna, Austria, April 2004.

[10] J. Barnes and R. Jafari. Locomotion Monitoring Using Body Sensor Networks. In *Proc. of Pervasive Technologies Related to Assistive Environments (PETRA)*, Athens, Greece, June 2008.

[11] M. Barry, J. Gutknecht, I. Kulka, P. Lukowicz, and T. Stricker. From motion to emotion: a wearable system for the multimedial enhancement of a butoh dance performance. *Journal of Mobile Multimedia*, 1(2):112–132, June 2005.

[12] Jennifer S. Beaudin, Stephen S. Intille, Emmanuel Munguia Tapia, Y Rockinson, and Margaret E. Morris. Context-sensitive microlearning of foreign language vocabulary on a mobile device, 2007.

[13] AK Bourke, JV OBrien, and GM Lyons. Evaluation of a thresholdbased tri-axial accelerometer fall detection algorithm. *Gait & Posture*, 26(2):194–199, January 2007.

[14] T. Choudhury, M. Philipose, D. Wyatt, and J. Lester. Towards Activity Databases: Using Sensors and Statistical Models to Summarize Peoples Lives. *IEEE Data Eng. Bull*, 29(1):49–58, March 2006.

[15] J. Christopher and C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, June 1998.

[16] Anind K Dey. Understanding and using context. *Personal Ubiquitous Comput.*, 5(1):4–7, February 2001.

[17] L. Feng, Z. Yueting, W. Fei, and P. Yunhe. 3D motion retrieval with motion index tree. *Comput. Vis. Image Underst.*, 92(2):265–284, 2003.

[18] H. Ghasemzadeh, J. Barnes, E. Guenterberg, and R. Jafari. A Phonological Expression for Physical Movement Monitoring in Body Sensor Networks. In *Proc. of Mobile Ad-hoc and Sensor Systems (MASS)*, Atlanta, GA, September 2008.

[19] A. Golding and N. Lesh. Indoor Navigation Using a Diverse Set of Cheap Wearable Sensors. In *Proc. of International Symposium on Wearable Computers (ISWC)*, San Francisco, CA, October 1999.

[20] J. Healey and R. W. Picard. Startlecam: A cybernetic wearable camera. In *Proc. of International Symposium on Wearable Computers (ISWC)*, Pittsburgh, Pennsylvania, October 1998.

[21] E. A. Heinz, K-S. Kunze, S. S. H. Junker, P. Lukowicz, and G. Troster. Experimental Evaluation of Variations in Primary Features Used for Accelerometric Context Recognition. In *Proc. of European Symposium on Ambient Intelligence (EUSAI)*, Eindhoven, Netherlands, November 2003.

[22] Duy Tam Gilles Huynh. *Human Activity Recognition with Wearable Sensors*. PhD thesis, TU Darmstadt, 2008.

[23] R. Jafari, W. Li, R. Bajcsy, S. Glaser, and S. Sastry. Physical Activity Monitoring for Assisted Living at Home. In *Proc. of International Workshop on Wearable and Implantable Body Sensor Networks (BSN)*, Germany, March 2007.

[24] Vineet K. Jolly. *Activity Recogniton using Singular Value Decomposition*. PhD thesis, Virginia Polytechnic Institute and State University, 2006.

[25] N. Kern, B. Schiele, and A. Schmidt. Multi-sensor Activity Context Detection for Wearable Computing. In *Proc. of European Symposium on Ambient Intelligence (EUSAI)*, Eindhoven, Netherlands, November 2003.

[26] Kristof Van Laerhoven. Medical healthcare monitoring with wearable and implantable sensors. In *Proc. of International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications*, September 2004.

[27] J. Lester, T. Choudhury, and G. Borriello. A practical approach to recognizing physical activity. In *Proc. of Pervasive Computing (PERVASIVE)*, Dublin, Ireland, May 2006.

[28] L. Liao, D. Fox, and H. Kautz. Location-based activity recognition. In *Proc. of Neural Information Processing Systems (NIPS)*, British Columbia, Canada, December 2005.

[29] K. Liszka, Michael A. Mackin, Michael J. Lichter, David W. York, Dilip Pillai, and David S. Rosenbaum. Keeping a Beat on the Heart. *IEEE Pervasive Computing*, 3(4):42–49, October 2004.

[30] B. Logan, J. Healey, M. Philipose, E. Tapia, and S. Intille. A Long-Term Evaluation of Sensing Modalities for Activity Recognition. In *Proc. of Ubiquitous Computing (UbiComp)*, Innsbruck, Austria, September 2007.

[31] P. Lukowicz, A. Timm-Giel, M. Lawo, and O. Herzog. Wearit@work: Toward real world industrial wearable computing. *IEEE Pervasive Computing*, 6(4):8–13, October 2007.

[32] D. Minnen, T.Westeyn, D. Ashbrook, P. Presti, and T. Starner. Recognizing Soldier Activities in the Field. In *Proc. of International Workshop on Wearable and Implantable Body Sensor Networks (BSN)*, Germany, March 2007.

[33] M. Muller, T. Rder, and M. Clausen. Efficient content-based retrieval of motion capture data. *ACM Trans. Graph.*, 24(3):677–685, 2005.

[34] R. Paradiso, G. Loriga, and N. Taccini. A wearable health care system based on knitted integrated sensors. *IEEE Transactions, Information Technology in Biomedicine*, 9(3):337–344, September 2005.

[35] C. Randell and H. Muller. Context awareness by analysing accelerometer data. In *Proc. of International Symposium on Wearable Computers (ISWC)*, Atlanta, GA, October 2000.

[36] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman. Activity Recognition from Accelerometer Data. In *Proc. of American Association for Artificial Intelligence (AAAI)*, Pittsburgh, Pennsylvania, November 2005.

[37] B.J. Rhodes. The wearable remembrance agent: a system for augmented memory. In *International Symposium on Digest of Papers Wearable Computers*, Cambridge, MA, October 1997.

[38] Matthias C. Sala, Kurt Partridge, Linda Jacobson, and James Begole. An exploration into activity-informed physical advertising using pest. *Pervasive*, 4480:73–90, May 2007.

[39] A. Schmidt, M. Beigl, and H. W. Gellersen. There is more to context than location. *Computer and Graphics Journal, Elsevier*, 23(6):893–902, December 1999.

[40] T. Stiefmeier, D. Roggen, G. Ogris, P. Lukowicz, and G. Trster. Wearable Activity Tracking in Car Manufacturing. *IEEE Pervasive Computing*, 7(2):42–50, October 2008.

[41] Tim Van Kasteren, Athanasios Noulas, Gwenn Englebienne, and Kr˙ Accurate activity recognition in a home setting. In *Proc. of International Conference on Ubiquitous Computing*, Seoul, South Korea, October 2008.

[42] K. VanLaerhoven and O. Cakmakci. What shall we teach our pants? In *Proc. of International Symposium on Wearable Computers (ISWC)*, Atlanta, GA, October 2000.

[43] K. VanLaerhoven, H. Gellersen, and A. Schmidt. Multi-sensor context aware clothing. In *Proc. of International Symposium on Wearable Computers (ISWC)*, Seattle, WA, October 2002.

[44] E. Villalba, M. Ottaviano, M.T. Arredondo, A. Martinez, , and S. Guillen. Wearable monitoring system for heart failure assessment in a mobile environment. *Computers in Cardiology*, pages 237–240, 2006.

[45] R. Want, B.N. Schilit, N.I. Adams, R. Gold, K. Petersen, D. Goldberg, J.R. Ellis, and M. Weiser. An overview of the parctab ubiquitous computing experiment. *Personal Communications, IEEE*, 2(6):28 –43, December 1995.

[46] Roy Want, Andy Hopper, Veronica Falcao, and Jonathan Gibbons. The active badge location system. *ACM Trans. Inf. Syst.*, 10(1):91–102, January 1992.

[47] J. A. Ward, Lukowicz, Trster P., and Starner T. G. Activity recognition of assembly tasks using body-worn microphones and accelerometers. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 28(10):1553–1567, October 2006.