Assessment and Classification of Movements in Bed Using Unobtrusive Sensors

Adriana Miorelli Adami
B. Sc. Mathematics Education, University of Caxias do Sul, 1995
M. Sc. Applied Mathematics, Federal University of Rio Grande do Sul, 1999

A dissertation submitted to the faculty of the
OGI School of Science & Engineering at
Oregon Health & Science University
in partial fulfillment of the
requirements for the degree
Doctor of Philosophy
in
Electrical Engineering

August 2006

The dissertation "Assessment and Classification of Movements in Bed Using Unobtrusive Sensors" by Adriana Miorelli Adami has been examined and approved by the following Examination Committee:

Dr. Misha Pavel
Professor
Thesis Research Advisor

Dr. Tamara Hayes
Assistant Professor
Thesis Research Advisor

Dr. Tran Thong
Assistant Professor
Department of Biomedical Engineering

Dr. Xubo Song
Assistant Professor
Department of Computer Science and Electrical Engineering

Dr. Clifford Singer
Associate Professor of Psychiatry
University of Vermont

Dedication

To my husband André,

my parents Ildo and Ivone, and

my sisters Nêmora and Ângela

Acknowledgments

Many people have contributed to the success of this work. I would like to thank my advisor, Dr. Misha Pavel, for giving me the opportunity to work on my doctorate under his guidance and for his support and advice. I also would like to thank my co-advisor, Dr. Tamara Hayes, for her guidance and encouragement during the last three years. I thank Dr. Xubo Song, Dr. Tran Thong and Dr. Clifford Singer for their assistance and direction as members of my doctoral committee.

I would like to thank the people at the Point of Care Laboratory for their help. In particular, I would like to thank Pavel Chytil for his help with the Elite Care study and John Hunt for sharing his knowledge of hardware.

This work would not be possible without all the subjects that participated in the studies. I would like to thank all subjects from OGI and from the community for their help and patience. I also would like to thank the residents of Elite Care that participated in the study. Special thanks to the administration and staff of Elite Care for allowing us to conduct a study at their facility, and for helping us with recruitment and technical issues. I would like to thank Dr. Clifford Singer and Eric Colling for their help with the Elite Care study and also for sharing their knowledge about sleep and sleep research.

I would like to express my gratitude to Bill Roberts, Melanie Erskine, Alena Tkacova, Mark Bouza, Pavel Chytil and Chelea Holdt from the Department of Biomedical Engineering, to Amy Johnson, Brenda Donin and Alison Roache Jones from the Department of Graduate Education, and to Carol Resco, Mary Hultine, Kristine Roley and Kathleen Steward from the Library for making my experience at OGI more memorable.

I would like to acknowledge the unconditional support, patience and love from my parents Ildo and Ivone Miorelli and from my sisters Nêmora and Ângela Miorelli. I thank my colleagues at University of Caxias do Sul for their support and friendship. Special thanks to my friends Pam, Clara, Circe, Magda, Tania, Adriana, Elaine, Monica, Rejane, Lunalva, Carmem and Wlamir for their constant support through these years. Last, but not the least, I would like to express my deepest gratitude to my husband André, for his love, encouragement, and help.

Finally, I would like to thank Intel Corporation and the Oregon Opportunity for providing financial support for my research.

Contents

List of Tables

List of Figures

Abstract

Assessment and Classification of Movements
in Bed Using Unobtrusive Sensors

Adriana Miorelli Adami, B.Sc., M.Sc.

Ph.D., OGI School of Science & Engineering
at Oregon Health & Science University

August 2006

Thesis Advisors: Dr. Misha Pavel and Dr. Tamara Hayes

The quality of sleep is an important attribute of an individual's health state and its assessment is therefore a useful diagnostic feature. Changes in the sleep-related behaviors as well as in the patterns of motor activities during sleep can be a disease marker, or can reflect various abnormal physiological and neurological conditions. Presently, there are no convenient, unobtrusive ways to assess and quantify the quality of sleep at point of care outside of a clinic.

This dissertation describes an approach and a system for unobtrusive assessment of activity patterns and movement in bed that uses load cells installed at the corners of a bed. The system focuses on identifying when a movement occurs and on determining the type of movement performed based on the forces sensed by the load cells. The feasibility and accuracy of the movement detection and classification is evaluated using data

collected in the laboratory and in a study with residents of an assisted-living facility (Elite Care, Milwaukie, OR).

The movement detection approach estimates the energy in each load cell signal over short segments to capture the variations caused by movement. The average equal error rate of the detector is 3.22% (± 0.54). The performance of the detector is invariant with respect to the individual's characteristics, e.g., weight, as well as those of the bed.

The dissertation describes several approaches to signal representation and discrimination techniques for clinically relevant classification of the type of movements with the goal of weight-invariant performance. The results of correct classification for an approach based on Gaussian Mixture Models using a time-domain representation and a wavelet-based time-frequency representation, as evaluated by laboratory experiments, are 84.6% and 82.2%, respectively. The simplicity of the resulting algorithms, the relative insensitivity to the weight and height of the monitored individual and the minimal training requirements make the resulting approaches practical and easily deployable in residential and clinical settings.

Chapter 1

Introduction

Quality of sleep is an important factor for a person's physical and emotional well being. The quality of our life is tied to the quality of sleep. People with sleep deficits may experience impaired daytime performance, irritability, and lack of concentration. Moreover, disruptions in sleep patterns adversely affect our resistance to diseases because they produce a reduction of the natural immune responses [1]. Therefore, how we sleep at night affects how we perform during the day and, unfortunately, people are not always aware of the consequences of sleep disruptions [2].

The clinical assessment of sleep and of the impact of sleep disruptions on health can be made at different levels. The framework presented in Figure 1.1 shows the range of functions and factors that are important in the assessment, and their time frames. The assessment of sleep can be based in a short-term assessment of the physiological functions during sleep to determine how well we sleep at a given night. Also, it can be based on a long-term assessment of daily rest-activity patterns to determine typical patterns of behavior. As illustrated in the framework presented in Figure 1.1, the short-term assessment may include, but is not limited to the following measures:

- Sleep-wakefulness cycle: assess the distribution of the sleep stages (rapid eye movement (REM) and non-rapid eye movement (NREM)). Sleep staging can be used to diagnose sleep disorders and to establish their severity [3].The assessment includes the time spent in each of the sleep stages as well as the time awake.

- Body movements: body movements are generally used as an indicator of sleep quality and depth [4]. Existing metrics for assessment of nocturnal motor activity include the frequency and distribution of movements.
- Respiration: breathing patterns change during sleep [5]. The analysis of breathing patterns may reveal changes in the depth and frequency of respiratory events that may disrupt sleep. The objective of respiratory monitoring is to detect breathing abnormalities and to evaluate physiologic alterations produced by abnormal breathing events [3].

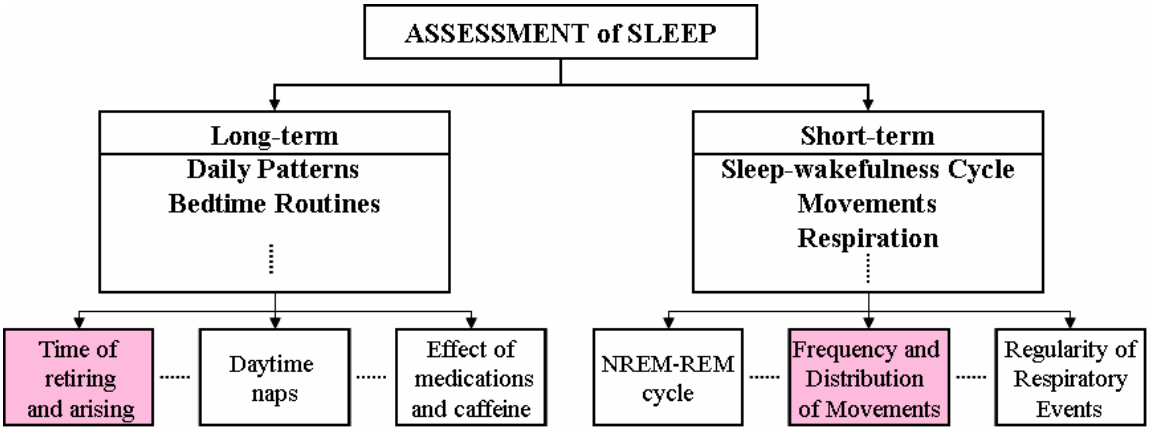


Figure 1.1: Framework for assessment of sleep: at each time scale, examples of clinically relevant measures are shown. The pink boxes correspond to the measures that we are primarily concerned with in this work.

The long-term assessment may include, but is not limited to the following measures:

- Daily patterns: a good sleep hygiene has as much importance as a balanced diet and the amount of exercise [5] for our health. The assessment of the regularity of, for example, bedtimes, get up times, and daytime naps is important because it helps promote effective sleep [5]. Lack of regularity in the sleep-wake schedule and the common practice of cheating hours from sleep can lead to chronic fatigue.
- Bedtime routines: routines such as the use of medications or caffeine in the period prior to bedtime represent one important aspect of a person's sleep to be assessed because they can interfere on sleep. The assessment of the effects of medications on sleep quality is thus helpful on clinical evaluation of sleep-related complaints.

Although all the measures cited above are clinically relevant, in this thesis we are primarily concerned with the assessment of daily patterns and the assessment of movement in bed. While the assessment of the NREM-REM cycle and of respiratory patterns has been done with well established technology in sleep clinics, the assessment of daily patterns and movement in bed can be greatly benefited by unobtrusive monitoring. The drawbacks of the traditional methods for assessment of daily patterns and movement in bed are described next.

1.1 Daily Patterns and Sleep-Related Parameters

A patient's sleep/wake schedule is an important step underlying clinical evaluation of sleep-related complaints [3]. Aspects related to timing of a person's sleep routine provide important clues regarding diagnosis and treatments [6]. Solutions for sleep complaints may sometimes rely solely on changes in habits and life style, based on what is learned from daily patterns. For example, assessing time in bed for treatment of insomnia is very important because restricting bedtimes is one important part of the treatment [7]. In clinical practice, the key measurements (also referred as sleep parameters) usually include some or all of the following:

- Bedtime: the time person retires to bed to sleep;
- Sleep time: the time person falls asleep;
- Wakeup time: the time person wakes in the morning;
- Get up time: the time person leaves the bed to start the day;
- Sleep latency: the amount of time it takes to fall asleep after going to bed; it is computed by subtracting bedtime from sleep time;
- Hours awake after sleep onset: amount of time person is awake after initially falling sleep;
- Number of awakenings;
- Sleep efficiency: it is computed by dividing the total sleep time by the total time in bed; total time in bed is computed by subtracting getup time from bedtime;
- Number of daytime naps;

- Total sleep time in a 24-hr period, including daytime naps.

1.1.1 Traditional Assessment of Daily Patterns

Sleep diaries represent a simple and inexpensive method for assessment of sleep parameters. The patient may be asked to record, on a daily basis, actual sleep times and daytime activities, as well as the occurrence of symptoms such as sleepwalking and nocturnal arousals [6]. This subjective account of daily patterns is valuable when symptoms are not easily accessible to laboratory testing, and has great value for assessing treatment effects and other factors that affect the consistency of a person's sleep. However, there is evidence to suggest that people have difficulties assessing their own sleep especially when suffering from insomnia [8] and depression [9]. A comparison between sleep diaries and polysomnography (PSG), which is considered the gold standard for sleep studies, in a study with depressed patients [9] showed a discrepancy in the estimation of sleep time equal or greater than 30 minutes in 78% and equal or greater than one hour in 52% of the patients. Sleep diaries have also limited usefulness for patients with frequent fluctuations in daytime vigilance, as seen in elderly persons [10]. In this case, a sleep diary may be kept by the staff in a long-term care facility or by the patient or a family member in the home setting, by documenting the sleep habits of the patient for several days to several weeks. According to Tractenberg et al. [11], caregivers may not report a sleep problem unless their own sleep is disturbed, and this is the major weakness of studies of sleep disturbance prevalence in Alzheimer's disease patients that use sleep diaries. In addition, it is very difficult to get patients or family caregivers to maintain good diaries for long periods of time.

Actigraphy has been used to study sleep/wake patterns for many years [12-14]. Actigraphs (or actimeters) are wristwatch-like devices that measure acceleration, and provide information on the activity level of the user. Actigraphs are usually placed on the non-dominant wrist, although they can also be placed at the site of movement to examine specific movements. Physical motion is translated to a numeric representation, sampled at a certain rate and aggregated at a constant interval usually referred as an epoch (e.g. 15 seconds), which varies according to the manufacturer [15]. The epoch-by-epoch samples are stored in the internal memory of the device for a prolonged period (1 or 2 weeks), and

later downloaded to a computer. The ability of actigraphy to detect sleep/wake events is based on the fact the wrist is moved more when the person is awake than when asleep. They provide the opportunity to conduct longitudinal sleep studies through the application of sleep/wake scoring algorithms on actigraph data recorded for a certain period of time. Data loss occurs when the person does not wear it. Therefore, the actigraph has to be worn all the times and patients have to keep records of the times when it is taken off.

In this work, we focus on determining typical patterns of the rest-activity cycle at nighttime and not the estimation of sleep and wake times. Therefore, the parameters of interest are called "sleep-related parameters", and refer to the following parameters:

- Bedtime: the time of retiring, when a person goes to bed with intent of sleeping;
- Get up time: the time when the person leaves bed with intent of starting the day;
- Total time in bed at nighttime (TIB): total time of bed rest derived from bedtime and get up time, and by subtracting the time spent out-of-bed between bedtime and get up time;
- Duration of longest "in-bed" period at nighttime;
- Number of times person gets up during the night;
- Total time "out-of-bed" at nighttime (TOB);
- Number of times person goes to bed during daytime;
- Total time in bed at daytime.

1.2 Movement in Bed

During sleep, major changes in motor activity depend on the sleep state [4, 7, 16]. Low motor activity levels and prolonged episodes of uninterrupted immobility associated with increasing sleep depth, whereas high activity levels are related to intermittent wakefulness during sleep, and arousals are often associated with movement [17]. Therefore, increased mobility in bed can be a sign of disrupted sleep because it is associated with arousals that may reduce sleep quality.

Movement in bed may itself be an indicator of health problems, i.e., the alteration of the pattern or amount of motor activity can be a disease marker [4, 7]. It can reflect illnesses ranging from flu to depression, pain, or the side effects of certain treatments [7, 18]. According to Lemke and Hyyppa's work [18, 19], depressed patients show increased motor activity at night associated with sleep disturbances. Sleep apnea[i] patients also have shown an increased in motor activity at night resulting from disrupted sleep [20]. According to Bennett's work [21], whole body movement is a marker of sleep fragmentation in patients with obstructive sleep apnea. Also, many neurological disorders are presented with abnormal movements during day and nighttime that may adversely affect sleep [7]. For example, normal body movements in Parkinson's disease patients may be repressed by motor daytime symptoms that persist during sleep such as a decreased ability to start and continue movements, and impaired ability to adjust body position. These symptoms worsen sleep quality, and can cause discomfort and pain [22].

There are also motor disturbances that are triggered by sleep such as restless legs syndrome (RLS) and periodic limb movements during sleep (PLMS). Patients with restless leg syndrome report feelings of discomfort in the legs, and they feel compelled to move (for example, tossing and turning in bed) to relief the discomfort [23]. Such symptoms disrupt sleep and cause daytime tiredness and sleepiness. Most of the time, these motor disturbances are ignored by the sleeping patient for a long time. Patients generally do not seek for medical advice, and therefore experience a reduced quality of life. According to the 2005 Sleep in America Poll conducted by the National Sleep Foundation [24], a survey of 1,506 adults living in the United States showed that 38% of the respondents reported that, within the past year, they had twitches or moved frequently in bed at night at least a few nights per week, with 29% reporting that they experienced restlessness every night or almost every night. Respondents were asked if they had ever experienced unpleasant feelings in the legs, a common report of those with restless legs syndrome, and 76% reported that they experience the symptoms at least three nights a week. PLMS are found in at least 80% of patients with RLS, and may provoke frequent arousals or even awakenings. PLMS are involuntary, repetitive movements, most typically seen in the lower limbs but sometimes seen in the arms.

[i] Sleep apnea is a disorder characterized by brief periods of recurrent cessation of breathing caused by airway obstruction with morbid or fatal consequences [5].

1.2.1 Traditional Assessment of Motor Disturbances in Sleep

The assessment of nocturnal motor disturbances in sleep is traditionally performed through one of the following methods: 1) by obtaining information about the nature of the movements from the patient, 2) overnight polysomnograph recording, and 3) actigraphy. The assessment of nocturnal motor disturbances in sleep is based on understanding of the nature of the movements, and the diagnostic is mostly based on the information provided by the patient. The existing metrics used for assessment of nocturnal motor disturbances include the type of movement [25], frequency, and duration [26-28]. The assessment also includes attention to time of onset and pattern of behavior of the movement to understand the nature of the movements. If only the number of movements or the movement time is taken into consideration, the description of the motor pattern during sleep is reduced to the evaluation of the arousal mechanism in patients [28].

Additional assessment may include overnight polysomnograph recording. Video-polysomnography (VPSG) is the gold standard to evaluate and study abnormal motor events occurring during sleep [7]. The VPSG combines the traditional PSG recording with simultaneous audiovisual monitoring and recording of the patient in the sleep laboratory. PSG consists of continuous recordings of several physiological measures including brain waves (electroencephalography), electrical activity of muscles, eye movement (electro-oculogram), breathing rate, blood pressure, blood oxygen saturation, and heart rhythm. Additional leads are applied to other parts of the body (for example, arms and legs) if there is a specific motor complaint [7]. It involves at least a full night's stay in a sleep laboratory attended by properly trained technicians [29]. Because it is a labor-intensive and costly technology, VPSG may not be indicated in all cases [29]. Sleep disorders that require PSG include sleep apnea, excessive daytime sleepiness, narcolepsy, and periodic limb movements during sleep (as described in Section 2.2.4).

Long-term assessment and behavior therapy require an inexpensive technique for which wrist actigraphy is a reasonably economical method that is commonly used [30]. Although it has the advantage that it can be used for extended periods of time, the exact nature and the number of movements that occur in an epoch are not recorded [7]. Activity monitors attached to a person's leg, ankle or feet [31-34] have been used for the assessment of nocturnal lower-limb activity. They have been used to monitor and define

treatment response for RLS and PLMS during successive nights. Because PLMS shows a high night-to-night variability [35, 36], the device has to be worn for many nights. The patient has to keep records of bedtimes and getup times, as well as times out of bed during the night, because the devices cannot differentiate between times in bed from out of bed.

1.2.2 Clinical Need

According to the 2006 report of the Committee on Sleep Medicine and Research [37], as many as 70 million people in the United States suffer from sleep disorders. The most common complaints of disrupted sleep include insomnia, excessive daytime sleepiness, and abnormal movements [3]. Insomnia affects approximately 10% of the US population, and conditions such as RLS and PLMS affect approximately 5% of the population [37].

Insomnia and excessive daytime sleepiness caused by pathologic events such as apnea are easily diagnosed with PSG. Insomnia and excessive daytime sleepiness can also be caused by inadequate sleep hygiene. However, when sleep hygiene is disrupted, PSG is not indicated [5], and the diagnoses relies on the use sleep diaries and actigraphs. Therefore, the assessment of daily patterns can be greatly benefited by the use of unobtrusive systems that do not present the subjectiveness and obtrusiveness problems of sleep diaries and actigraphs.

It has been suggested by Aaronson and others [4, 16] that a sufficiently detailed record of nightly movement, preferentially obtained from non-invasive monitoring devices, could provide valuable information in the evaluation of motor disturbances that affect sleep quality. Since VPSG is not indicated in all cases, and actigraphs are obtrusive, the diagnosis and treatment of health and sleep problems related to the changes in body movements in bed may be greatly benefited by the use of unobtrusive systems that assess movement in bed over many nights.

This thesis focus on the assessment of daily patterns and the assessment of movement in bed using load cells installed at the corners of a bed. Load cells are strain gauge transducers that convert applied force into a resistance change, and they are commonly used in electronic scales and widely deployed in industrial systems. It addresses the question of what information can be inferred about movement in bed from

the weights sensed by the load cells. The technical problem of estimating aspects of the movements from load cell data is described next.

1.3 Problem Definition and Research Contributions

As stated in Section 1.2.1, the clinically relevant information for the evaluation of motor disturbances involves the movements of the person in the bed: when the movements occur, how much a person moves and which parts of the body move. The instantaneous distribution of the mass of the body determines the weights at the corners of the bed sensed by the load cells when someone is lying on bed. Assuming that the mechanical properties of the bed can be characterized by a linear system, and that the human body can be modeled by a finite number of components representing different parts of the body, the load cell responses could be described as

$$\vec{w}(t) = [x * h](t), \tag{1.1}$$

where $x(t)$ represent the bed coordinates of the center of mass of each component (assuming that the mass of each component is taking into account in $x(t)$, for simplicity), and $h(t)$ represents the system impulse response. Therefore, the body movements "convolved" with the bed-human system response generates time-varying weight signals. The technical problem is whether it would be possible to estimate aspects of the movements from load cell data, for example, by inverting the bed-human system response.

There are two major problems with any attempts to invert the system response:

(1) Inadequate signal dimensionality: although the human body parts are not independent, the distribution of the human body mass greatly exceeds that of the four-dimensional load cell signals. The mapping from the body mass distribution to the load cell signals, therefore, results in a substantial loss of dimensionality. This problem can be illustrated by examining a simple 6-component (head, arms, legs and torso) model of human body shown in Figure 1.2. To model the effect of movement in bed, we could use a 6-component model that greatly simplifies the distribution of body mass. The problem with this model is that a large number of parameters (6 x 2) related

to the estimation of the center of mass of each component needs to be estimated, and only the weights at the corners of the bed are known.

(2) **Nonlinearity**: the second critical issue is the highly nonlinear nature of the bed-human system. In preliminary experiments undertaken to characterize the dynamic properties of the bed-human system, we demonstrated the highly nonlinear nature of the system. In particular, since the mass of an individual is comparable to that of the bed, as the individual moves from one position to another, the system properties change significantly. It is therefore not possible to invert the system response.

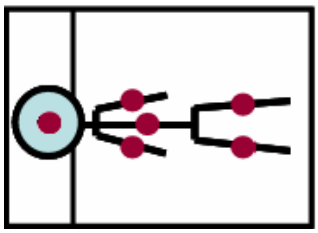


Figure 1.2: A 6-component model of the human body posture on bed: the distribution of the body mass is simplified by determining the center of mass of the head, arms, legs and torso.

The main problem addressed by this dissertation is the degree of inference that can be made to overcome the information is lost due to the inadequate signal dimensionality combined with the high degree of nonlinearity. Although the exact inference from the static or instantaneous information is impossible, additional constraints, imposed by the kinematics and kinetics of the human body movements may lower the effective dimensionality of the assessment problem. For example, it is an empirical question whether the time course of getting on the bed provides sufficient constraints to enable inference. To resolve this question, we turned to a pattern recognition approach treating the system as a black box and the signals as stochastic processes whose parameters depend on the movements.

In this thesis, mobility in bed is characterized by the periods of postural immobility, movement times, and the types of movement performed. The problem of assessing movement in bed with load cells is framed as a problem of *detection* that consists of estimating the time intervals when a movement occurs, and a problem of *classification* that consists of determining the type of movement performed in a given time interval. The term ‘movement in bed’ is used instead of ‘sleep movement’ to refer to

body movements occurring during sleep and awake periods [28]. This work focuses on the detection of major posture shifts and smaller movements (which involve the head, torso and limbs).

The major contributions of this thesis are summarized by the following points:

- **Development of a System for Unobtrusive Assessment of Movement in Bed**:

  We developed a system for assessment of movement with load cells that extends previous work to classify, and not only detect, movements in bed. Unlike many existing mattress pad-based approaches, the well-proven load cell technology, based on strain gauge sensors, provides s and reliable data, and it is a practical solution for long-term monitoring. Also, in particular:

  - **Approximation of Detection Threshold:** we showed that the movement detection threshold for a new subject can be approximated by linear regression of the thresholds of a known set of subjects.

- **Derivation of a Principled Approach for Classifying Body Movements in bed:**

  We developed a 3-class movement classification approach that discriminates posture shifts, medium amplitude movements and leg movements. Also, in particular:

  - **Invariant Feature Representation**: we found a feature representation that would be relatively invariant across subjects, and in particular to their weight and height.
  - **Additive Functional Representation**: despite the essential nonlinearity of the bed-human system, we showed that, for the purposes of movement classification, the system can be viewed in terms of additive representation, as a locally linear system with respect to small movements.
  - **Feasibility of the System**: we showed that the system does not require extensive training, which is particularly valuable for the viability of the system in sleep studies with different populations.

- **Development of a Methodology for Determining Typical Patterns of Sleep Behavior, such as Bedtime and Get Up Time Behaviors:**

  We created a metric called *in-bed profile* for characterization of a person's typical patterns of sleep behavior. This metric could be used with other types of sensors that provide similar information (for example, motion sensors). Also, it could be used in other applications besides sleep hygiene, where the knowledge about a person's habits could be helpful, for example, for localization.

1.4 Organization of the Thesis

The remainder of the thesis is organized as follows. Chapter 2 reviews the characteristics of motor activity in bed, and presents the related work on unobtrusive assessment of sleep-related parameters and on bed sensing technology.

Chapter 3 introduces the system for unobtrusive assessment of mobility in bed with load cells. The inference problem of deriving movement information from load cell signals is described in details. A description of the load cells and system setup used in this work is presented. Also, an introduction to the approach for estimation of sleep-related parameters is presented.

Chapter 4 describes the experimental setup for collection of load cell and actigraphy data used in this work. Data were collected from three experiments: two performed in a laboratory with different bed sizes and one in an assisted-living facility.

Chapter 5 describes a methodology for determining typical patterns of sleep behavior. Bedtime and get up times are estimated by using the person's in-bed profile and information about the use of the bedroom‘s light (times when the light is turned on/off). The estimated parameters are compared with the estimates from sleep diaries completed by caregivers for a period of 2 weeks.

Chapter 6 describes the framework for movement detection. Movement detection is formulated as a problem in statistical hypothesis testing. The movement detection approach is evaluated on load cell data collected from 15 healthy subjects. Evaluation data were collected in the laboratory, under controlled conditions with subjects awake, and a video technique is used as the ground truth. We demonstrate that the bed size has

no effect on the performance of the proposed detector. In addition, in-home assessments conducted with residents of an assisted-living facility are included to demonstrate its feasibility in a realistic setting. We also show that actigraphy under-reports movement in comparison with our method.

Chapter 7 presents a framework for movement classification. The approach for subject-dependent classification of movement in bed uses Gaussian mixture models, and two feature representations are tested: a time-domain representation that is based on simple descriptors of the trajectory of the body center of mass during movement, and a time-frequency domain representation based on wavelets and multiresolution analysis. We show the performance results on the laboratory data.

Chapter 8 summarizes the results of this work, and discusses future directions.

Chapter 2

Background

In the previous chapter, we discussed the importance of rest-activity patterns and movements in bed to the assessment of sleep disturbances. In this chapter, we present the related work on unobtrusive assessment of sleep parameters in Section 2.1. We also review the characteristics of motor activity during sleep in Section 2.2, and present the related work on bed sensing technology in Section 2.3.

2.1 Related Work on Unobtrusive Assessment of Sleep Parameters

Given the drawbacks of sleep diaries, researches have looked for alternative ways to obtain information about rest-activity patterns at nighttime from unobtrusive sensors installed in the bedroom. Chan et al. [38, 39] propose a system that uses motion sensors installed in areas such as the bedroom and bathroom to monitor activity during the night. The system consists of 10 infrared motion sensors installed on the ceiling that include one above the bed and in areas adjacent to the bed. A number of different activities such as going to bed, restlessness in bed, getting out of bed, and getting out of the room are monitored. The activities are detected by the pattern of the sensors activations, and by setting thresholds. They found good agreement with the nurse staff annotations in an 8-month study that monitored 4 subjects. Although motion sensors represent a cheaper technology than load cells, the proposed system has to be reconfigured every time the environment changes. With the use of load cells installed under a person's existing bed, no changes have to be made when the environment changes. In addition, the proposed

system cannot discriminate the patient being in bed and standing near the bed, and it recognizes the later as restlessness in bed. It is straightforward to detect if the person is in and out of bed with load cells due to the drop in the total weight when someone gets up.

Rachwalski et al. [40] use a pressure pad, placed between the mattress and the bed frame and below a person's hips, to measure activity in bed and study the sleep patterns of Alzheimer's sufferers. Rest-activity is used as a surrogate for sleep/wake patterns in a similar way that is done with actigraphy, and sleep patterns are characterized by the sleep latency and the number of awakenings at night. Activity in bed is measured by changes in the pressure measurements. The pressure measurements are sampled at 1 Hz, and aggregated into 30-second epochs by averaging the data every 30 seconds. Sleep latency is determined by the first period when 5 consecutive minutes of inactivity is recorded. Awakenings (or periods of restlessness) are defined by 3 consecutive minutes of body movements. Caregivers reported the level of nighttime restlessness and daytime behaviors in a questionnaire. Results showed that the sensor data were corroborated by the subjects or caregivers' reports in a one-month study that monitored 3 subjects, with one subject in advanced Alzheimer's disease stage. As mentioned earlier in this section, sleep diaries have limited usefulness due to their subjectivity and may not correspond well with objective sleep estimates [41]. Also, even for normal subjects, there is obviously some difficulty in recollecting exact sleep times or the number and length of awakenings during the night [42]. Therefore, the evaluation of the proposed system is not adequate because sleep diaries should not be used as ground truth for sleep.

The solution we are proposing for estimation of sleep-related parameters with load cells cannot determine sleep and wake times, and movement in bed is not used to derive information about sleep and wake times. However, the information derived from the load cells is still valuable as an objective and continuous measure of daily patterns, and it is particularly valuable in sleep studies in populations who would not be able to remember specific hours to complete sleep diaries or who would depend on subjective reports from caregivers or family members, as we demonstrate in Chapter 5.

2.2 Motor Activity during Sleep

The normal motor activity throughout a night is diverse: it varies from complex, determined activity to simple movements that are primitive in character [43]. Human sleep is characterized by episodes of immobility interrupted by episodes with different activity levels. Because the organization of the motor activity during sleep is related to the sleep macrostructure (cyclic patterns involving two sleep states), a brief overview of the normal sleep cycle is presented next.

2.2.1 The Sleep Cycle

There are 2 primary states of sleep: rapid eye movement (REM) and non-rapid eye movement (NREM). In normal sleep in adults, there is an orderly progression from wakefulness to sleep onset, and then to NREM and REM sleep [7]. Approximately 20% of sleep is spent in REM sleep and 80% in NREM sleep. The entire sleep cycle usually lasts about 70-120 minutes, and it is repeated four to six times before awakening [44].

NREM sleep is divided into 4 stages. Sleep progresses from stage I (the lightest level, during which the sleeper can be easily awakened) to stage IV (deepest level). There is a progressive decrease of motor activity including muscle tone [7]. In stage IV, blood pressure is at its lowest, and heart and breathing rates are at their slowest. During NREM sleep, many of the restorative functions of sleep occur. Hormones are released which help the body rebuild itself from damage done during the day.

REM sleep is characterized by a decrease or absence of motor activity. It is the period where dreams occur, and the normal loss of muscle activity prevents the sleeper from acting out his/her dreams. The eyes move rapidly, and muscles may jerk involuntarily. REM sleep is characterized by changes in heart rate, blood pressure, respiration, and perspiration [7].

2.2.2 Normal Body Movements during Sleep

The normal motor phenomena during sleep includes hypnic jerks at sleep onset, postural shifts, body and limb movements, and sleep myoclonus (small bursts of movements) [7]. Hypnic jerks are very common movements occurring at the transition to sleep, and consist of abrupt flexion movements that are often felt as an illusion of falling. It is a

benign movement that has little effect on sleep, and it probably occurs in most people. Arousals, brief periods of interrupted lighter sleep that may lead to full awakening, may both follow and lead movements such as body shifts [7].

In NREM sleep, motor activity is reduced in comparison to the waking state. Postural shifts, which may signal stage changes (into or out from wake or REM), occur. The frequency of all movements decreases with depth of sleep, with progressive decrease in the number of movements from stage I to stage IV [7]. In REM sleep, muscle tone is reduced, but small bursts of movements known as sleep myoclonus may occur. Sleep myoclonus consists of phasic muscle bursts that are typically seen in REM sleep but can also occur during NREM sleep. They are small flickering movements that may not cause apparent movement [7].

2.2.3 Distribution of Body Movements

According to the work of Kleitman and others [45], good sleepers (people with a sleep latency of less than 15 minutes and that sleep for at least 7 hours) have 3 to 5 minutes of body mobility during the night, for a total of 20 to 60 movements per night. In those studies with good sleepers, they took about 5 to 10 seconds to move from one position to another. The ratio of major (major posture shifts) to minor movements (small changes in body position) varied from person to person, but was usually less than one, i.e., more than half of the movements were found to be minor. However, according to Alihanka [28], there is a marked discrepancy between the total number of body movements and other measurements found in these sleep studies due to the fact that different techniques, which includes video, human observation, and time-lapsed photography, were used to measure movements at different sampling rates.

There are also contradictory theories about the distribution of body movements in normal sleep. According to Johnson's work [4], the distribution of movements though the night is a personal characteristic of the sleeper. Even when exposed to environmental noises during sleep, the rate of body movements remains generally unchanged. In the presence of noise, the movements are temporally redistributed to the proximity of the noise, without any increase in their total number. However, according to the work of Kleitman and others [45], in general, there is greater mobility during the second half of

the night as compared to the first half, with the ratios varying from 2:1 to 3:2. It thus suggests a gradual increase in movements through the night instead of a distribution that is typical of each sleeper.

Body movements impressively change with age [26]. A decrease in the number of body movements is observed from infancy to adulthood also continues at later ages. However, actigraphy studies have reported contradictory results indicating no change or even increase in the activity level with aging [46]. In the elderly, knowledge on the pattern of body movements is still largely incomplete. According to the work of Gori et al [26], a study with healthy elderly subjects showed that when body movements occur in elderly individuals, they are likely to be followed in the next 60 seconds by a sleep stage change or by a spontaneous awakening. This suggested that body movements may act as co-factor in a process leading to sleep state shifts. A method of unobtrusively evaluating the dynamics of body movements across different populations would be an extremely valuable tool for increasing our general understanding of sleep behaviors.

An overview of some of the most common motor disturbances that can occur at sleep onset or during sleep is presented next. We also discuss how the diagnosis and treatment of these illnesses would be benefited by the use of a continuous and unobtrusive system for assessment of body movement in bed.

2.2.4 Common Motor Disturbances

According to Phillips's work [22], the most common movement disorders are Parkinson's disease (PD), restless legs syndrome (RLS), and periodic limb movements during sleep (PLMS). PD is not a primary sleep disorder, but it has a major impact on sleep quality [22] because nocturnal motor problems are common in PD patients [7]. Normal body movements in bed may be repressed by motor daytime symptoms that persist during sleep such as akinesia (inability to initiate voluntary movements [47]) and bradykinesia (slow execution of movement [47]). These symptoms worsen sleep quality, and can cause discomfort and pain [22]. The conventional methods for clinical assessment of nocturnal disability in PD are based on PSG, patients and caregivers diaries, and a scale called Parkinson's Disease Sleep Scale (PDSS) [48]. The PDSS is a 15-item measure that assesses aspects of nocturnal disability such as nocturnal restless legs and akinesia, and is

based on patients and caregivers experiences in the past week to respond the questions. The assessment of the symptoms has been done with sensors attached to the body [47] during daytime periods. Therefore, patients and clinicians could be benefited by an in-home objective and unobtrusive nocturnal assessment of movement in bed that can examine aspects such as slowness of motion and long periods of immobility.

Restless legs syndrome (RLS) symptoms occurs while the person is awake (at sleep onset), and may also occur on awakenings in the middle of the night [7]. RLS is activated by relaxation, and it is diagnosed by the presence of voluntary leg movements due to evening discomfort that forces the person to move to get relief [29]. It can also be manifested as tossing and turning in bed or stretching the legs, and it can profoundly disrupt a person's sleep. RLS can start at any age, and the severity of the disease appears to increase with age such that older patients experience symptoms more frequently and for longer periods of time [49]. In the absence of treatment, patients with RLS experience a reduced quality of life. RLS is diagnosed by obtaining a history [22], i.e., the clinician asks the patient about detailed information on sleeping habits and sleep history. Therefore, because people with RLS often do not complain or seek medical attention [22], and because the diagnostic is based primarily on the patient's subjective account of the alterations in the amount and type of body movements, the use of a continuous and unobtrusive system for assessment of movement in bed at home can be very valuable to validate the patient's complaints. A study that employed an activity sensor on both ankles showed that the nocturnal lower limb activity per minute discriminates RLS patients from healthy patients [32].

Periodic limb movements during sleep (PLMS) are involuntary, repetitive movements, most typically seen in the lower limbs but sometimes seen in the arms. Typical PLMS consist of dorsiflexions of the ankles (bending the ankle upward to bring the toes closer to the knee), and sometimes accompanied by flexions of the ankles, knees, and thighs, and may be unilateral or bilateral [23]. They occur about every 20-40 seconds, primarily in NREM stages I and II, and last approximately 0.5-5 seconds [7]. At least 4 consecutive leg movements must fulfill these criteria to be scored as PLMS [34]. Because PLMS may lead to awakenings, it is considered as a contributing factor for daytime fatigue and chronic insomnia [49]. Approximately 35% of the elderly population

experience PLMS. PLMS and RLS are also common findings among PD patients [35]. Unlike RLS, which is diagnosed by taking a patient's history, PLMS is diagnosed in a sleep laboratory by the recording of bilateral surface electromyogram (EMG) of the anterior tibialis muscles [7]. Although actigraphy has been used for diagnostic of PLMS [33], it requires patient's compliance because the person has to wear it every night. Also, the patient has to keep a diary of the bedtimes and get up times to determine time in bed. The use of an unobtrusive system, which enables classification of movements in bed that include leg movements, could be very valuable for evaluation of the efficacy of medications to treat PLMS.

Many bed sensing devices have been proposed for unobtrusive assessment of body movement in bed [28, 40, 50-53]. A description of the most representative work on assessment of body movement with bed sensors is described next.

2.3 Related Work on Bed Sensing Technology

Most of the current work on assessment of body movement is only focused on the detection of movements, and a number of those require special modifications to the bed. For example, Lu and Tamura [50, 51] propose a bed temperature measurement system for detection of body movement. An array of 15 thermistors is placed between a waterproof sheet and the mattress, perpendicularly to the body position. A body movement is detected when the sum of the square of the temperature differences for all sensors is greater than a certain threshold. Data are collected every five seconds, and the system is able to detect torso and leg movements using an array at the waist and another at the lower limbs. Their system only estimates the frequency of movements and time in bed.

Van der Loos et al. [52] propose a system called SleepSmart™, composed of a multi-sensor mattress pad that is placed on the top of a mattress, to estimate body center of mass and index of restlessness. The mattress pad is composed by one array of 54 force sensitive resistors and one array of 54 resistive temperature devices. The arrays are denser under the torso than under the legs. Data are sampled at 100 Hz, and the body center of mass is estimated from moment calculations of the force and temperature signals. An index of restlessness is calculated by integrating the absolute changes in body center of mass at 25-second intervals. It records posture and temperature maps, and

reports the body center of mass every five seconds. The proposed system does not classify the type of movement. In addition, the pad must be aligned with bed, and should not move to avoid errors in the measurements.

Harada et al. [54] use a 76.8" x 30.4" x 0.68" (192 x 76 x 1.7 cm) array of 210 pressure sensors to estimate body position in bed. Three sleep positions are considered - supine (lying on the back), and right and left lateral positions, and the estimation of body position is based on detection of the position of the following body parts: 1) back side of the head, center of the blade bone, sacrum, right and left heels for the supine position, and 2) side of the head, shoulder, base of the thigh, knee and ankle for right and left positions. The position of each part of the body is estimated by cross-correlation between the pressure distribution image from the sensor and pressure distribution templates. The templates are created from models based on different positions of the neck, thigh, knee, calf and ankle joints. They created 180 templates for each position based on variations of the rotation angles. Posture is displayed through a computer image at 10 Hz using a 3-D model. This system requires that the joints model must be adjusted for each person before use. Also, the pressure mat must to be aligned with the bed and should not move.

Rachwalski et al. [40] use a pressure pad placed between the mattress and the bed frame, below a person's hips, to measure activity in bed. The pad is constructed of plastic fiber embedded in foam, and the sensing element is comprised of a 'transmit' and a 'receive' fiber. A red light shines light through a 'transmit' fiber, and when an external force compresses the foam, there is an increase in the intensity of the backscattered light, which is monitored by a 'receive' fiber [55]. The pressure measurements are sampled at 1 Hz, and aggregated into 30-second epochs by averaging the data every 30 seconds. A body movement is detected when the average of the rectified highpass filtered pressure measurements is greater than a certain threshold. The system only reports the frequency of movements and time in bed, and it does not classify the type of movement. Even though the pad allows for more measurement points than the load cells, it is not clear that leg movements could be easily classified because the pad only covers an area under the hips. In addition, informal experiments with this sensor in our laboratory suggest that there is an inertia associated with the change in the intensity of the light when a sensing element is pressed for a long period.

Several authors [27, 28, 53, 56, 57] have employed the static charge sensitive bed (SCSB) for monitoring of motor activity. The SCSB is composed of two metal plates with a wooden plate in the middle that must be placed under a special foam plastic mattress (2"- 4"), which operates like a capacitor. When a person moves, static charges are formed in the clothing and in the mattress, and these charges induce potential differences between the plates. The potential differences between the capacitor plates produce the raw signal (sampled at 250 Hz), which is filtered into different frequency bands to detect body and respiratory movements. Alihanka [27] used the SCSB sensor and automatically classified the detected movements by the SCSB into 4 classes based on their duration: movements lasting less then 5 seconds, movements lasting between 5 and 10 seconds, movements lasting between 10 and 15 seconds, and movements lasting more than 15 seconds. He analyzed the distribution of the different types of movements and its relationship with the duration, and later concluded that duration is not sufficient to completely describe classes of movements. Rauhala et al. [53] used this system to detect periodic leg movements, and showed that it has high sensitivity to detect any periodic movement in the body. However, no automated method that can recognize patterns has been developed to recognize PLMS and discriminate these movements from other movements also detected by the sensor. Also, problems with the calibration of the sensor and with the electromagnetic fields in the room of the recordings have been reported [7].

Wilde-Frenz and Schulz [25] use a combination of PSG and an actogram placed under the mattress, in the lumbar area, to detect and classify movements. The actogram consists of a mechanoelectrial transducer with direct input to the PSG. The following classes of movements are discriminated based on actogram and artifact information from the head leads (EEG, EOG and EMG):

- Type I movement (isolated movement of the head, trunk or the limbs): a 30-second epoch containing either motor artifacts in the head leads or a signal in the actogram.
- Type II (an integrated activity of more distant parts of the body such as head and the trunk or the limbs): a 30-second epoch containing motor artifacts in the head leads and a signal in the actogram.

- Type III (major posture shifts): a 30-second epoch containing a signal in the actogram and motor artifacts in the head leads for more than 15 seconds.

Although this method provides a good discrimination between different classes of movements, it becomes an impractical solution because it employs PSG in combination with the bed sensor.

While many pad-based solutions have been proposed, in general, these sensor signals are very sensitive to the position of the pad on the bed, the material of the pad, the material that the mattress is made of, and the construction of the bed. These are all factors that are less important when using load cells. Thus, the use of four load cells under the bed is a practical solution for bed sensing. It provides s and reliable data, and it can be installed on a person's exiting bed. It allows both detection of body movement and classification of the type of movement. The system for unobtrusive assessment of mobility in bed with load cells is introduced in the next chapter.

Chapter 3

Assessing Mobility in Bed Using Load Cells

The purpose of this chapter is to introduce a system for unobtrusive assessment of mobility in bed with load cells. The system is focused on detection of body movement and classification of the type of movement. In Section 3.1, the system is introduced, and the load cells and the system setup used in this work are described. A description of the inference problem of deriving movement from load cell signals, and an introduction to the approach used in this thesis are presented in Section 3.2. An introduction to the approach for estimation of sleep-related parameters is presented in Section 3.3.

3.1 Unobtrusive Assessment of Mobility in Bed

The system for unobtrusive assessment of mobility in bed focuses on identifying the time intervals when a movement in bed occurs and on determining the type of movement performed in a given time interval. The term 'mobility in bed' has been defined as the ability to move to/from lying position, turning from side to side, and positioning body while in bed, and it is listed as an activity of daily living (i.e., it is included on a list of basic activities that support survival, and are designed to measure functional ability) by the Canadian Institute for Health Information [58]. In this work, mobility in bed is characterized by the periods of postural immobility, movement times, and the types of movement performed. The system consists of 2 parts, as depicted in Figure 3.1: detection of body movement and classification of the type of movement. The problem of *detection of movement in bed* consists of estimating the time intervals when a movement in bed

occurs. The problem of *classification of movement in bed* consists of determining the type of movement performed in a given time interval.

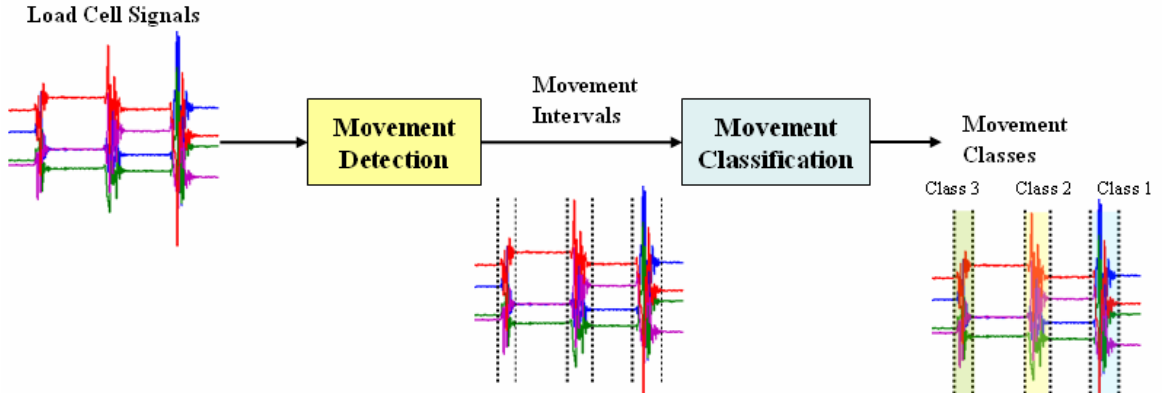


Figure 3.1: Diagram of the system.

The system is designed to detect postural shifts, smaller position changes and limb movements based on changes in the magnitude of the load cell signals. It provides a detailed description of the motor pattern during sleep in terms of the time of onset, duration, frequency, and the type of movement. By determining the movement times, one can also assess the amount of postural immobility per night. The number of postural immobility episodes (longer than 15 minutes) has been shown to be positively related to subjective estimates of the goodness of sleep [59]. The system utilizes four load cells placed under a bed, one at each corner. The load cells and the setup used in this work are described next.

3.1.1 Load Cells

Load cells are strain gauge transducers that convert applied force into a resistance change. They are widely deployed in industrial systems and also commonly used in electronic scales. They can be manufactured to measure loads on nearly any scale, ranging from measuring ingredients for pharmaceutical productions in milligrams, to the weight of a freight train with several hundred tons [60]. They are of relatively low cost, and represent a simple and durable technology.

Given that the changes in resistance are small, the transducer is typically configured as a Wheatstone bridge made up of four resistive elements R_i, $i = 1, 2, 3, 4$ arranged in a diamond orientation, as shown in Figure 3.2. An input DC voltage V_{in}, or

excitation voltage is applied to terminals A and C. The output voltage V_g is measured across the terminals B and D [61], and it is given by

$$V_g = \left(\frac{R_3}{R_2 + R_3} - \frac{R_4}{R_1 + R_4} \right) V_{in}. \tag{3.1}$$

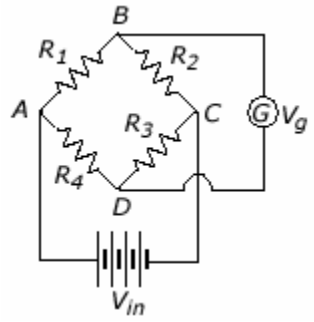


Figure 3.2: The configuration of a Wheatstone bridge.

When the bridge is balanced (no load), the output voltage V_g is zero and, from Equation 3.1,

$$\frac{R_1}{R_2} = \frac{R_4}{R_3}.$$

If, for example, we replace R_1 in Figure 3.2 with a strain gauge, any changes in the strain gauge resistance will unbalance the bridge and will causes a voltage to appear across the middle of the bridge.

Although forces are not directly measured, but inferred from the resultant strain, the output is linearly proportional to force, with the relationship determined by calibration. In the calibration procedure, we employ linear regression based on known test weights to transform the raw output into weight values. The calibration process is described in details in Appendix A.

3.1.2 Load Cell System Setup

For compatibility, we used a similar setup as that used at Elite Care because a study was also conducted at that assisted-living facility. The load cells used in this work are single point load cells, model AG100 C3SH5eF (Scaime™, France). Figure 3.3a) shows the setup used for the experiments. Four load cells are placed under a bed, one at each corner. The load cells/corners are numbered 1 through 4 as shown in Figure 3.3b).

(a) (b)

Figure 3.3: (a) A load cell (middle) is placed in each corner of the bed. (b) Load cells are numbered 1 through 4.

After calibration, the weight measurements from the 4 load cells are composed of the bed weight values and the bed-and-person weight values. When the person is not lying on the bed, only the bed weight is measured. For simplicity, the resultant weight measurements are reported in kilograms or pounds instead of force/weight units. The individual readings from the load cells are given by $w_i(t)$, for $i = 1, 2, 3, 4$, where $w_i(t)$ denotes equally spaced samples sampled at each discrete time t. They can be summed to generate one output signal $w(t)$ at each time t, which is the total weight in pounds:

$$w(t) = \sum_{i=1}^{4} w_i(t). \tag{3.2}$$

The load cells specifications [60] are show in Table 3.1. The nominal load or capacity is the maximum load a load cell is designed to measure within its specifications. They are driven with 5 V, and return 1.9 mV per Volt of excitation, resulting in an output ranging from 0 to 9.5 mV. For example, if a 25-kg load is placed on a load cell with capacity of 100 kg with a 5 V excitation voltage, it will have an output of 2.375 mV. The combined error in Table 3.1 refers to errors due to non-linearity and hysteresis measured by the manufacturer [62]. Hysteresis error refers to differences between an upscale sequence calibration and a downscale sequence calibration of a sensor [61]. The minimum division or resolution is the smallest change in load which produces a detec change in the signal.

An acquisition board from Elektrika Inc., model 335-2001 Rev. C, is connected to the load cells. The output signal from the load cells is amplified by 200, resulting in an output voltage between 0 and 1.9 V. The amplified output is then sampled by a 13-bit

analog to digital converter and transmitted to a PC trough a USB cable. The board has an anti-aliasing filter set on 50 Hz for a 200 Hz data collection rate. However, data can be collected at higher rates.

Table 3.1: Load cells specifications.

Material	aluminum
Nominal load (capacity)	100 kg
Nominal Sensitivity	1.9 mV/V
Combined error	0.017 kg
Minimum division (resolution)	0.020 kg
Compensated temperature range	-10°C to 40 °C

A metal plate under each load cell is used to serve as a base for each load cell on the floor, and a 2.4" (6-cm) pin is placed on one of the top holes to allow the bed to stabilize on one point on each load cell. The metal plate has a size of 7.2 x 4 x 0.4" (18 x 10 x 1 cm). A hole with a 0.8" (2-cm) diameter and 3.2" (8-cm) length was made under each bed post to place the pins, allowing the bed to stabilize on one point on each load cell. Figure 3.4 shows the metal plate under the load cell, and the pin on the right side of the load cell.

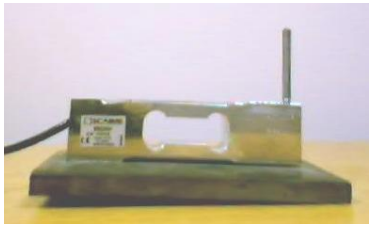


Figure 3.4: Metal plate under the load cell, and pin on the right side of the load cell.

3.2 Inference of Body Movements in Bed

The weights sensed by the load cells at the corners of the bed $w_i(t)$ at each discrete time t represent the instantaneous distribution of the mass of the body when someone is lying on bed. These time-varying weight signals are generated by the body movements

"convolved" with the bed-human system response. Therefore, we need to determine if the system response due to the bed can be inverted to infer about the movement.

Consider the problem of modeling the human body pose in bed, and estimating the output of each load cell from the sleeper's position. For simplicity, let's assume that the human body can be modeled by N components representing different parts of the body (a 6-component model was illustrated in Figure 1.2), where each component k has a certain mass m_k, and with its center of mass located at (x_k, y_k) bed coordinates. The distribution of the mass of the body determines the weights at the corners of the bed sensed by the load cells. In an ideal situation where the bed could be represented by a weightless plate and had no dynamics, the output from load cell i, for $i = 1, 2, 3, 4$, at each discrete time t, would be represented by

$$w_i(t) = b_i + \sum_{k=1}^{N} m_k \theta_{ik}(t) + n_i(t), \tag{3.3}$$

where

- $w_i(t)$ corresponds to the weight measured by load cell i, at time t,
- b_i corresponds to the proportion of the bed weight measured by load cell i (for simplicity , it is assumed to be constant),
- $\theta_{ik}(t)$ is a time-dependent representation of the position of the center of mass of each component of the body k based on its proximity from load cell i, i.e., θ_{ik} is a function of time and of the locations of each load cell i (x_i, y_i) (which are known) and each component k (x_k, y_k), and it can be represented as a matrix with elements $\theta_{ik}(t) = f(x_i, y_i, x_k, y_k, t)$, and
- $n_i(t)$ corresponds to the noise term from load cell i at time t, due to vibration on the floor and A/D conversion.

Equation 3.3 defines the forward model for estimating the load cell outputs from the position of each body component. According to Equation 3.3, the response of the bed-human system during a movement depends on the mass and on the location of each body component, represented by each component k, on the bed.

The inference problem is further complicated by the physics of the bed and the dynamics of the bed-human system. Using a lumped[ii] model of a human body as in Equation 3.3, the input into the system is a 2N-dimensional vector. To the extent that the sleeper is moving, this input then generates time-varying weight signals as the output from the load cells. Assuming that the mechanical properties of the bed can be characterized by a linear system, the load cell responses are described according to Equation 1.1.

In order to infer the sleeper's movements, we therefore need to solve the inverse dynamic problem of deriving (x_k, y_k) from the load cell outputs. There are two major problems with the inversion of the system response: inadequate observed signal dimensionality and nonlinearity to be considered in turn. The mapping from the body mass distribution to the load cell signals results in a substantial loss of dimensionality. The desired positions of the center of mass of the N components representing different parts of the body require the estimation of N x 2 parameters, and only 4 load cell measurements are known. Also, if even a simple 6-component model is used (as illustrated in Figure 1.2) to reduce the number of parameters to be estimated, this simple model underestimates the amount of information required to adequately model the movements.

The second problem concerns the linearity of the system. At the outset, the system consisting of a mass, springs, and friction would appear to be linear. If the mass of the human was insignificant relative to the mass of the bed, the system could be well be approximated by a linear system, and the inverse problem could be addressed by blind deconvolution or similar techniques. Since the mass of a human is comparable to the mass of the bed, a change in the location of the body changes the dynamics of the system, thereby introducing nonlinear behavior.

In order to illustrate the nonlinear behavior of the system, we analyze a grossly oversimplified model of the system shown in Figure 3.5. First, we reduce the 3-dimensional physical system to a 2-dimensional problem. We replace the mattress, bedspring and the bed frame by a mass m_0 attached at the end of a weightless rod of length L. To further simplify the problem, we assume that the rod is attached to a fixed

[ii] The term 'lumped' refers to lumped-parameter analysis, as opposed to distributed-parameter analysis. Lumped parameter systems are systems with a finite number of degrees of freedom [60].

pivot point. The mass m_0 represents the inertia of the bed. In a similar manner the spring and friction are represented by the corresponding lumped parameters. The mass m_1 represents the proportion of the human body mass projected to one corner of the bed. Let $x(t)$ represent the position of m_1 on the rod at time t, and $s(t)$ represent the vertical displacement of m_0 at time t as m_1 moves from its initial position. Assuming that the system starts at rest, and that $x(t) = x_0$ for $t = 0$ (i.e., the person is in bed, but lying still at x_0), the differential equation representing the system is given by

$$\left[m_0 + m_1 \left(\frac{x}{L} \right)^2 \right] \ddot{s} + c\dot{s} + ks = \frac{m_1 g}{L} (x - x_0), \tag{3.4}$$

where c represents the damping constant, k represents the spring constant and g represents the acceleration due to gravity. Assuming that $x(t) = x_0 + d(t)$ for $t > 0$, where the function $d(t)$ represents the horizontal displacement of m_1 at time t, and using a test of homogeneity, it is possible to show that the system represented by Equation 3.4 is nonlinear.

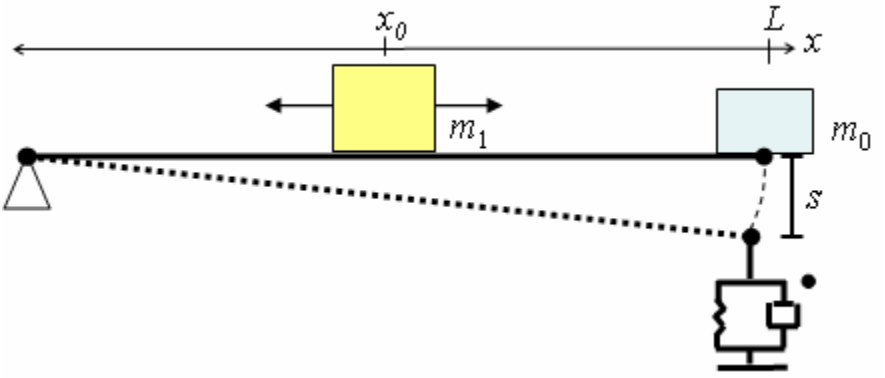


Figure 3.5: A simplified version of the bed-human system.

If m_1 is very small in comparison to m_0, the system behaves closely to being linear. Since the mass of the human m_1 is comparable to that of the mass of the bed m_0, as the individual moves from one position to another, the system properties change significantly. It is therefore not possible to invert the system response.

The inverse static problem of estimating (x_k, y_k) is underdetermined because the 4 load cell measurements can determine up to 4 unknowns. Additional constraints imposed by the kinematics and kinetics of the human body movements may lower the effective dimensionality of the problem. Due to these problems, we turned to a pattern recognition

approach treating the system as a black box and the signals as stochastic processes whose parameters depend on the movements.

3.2.1 Characteristics of the Load Cell Outputs

There are 2 characteristics of the load cell outputs, which are intrinsic to the system setup, to the weight of the person moving in bed, and to the part of the body being moved, that have been incorporated in the detection approach. These characteristics are:

- Transients corresponding to the dynamic forces generated by the movement.
- The repositioning of the center of mass of the body due to movement can result in changes in the weight measured at each corner after a new position is reached.

The plot in Figure 3.6 illustrates such characteristics from an example of the load cell data collected during a sequence of 3 movements: a leg movement (around 10 seconds), a posture shift (between 30 and 40 seconds), and an arm movement (around 52 seconds).

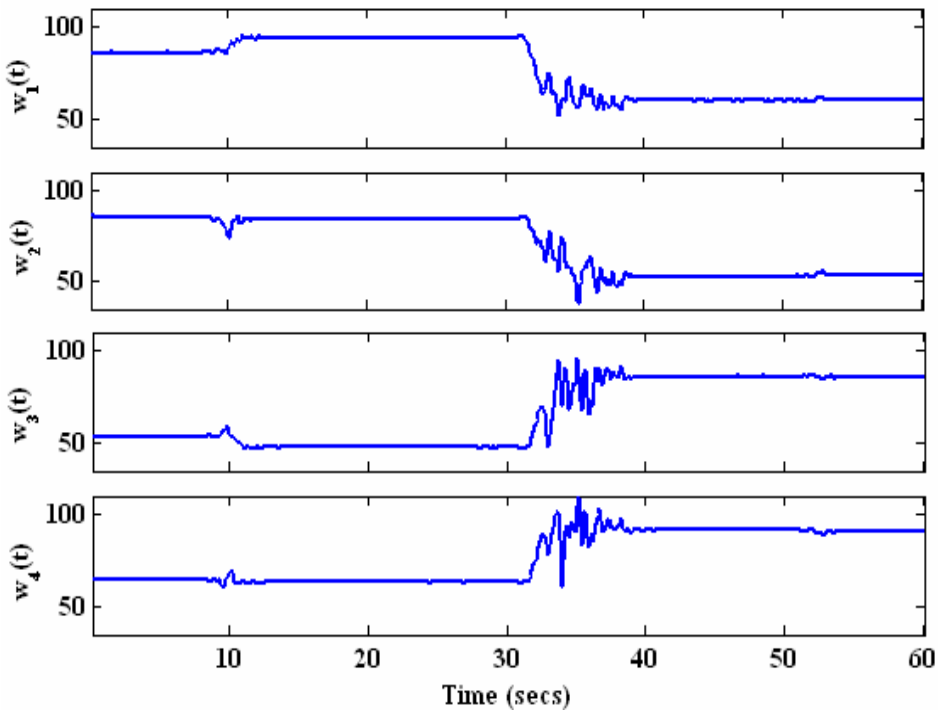


Figure 3.6: Load cell outputs $w_i(t)$, for $i = 1, 2, 3, 4$, in pounds, during a sequence of 3 movements: a leg movement (around 10 secs), a posture shift (between 30 and 40 secs), and an arm movement (around 52 secs).

The plot shows the individual outputs $w_i(t)$, for $i = 1, 2, 3, 4$, in pounds. While changes in the signal caused by large movements (posture shifts) are easier to visualize,

small movements cause smaller oscillations on the data, since the amplitude of the transients depends on the body part being moved, and on the weight of a person. The intensity of the movement is also reflected by the difference between the total weight measured at each corner before and after a movement. The difference is due to the repositioning of the center of mass of the body due to movement.

3.3 Estimation of Sleep-Related Parameters

Since the total weight on the bed drops when someone gets up from the bed, this information can be used to estimate the sleep-related parameters listed in Section 1.1. Figure 3.7 shows an example of the load cell data collected for two consecutive 24-hour periods (4:00 PM – 4:00 PM): day 1 (Figure 3.7a) and day 2 (Figure 3.7b). The vertical axis shows the total weight on the bed, in pounds. This person gets up at night, as evidenced by the drop in weight for short intervals during the nighttime hours. The variation seen in the data when there is a person lying on the bed is due to movement.

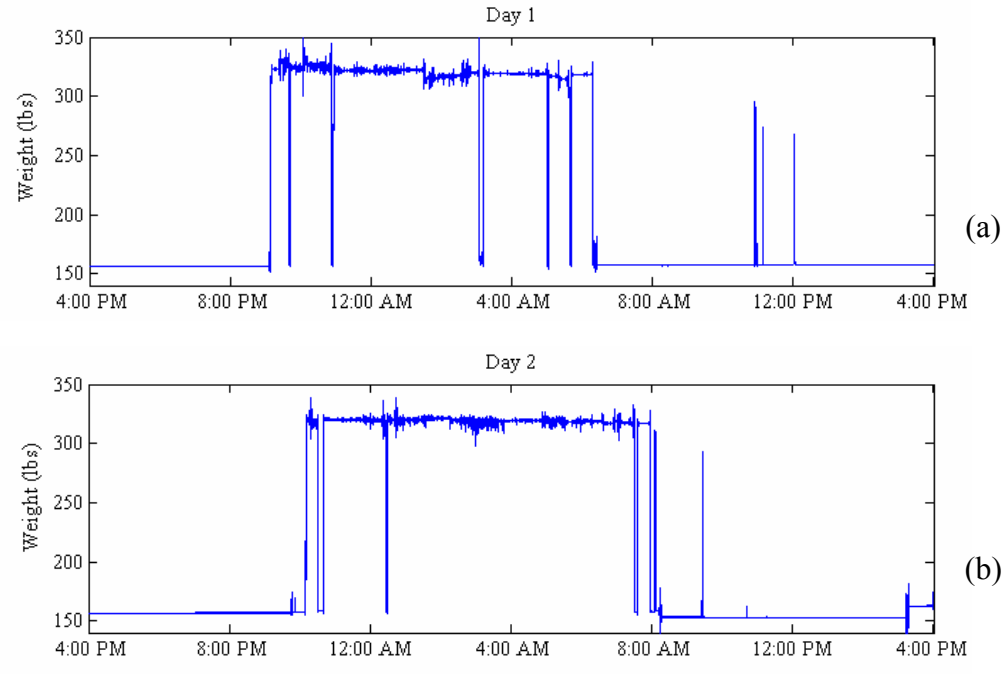


Figure 3.7: Output of load cells for two 24-hour periods.

The time periods when there is a person lying or sitting on the bed are referred as *in-bed* events; otherwise, they are referred as *out-of-bed* events. A visual inspection of Figure 3.7 shows some of the issues associated with estimating the in-bed and out-of-bed

events. There are several possible times at which the person may be considered to be in bed (e.g., around 9:00 PM or 9:30 PM in Figure 3.7 (a), and around 10:00 PM or 11:00 PM in Figure 3.7 (b), and therefore it is difficult to decide about the bedtime. Also, there are several possible times at which the person may be considered to be out of bed (e.g., around 5:00 AM, 6:00 AM or 7:00 AM in Figure 3.7 (a), and around 7:30 AM or 8:00 AM in Figure 3.7 (b), and therefore it is difficult to decide about the get up time.

In order to interpret such data, we develop an algorithm for computation of sleep-related parameters. Since inter-individual variability in rest-activity patterns is high, separating nighttime and daytime data on a per-person basis simplifies the computation of the sleep-related parameters. A person's *in-bed profile* that gives a good approximation of the person's habits in terms of being in bed or not during any time of the day is used to estimate the nighttime and daytime periods in an individual basis. Because the time of lights off is often used as an indicator of the time a person decides to sleep [63], we also incorporate the information about the use of the bedroom‘s light into the decision about bedtime and get up time. The algorithm is described in Chapter 5.

3.4 Summary

This chapter introduced a system for unobtrusive assessment of mobility in bed with load cells. Mobility in bed is characterized by the periods of postural immobility, movement times, and the types of movement performed. The goal of the system is to identify the time intervals when a movement in bed occurs and to determine the type of movement performed in a given time interval. The presented system provides a description of the motor pattern during sleep in terms of the time of onset, duration, frequency, and the type of movement. Earlier work in this area is mostly focused on the detection of movement, and simpler classifications systems have been proposed. The advantages of this approach are that the measurements can be done using a person's existing bed, and that it does not require a special mattress or sheet. In addition, it utilizes existing technology with relatively low cost.

The system uses four load cells under the bed, one at each corner. Load cells are strain gauge transducers that convert applied force into a resistance change. Forces are

not directly measured, but inferred from the resultant strain. The output is linearly proportional to force, and the relationship is determined by calibration.

The weights sensed by the load cells at the corners of the bed represent the instantaneous distribution of the mass of the body. To infer about movement, we need to determine if the system response due to the bed can be inverted. The mapping from the body mass distribution to the load cell signals results in a substantial loss of dimensionality. In addition, the system is nonlinear, and the inverse problem cannot be addressed by blind deconvolution or similar techniques. Therefore, we used a pattern recognition approach treating the system as a black box and the signals as stochastic processes whose parameters depend on the movements.

Since the total weight on the bed drops when someone gets up, this information can be used to estimate sleep-related parameters related to a person's sleep habits. The algorithm is described in Chapter 5. It incorporates information about a person's in-bed profile with information about the use of the bedroom‘s light into the decision about bedtime and get up time. The experimental setup for collection of load cell, light and actigraphy data used in this work is described in Chapter 4.

Chapter 4

Data Collection

This chapter describes the experimental setup for collection of load cell, light and actigraphy data used in this work. Data were collected from three experiments: two performed in a laboratory and one in an assisted-living facility. The chapter is organized as follows. Sections 4.1 and 4.2 describe the experimental setup for laboratory data collection with a twin bed size and a full bed size, respectively. Different beds were used to provide data for testing of the approaches for detection (Chapter 6) and classification (Chapter 7) of movements. In Section 4.3, we describe the experimental setup for data collection at an assisted-living facility. These data are used for estimation of the sleep-related parameters and a comparison with sleep diaries, and also for demonstrating the feasibility of the approach in a realistic setting, as presented in Chapter 5.

4.1 Experiment 1: Twin Size Bed

The goal of this experiment is to provide data for evaluating the proposed system for detection and classification of movements in bed. Data were collected in the laboratory, using a twin size bed (size 39”x 75”) with a box spring mattress (as shown in Figure 3.3). This dataset is going to be referred as LAB1_TWINSIZE. The details about the subjects involved in the study, the protocol, and the sensors used are described next.

4.1.1 Subjects

Nine adults (4 men and 5 women), age ranging from 24 to 45 years (mean age 32.1 ± 6.5 years old) participated in the study. The inclusion criteria for the study are: 21 years old or older, and with no mobility problems. Each subject signed a consent form, and received a compensation of $20.00 for his/her participation. Table 4.1 shows the description of the subjects. The experimental session lasted approximately 2 hours.

Table 4.1: Description of Subjects

Subject Number	Age	Sex	Weigh (lbs)	Height
1	31	M	275	6' 1"
2	33	F	115	5' 1"
3	27	F	85	5'
4	27	F	108	5'
5	24	M	178	6'
6	39	M	175	5' 11"
7	45	F	150	5' 9"
8	30	F	125	5' 4"
9	33	M	160	5' 11"

4.1.2 Data Collection Protocols

Because the subjects were awake during the experiment, data were collected using two different protocols, *free movement* and *fixed movement,* to allow both diversity and uniformity of movements. The main difference between these two protocols is that the latter requires the subject to perform a pre-determined set of movements in bed.

In the *free movement protocol,* each subject was asked to lie in bed and freely move 10 times. Subjects were instructed to move accordingly to the types of movements typically seen during sleep. The movement should start only after hearing a beep sound. After the beep sound, subjects had approximately 15 seconds to perform a movement and then to rest in a still position. This part of the experiment lasted about 6 minutes.

In the *fixed movement protocol,* each subject performed 5 trials composed of 20 pre-defined movements each, done in different order in each trial. Before each trial, the subjects were instructed to follow the instructions played through a speaker and to move

only after hearing a beep sound. During the trial, every 15 seconds, a recording instructed the subject to perform a movement at the beep sound. After the beep sound, subjects had approximately 15 seconds to perform a movement and then to rest in a still position. Each trial lasted about 8.5 minutes.

To reduce errors during the experiment, subjects were instructed about how the final positions should look like by means of sketches or by having the researcher showing the final positions in the bed. The subjects were also encouraged to move as they typically would move in bed (i.e., make the movements in a natural and comfor way), but follow closely the sequence played in the recording.

Each of the 5 trials performed is composed by a different combination of the movements set in Table 4.2. The chosen set of pre-defined movements is composed by 6 large movements of torso and limbs (posture shifts) and 14 small movements (6 isolated movements of the head or arms, and 8 leg movements), as shown in Table 4.2. According to the second column of Table 4.2, each large movement is followed by one or more small movements, and the complete description of each group of movement is shown in the third column.

Table 4.2: Set of small and large body movements chosen for the study.

Group Number	Posture Shift and Small Movement	Group of Movements
1	Back (supine) to right (lateral position), leg movements	Move from back to right, with arms and legs bent. Straighten legs. Bend left leg. Straighten left leg. Bend legs.
2	Back (supine) to left (lateral position), leg movements	Move from back to left, with arms and legs bent. Straighten legs. Bend right leg. Straighten right leg. Bend legs.

Group Number	Posture Shift and Small Movement	Group of Movements
3	Left (lateral position) to right (lateral position), arm movements	Move from left to right, with arms bent and straight legs. Straighten left arm. Bend left arm.
4	Right (lateral position) to left (lateral position), arm movements	Move from right to left, with arms bent and straight legs. Straighten right arm. Bend right arm.
5	Right (lateral position) to back (supine), head movement	Move from right to back, with straight arms and legs. Turn head to the right.
6	Left (lateral position) to back (supine), head movement	Move from left to back, with straight arms and legs. Turn head to the left.

Not all permutations of the above set were allowed because most of them cannot be physically performed. Since every trial started from the back (supine) position, and that no repetitions were allowed, there were only 6 permutations that can be executed. To generate 5 trials for each subject based on 6 permutations, we randomly permuted the order in which the subject would do the trials and remove the sixth trial, which protects against possible bias. The determination of the trials performed by each subject was done ahead of time for all subjects, and it removed any ordering effect.

The selected set of movements performed in the data collection is supported by different movement descriptions found in the literature [4, 16, 25]. As shown in Table 4.3, all the movement descriptions found in the literature include major postural shifts because these gross body movements occur quite consistently throughout the night [4]. They also include isolated movements of the head, arms, and legs.

Our set of movements is also supported by Carlson's work [64]. In her work, Carlson introduced a series of movements in bed designed to simulate those typically observed during sleep. Some of such movements consist of turning to the right or left and then supine (back), bending and straightening the legs, and bending and straightening the arms.

Table 4.3: Movement Descriptions

Study	Movement Description
Aaronson [16]	Major posture shift – trunk rotation of at least 45° or a displacement of at least 3 limbs Postural immobility – absolute motionlessness or small movements limited to 1 or 2 limbs
Muzet [4]	Type I – movements without displacement that affect extremities of the body Type II – movements that affect only one part of the body, and modifies its position Type III – large postural body changes
Wilde-Frenz [25]	Type I – isolated movement of either the head, trunk or the limbs Type II – an integrated activity of more distant parts of the body such as head and the trunk or the limbs Type III – major posture shifts

4.1.3 Sensors

Two types of sensors were used to collect data from each trial: load cells and actigraph. Data from load cells under the bed were collected at 200 Hz (as described in Section 3.1.2). Load cell data were downsampled to 10 Hz because the energy of the load cell signal for the set of movements performed is most concentrated below 5 Hz, as shown in Figure 4.1. Also, because this work only included assessment of voluntary movements, which rarely exceeds 3-4 Hz [12], this choice is appropriate but does not impede that higher sampling rates can be used when analyzing involuntary movements.

Actigraphy data were collected simultaneously with load cell data, and an analysis of the data is presented in Section 6.3.6. Because it is a common practice to place the actigraph on the non-dominant wrist in adults [65], subjects wore a wrist-actigraph on the non-dominant wrist. The actigraphy monitor used in this study is Actiwatch64 (Mini-Mitter Company, Inc., Bend, Oregon [66]). With each movement of the wrist, an accelerometer inside the actigraph generates a variable voltage that is sampled at a frequency of 32 Hz. The signal is integrated over a user-selected epoch, and a value expressed as “activity counts” is recorded on local memory. The epoch used in the study is 15 seconds. For an epoch of 15 seconds, the recording time is 11.3 days, i.e., the

actigraph fills its memory with samples after that period. A start time for data collection must be set, based on PC clock. Therefore, for time synchronization of the load cells and actigraph data, the same computer was used for load cell data collection and actigraph setup.

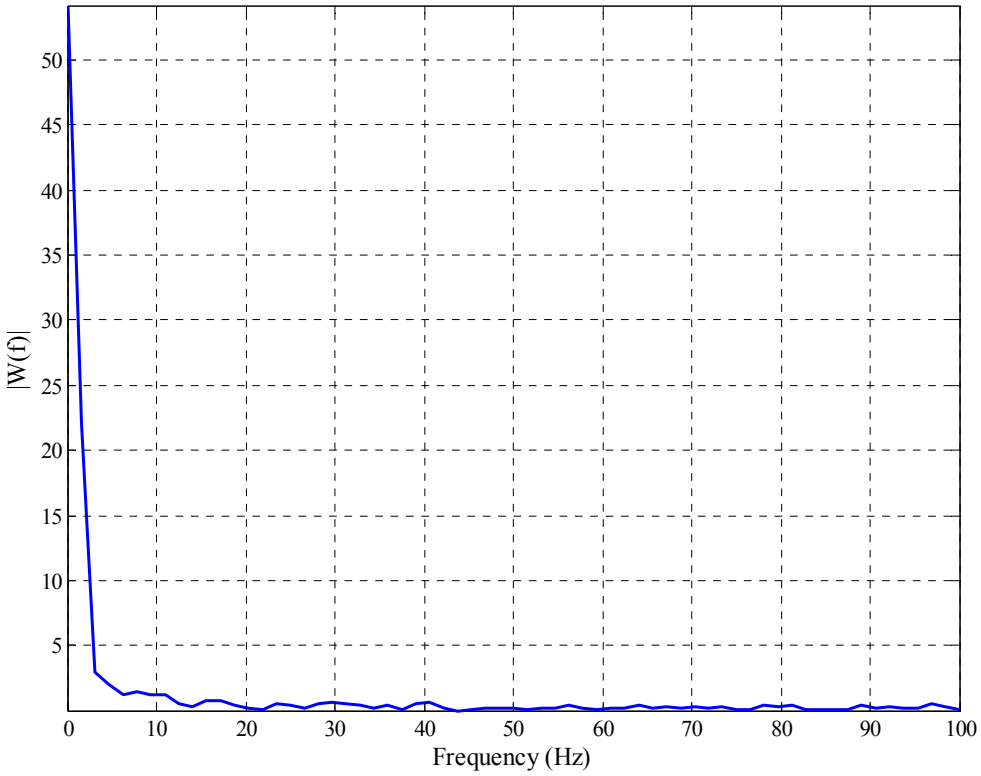


Figure 4.1: Average magnitude spectrum of the total weight |W(f)|. Horizontal axis shows the frequency from 0 to 100 Hz, and vertical axis is the magnitude of the average Fourier transform.

4.1.4 Assessment of Actual Movements

We used a video technique as the ground truth for this experiment. A Creative web cam NX Ultra camera was mounted on the ceiling, 2 meters above the bed, to record images of the whole bed. Uncompressed 480 x 640 RGB images were recorded at a rate of 10 frames per second, simultaneously with the load cells and actigraph data, for offline analysis.

To allow a quantitative measure of body movement using video, subjects wore cloth bands of different colors on the head, arms, legs, and torso. The actual movement intervals were estimated by tracking the trajectories of the cloth bands. Figure 4.2 shows an example of the camera's view, and the cloth bands used for the experiment.

Figure 4.2: Setup used during experiment: view of the camera. Subject worn cloth bands on the head, torso, arms, and legs.

The location of every cloth band, consequently the respective part of the body, is estimated using template matching [67]. The normalized cross-correlation is computed to find the closest match between a given template and an image. The region of the image with highest correlation is defined as the closest match. First, templates from each cloth band were extracted from the first frames of the video, when the subject is lying on his/her back with straight legs and arms. Then, the template matching is performed for each type of cloth band. The location of each cloth band is determined by the location in the image with highest cross-correlation. Details about the analysis of video data are presented in Appendix B.

4.2 Experiment 2: Full Size Bed

The goal of this experiment is to provide data from a different bed setting for evaluation of the proposed system for detection and classification of movements in bed. Data were collected in a laboratory, using a full size bed (size 54”x 75”) with a box spring mattress. This dataset is going to be referred as LAB2_FULLSIZE. The experimental setup and procedure were identical to the previous experiment.

4.2.1 Subjects and Data Collection Protocol

Data was collected from five subjects from the LAB1_TWINSIZE dataset and six new subjects. The goal for collecting data from previous subjects on a new bed is to test the generalizability of the approaches proposed in this work. The previous subjects (1, 2, 3,

4, and 8, as shown in Table 4.1) repeated the 5 fixed-movement trials on the full-size bed.

The new subjects are six adults (3 men and 3 women) with age ranging from 22 to 35 years (mean age 27.8 ± 4.7 years old), as shown in Table 4.4. They performed one trial of the free-movement protocol plus ten trials of the fixed-movement protocol. Therefore, this group of subjects has more data than the group of subjects described in Section 4.1.1. The reason for increasing the number of trials in this experiment was that we wanted to have a group with a larger dataset to determine the minimum training set size necessary for movement classification, as explained in Section 7.4.4.

Table 4.4: Description of Subjects.

Subject Number	Age	Sex	Weigh (lbs)	Height
10	22	F	150	5'5''
11	25	F	175	5'4''
12	30	F	125	5'6'
13	35	M	140	6'
14	30	M	140	5'10''
15	25	M	150	5'10''

4.3 Elite Care Study

In this section, we present the experimental setup for actigraph, light and load cell data collection in an assisted-living facility. The goal of this experiment is to provide data for estimation of the sleep-related parameters, and for demonstrating the feasibility of the movement detection approach in a realistic setting.

4.3.1 Subjects

The subjects were residents of Elite Care, an assisted-living facility located in Milwaukie, Oregon[iii]. Each subject and/or guardian his/her signed a consent form, and received a compensation of $20.00 for his/her participation. The inclusion criteria for the study are:

- 60 years old or older.
- Living alone.

[iii] http://www.elite-care.com

- With no mobility problems such as paralysis of the legs.
- Not acutely ill and independent ambulatory.
- With no physical condition that affects the use of an actigraph on one wrist.

Despite the efforts for recruiting subjects, only 2 subjects completed the study. A total of 5 candidate subjects met the inclusion criteria but did not contribute data because of the following reasons: withdraw from the study, moving, and cognitive impairment that interfere in the compliance with the use of the actigraph.

Each subject was assessed at the beginning of the study. The assessment included a brief physical examination and medical history review to verify that the subject is not acutely ill and independent ambulatory, and that he/she has not a physical condition that affects the use of an actigraph on one wrist. It also included the following tests: the Mini-Mental State Examination (MMSE) [68], the Cornell Depression Scale [69], and the Sleep Disorders Inventory [70]. These tests may help in making distinctions of diagnosis in case one needs to study on reasons for changes on level of activity that may be seen in the data. The Mini-Mental State Exam, ranging from 0 (worst) to 30, describes general cognitive functioning. The Cornell Depression Scale, a depression measure that has been validated in both demented and non-demented subjects, includes 19 items covering mood or affect signs, behavioral disturbances, physical symptoms of depression, sleep and diurnal symptoms, and depressed thinking. Higher scores indicate greater levels of depression. The Sleep Disorder Inventory is a 2-week history of seven sleep disturbance symptoms, occurring at least once weekly, as reported by the caregiver. Table 4.5 shows the description of the subjects, and their scores.

Table 4.5: Subjects Description and Scores

Subject Number	Age	Sex	Weight (lbs)	Height	MMSE Score	Cornell Depression Scale
1	92	M	202	5’ 8’’	17	4
2	88	F	220	5’ 2’’	21	6

Each subject lives in a one-bedroom apartment. There are two caregivers during the day shift that take care of up to 12 residents, and one caregiver at the night shift.

Because the subjects are very independent and usually do no need help at night, the caregiver does not make regular nighttime checks at night for these residents.

4.3.2 Sensors and Protocol

Three types of sensors were used to collect data for this experiment: load cells, actigraph, and light switch. Data were collected over a period of three weeks. Load cell data were collected for the whole 3 weeks, but actigraph and light switch data were collected only for the last 2 weeks. The choice of 2 weeks was based solely on compliance of the use of the actigraph. The first week of load cell data was used for estimating the in-bed profiles, according to Section 5.1.1.

At Elite Care, the setup in place for load cell data collection is the following: load cells are connected to a Programmable Logic Controller (PLC), which is connected to a computer through a serial port. Data are collected at 10 Hz from 4 load cells placed under the beds. The subjects worn a wrist-actigraph (the same one as described in Section 4.1.3), and the epoch used in the experiment was 15 seconds. Subjects were instructed to wear the actigraph at all times, including when taking a shower. Caregivers were asked to monitor compliance of actigraph use, by writing the times the actigraph was taken off.

The use of the bedroom light, by means of the times when it is turned on and off, is recorded by the facility through the switch located in the entrance of the bedroom. Such information was also available, even thought one of the subjects only used the lights in the bedroom a few times.

Caregivers were asked to complete a daily sleep diary during the study, including observations of bedtime, lights out, lights on and morning arising, in addition to comments regarding nighttime awakening and sleep quality. They received detailed instructions about the completion of the sleep diary from a nurse, prior to the beginning of the study. The sleep diary created for the study is shown in Appendix C. Because the caregivers work in shifts, the caregiver responsible for the evening shift completed the questions related to bedtime (questions 1 to 4) , and one of the caregivers responsible for the day shift completed the questions related to wake up time (questions 5 to 8). They were instructed to include in the ‘Comments’ section any comments that subjects made about their sleep, and any unusual situation that may have occurred during the night.

Chapter 5

Sleep-Related Parameters Estimation

This chapter describes the algorithm (Section 5.1) for estimation of sleep-related parameters. Section 5.2 shows the results on data collected from 2 subjects in an assisted-living facility (as described in Section 4.3). The estimated parameters are compared with the estimates from sleep diaries completed by caregivers for a period of 2 weeks.

5.1 A Methodology for Estimation of Sleep Behaviors

As previously discussed in Section 3.3, there are issues associated with estimating bedtimes and get up times from the load cell outputs. For example, as it was illustrated in Figure 3.7, there are several possible times in Figure 3.7 (a) at which the person may be considered to be in and out of bed, and therefore it is difficult to decide about bedtimes and get up times. One way to resolve the situation is to first determine, based on a person's habits, the periods corresponding to nighttime and daytime.

The algorithm for estimating the sleep-related parameters is composed of two parts. In the first part, the algorithm computes the *in-bed profile* of a person, and extract nighttime and daytime periods. In the second part, it estimates the parameters by analyzing the information at nighttime and daytime separately. To recapitulate, from Section 1.1, the parameters are:

1. Bedtime: the time of retiring, when a person goes to bed with intent of sleeping;
2. Get up time: the time when the person leaves bed with intent of starting the day;

3. Total time in bed at nighttime (TIB): total time of bed rest derived from bedtime and get up time, and by subtracting the time spent out-of-bed between bedtime and get up time;
4. Duration of longest "in-bed" period at nighttime;
5. Number of times person gets up during the night;
6. Total time "out-of-bed" at nighttime (TOB);
7. Number of times person goes to bed during daytime;
8. Total time in bed at daytime.

Section 5.1.1 describes the algorithm for estimation of nighttime and daytime periods. The estimation of bedtime and get up time parameters is described in details in Section 5.1.2. The estimation of the remaining parameters is straightforward after estimating bedtime and get up time.

5.1.1 Daytime and Nighttime Periods Estimation

Separating nighttime and daytime data simplifies the computation of the sleep-related parameters. Thus, the in-bed profile is used to estimate the nighttime and daytime periods in an individual basis. The in-bed profile is defined as a function that represents the proportion of times the person is in bed at each instant *m* of the day. We hypothesize that, over a representative period of time, *N* (in days), the profile $P_d(m)$ for a given day *d* gives us a good approximation of the person's habits in terms of being in bed or not at each instant *m* of the day. Since the goal of this work is to estimate parameters related to the nighttime period, we define day as the period between 4:00:00 PM of the current day and 3:59:59 PM of the next day (a total of 24 hours). This way we guarantee that each day includes the nighttime period. Given the within-subject variability in the sleep/wake cycle due to factors such as amount of prior sleep [71], a period of one week is frequently used as a representative period for studies on sleep patterns and on changes in circadian rhythms [3, 71]. Therefore, for each day *d*, the in-bed profile $P_d(m)$ is estimated as the average of the in-bed events at each instant *m* over the previous 7 days (i.e., $N = 7$). Therefore, each element of P_d is a value inside the interval [0, 1] that represents the proportion of times the person is in bed at each instant *m*, and is given by

$$P_d(m) = \frac{1}{N} \sum_{i=d-N}^{d-1} v_i(m), \tag{5.1}$$

where v_i, *a* binary vector, represents the state (out-of-bed and in-bed) of a person at each instant *m*. The v_i vector is estimated by a thresholding operation on the load cell data for a day *i*, where a zero value indicates that a person is out of bed, and one indicates that a person is in bed at each instant *m*. Since load cell data (Equation 3.2) are basically composed of the bed weight values and the bed-and-person weight values, the threshold is computed as the midpoint between the averages of the bed weight and bed-and-person weight distributions over the day *i*.

The in-bed profile is used to estimate the nighttime and daytime periods in an individual basis as follows. To estimate these periods, let's assume that 4:00 AM represents nighttime for all individuals, and calculate as follows:

1. **Find the time interval when profile P_d is above a certain threshold and that contains the 4:00 AM event:** the time corresponding to the beginning of the interval is assumed to be the beginning of the nighttime period. The end of the nighttime period corresponds to the time when the interval ends.
2. **The daytime period is defined as the period, between 24-hour periods (4:00:00 PM - 3:59:59 PM of the next day), that does not include the nighttime period:** for example, when a person's nighttime period corresponds to 9:02 PM until 7:34:59 AM, then the daytime period corresponds to 4:00:00 PM until 9:01:59 PM and from 7:35 AM until 3:59:59 PM.

5.1.2 Bedtime and Get Up Time Estimation

The decisions about the sleep-related parameters are always difficult in the absence of a ground truth. Logical, but arbitrary rules are necessary to make the method a consistent approach, avoiding the biases of subjective guesses. Because the time of lights off is often used as an indicator of the time a person decides to sleep [63], bedtime and get up times are estimated by using the person's profile and information about the use of the bedroom's light (times when the light is turned on/off). In the data used for development of the algorithm, the light switches are located in the entrance of the bedrooms (not accessible when the person is in bed).

The algorithm for estimating the bedtime and get up times can be divided into 4 steps:

1. **Find the set of in-bed events for a given night**: this step is the same as the one used to estimate the in-bed events for the profile. The number of in-bed events varies per night (depending on how many times the person gets up at night), and a rule is used to eliminate very short in-bed events. Since short in-bed events may not be reliable for estimating the rest period because they may not correspond to events where the person had the intent of falling sleep, all in-bed events that last less than 20 minutes are discarded. Our motivation for the "20-minute rule" comes from Cole's work [72], where the sleep onset, in actigraph records, is defined as the beginning of the first interval containing 20 minutes scored as sleep with no more than one minute of wakefulness intervening.
2. **Select the candidate events for nighttime rest period from the set of in-bed events**: because the person's profile provides an approximate estimation of the average in-bed period, the person's profile is used to select the candidate events. The candidates are the events that intersect the person's profile.
3. **Refine the set of candidate events by incorporating daily changes in the rest period**: even though the person's profile represents the most common time period of rest, the person may change the time she/he goes to bed or to gets up in a particular day. Therefore, such changes must be incorporated in the selection of candidate events for the rest period. Two factors are used to add in-bed events that occur around of the set of candidate events: time proximity (i.e., the closer the event is to the set of candidates, the more likely the event belongs to the set of candidates) and the light sensor (i.e., the time when the light is switched off can indicate intent of going to bed, or the time when the light is switched on can indicate intent of getting ready for the day). The set of candidate events changes as long as the rules are satisfied:
    a. Insert the in-bed event that occurs before the current candidate set into the beginning of the set if either one of the following rules is satisfied:

i. If the time elapsed between the event in question and the earliest in-bed event of the candidate set is equal or less than 5 minutes. For example, in Figure 5.1 (a), the in-bed event before the set of candidate events is inserted in the beginning of the set because the time elapsed is less than 5 minutes. However, in Figure 5.1 (b), the same does not happen for the in-bed event preceding the candidate events because the time elapsed is greater than 5 minutes.

ii. If there is a light event before the event in question and the time elapsed between the event in question and the earliest in-bed event of the candidate set is equal or less than 10 minutes.

b. Insert the in-bed event that occurs after the candidate set into the end of the set if either one of the following rules is satisfied:

i. If the time elapsed between the event in question and the latest in-bed of the candidate set is equal or less than 5 minutes. For example, in Figure 5.1 (a), the in-bed event after the set of candidate events is not inserted into the set because the time elapsed is greater than 5 minutes.

ii. If there is a light event after the event in question and the time elapsed between the event in question and the latest in-bed of the candidate set is equal or less than 10 minutes. For example, in Figure 5.1 (b), the in-bed event after the set of candidate events is inserted into the end of set of candidates because there is a switching on lamp event after the in-bed event in question, despite the fact that the time elapsed is greater than 5 minutes.

The 5-minute threshold was determined empirically from the data based on the average elapsed time between in-bed events without a light event. The use of light gives information about the person's intention to go to bed to try to fall asleep or the intention of getting up to start the day. Therefore, the threshold can be higher than the threshold used in the absence of this additional piece of

information. In the presence of light information, we chose to use a threshold that is twice the threshold used when there is no light event.

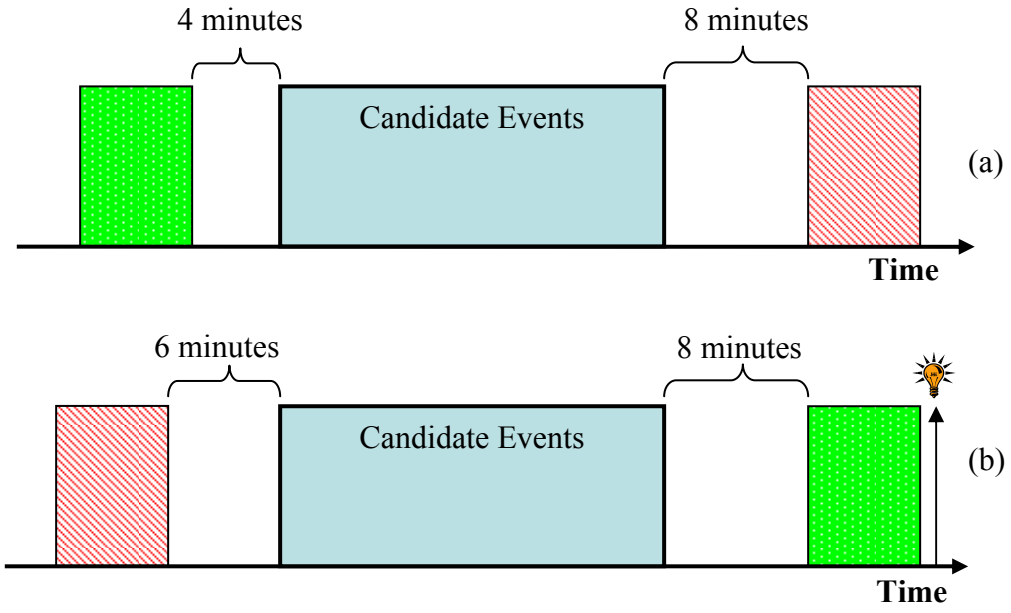


Figure 5.1: Example of how rules are applied to refine the set of candidate events for the rest period. The large rectangle on the middle of the time axis represents the set of candidate events. The blocks before and after the candidate events represent the in-bed events being analyzed by the algorithm. A switching on lamp event is represented by a lamp in (b).

4. **Estimate the bedtime and get up time from the set of candidate events.** The bedtime is assumed to be the beginning of the earliest in-bed candidate event. The get up time is assumed to be the ending time of the latest in-bed event candidate.

All the remaining sleep-related parameters can be directly obtained at this point.

5.2 Results

In this section, we show the results of the sleep-related parameters estimation on load cell and light data collected from 2 subjects in the study described in Section 4.3. The subjects are two residents of an assisted-living facility, and sleep diaries (see Appendix C) were completed by the caregivers. We show the results of each step of the algorithm and compare the estimates of some of the sleep-related parameters derived from the algorithm with the caregiver's reports.

The in-bed profiles of the two subjects, obtained by using Equation 5.1 and estimated over the first week of the study, are shown in Figure 5.2. The horizontal line shows the value of the threshold. In this work, the threshold is defined as 0.5. That is, according to Equation 5.1, for $N = 7$ a threshold of 0.5 indicates that the person was in bed, at each instant m, for at least 3.5 of the previous 7 days, and thus the time period is considered nighttime. The profiles illustrate very well how the rest-activity patterns differ among people. For subject #1, the profile starts at 10:24 PM, and ends at 7:18 AM. It also shows that the subject frequently went to bed at daytime during the first week of the study. For subject #2, the profile starts at 9:02 PM, and ends at 6:45 AM. It shows that this subject never went to bed at daytime during this period.

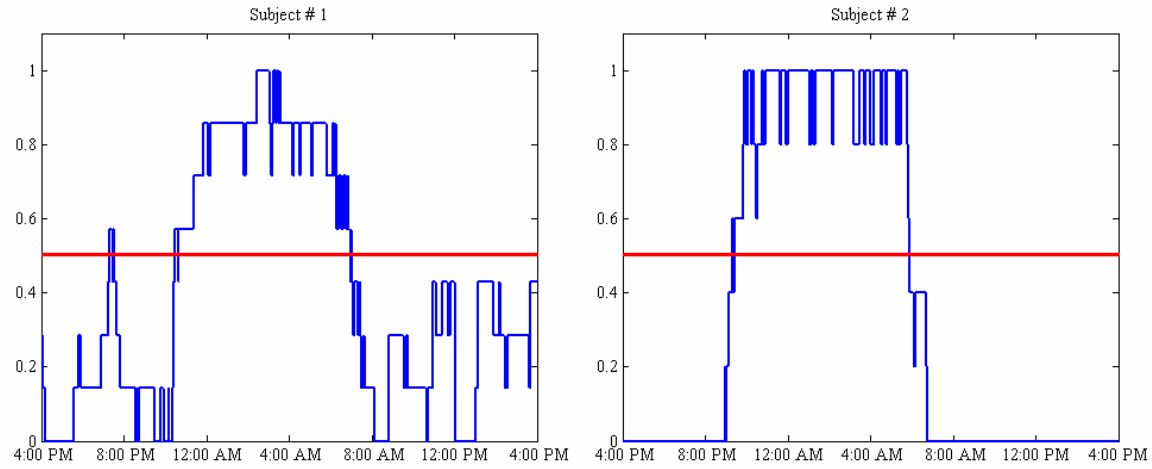


Figure 5.2: In-bed profiles $P_d(m)$ of the subjects estimated over the first week of the study. The horizontal line shows the value of the threshold. From the profile, it is visually clear that subject # 1 has more variability in his/her patterns than subject #2.

Using the algorithm described in Section 5.1.2, sleep-related parameters are estimated for the last 2 weeks of the study. Figure 5.3 shows plots of the sleep-related parameters, numbers 3 to 8, for subject #1. In this example, the average time in bed at nighttime is 6.89 (± 2.19) hours, and the average time spend in bed at daytime is 1.86 (± 1) hours. For the period of the study, this person spent several hours in bed during daytime, and got up from bed during nighttime at least once on 11 (out of 14) days. According to the sleep diaries reports completed by the caregivers, these two weeks were typical, with no noticeable changes or complaints.

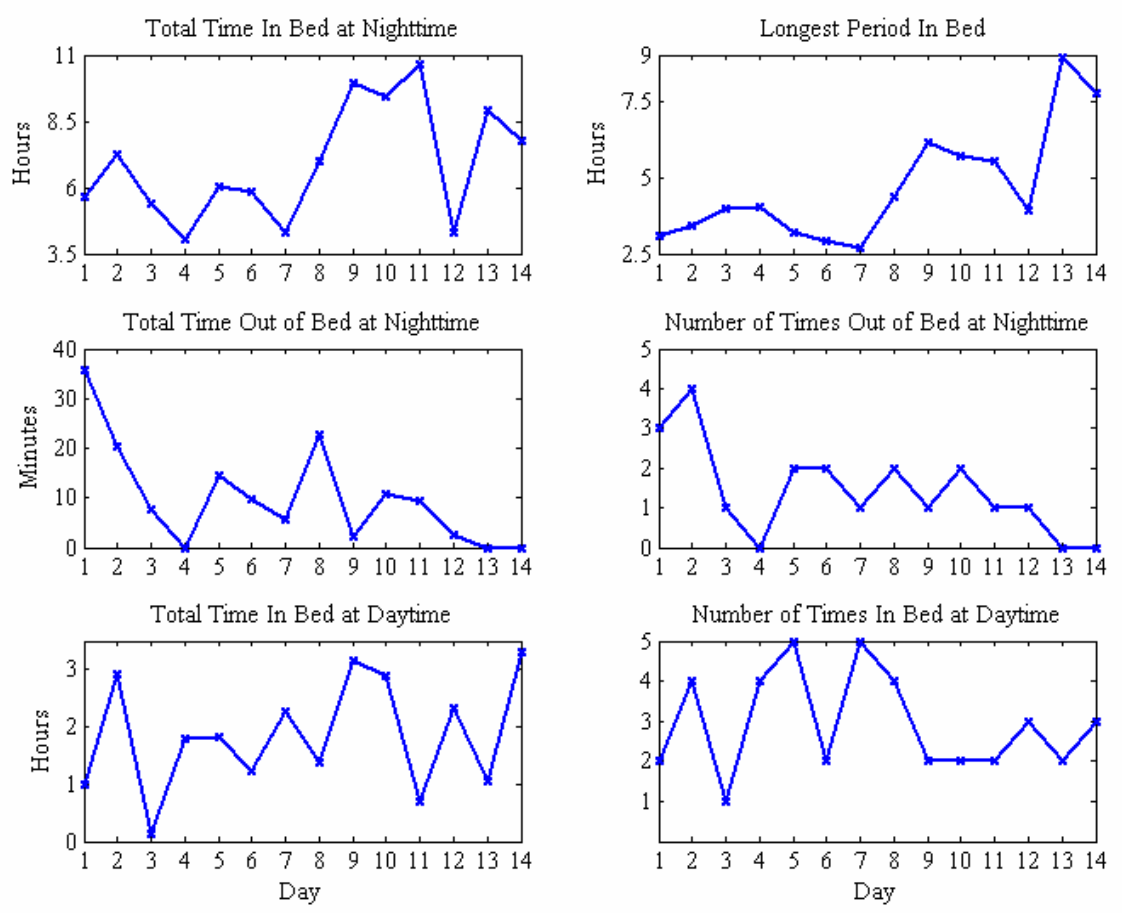


Figure 5.3: Plot of sleep-related parameters for subject # 1, for 14 consecutive days.

Unfortunately, it is difficult to have caregivers completing sleep diaries for long periods, and a lot of information is missing in the diaries. Table 5.1 shows a summary of the sleep diary completion for each subject.

Table 5.1: Summary of completion of sleep diaries.

		Subject	
		#1	**#2**
Bedtime (Question # 1)	Completed	4	8
Get up Time (Question # 6)	Completed	5	10
Times Out of Bed at Nighttime (Question # 4)	Completed	2	1
	Don't Know	1	7

According to Table 5.1, from a total of 14 answers (corresponding to 14 nights) for each question, only 28.6% of the answers were provided for bedtime (question #1), for subject #1, and 57.1% for subject #2. A similar trend was shown for the getup time

(question #6). For the number of times out of bed at nighttime (question #4 in the diary), the caregivers had also the option to answer 'Don't Know', and the answers for this question were very poor.

A boxplot of the differences, in minutes, between sleep diary and the algorithm estimates of bedtime and getup times for the two subjects in the study is shown in Figure 5.4. The discrepancy between the objective measures from the sensors and the subjective measures from diaries is larger for subject #1. However, because of the low compliance in the completion of the diaries for subject #1 (only 28.6% of the questions were answered), and because we only have 2 subjects, we can only speculate that the discrepancy is due to the fact that subject #1 has less consistent sleep habits. The lack of a sleep routine is consistent with bedtime algorithm estimates varying from 10:02 PM to 1:59 AM and get up times varying from 4:01 AM to 9:22 AM, which makes more difficult to caregivers to reliably estimate bedtimes and get up times without constantly supervising the person. Subject #2 has a more consistent bedtime routine, which makes easier to caregivers to reliably estimate bedtimes and get up times.

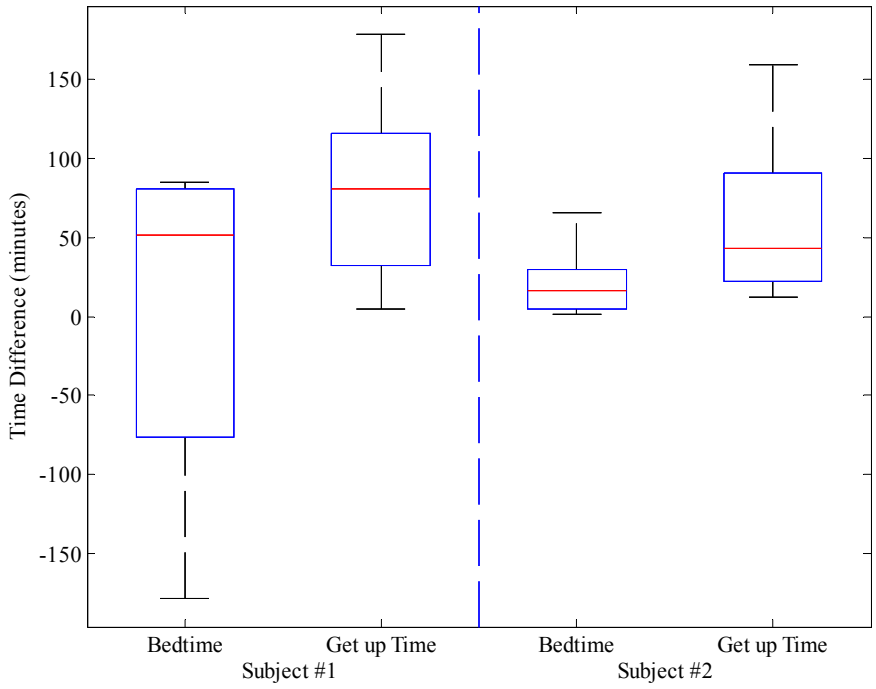


Figure 5.4: Boxplot of the time differences, in minutes, between sleep diary and algorithm estimates of bedtime and getup times for subjects #1 and #2.

We did not compare the algorithm and sleep diaries estimates of the number of times out of bed at nighttime (question #4 in sleep diary) because there were only a few answers (only 10%). From the algorithm estimates, subject #1 got up from the bed at nighttime, on average, 1.4 times (±1.2), and subject #2 got up 3.4 times (±1.7) on average. This suggests that many out-of-bed events were unnoticed by the staff.

The discrepancies between the objective measures from the sensors and the subjective measures from diaries are expected because there is always some difficulty in recollecting exact bedtimes or the number and length of periods out of bed during the night [42]. In addition, there may be errors in the algorithm estimates because of the thresholds (5, 10 and 20-minute rules in Section 5.1.2) used.

The results thus indicate that the use of this methodology for determining sleep-related parameters estimates is valuable, especially in facilities where people live almost independently and a lot of information may be missing in the sleep diaries. It is also valuable in sleep studies that include populations who would not be able to remember specific hours of sleep to complete sleep diaries, and in other situations where the use of sleep diaries or actigraphs is not feasible.

5.3 Summary

This chapter presented an algorithm for estimation of sleep-related parameters. Such algorithm is composed of two parts. In the first part, we compute the in-bed profile of each person. The profile represents the proportion of times the person is in bed at each instant of the day over a given period. Over a given period, we hypothesize that the profile gives us a good approximation of the person's habits in terms of being in bed or not during any time of the day. A period of one week is frequently used as a representative period for studies on sleep patterns and on changes in circadian rhythms [3, 71]. For any given day, the in-bed profile is estimated as the average of the in-bed events over the previous 7 days. Using a thresholding operation, the algorithm extracts the nighttime and daytime periods from the profile. Separating nighttime and daytime data on a per-person basis simplifies the computation of the sleep-related parameters. In the second part, the algorithm relies on simple rules to approximate the bedtime and get up times, by using the person's profile and information about the use of the bedroom‘s

light (times when the light is turned on/off). The estimation of the remaining parameters is straightforward after the bedtime and get up time are approximated.

We compared the estimates of bedtime and get up time from load cell and light data with the sleep diaries reports in a 2-week study with 2 subjects. The subjects are residents of an assisted-living facility, and sleep diaries were completed by the caregivers. The discrepancies between the objective measures from the sensors and the subjective measures from diaries are expected because there is always some difficulty in recollecting exact bedtimes or the number and length of periods out of bed during the night [42]. In addition, there may be errors in the algorithm estimates because of the imposed thresholds. The threshold values were based on the data of 2 people only, and based on a rule used in actigraphy for determining sleep onsets.

The importance of this type of continuous assessment is reinforced by the fact that a large amount of information can be missed in the sleep diaries. It is more difficult to the staff to know the sleep routines of residents that very independent and do not need constant attention. However, the fact that these residents cannot be under constant supervision may result in problems being unnoticed by the caregivers. It can also be valuable in sleep studies in populations who would not be able to remember specific hours of sleep to complete sleep diaries.

In addition, bedtime and getup times can be assessed for long periods of time to examine changes in sleep habits due to seasonal variations in the duration of daylight. Also, studies about changes in circadian patterns of rest and activity in the elderly population, which have been done with the use of actigraphs [41], could be benefited from this approach because the subject does not have to wear any device.

Chapter 6

Detection of Body Movements in Bed

This chapter describes the framework for movement detection from load cell data. Section 6.1 formulates the movement detection framework as a problem in statistical hypothesis testing, and the approach is described in Section 6.2. Section 6.3 presents results related to parameter optimization and detection performance evaluated on laboratory data. We also show results of the detection approach on the load cell data collected at the Elite Care study. In addition, a comparison of load cell and actigraphy with respect to the detected movements is also presented in Section 6.3.

6.1 Statistical Framework

The problem of detection of movement in bed consists of estimating the time intervals when a movement in bed occurs. Although the task consists of segmenting the load cell data into time intervals corresponding to periods of movement and no movement, it can be seen as a detection task in the sense that it determines whether someone is moving or not at a given time t. Given a measurement f from the load cells under the bed sampled at time t, which is referred to as a feature value, the goal is to decide if a movement has occurred or not at time t. The problem can be formulated as a statistical hypothesis testing problem with two mutually-exclusive hypothesis:

H_0: a movement has not occurred at time t,

H_1: a movement has occurred at time t.

This task can be performed with the likelihood ratio (LR) test [73]. The LR test is a comparison between the likelihood ratio (or difference in the log domain) of two hypotheses and a threshold, and it is given by

$$\frac{p(f \mid H_1)}{p(f \mid H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda,$$

where $p(f \mid H_j)$ is called the likelihood function for the "j^{th}" hypothesis, for $j = 0,1$, evaluated for an observed feature value f at time t, and λ is the decision threshold. $p(f \mid H_0)$ and $p(f \mid H_1)$ are the likelihoods that the feature f is generated under the two competing hypothesis, and H_1 is accepted if the ratio is larger than a threshold λ. In general, this ratio can range between zero and infinity [74].

If the probability density functions of the features under the two hypothesis $p(f \mid H_0)$ and $p(f \mid H_1)$ were known, several techniques for designing an optimal test are available in the literature [73]. In general, these probability density functions are not known, and have to be estimated from a limited number of data samples.

The movement detection approach based on the LR test can be divided into three modules: feature extraction, likelihood ratio estimation, and decision, as illustrated in Figure 6.1. In the first module (feature extraction), the problem amounts to finding a new representation for the load cell data, *f*. In order to apply the LR test in the second module (likelihood ratio estimation), we first need to estimate the likelihood functions. This estimation depends on the distribution used to model the respective feature space. In the last module (decision), a decision is made based on the value of the LR and the value of the decision threshold. The decision output is a binary value depending on the decision that H_1 or H_0 is true.

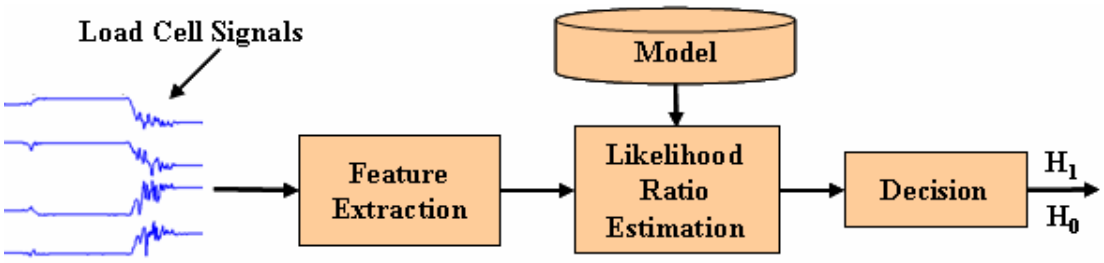


Figure 6.1: Movement detection framework based on likelihood ratio test.

6.1.1 Performance Measure

There are two types of errors that may occur in a detection system:

- *Miss Detection*: a movement is not detected when it did occur. This type of error is also often referred to as type I error.
- *False Alarm*: a movement is falsely detected when it did not occur. This type of error is also often referred to as type II error.

Therefore, the performance of a detector is usually characterized in terms of 2 error measures: the miss detection rate (MDR) and the false alarm rate (FAR). The MDR corresponds to the probability of missing a movement given that a movement has occurred $p(miss \mid H_1)$, and the FAR corresponds to the probability that a movement is falsely detected when it did not occur $p(false\ alarm \mid H_0)$. Also, the costs associated with the errors C_{Miss} and C_{FA} have to be considered when making a decision rule (i.e., selecting the decision threshold).

In this dissertation, detection performance is evaluated using the equal error rate (EER). The EER essentially combines misses and false alarm rates into a single number by finding the decision threshold at which both rates are equal. Therefore, the EER is the point where the false alarms and miss detection rates are equal, and both errors have the same cost. The EER results are reported at the correspondent EER decision threshold estimated on the testing data. The performance measure focuses on the EER because of the lack of knowledge about the typical values of the prior probabilities of a movement $P(H_1)$ and non-movement $P(H_0)$. In addition, the costs associated with missing and falsely detecting movements depend on the application. The costs may be different for clinical and research settings (screening purposes in a research context versus assessment in a clinical context). For example, an inadequate sensitivity in the detection of movement could lead to non-ill or non-affected people entering into a clinical research study that uses the detector for recruitment of subjects. In other words, the type of application dictates whether a low FAR or low MDR is more critical.

The decisions of the detector are discrete, and FAR and MDR are defined by the counts of correct and incorrect sample decisions. Let $GT(t)$ represent the binary decision outputs for the ground truth at each discrete time t, i.e.,

$$GT(t) = \begin{cases} 1, \text{ if an actual movement occurred at time } t \\ 0, \text{ if no movement occurred at time } t \end{cases},$$

and $D(t)$ represent the binary decision outputs of the detector at each discrete time t

$$D(t) = \begin{cases} 1, \text{if a movement was detected at time } t \\ 0, \text{if no movement was not detected at time } t \end{cases},$$

for $0 \le t \le T$, where T corresponds to the total testing period. Based on these definitions,

$$MDR = 1 - \frac{\sum_{t=0}^{T} GT(t)D(t)}{\sum_{t=0}^{T} GT(t)} \quad \text{and} \quad FAR = \frac{\sum_{t=0}^{T} (1 - GT(t))D(t)}{\sum_{t=0}^{T} (1 - GT(t))}.$$

The detection performance can also be described in terms of a graph representing the probabilities involved in the task. Detection tasks involve a tradeoff between miss and false alarm errors, and the capabilities of the system are best represented by a receiver operating characteristic (ROC) curve. The ROC curve is estimated by continuously changing the value of the decision threshold, i.e., a given threshold defines an operating point on the ROC curve. The ROC curve uses the FAR and the correct detection or hit rate (HR), which corresponds to 1 – MDR, as the *x*- and *y*-axes, respectively [73].

6.2 An Approach for Detection of Movement in Bed

In this section, we present an approach for detection of movement in bed from load cell signals. The general idea is to assess the energy in each load cell signal in short segments, and then form a weighted combination with scaling coefficients that are inversely proportional to the distance of each load cell to the center of mass of the body.

The load cell data stream contains periods when the person is in bed (in this case, the total weight $w(t)$ is a sum of the weights of the bed and of the person), and periods when the person is out of bed (in this case, the total weight $w(t)$ is only the weight of the bed). In order to apply the approach for detection, it is necessary to preprocess the data to select only the periods when the person is in bed. A thresholding operation on $w(t)$ is used to determine such periods. The threshold is computed as the midpoint between the averages of the distributions of the weights when the person is in bed and out of the bed over nighttime. After this preprocessing step, the movement detection framework can be applied according to the steps described next.

6.2.1 Feature Extraction

The feature extraction processing consists of three parts, as depicted in Figure 6.2. The *first part* estimates the short-term mean-square differences. A movement of any part of the body in bed is reflected into $w_i(t)$ (Equation 3.3). The load cell signal $w_i(t)$ varies the most during movement, and the estimation of the short-term mean-square differences per load cell signal offers one way of capturing those variations. The short-term analysis of the mean-square difference for each $w_i(t)$ is given by

$$s_i^2(t) = \frac{1}{L-1} \sum_{k=-\left(\frac{L-1}{2}\right)}^{\frac{L-1}{2}} \left(w_i(t-k) - \overline{w}_i(t)\right)^2, \tag{6.1}$$

where L is an odd number and represents the length of the analysis window, and

$$\overline{w}_i(t) = \frac{1}{L} \sum_{k=-\left(\frac{L-1}{2}\right)}^{\frac{L-1}{2}} w_i(t-k)$$

represents the mean calculated over the analysis window. The mean is not computed once for the whole signal $w_i(t)$ because shifts of the body center of mass may occur because of movements, causing the mean to change many times (as described in Section 3.2.1).

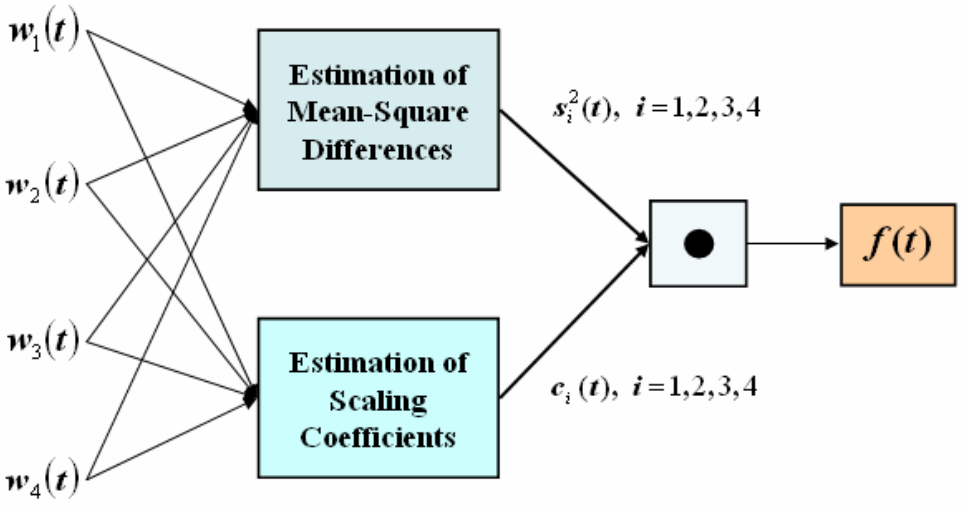


Figure 6.2: Sequence of steps to extract features from the load cell signal.

An example of load cell data $w_i(t)$ and the correspondent $s_i^2(t)$, for $i = 1$ is shown in Figure 6.3. The data corresponds to a sequence of movements composed of small

(solid circles) and large (dashed circles) movements. The small movements include movements of arms, legs and head; large movements correspond to posture shifts.

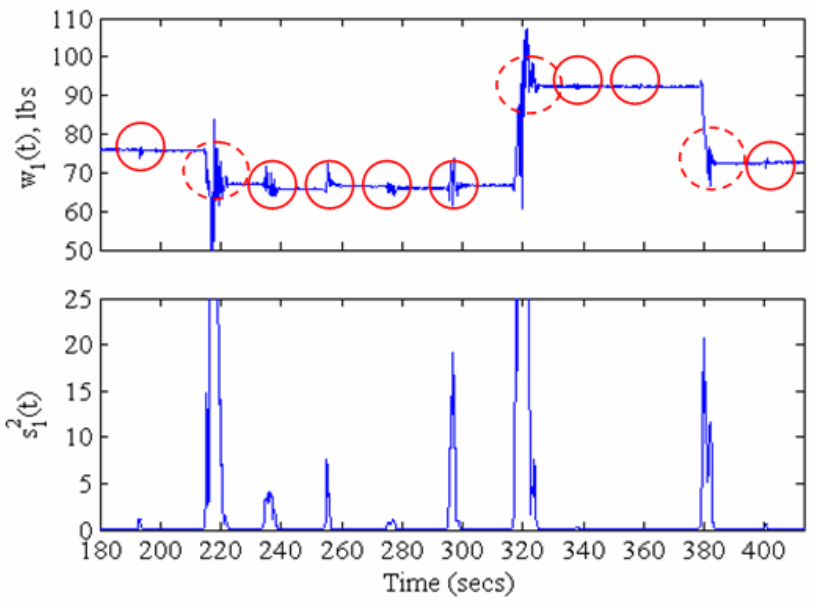


Figure 6.3: Load cell data $w_1(t)$, in lbs, and the correspondent $s_1^2(t)$ collected during a sequence composed of small (solid circles) and large (dashed circles) movements.

The *second part* of the algorithm consists of estimating the scaling coefficients for a weighted combination of the mean-square differences $s_i^2(t)$. The goal is to assign degrees of relevance to each load cell signal based on its distance from the estimated center of mass of the body. The positions of the load cells are fixed and known a priori, and the position of the center of mass of the body is estimated using the proportion of the bed weight measured by each corner. The scaling coefficients $c_i(t)$ are given by

$$c_i(t) = \frac{1}{d_i(t)+1},$$

where $d_i(t)$ is the distance, at time t, between load cell i and the center of mass of the body. The coefficients are inversely proportional to the distances between the load cells and the center of mass of the body. In addition, $c_i(t)$ satisfies the following conditions:

$$\lim_{d_i(t)\to 0} c_i(t) = 1 \qquad \text{and} \qquad \lim_{d_i(t)\to \infty} c_i(t) = 0.$$

The *third part* estimates the feature value *f(t)*. The feature space is represented by a one-dimensional vector given by a weighted sum of the mean-square differences, and it is denoted by

$$f(t) = \sum_{i=1}^{4} c_i(t) s_i^2(t) \ . \tag{6.2}$$

Estimation of the Body Center of Mass

For a collection of particles, the center of mass can be found using the law of moments or law of levers [75] as follows. Consider two particles of mass m_1 and m_2 that fall on the *x*-axis, which represents a lever, and where x_1 and x_2 are the distances between the position of the masses and a pivot point located at the origin of the x-axis. The balance point of this system, also called the center of mass, is defined as

$$x_{CM} = \frac{m_1 x_1 + m_2 x_2}{M},$$

i.e., it is given by the ratio of the total moment to the total mass $M = m_1 + m_2$.

We can generalize this definition for a system of *n* particles in a random arrangement. In this case, the coordinates of the center of mass are given by

$$\begin{aligned} x_{CM} &= \frac{m_1 x_1 + m_2 x_2 + ... + m_n x_n}{M} \\ y_{CM} &= \frac{m_1 y_1 + m_2 y_2 + ... + m_n y_n}{M} \end{aligned} \ . \tag{6.3}$$

Based on Equation 6.3, the estimation of the body center of mass when someone is lying on bed is derived as follows. Let the coordinates of the bed corners be arranged in a two-dimensional Cartesian system as illustrated in Figure 6.4.

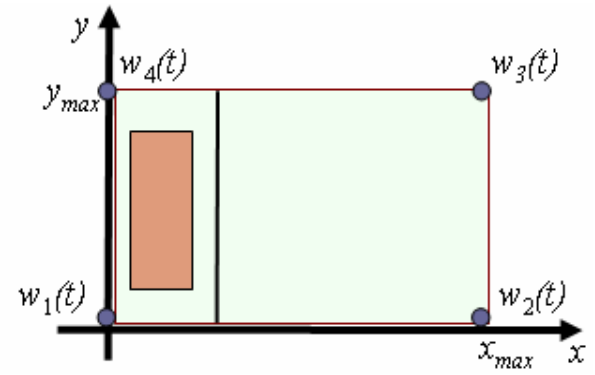


Figure 6.4: Representation of the bed coordinates in a Cartesian system.

Given the weights measured at each corner of the bed, the length and width of the bed (x_{max} and y_{max}, respectively), and according to the two-dimensional Cartesian system illustrated in Figure 6.4, Equation 6.3 can be rewritten as

$$x_{CM}(t) = \frac{[w_1(t) - w_1(t_0)]0 + [w_2(t) - w_2(t_0)]x_{max} + [w_3(t) - w_3(t_0)]x_{max} + [w_4(t) - w_4(t_0)]0}{\sum_{i=1}^{4}(w_i(t) - w_i(t_0))}$$

$$y_{CM}(t) = \frac{[w_1(t) - w_1(t_0)]0 + [w_2(t) - w_2(t_0)]0 + [w_3(t) - w_3(t_0)]y_{max} + [w_4(t) - w_4(t_0)]y_{max}}{\sum_{i=1}^{4}(w_i(t) - w_i(t_0))}$$

or simply

$$x_{CM}(t) = x_{max}\frac{[w_2(t) - w_2(t_0)] + [w_3(t) - w_3(t_0)]}{\sum_{i=1}^{4}(w_i(t) - w_i(t_0))} \quad (6.4.1)$$

$$y_{CM}(t) = y_{max}\frac{[w_3(t) - w_3(t_0)] + [w_4(t) - w_4(t_0)]}{\sum_{i=1}^{4}(w_i(t) - w_i(t_0))} \quad (6.4.2)$$

where $x_{CM}(t)$ and $y_{CM}(t)$ are the coordinates of the body center of mass when someone is lying in bed, at a given time t. The constant terms $w_2(t_0), w_3(t_0)$, and $w_4(t_0)$ correspond to the proportion of the bed weight measured by corners 2, 3 and 4, at time t_0, just before the person goes to bed. Since the total weight of the bed may not be equally distributed in the four corners in most cases, which can affect the location of the center of mass of the system, the weight of the bed is removed from the measurements. The values of $x_{CM}(t)$ and $y_{CM}(t)$ are reported in centimeters.

To make an analogy to the mechanics of a lever in static equilibrium [75], in Equation 6.4.1, it is assumed that there is a pivot point at the side of the bed that corresponds to the coordinates (0, 0) and (0, y_{max}), which is the y-direction. We want to calculate the weights that are required at (x_{max}, 0) and (x_{max}, y_{max}) to counterbalance the total weight put on the bed. Since the weights and the distances are all known, we can estimate x_{CM}. In Equation 6.4.2, the same is done for the x-direction, assuming a pivot point at the side of the bed that corresponds to the coordinates (0, 0) and (x_{max}, 0).

6.2.2 Likelihood Ratio Estimation

A kernel density estimation (KDE) procedure is used to estimate the likelihood functions for the hypotheses $p(f|H_j)$, for $j = 0,1$. The KDE, also know as Parzen-window approach, is a nonparametric technique which allows for less rigid assumptions about the distribution of the observed data since no parametric structure is assumed for the density function $p(f|H_j)$ [76].

In general, in the KDE procedure, we are given a sample of n real observations $X = (X_1, X_2, ..., X_n)$ whose underlying density is to be estimated. It is assumed that the observations are independent and identically distributed (i.i.d.). The probability density function is approximated by the kernel estimator with kernel K defined by

$$p_n(x) = \frac{1}{nh_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n} \right), \tag{6.5}$$

where $p_n(x)$ denotes the density estimator of the random variable X, and h_n is the window width (also known as smoothing parameter) [76]. The kernel estimator can be considered as a sum of 'bumps' placed at the observations. The kernel function K determines the shape of the bumps, and the window width h_n determines their width and thereby it controls the smoothness of the estimated density. The larger values of h_n yield lower variance of the estimates at the cost of filtering out the fine structure. In particular, as h_n approaches zero, the estimate becomes a sum of Dirac delta function spikes at the observations, while as h_n becomes large, the details are smoothed.

For the problem at hand, let $F_{H_0} = (f_1, f_2, ..., f_n)$ be a training dataset consisting of n features estimated from Equation 6.2, corresponding to periods without a movement in bed. Also, let $p_{n,H_0}(f) = p(f \mid H_0)$ be the likelihood function for the hypothesis H_0. Then, from Equation 6.5,

$$p_{n,H_0}(f) = \frac{1}{nh_n} \sum_{j=1}^{n} K\left(\frac{f - f_j}{h_n} \right),$$

and the kernel used is a Gaussian probability density function, namely:

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \text{ for } -\infty < x < \infty.$$

The selection of the window width h_n is determined by a compromise between representing the details suggested by Silverman [76], $h_n = 1.06n^{-1/5}\sigma$, where n is the number of observed data points, and σ is an estimate of the standard deviation computed from all observations. The same procedure is repeated to estimate the likelihood function for the hypothesis H_1 based on features estimated from training data corresponding to periods with movement in bed.

6.2.3 Decision and Post-Processing

In a decision process that minimizes errors, the LR of each sample is compared to a threshold, producing a sequence of decisions that reflect the time periods when the subject is either moving or not. Since the threshold estimation depends on information (i.e., priors and costs of false alarm and miss detections) that is application-dependent, the decision threshold is estimated *a posteriori* by searching a value that produces the EER when applied to the likelihood ratio estimated from the testing data. In practice, the threshold is obtained by varying its value across all available values of LR and determining which value better satisfies the EER condition.

Following each decision, detected adjacent movements that are close to each other are concatenated into one single movement because a movement can be composed of a sequence of smaller movements apart from each other. Assuming that a person does not stop moving within a second interval, consecutive movements that are apart by less than this amount are concatenated into one single movement. Also, short movements are discarded to eliminate spurious movements due to noise to further improve performance. Detected movements shorter than one second are discarded, and the one-second minimum was determined empirically from the ground truth video movement analysis, since no movement shorter than one second was found.

6.3 Results

In this section, we present the detection performance evaluated on the laboratory data described in Sections 4.1 and 4.2, and discuss the effect of weight and bed size on the performance of the detector. We also report clinical metrics derived from the detection approach on the load cell data collected at the Elite Care study described in Section 4.3.

A comparison of load cell and actigraphy with respect to the detected movements is also presented. Before the results, we present the selection of window length, parameter *L*.

6.3.1 Data Preparation

The detection system is evaluated individually for each of the 15 subjects described in Sections 4.1.1 and 4.2.1. As explained in Section 4.2.1, there are two groups of subjects, and the second group (subjects 10 to 15) has more data than the first group (subjects 1 to 9). Each subject's model is estimated from the first 3 fixed-movement trials for subjects 1 to 9 (approximately 21 minutes) and from 6 fixed-movement trials for subjects 10 to 15 (approximately 42 minutes). The remaining fixed-movement trials (2 for subjects 1 to 9, and 4 for subjects 10 to 15) and the free protocol trial are used for evaluation (testing data), corresponding to approximately 17 minutes of data from subjects 1 to 9 and 34 minutes from subjects 10 to 15. On average, 30% of the time is spent moving in a trial. Testing data contain approximately 50 movements per subject for subjects 1 to 9, and approximately 90 movements per subject for subjects 10 to 15. There are individual differences in the number of movements due to extra movements performed when the subject was supposed to be still.

6.3.2 Parameter Optimization: Length of Analysis Window

The estimation of the mean square-differences for each load cell given by Equation 6.1 involves the choice of the parameter *L*, which corresponds to the length of the analysis window. The goal is to determine a value for *L* that would provide optimal estimates, maximally independent of the person or bed type. For this purpose, we assumed equal utility of errors and used the EER as the evaluation measure. The values tested for *L* vary from 5 to 31 samples, corresponding to 0.5 to 3.1 seconds.

The average EER (in percent) across all subjects, and corresponding standard deviations, for the tested values of *L* are shown in Figure 6.5. The ROC curves and the values of the EER for each subject are shown in Appendix D. The best performance is achieved for $L = 11$ (EER = 3.22% ± 0.54) and $L = 15$ (EER = 3.45% ± 0.55). Using a test for the differences between two means [77], the performances for 11 and 15 are not

significantly different (significance level at 0.05), but the difference between them and the remainder of the window lengths is statistically significant. We chose to use $L = 11$.

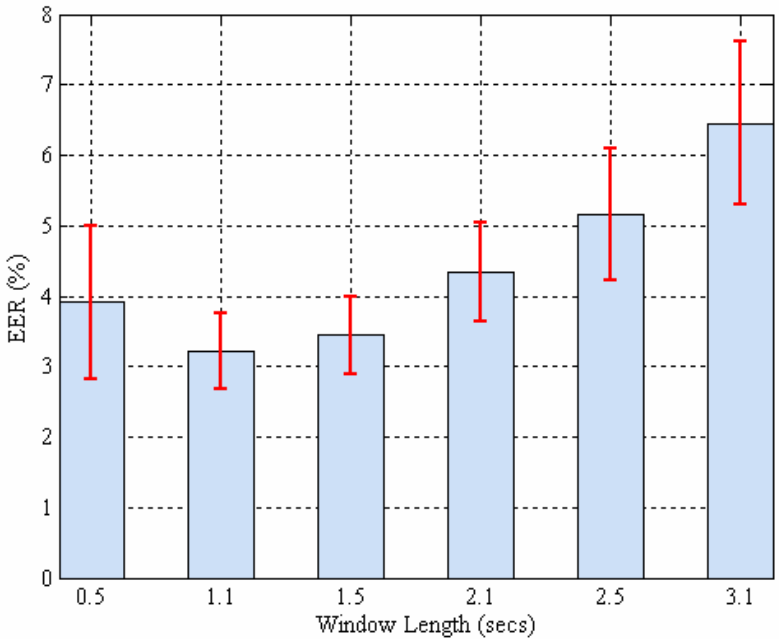


Figure 6.5: Average EER, in percent, and correspondent standard deviation across all subjects, for the window lengths *L*= 5, 11, 15, 21, 25, and 31, which are showed in seconds.

The length of the analysis window L affects the estimation of the onset and offset times of a movement. Unlike in other detection problems (e.g., detection of muscle contraction in studies of motor control [78]), the exact determination of the onset and offset times is not the goal here. However, we need to be aware of this source of error, and take it into account when using the estimated intervals to, for example, classify a movement. The onset error is defined as $\varepsilon_0 = t_0 - \hat{t}_0$, where t_0 is the true onset time of a movement and $\hat{t}_0$ is the estimated onset time. The offset error is defined as $\varepsilon_1 = t_1 - \hat{t}_1$, where t_1 is the true offset time of a movement and $\hat{t}_1$ is the estimated offset time. Therefore, negative onset/offset errors represent delayed onset/offset detections, and positive onset/offset errors represent anticipated onset/offset detections. Figure 6.6 shows the average onset and offset errors across all subjects, and corresponding standard deviations, in milliseconds, for several values of L. For $L = 11$, the onset time is, on average, anticipated by 135 milliseconds (± 230 milliseconds), and the offset time is, on average, delayed by 169 milliseconds (± 520 milliseconds). For those applications where

the onset and offset estimates are critical, it would be possible to use these results to minimize the average errors.

In general, change detection approaches that use a window to estimate features have a time resolution problem as the window length increases, i.e., the onset of events are anticipated and the offset of events are delayed. For this problem, onsets are anticipated as the window length increases as illustrated in Figure 6.6 (top). However, as shown in Figure 6.6 (bottom), the offsets are delayed for L smaller or equal to 15, anticipated for windows between 21 and 35 samples long, and then delayed again for L larger than 35 samples.

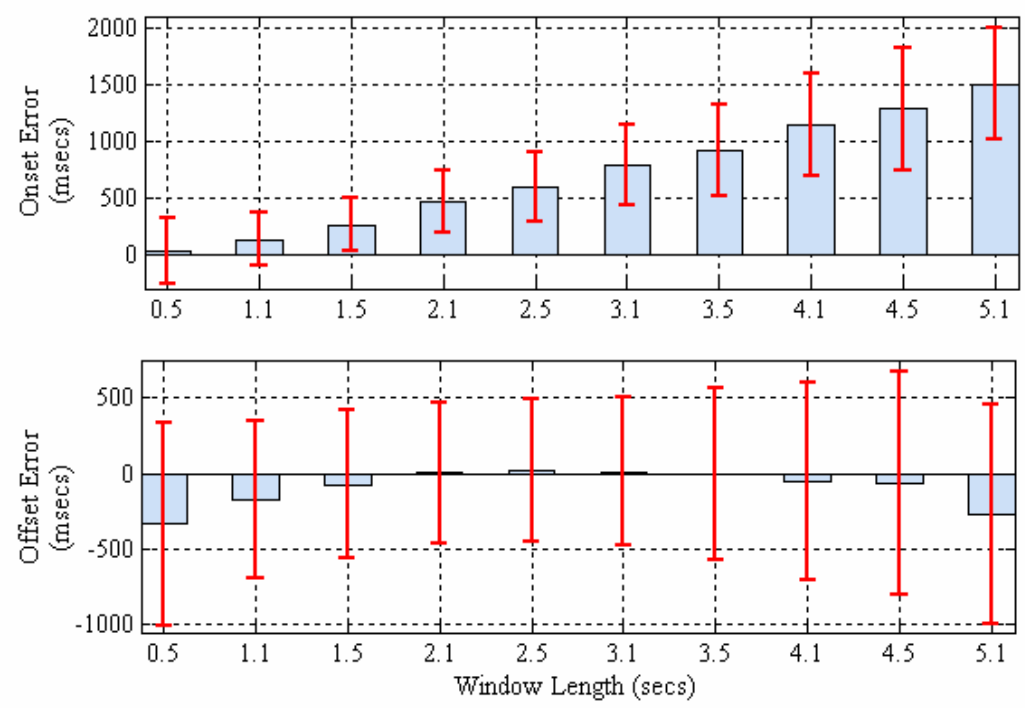


Figure 6.6: Average onset (top) and offset (bottom) errors, in milliseconds, and correspondent standard deviations across all subjects, for the window length *L* varying from 5 to 51, which is showed in seconds.

The behavior of the offset error is caused by a typical characteristic of the dynamic of many movements, i.e., the beginning of the movement concentrates most of the intensity (causing more variation in the load cell signal) of the movement and the latter part of the moment has the least intensity due to body adjustments to reach a resting position. Such behavior is illustrated in the top graph of Figure 6.7 for a leg movement (the vertical bars represent the movement boundaries estimated using the video data).

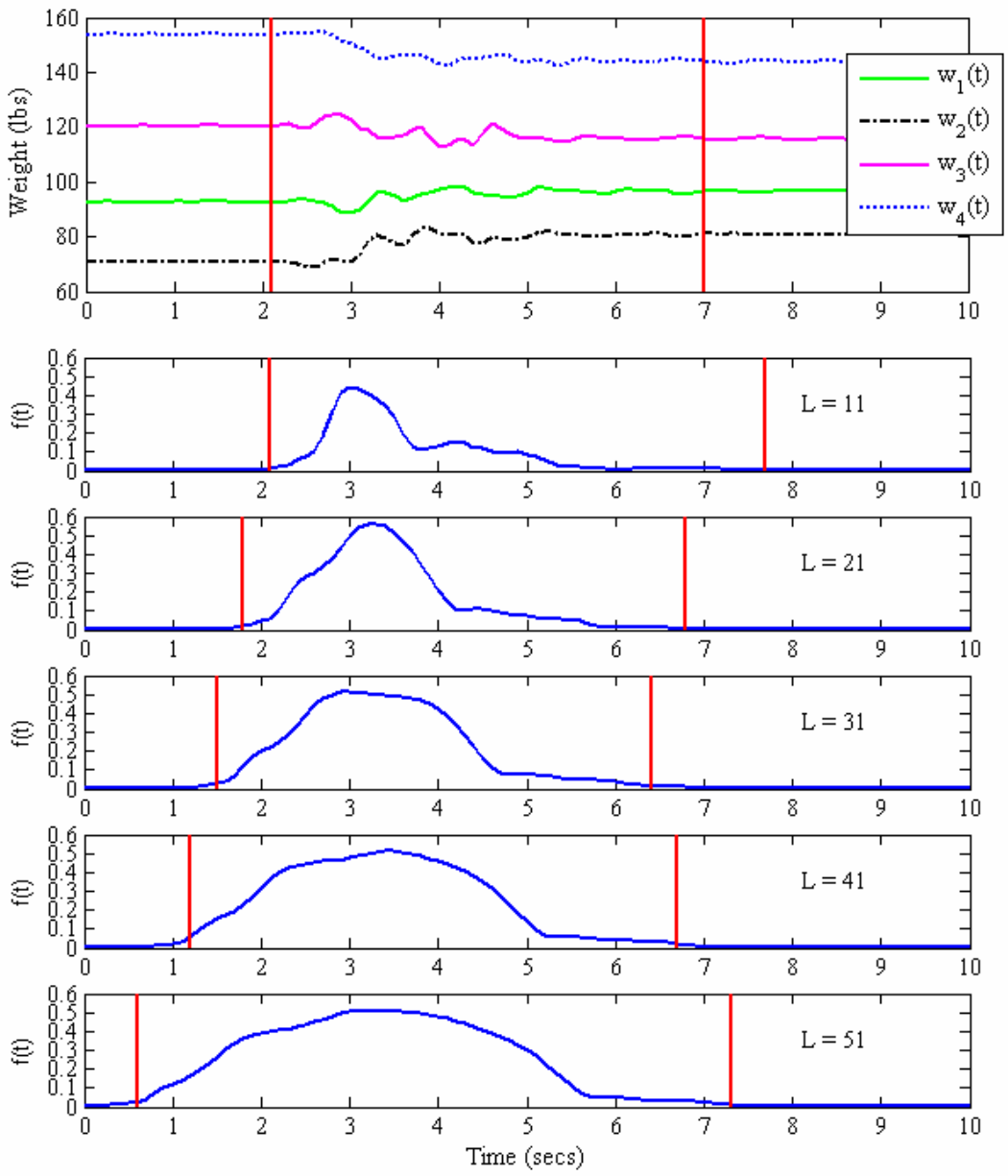


Figure 6.7: Individual load cell outputs $w_i(t)$, for i = 1, 2, 3, 4, during a leg movement (top). Vertical lines indicate the true onset and offset times of the movement. Remaining plots show the correspondent feature value $f(t)$, for L = 11, 21, 31, 41, and 51. The vertical lines indicate the estimated onsets and offset s times of the movement for each value of L.

Most of the intensity of the movement occurs earlier (2 to 5 seconds), and the load cell data from the remainder of the movement shows only a small variation. The small variation of the load cells signals is very similar to the variation caused by noise (seen outside the movement boundaries), which increases the confusion between time periods with movement and without movement. In the second plot from the top of Figure 6.7,

which shows the estimated boundaries for the features estimated using a window length equal to 11, note that the decision threshold estimation becomes very difficulty due to small difference between the feature *f(t)* before and after the true offset boundary.

As the analysis window increases, the regions with largest variation are incorporated in the feature estimation of the latter part of the movement, causing an increase in the amplitude of the features *f(t)*. The amplitude increase of the feature *f(t)* during movement period improves the discrimination due to the shortening of the overlapping tails (which are responsible for the false alarm and missed detection rates) of the distributions of the feature *f(t)* for the time periods with and without movements. Such improvement in the discrimination causes an anticipation of the offset boundary of the third ($L = 21$) and fourth ($L = 31$) plots from the top of Figure 6.7. However, the increase in variance is also incorporated into the features from the regions around the movements causing the distributions to overlap and, consequently, to delay the movements offset. The fifth ($L = 41$) and the sixth ($L = 51$) plots from the top of Figure 6.7 illustrate the delaying of the offset. Note that the amplitude of the feature *f(t)* for such plots from second 6 to 7 is greater than the ones for the same period on the second and third plots.

6.3.3 Performance Results

The performance for each subject for $L = 11$ is shown in Figure 6.8. The average EER across all subjects is 3.22% (± 0.54). This performance is certainly within the reliability of most clinical tests. For a total of 890 movements in the testing data, only 11 movements are missed and 14 are falsely detected. The missed movements include 7 head movements, 2 arm movements, and 2 medium amplitude movements that include arm and leg movements to adjust position. For an average miss detection rate of 3.22%, 2.69% accounts for miss detections at the onsets and offsets, and only 0.53% accounts for missed movements. For the false alarm rate, 2.77% accounts for false alarms at the onsets and offsets, and only 0.45% accounts for falsely detected movements. This shows that most of the errors come from errors at the estimation of the onsets and offsets. Depending on the application, these errors may be insignificant if the objective is the number of movements.

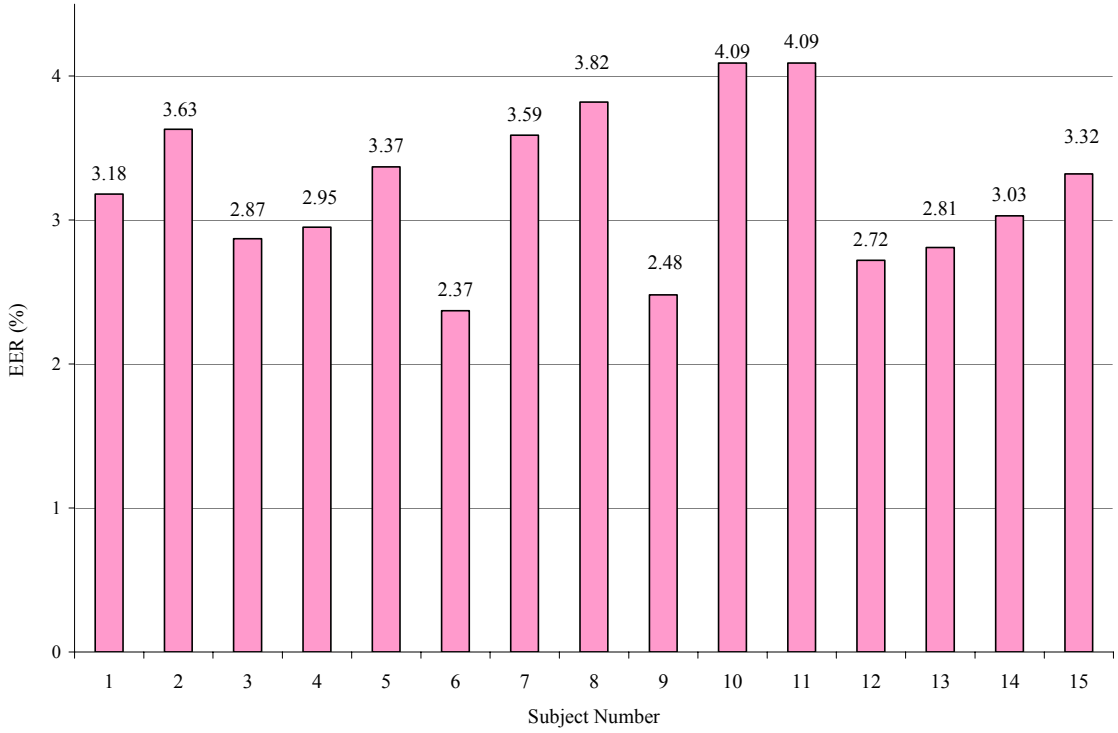


Figure 6.8: Individual EERs in percent, for L = 11. The average EER is 3.22 (±0.54).

An interesting result is that the person's weight has no effect on the performance of the proposed detector. Despite the fact that moving an object with a greater mass should provide a stronger force signal, our signal processing yields weight-invariant detection. In particular, the data suggest that there is no correlation between the subject's weight and the EER. A linear relationship between the EERs and the subjects weight (shown in s Table 4.1 and 4.4) is weak ($r = 0.047$). This shows that the approach does not impose any constraints on its use with very light people.

Another useful invariance is that the bed type (i.e., the size of the bed) has no effect on the performance of the proposed detector. We compare the performance results on the subjects that were tested in the LAB1_TWINSIZE and LAB2_FULLSIZE experiments. The detection system for each bed type is evaluated individually for each of the 5 retested subjects using the 5 fixed-movement trial data from each bed. Each subject's model is estimated from the first 3 fixed-movement trials (lasting approximately 21 minutes), and the remaining fixed-movement trials are used for evaluation (lasting approximately 14 minutes). Table 6.1 lists the individual EERs in percent for the 2 bed sizes tested. The binomial test for differences in proportion [79] shows that there is not a

difference in the performances of the 2 bed systems for all subjects, at a significance level of 0.05. The binomial test, also referred as Z-test for the equality of two proportions, verifies the assumption that the proportions p_1 and p_2 of two populations are equal, based on n_1 and n_2 samples of each population, and for $n_1, n_2 \geq 30$. This result provides some evidence of the generalizability of the proposed detector for different bed sizes.

Table 6.1: Individual EERs, in percent, for the subjects tested in 2 beds.

Subject Number	EER (%)	
	Twin Size Bed	Full Size Bed
1	3.47	3.42
2	3.24	3.58
3	3.18	2.70
4	3.20	2.76
8	3.90	3.46

6.3.4 Approximation of the Decision Threshold

In the estimation of the decision threshold when using the LR test, training data are usually required to estimate the likelihood functions. However, a new subject may be presented to the system with no training data being available. To mitigate this problem, we developed a technique that approximates the threshold for the new subject by linear regression of the thresholds of a known set of subjects.

Let λ_{EER} correspond to the EER threshold. By definition, there is a decision point f^* such that

$$\frac{p(f^* \mid H_1)}{p(f^* \mid H_0)} = \lambda_{EER}.$$

Because the features f (as defined in Equation 6.2) are proportional to the person's weight, the decision point f^* varies linearly with the weight, with a correlation coefficient $r = 0.906$ ($p < 0.0001$), as illustrated by Figure 6.9. The detector decision can be estimated by applying a thresholding operation on the features f calculated from a test dataset, using the predicted value of f^* (from the fitted regression equation) as threshold.

This method for approximating the decision threshold was tested using a leave-one-out approach [80] on the 15 subjects. The value of f^* is predicted for each subject based on the least square linear regression fit of the remaining subjects. Since the EER

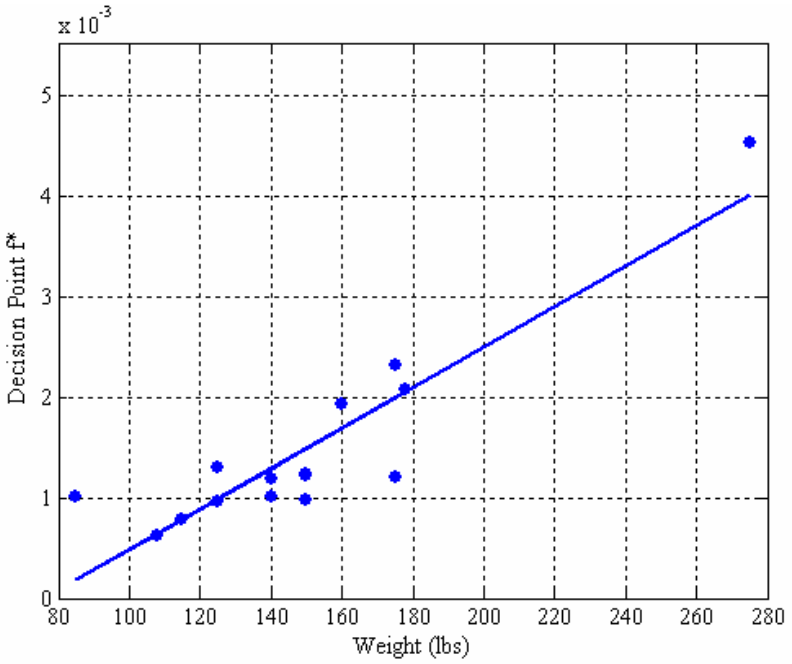


Figure 6.9: Plot of the decision point f* versus weight, in pounds (for L = 11 and for all 15 subjects) with the fitted least square regression line.

was used to measure the performance and to estimate the decision threshold in the previous results, a detection cost is used to compare the results of both approaches (EER-based estimation and least-squares based estimation). The detection cost function C_{DET} [74] is given by

$$C_{DET} = C_{Miss} MDR\, P(H_1) + C_{FA} FAR\, P(H_0). \quad (6.6)$$

The parameter values in Equation 6.6 (as defined in Section 6.1.1) are $C_{Miss} = C_{FA} = 1$ and $P(H_1) = P(H_0) = 0.5$, and correspond to the values used to define the EER.

Although this method only provides an approximation for the threshold, we have not found a statistically significant difference in the performance of the detector between the EER thresholds (thresholds derived from the EER measure) and the approximated thresholds. Figure 6.10 shows the boxplots of the detection costs for all 15 subjects corresponding to (a) the cost associated with decision threshold based on the EERs, and (b) the cost associated with the approximated threshold value. Despite the larger variability in the detection cost C_{DET} in Figure 6.10 (b), there are no statistically significant differences in the median values of C_{DET} for the dataset of 15 subjects tested. The median values of C_{DET} are 0.0325 and 0.03628 for the estimated and approximated

cases, respectively. Even though, in this work, this result applies to the EER criterion, it is still possible to use regression when other criteria are chosen.

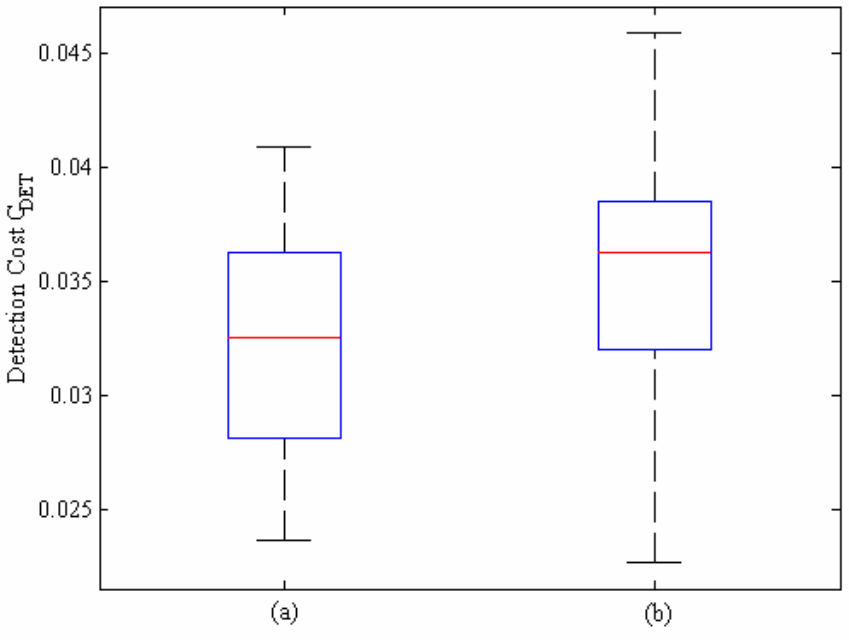


Figure 6.10: Boxplots of the detection costs for all 15 subjects corresponding to (a) the cost associated threshold estimated from the EERs, and (b) the cost associated with the approximated threshold values.

6.3.5 Application to Elite Care Data

In this section we apply the detection approach to the load cell data collected at the Elite Care study, and estimate clinically relevant measures based on the detected movements. The clinical measures are: the number of postural immobility periods (number of immobility periods longer than 15 minutes, as defined by [59]), the number of movements and the frequency of movements (number of movements per minute). As suggested by Kronholm [81], the frequency of movements for each third of the night (based on total time in bed) can be used to assess differences in the temporal distribution of movements throughout the night and across nights. Motor episodes occurring repeatedly during the first third of the night strongly suggest a link to NREM stages III and IV, whereas episodes appearing in the last third suggest association with REM sleep [5].

As described in Section 4.3.2, load cell and actigraph data were simultaneously collected from 2 residents of an assisted-living facility for a period of 14 days. Since the

data were collected continuously, we first determined the bedtime and getup times for each day using the algorithm described in Section 5.1. Using the least square regression approach described in Section 6.3.4, we computed the decision threshold of the detector for each subject based on knowledge of his/her weight. An example of load cell data collected from subject #2 is shown in Figure 6.11(a), where the dashed vertical lines indicate the intervals where a movement was detected. The time period shown in Figure 6.11 corresponds to the beginning of an in-bed event, and the first movement detected (between 1:16:57 AM and 1:18:27 AM) corresponds to a movement related to getting into the bed that lasted approximately 1.5 minutes. A total of 16 movements were detected during that period, and the remaining movements lasted, on average, 1.6 (± 0.6) seconds. As illustrated in Figure 6.11 (b), the actigraph data collected during the same period measured large activity counts at the beginning of the period corresponding to the large movement related to getting into the bed, and some of the movements detected by the load cells were missed by the actigraph (for example, between 1:19:21 AM and 1:21:45 AM).

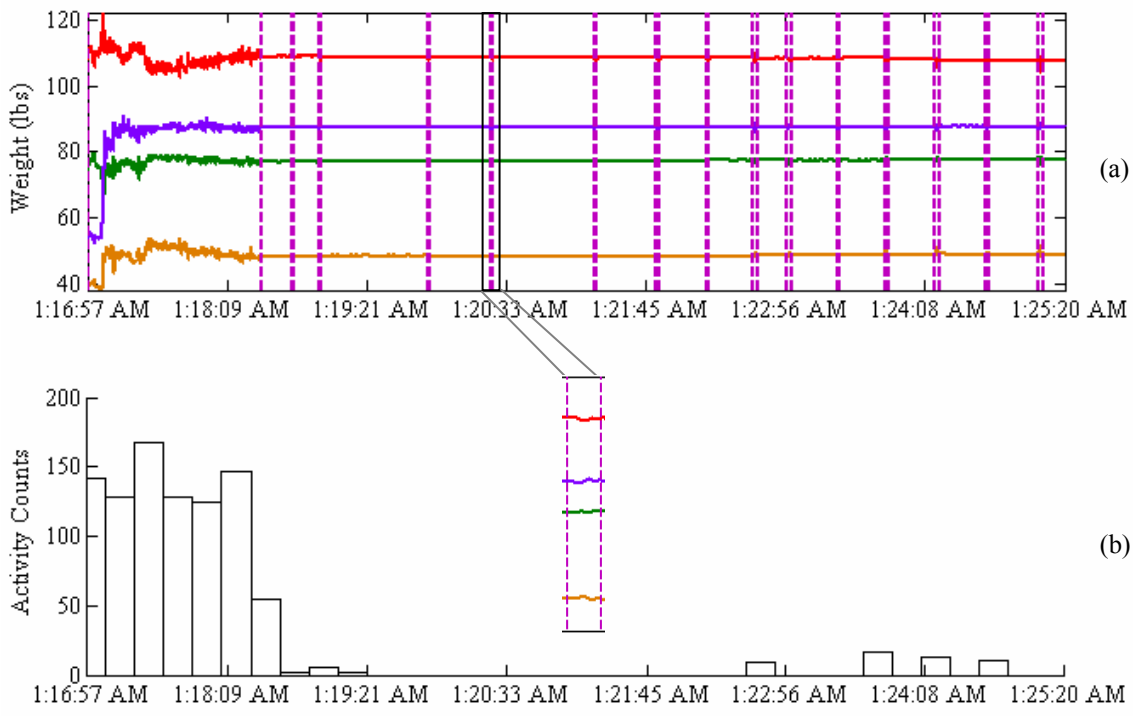


Figure 6.11: (a) An example of load cell data collected from subject #2 is shown, where the dashed vertical lines indicate the intervals where a movement was detected. The emphasized rectangle shows a zoomed view of a movement detected by the load cell data. (b)Actigraph data collected during the same period.

Given the detection results and the total time spent in bed, the number of postural immobility periods (number of immobility periods longer than 15 minutes) and the frequency of movements (number of movements per minute) for each third of the night were calculated for Subject #1 (as shown on the left column of Figure 6.12) and Subject #2 (as shown on the right column of Figure 6.12). As illustrated in Figure 6.12, over these 2 weeks, Subject #1 experienced higher frequency of movements and less periods of nocturnal immobility than Subject #2, especially during the first 7 days. The higher level of activity in Subject #1 is confirmed by the caregiver's report in the Sleep Disorder Inventory (described in Section 4.3.2). According to the inventory, this subject gets up at night frequently, and sometimes gets involved in activities at night, thinking that is daytime.

In the last seven days, a decreased in activity occurred for Subject #1, as illustrated by longer periods in bed with an increased number of postural immobility periods and less activity in the first 2 thirds of the night. Subject #1 suffers from dementia, which may justify the high variability in the level of activity during the night. Subject #2 showed a decreased in nocturnal activity in the last third of the night during the last 7 days. Unfortunately, these objective measures could not be compared with the subjective reports of sleep quality, given by Question #7 in the sleep diary completed by caregivers, because of the low compliance in the completion of the daily diaries.

6.3.6 Analysis of Actigraphy Data

Using actigraph data that were collected simultaneously with load cell data, we show that the actigraph under-reports movements as compared to load cells. Before we can perform any comparison, the load cell detection results (sampled at 10 Hz) and the ground truth information (sampled at 10 frames per second, as described in Section 4.1.4) are converted into 15-second epochs as the actigraph reports every 15 seconds.

The actigraph used in the study (as described in Section 4.3.2) has the following data acquisition and recording methods:

- Data acquisition: data are sampled at 32 Hz and, for every sample, its value is compared to a baseline value by computing the difference. The difference is kept as one activity value.

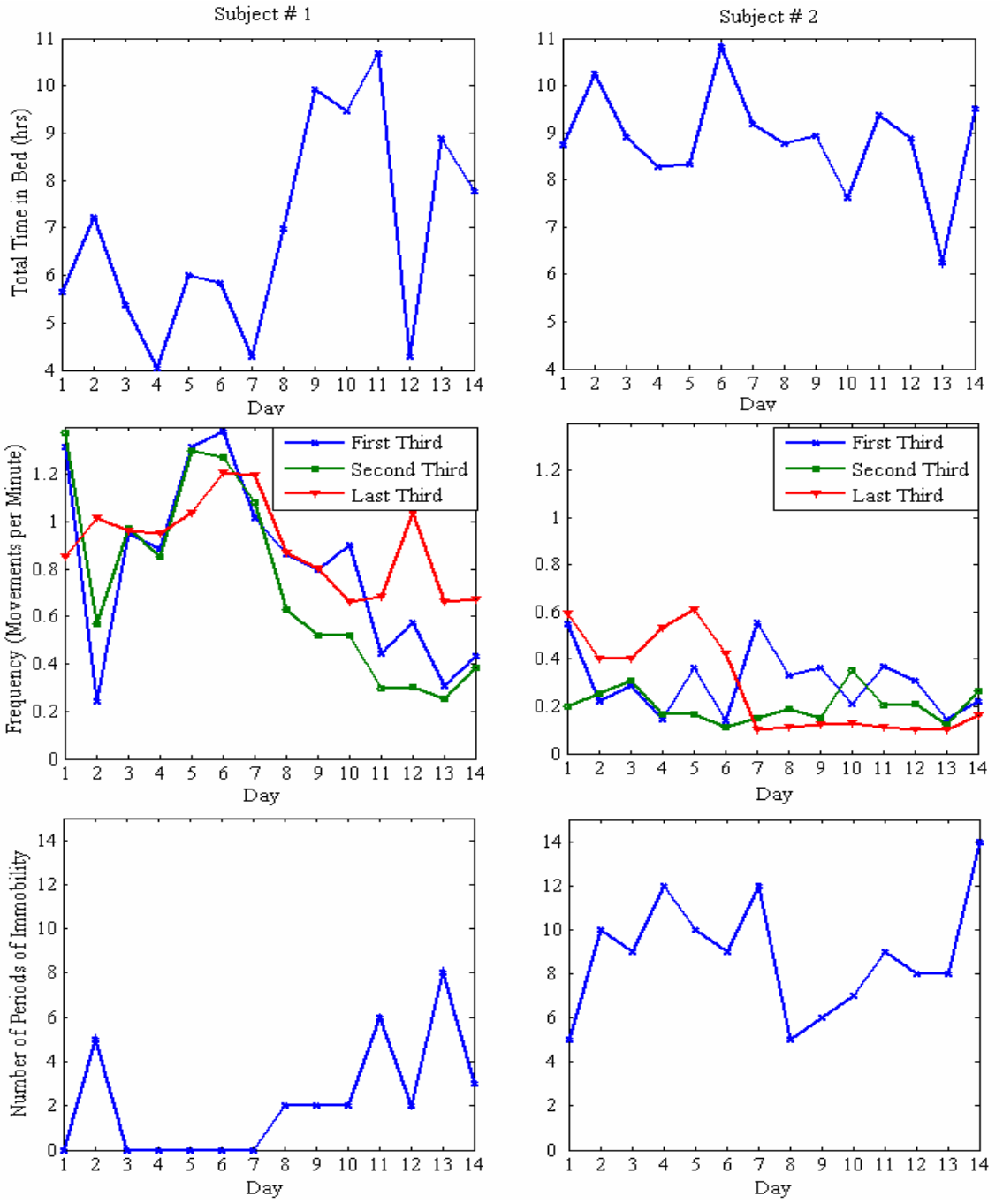


Figure 6.12: Total time in bed, frequency of movements (number of movements per minute) for each third of the night, and the number of postural immobility periods (number of immobility periods longer than 15 minutes) for subject #1 (left column) and subject #2 (right column), for a period of 14 days.

- Recording method: the resulting one-second value is the peak value obtained from the 32 resulting activity values from previous step. The peak value over a one-second interval is used to make the actigraph sensitive to small

movements and to rapidly changing movements often experience in sleep. The one-second activity values are finally accumulated depending on the epoch chosen. For example, for an epoch length of 15 seconds, 15 one-second activity values are summed to generate the final recorded value called activity count.

The final activity count is a 15-bit unsigned integer. As defined by the manufacturer, the activity count is zero if no movement has been detected in an epoch. An activity count larger than one corresponds to movement, and its magnitude monotonically increases with the amount of wrist activity.

Given that actigraphs measure the amount of movement based on movement of the site of placement of the sensor (e.g., wrist and ankle), the actigraph under-reports movements as compared to load cells because it cannot detect motion of the other limbs unless an actigraph is placed on every limb. To compare the actigraph result with the ground truth, every epoch from the actigraph with an activity count larger than zero is said to have a movement when an actual movement has occurred within the time interval of the respective epoch. The actigraph hit rates, i.e., the probability that a given movement is detected by the actigraph given that the movement has occurred according to the ground truth, are shown in Table 6.2 for 3 types of movement. These probabilities show that, whereas most of the posture shifts are detected by the actigraph, head and arm movements (when moving the arm that does not have an actigraph) are missed. Approximately 44% of the leg movements are detected, with corresponding median activity counts equal to 3, i.e., the actigraph detected very small movements of the wrist that cannot be seen from the video during leg movements. The load cells hit rates are calculated over 15-epochs by comparing the movements detected by load cells with the actual movements in a similar manner: every epoch is said to have a movement when an actual movement that has been detected by the algorithm has occurred within the time interval of the respective epoch. The load cells hit rates are shown in Table 6.2. Because these probabilities are calculated over 15-epochs, onset and offset errors of the algorithm are not considered here, which makes the error smaller (as stated in Section 6.3.3, most of the algorithm errors come from errors at the estimation of the onsets and offsets, and only medium amplitude movements were missed). Therefore, the load cells yield more

accurate detection of movements in bed than actigraphs, and in particular the load cells can detect a wider range of movements in bed rather than the movement of a specific limb.

Table 6.2: Actigraph and Load Cells Hit Rates for Three Types of Movements.

	Posture Shifts	Medium Amplitude Movements[iv]	Leg Movements
Actigraph	1	0.732	0.438
Load Cells	1	0.974	1

6.4 Summary

This chapter presented an approach for detection of movement in bed from load cell signals. Since the load cell signal varies the most during movement, the approach uses a weighted combination of the short-term mean-square differences of each load cell signal for capturing the variations caused by movement.

The approach, based on likelihood ratio test, can be divided into three modules: feature extraction, likelihood ratio estimation, and decision. In the first module, for each time t, the short-term mean-square difference is estimated using a sliding window on each load cell signal $w_i(t)$. Then, the output from each load cell is combined using a weighted combination with scaling coefficients that are inversely proportional to the distance of each load cell to the center of mass of the body at time t. In the second module, the kernel density estimation method is used to estimate the likelihoods functions of the two hypotheses: a movement has not occurred at time t, and a movement has occurred at time t. In the last module, a decision is made based on the value of the likelihood ratio test and the value of the decision threshold. The detection performance was evaluated using the EER on the testing data because no information about the prior probabilities and costs is available. Of course, if any information about the prior probabilities and/or costs associated with falsely detecting or missing a movement is available, these can be easily incorporated in the determination of the decision criterion using, for example, the detection cost function in Equation 6.6.

[iv] Smaller movements to adjust, not change, position and that do not include isolated leg movements, such as isolated head or arm movements.

The estimation of the mean square-differences for each load cell involves the choice of the parameter L, which corresponds to the length of the analysis window. We showed that the best performance is achieved when $L = 11$ (EER = 3.22% ± 0.54) and $L = 15$ (EER = 3.45% ± 0.55). The performances for 11 and 15 are not significantly different, and the window length used in this work is $L = 11$.

We analyzed the performance of the proposed approach for 15 subjects, and showed that the approach reliably detects movements. Most of the errors come from errors at the estimation of the onsets and offsets, and not from missed or falsely detected movements.

One of the major contributions of the proposed approach is that the person's weight has no effect on the performance of the proposed detector. In particular, the data suggested that a linear relationship between the EERs and the subjects weight is weak ($r = 0.047$). It thus shows that the approach does not impose any constraints on its use with very light people. We also provide some evidence of the generalizability of the proposed detector for different bed sizes. We compared the performance results on subjects that were tested in 2 different beds. The binomial test for differences in proportion shows that there is not a difference in the performances of the 2 bed systems for all subjects, at a significance level of 0.05.

We also showed that, in the absence of training data to estimate the likelihood functions, least square regression can be used for obtaining an approximation of the decision threshold. The threshold value was approximated for each subject based on the least square linear regression fit of the remaining subjects. We have not found a statistically significant in the performance of the detector for the EER thresholds (thresholds derived from the EER measure) and the approximated thresholds.

We also applied the detection approach on the load cell data collected at the Elite Care study. The number of postural immobility periods, the number and duration of movements and the frequency of movements were estimated from data from the 2 subjects, over a period of 2 weeks. The examples illustrate the common intra-individual and inter-individual differences in nocturnal mobility that are frequently observed in multiple-night recordings [56].

The final section reported a comparison between the load cell detector and the actigraph. Given that actigraphs measure the amount of movement based on movement of the body location of placement of the sensor (e.g., wrist and ankle), the actigraph under-reports movements as compared to load cells because it cannot detect motion of the other limbs unless an actigraph is placed on every limb. Therefore, for the detection of a wider range of movements in the bed rather than the movement of a specific limb, the load cells provide more accurate results than actigraphs.

Chapter 7

Classification of Body Movements in Bed

This chapter describes an approach for movement classification from load cell data. Following a description of the problem framework (Section 7.1) and a brief introduction to pattern recognition (Section 7.2), we describe (Section 7.3) and evaluate (Section 7.4) the approach for subject-dependent classification of movements in bed using a time-domain representation of the body center of mass. In Section 7.5, we present an alternative feature representation based on wavelets, and compare the performances of both types of representation. In the final section, we present the performance of the integrated system combining movement detection and classification.

7.1 Movement Classification Framework

The problem of classification of movements in bed consists of determining the type of movement performed in a given time interval. As discussed earlier (Section 4.1.2), different movement descriptions have been adopted to analyze the distribution of movements during sleep. Accordingly, in this work, movements in bed are divided into 3 classes:

- **Class 1 ─** Major posture shifts: changes in body position that involve a torso rotation larger than 45 degrees. An example of this class is shown in Figure 7.1. These large movements may represent movements related to getting into or out of bed, or large movements associated with wakefulness.

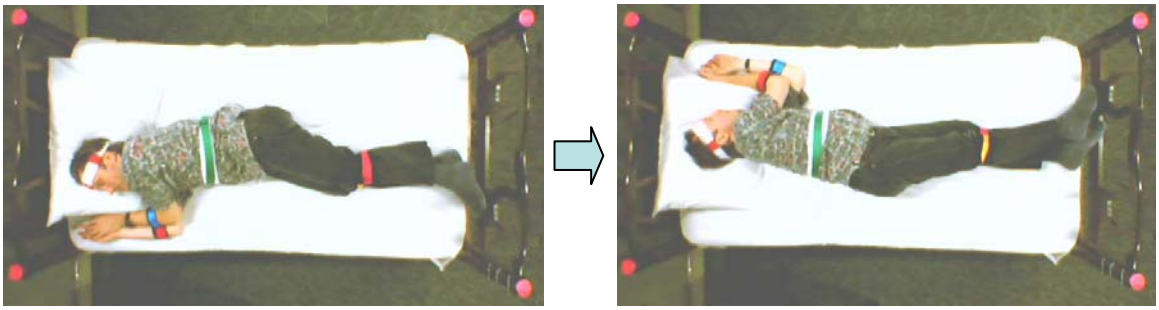


Figure 7.1: A class-1 movement example: turning from right to the left.

- **Class 2 —** Small and medium amplitude movements: changes in body position involving the head, arms, torso rotations smaller than 45 degrees, any combination of upper and lower limbs, and any combination of limbs and torso rotations smaller than 45 degrees. An example of this class is shown in Figure 7.2. These medium amplitude movements may represent restlessness or position changes associated with NREM stage I, as discussed in Section 2.2.2.

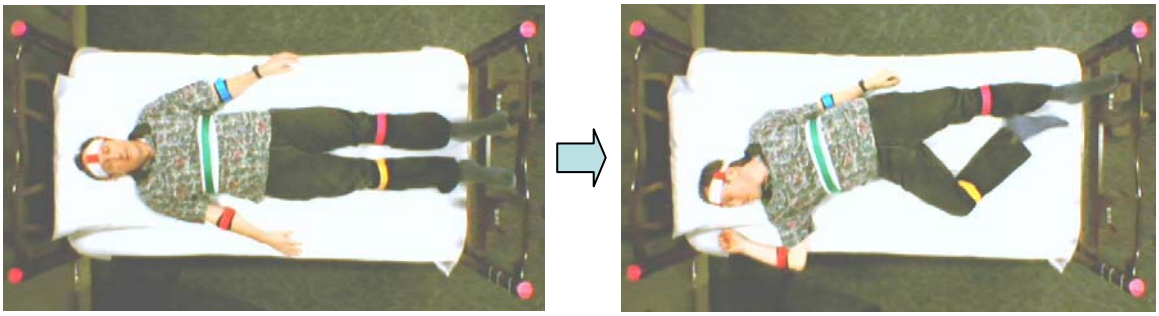


Figure 7.2: A class-2 movement example: small change in position involving movement of the head, legs and arms.

- **Class 3 —** Leg movements: isolated movement of lower limbs (thighs, legs and feet). An example of this class is shown in Figure 7.3. These leg movements can be associated with periodic limb movements in sleep (PLMS) or restless leg syndrome, as presented in Section 2.2.4.

As we discussed in Section 3.2, a statistical approach is used for classification. A brief review of statistical pattern recognition systems is described next.

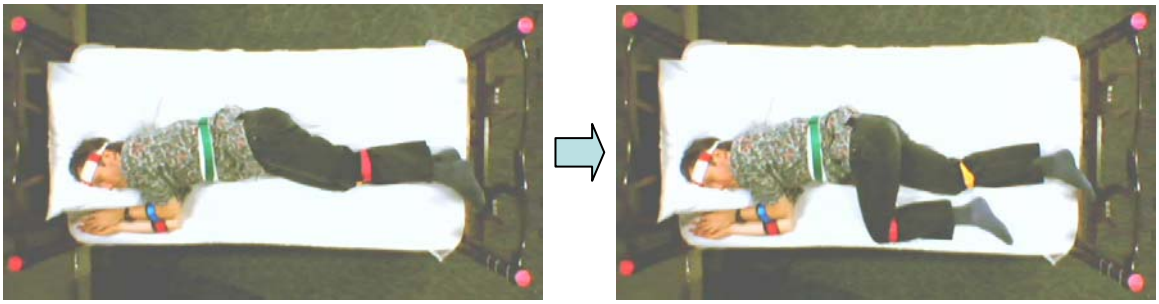


Figure 7.3: A class-3 movement example: bending one leg.

7.2 Statistical Pattern Recognition Systems

The classification of patterns under uncertainty requires the design of a statistical pattern recognition system [82]. We start by defining the terminology and the notation used to describe a statistical pattern recognition system. A pattern can be a fingerprint image, a speech signal, or load cell signals in this case. Each pattern is characterized by a vector of D feature values $x = (x_1, x_2,\ldots, x_D)$, also called feature vector. The features are the variables used to represent the pattern for the purpose of classification [83]. In a multicategory classification problem there are N classes, denoted $c_1, c_2, \ldots, c_N$, and we define a categorical variable z associated with each pattern that denotes the class membership. That is, if $z = k$, then the pattern belongs to c_k, $k \in \{1, 2,\ldots, N\}$. Given a pattern, the classification task consists of identifying the class to which the pattern belongs. A decision rule partitions the measurement space into N regions, Ω_k , k = 1, 2, …, N. If a feature vector is in Ω_k, then it is assumed to belong to c_k. The boundaries between the regions Ω_k are the decision boundaries or decision surfaces.

In a typical approach, e.g., as suggested by Jain [82] and shown in Figure 7.4, a statistical pattern recognition system has 2 modules: training and classification. The *training module* can be divided into 3 steps: pre-processing, feature extraction/selection, and learning. Assuming that we have a set of training patterns (a set of patterns used to design the classifier), the pre-processing step includes operations that contribute to defining a compact representation of the pattern (for example, removing noise and normalization). The feature extraction selection step finds the appropriate features for representing the training patterns by transforming them to an appropriate form for subsequent classification, and it reduces the dimension of the feature vector to remove redundant or irrelevant information. The last step of the training module, referred as

learning, consists of building a classifier. The objective is to establish decision boundaries in the feature space that separate patterns belonging to different classes. The features are assumed to have a probability density function conditioned on the class. Therefore, a feature vector x belonging to class c_k is viewed as an observation drawn randomly from the class-conditional probability function $p(x \mid c_k)$ [82]. Different strategies can be used to build a classifier, depending on the information available about the class-conditional densities of the features. These strategies include the optimal Bayes decision rule (when the class-conditional densities are completely specified), parametric methods (when the form of the class-conditional densities is known), and nonparametric methods (when the form of the class-conditional densities is not known) [82].

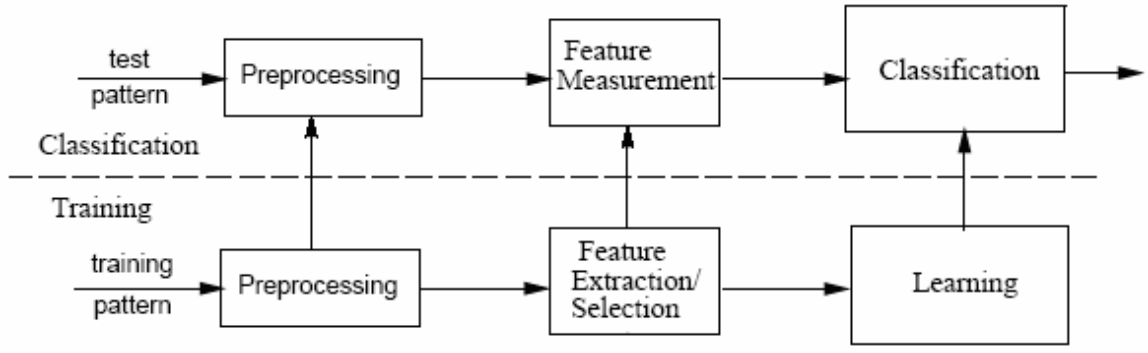


Figure 7.4: A statistical pattern recognition system adapted from [82].

The *classification module* is also divided into 3 steps: pre-processing, feature measurement, and classification. After pre-processing and extracting the feature corresponding to a test pattern (unseen in the training module), in the same way it was done in the training module, the trained classifier assigns a test pattern to one of the N classes based on the measured feature.

7.3 An Approach for Subject-Dependent Movement Classification

The high variability among human physical characteristics provided a motivation for subject-dependent classification of movements in bed. However, because subject-independent models are especially important in situations where training is difficult or impossible to conduct, we also present the results of a subject-independent model in Section 7.4.4.

To capture the *subject-dependent* feature distribution, we use Gaussian Mixture Models (GMMs) to represent each class of movement. More specifically, the distribution of feature vectors extracted from a subject's movement data, for each class of movement, is modeled by a Gaussian mixture density. Following the processing flow as shown in Figure 7.5, the first step consists of pre-processing the load cell data. Although this approach is subject-dependent, the attempt is to use features that would be as much as possible independent of the subjects' height and weight. We therefore represent the raw load cell signal by the trajectory of the body center of mass, in terms of the coordinates $x_{CM}(t)$ and $y_{CM}(t)$ as in Equations 6.4.1 and 6.4.2.

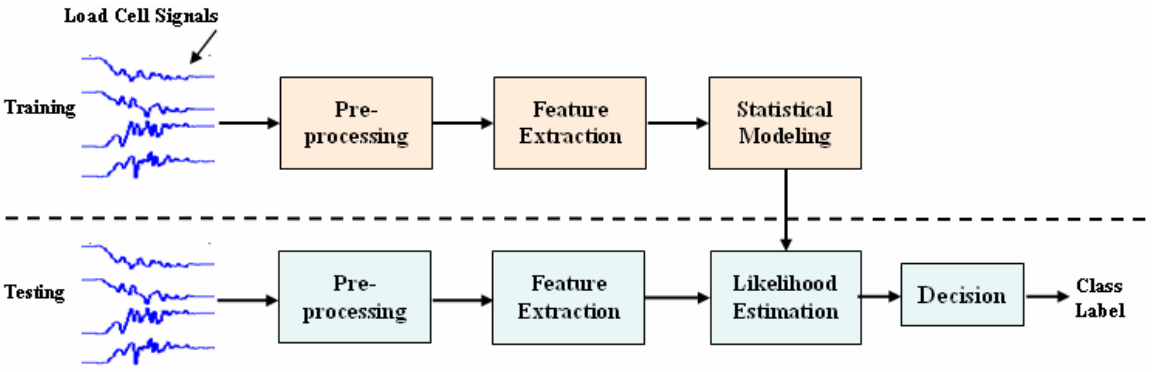


Figure 7.5: Movement classification framework.

In the feature extraction step, the following features are extracted from the trajectory of the center of mass: the distance between initial and end points of the trajectory, the length of the trajectory, and the variance of the trajectory in the *y*-direction. The *y*-direction corresponds to motion from one side of bed to the other. These features are descriptors of the trajectory of the center of mass during movement, and provide a simple characterization of the spatial and (indirectly) temporal aspects of the movements in bed. In the statistical modeling step, the goal is to estimate the parameters of each GMM that represents a certain movement class. In the testing module, after the testing data are pre-processed and features are extracted as was done in the training module, the system estimates a likelihood score for each class for every test pattern, based on the model parameters estimated during the training module. A class label is assigned based on the maximum likelihood (ML) rule [84]. Details of each step are described next.

7.3.1 Pre-Processing

The pre-processing step consists of estimating the trajectory of the center of mass during movement from load cell data from each movement. In studies of posture or movement, the trajectory of the body center of mass has often been a parameter of interest [85]. Lee [86] uses the (x, y) coordinates of the center of mass as part of the gait silhouette image representation in a problem of gait classification. In this application, for a movement defined over a time interval [t_0, t_1], the trajectory of the body center of mass is given by the coordinates of the body center of mass $x_{CM}(t)$ and $y_{CM}(t)$, over the interval [t_0, t_1], as in Equations 6.4.1 and 6.4.2. The trajectory is represented by a finite number of segments that connect all the points (in this case, bed coordinates) representing all the positions taken by the center of mass during a movement.

7.3.2 Feature Extraction

The feature set consists of 3 features that are extracted from the trajectory of the body center of mass: the distance between initial and end points of the trajectory, the length of the trajectory, and the variance of the trajectory in the y-direction perpendicular to the sleeper's body axis. The choice of features was motivated by consideration of the nature of the movements to be discriminated, including our understanding of the kinematics involved and our observation of free movements during the trials.

The distance between the initial and end points of the trajectory is the length of the path connecting the last position of the body center of mass just before a movement has started and the final position of the body center of mass after a movement ends. This feature provides spatial information about a movement in terms of the displacement of the body center of mass as a result of the movement.

The length of the trajectory is the total length of all the segments that constitute the trajectory. This feature indirectly incorporates temporal information about the movement through the number and length of the segments that constitute a given trajectory. Intuitively, it is reasonable to assume that the length of the trajectory is considerably larger for class-1 movements, and smaller for class-3 movements. The trajectory is very short for arm and head movements in class 2 because there is not a considerable displacement of mass during these movements. The length of the trajectory

is not necessarily proportional to the duration of a movement because it depends on the parts of the body that were involved in the movement (different body parts affect the center of mass differently). For example, a movement can only involve arms and head movements and last as long as other movements, but the observed changes in the position of the body center of mass during movement are likely to be small.

The sample variance of the trajectory in the *y*-direction, which corresponds to motion from one side of bed to the other (Figure 6.4), also provides spatial information about a movement. The sample variance of the trajectory in the *y*-direction $\hat{\sigma}_y^2$ is calculated as follows:

$$\hat{\sigma}_y^2 = \frac{\sum_{i=1}^{N} (y_{CM}(i) - \bar{y}_{CM})(y_{CM}(i) - \bar{y}_{CM})^T}{N-1},$$

where N corresponds to the number of observations over the interval of the movement and $\bar{y}_{CM}$ corresponds to the sample mean. This feature is particularly useful in discriminating movements involving upper and lower body in class 2 from lower body movements in class 3. A displacement of the torso in the *y*-direction occurs more frequently when adjusting position for class 2 (medium amplitude) than when performing a leg movement. This is consistent with observation of the movements made by the subjects during the free protocol trial (when they were asked to move freely, as described in Section 4.1.2). Therefore, the fact that the variation in the *y*-direction is smaller when someone only moves the legs facilitates the discrimination.

7.3.3 Statistical Modeling

As noted above, the bases for the movement classification system are the GMMs used to model the underlying time-domain representations that characterize a person's motion patterns. The goal of the statistical modeling step is to estimate the parameters of each GMM that represents a certain movement class c_k, using training data from movement k.

Finite mixtures are a flexible and powerful probability modeling tool [83]. A linear combination of Gaussian basis functions is capable of forming smooth approximations of arbitrarily shaped densities. GMMs have been used successfully for

similar problems such as the task of discriminating six classes of limb motions from myoelectric signals [87].

A Gaussian mixture model describes the probability distribution of a given data set as a linear combination of several Gaussian densities [74]. In this model, each *d*-dimensional random vector *x* is assumed to be drawn independently from a mixture density given by the equation

$$p(\mathrm{x}|\Theta) = \sum_{i=1}^{M} \varpi_i \, p(\mathrm{x}|\mu_i, \Sigma_i), \qquad 0 \le \varpi_i \le 1 \quad \text{and} \quad \sum_{i=1}^{M} \varpi_i = 1 \tag{7.1}$$

where ϖ_i defines the mixing weight of the i^{th} Gaussian component (for all i = 1, 2, …, M) given by the relative importance of each component in the density function, $p(\mathrm{x}|\mu_i, \Sigma_i)$ represents the i^{th} d-dimensional Gaussian component with mean $\boldsymbol{\mu}_i$ and covariance $\boldsymbol{\Sigma}_i$ given by

$$p(\mathrm{x}|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} exp\left\{ -\frac{1}{2} (\mathrm{x} - \mu_i)' \Sigma_i^{-1} (\mathrm{x} - \mu_i) \right\},$$

and $\boldsymbol{\Theta} = \{\varpi_1, \ldots, \varpi_M, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_M, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_M\}$ represent the mixture density parameters.

The mixture density parameters are estimated using a maximum-likelihood approach. The expectation maximization (EM) [74] is the algorithm used for finding the maximum-likelihood parameter estimates, and it is described in Appendix E. The mean of each component in Equation 7.1 is initialized to a centroid derived from the Linde, Buzo and Gray (LBG) algorithm [88]. The LBG algorithm is also described in Appendix E.

Using this notation, each class c_k is represented by a GMM, and it is referred to by its model $\Theta_k = \{\varpi_1, \ldots, \varpi_M, \mu_1, \ldots, \mu_M, \Sigma_1, \ldots, \Sigma_M\}$. The degrees of freedom or the number of parameters that must be estimated to approximate the model for a class using a *d*-dimensional feature vector using *M* mixtures are: *M* mixing weights, *Md* means, and M*d* (*d* + 1)/2 covariance elements (since there are *d* diagonal elements and *d* (*d* - 1)/2 off-diagonal elements). In order to reduce the number of parameters to *M* (2*d* + 1), we use GMMs with diagonal covariance matrices. Although the use of a diagonal covariance matrix has the underlying assumption that the features are uncorrelated, this simplified representation is sufficient for the purpose of the classification task, since we have

observed that the performances of the diagonal matrix GMMs are not significantly different that the performances of the full matrix GMMs.

7.3.4 Likelihood Estimation and Decision

This step consists of estimating the likelihood of a class, and assign a class label based on the maximum likelihood (ML) decision rule [84] as follows:

$$\hat{c}_k = \underset{c_k}{arg\,max}\, p(x \mid c_k)$$

The difference between choosing the ML rule and the maximum a posteriori (MAP) rule lies in the knowledge about the priors $p(c_k)$, $k = 1, 2, 3$. For the ML rule, the decision is entirely based on the likelihoods $p(x \mid c_k)$. For the MAP rule, the decision rule is

$$\hat{c}_k = \underset{c_k}{arg\,max}\, p(x \mid c_k) p(c_k),$$

where both likelihoods and priors are important in making a decision. Because we have no knowledge about the priors, and because the priors may vary from person to person, it is reasonable to assume to use the ML rule in the context of this work.

7.4 Results

In this section, we present the classification performance evaluated on the laboratory data described in Sections 4.1 and 4.2. We also analyze the effect of training set size on the classification performance, and then we evaluate a subject-independent model. Prior to presenting the performance results, we discuss our approach to the optimal choice of the number of mixtures in the GMMs.

7.4.1 Data Preparation

The classification approach is evaluated individually for each of the 15 subjects. For each subject, movement data from the trials are randomly split into 2 sets: training (3/5 of the dataset) and testing (2/5 of the dataset). The classifier is designed using the training set, and the performance is evaluated on the test set. Table 7.1 shows the number of movements per class for each subject, and the total number of movements per class in the training and testing data. The training data contain 1711 movements and the testing data contain 1107 movements. The average duration of class-1 (large) movements is 7.98 (±

2.29) seconds, 4.06 (± 1.8) seconds for class-2 (medium) movements, and 4.94 (± 1.23) seconds for class-3 (legs) movements.

Table 7.1: Evaluation Data

Subject Number	Training Data			Testing Data		
	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
1	40	56	35	26	37	22
2	39	66	35	25	44	23
3	42	65	21	28	43	14
4	38	60	30	24	40	19
5	17	30	16	11	20	8
6	21	25	23	14	16	14
7	23	26	20	15	17	13
8	39	50	39	26	33	26
9	20	39	18	13	26	4
10	39	49	39	26	32	25
11	38	68	26	25	44	16
12	39	63	25	26	41	16
13	41	45	42	26	30	28
14	38	72	17	25	48	10
15	40	53	44	26	34	28
Totals	514	767	430	336	505	266

The individual differences in the number of movements per class occurred because:

1) Some subjects performed more trials than others.

2) Even though each trial consisted of 6 class-1 movements, 8 class-2 movements and 6 class-3 movements, every time someone moved another body part besides the legs during isolated leg movements, the movement had to be classified as class 2 instead of class 3. Even though such modification introduced an uneven distribution of movements between classes across subjects, it provided data from movements that are very common in natural settings that were not introduced in the experiment in the first place.

3) Extra movements not included in the protocol, which were performed when the subject was supposed to be still, were also included.

4) The number of movements per class in the free-protocol trial varies per person.

Also, subjects 1, 2, 3, 4, and 8 have data from 2 different bed sizes. Since we have not found a significant difference in the classification performances across beds, the results shown are calculated based on combination of these datasets.

7.4.2 Performance Measure

The performance measure used in this work is the classification rate across all subjects, which is the proportion of test samples from all subjects that are correctly classified. The classification rate across all subjects is used because we want to measure the overall performance of the classifier independently of the subject. The classifier performance is reported based on knowledge of the true movement intervals, i.e., load cell data from each movement are extracted using the true intervals, and not the detected intervals from the detection step. However, the effect of the detector's boundary errors on the classifier's performance is investigated.

Unless indicated otherwise, all the future comparisons between different classifier conditions are done with the McNemar's test [83]. This test is used to determine whether the difference between 2 classifiers is significant or not, and it is computed as following. Suppose that we have 2 classifiers, A and B. Let

n_{00} = number of samples misclassified by both A and B,

n_{01} = number of samples misclassified by A but not by B,

n_{10} = number of samples misclassified by B but not by A, and

n_{11} = number of samples misclassified by neither A nor B.

The idea is that there is little information in the number of samples on which both systems under consideration get the correct result, or for which both get an incorrect result. These may be excessively easy or excessively difficult samples. So, the test is based entirely on the values of n_{01} and n_{10}. The null hypothesis that the classifiers have the same error can be rejected with probability of incorrect rejection of 0.05 if $|z| > 1.96$, where $z = \frac{|n_{01} - n_{10}| - 1}{\sqrt{n_{10} + n_{01}}}$. The quantity z^2 is χ^2 distributed with one degree of freedom. Unless specified, the level of significance is set to $\alpha = 0.05$.

7.4.3 Number of Mixtures Components

The number of mixture components in the GMMs is estimated using 3-fold cross-validation of the training set [83]. Training data from each subject are randomly split into 3 disjoint sets, each containing roughly the same number of data samples. Each set is used in turn as an independent test set while the remaining 2 sets are used for training. The classification rate is estimated over all sets. Figure 7.6 shows the effect of increasing the number of components in the Gaussian mixture model. Increasing the number of components beyond 3 increases the error, and models with as few as 2 components give reasonable performance. The difference in performance for 1 and 2 components is statistically significant with a significance level at 0.01. The difference in performance for 2 and 3 components is not statistically significant, and the difference in performance for 2 and 4 components is statistically significant with a significance level at 0.01. We chose to use 2 mixtures.

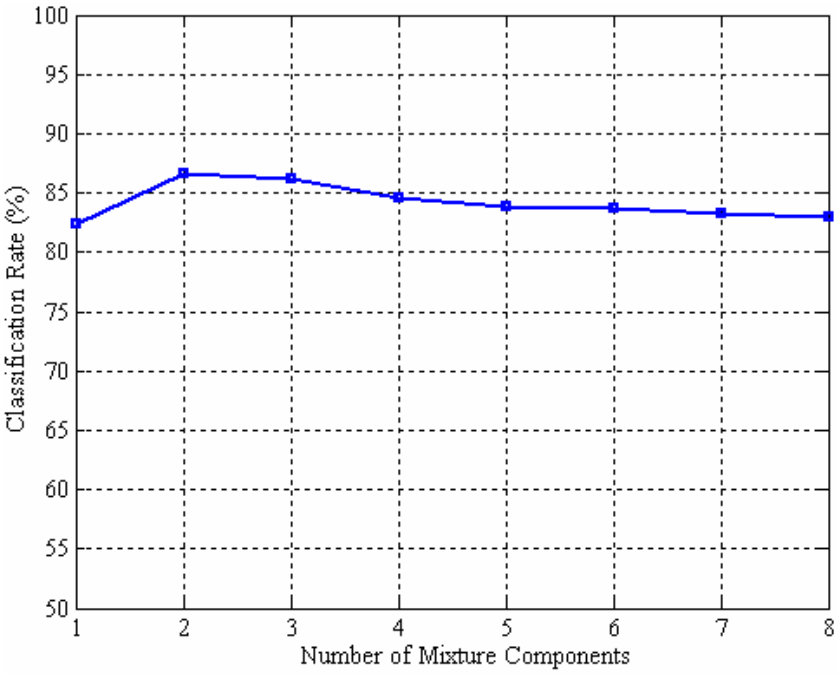


Figure 7.6: Classification rate as a function of the number of mixture components.

7.4.4 Performance Results

The overall classification rate on the test data is 84.6%, and the corresponding confusion matrix is presented in Table 7.2. The most frequent errors are between classes 2 and 3. A closer examination of the errors show that, in many cases, the classifier mistakenly

classified movements consisting of leg movements and very small adjustments of head or torso (class 2) as leg movements (class 3). In such cases, the small movements in the upper body do not substantially affect the overall trajectory of the center of mass. We speculate that this type of mistake may also be common when the same task is done by a human scorer because, depending on the conditions for visualization of the movements, small movements in the upper body may be missed. The classifier also mistakenly classified arm movements (class 2) as leg movements (when one leg was moved) a few times, and these mistakes were more common among short and light subjects (subjects 2, 3 and 4).

Table 7.2: Confusion matrix for the 3-class movement classification problem: large, medium and leg movements.

		Estimated		
		Class 1:Large	**Class 2: Medium**	**Class 3:Legs**
True	**Class 1: Large**	325	9	2
	Class 2: Medium	13	391	101
	Class 3: Legs	2	44	220

When the variance of the trajectory in the *x*-direction is included in the feature vector, the overall classification rate is 85.8%. Although we had previously expected that such feature would help discriminating leg movements (because the center of mass moves up/down when someone bends/straightens a leg), the difference in performance between the 3-dimensional feature vector and the 4-dimensional feature vector is not statistically significant.

The individual performances are shown in Figure 7.7. The χ^2 test for differences among proportions [79] shows that there is a difference between the performances across subjects, at a significance level of 0.05 and with 14 degrees of freedom. The degree of difficulty of a classification problem depends on the variability of the features in a class. We speculate that there are differences in the classification performance across subjects because the intrasubject movement variability may be larger in some subjects, which results in a larger intraclass variance. The intrasubject movement variability is due to the fact that voluntary movement is continuously influenced by human neuromusculoskeletal system constraints (biomechanical and anatomical) [89].

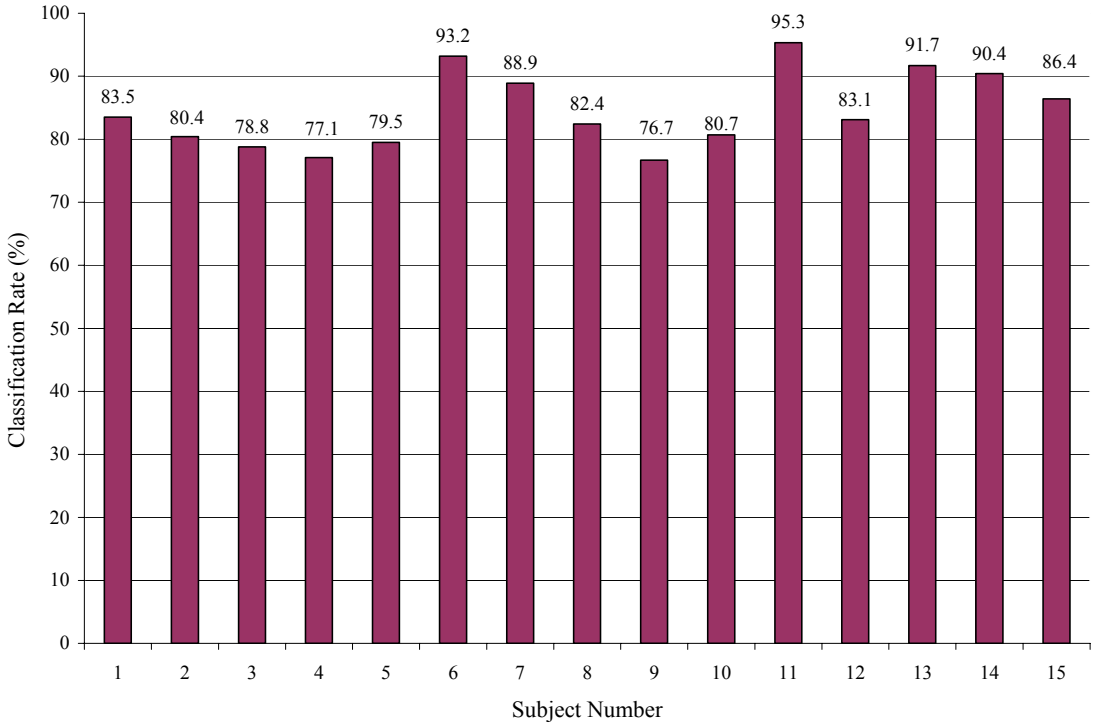


Figure 7.7: Individual classification performances.

The detection errors in the onset and offset of the movements have no effect on the performance of the movement classifier. The performance of the classification using the estimated boundaries is 83.4%. Such performance is not statistically different from the performance using the ground truth boundaries. This result shows the robustness of the features with respect to small differences in the detected boundaries.

Effect of Training Set Size

The use of subject-dependent models requires learning parameters for each subject with data of each subject. The disadvantage is that it takes time to collect subject-dependent data. To select the most appropriate parameterization, it is important to know the minimum amount of data necessary to train the model for each person.

We examined how the classification rate on the testing data changes as we progressively increased the amount of training samples per class. Because the number of samples per class is different for each subject, we only included in this analysis subjects that had at least 30 movements per class. The seven subjects included were 1, 2, 4, 8, 10, 13, and 15. The overall classification rate, which was calculated on the test data previously selected to each subject, was computed as a function of the number of training

samples, with the number of samples increasing from 5 to 30 in increments of 5. Figure 7.8 shows the classification rate as a function of the number of training samples per class for the chosen test set. The term 'All' in the plot corresponds to the value of the classification rate when all the available training samples were used.

Using all samples available in the training set, the overall classification rate is 83.2%. The classification rate is 81.5% when using only 10 samples per class, and the difference between these performances is not statistically significant. The difference in performance for 5 and 10 samples is statistically significant with a significance level at 0.01. Therefore, to achieve comparable results for the data examined in this work, at least 10 samples per class are necessary to train the model for each person.

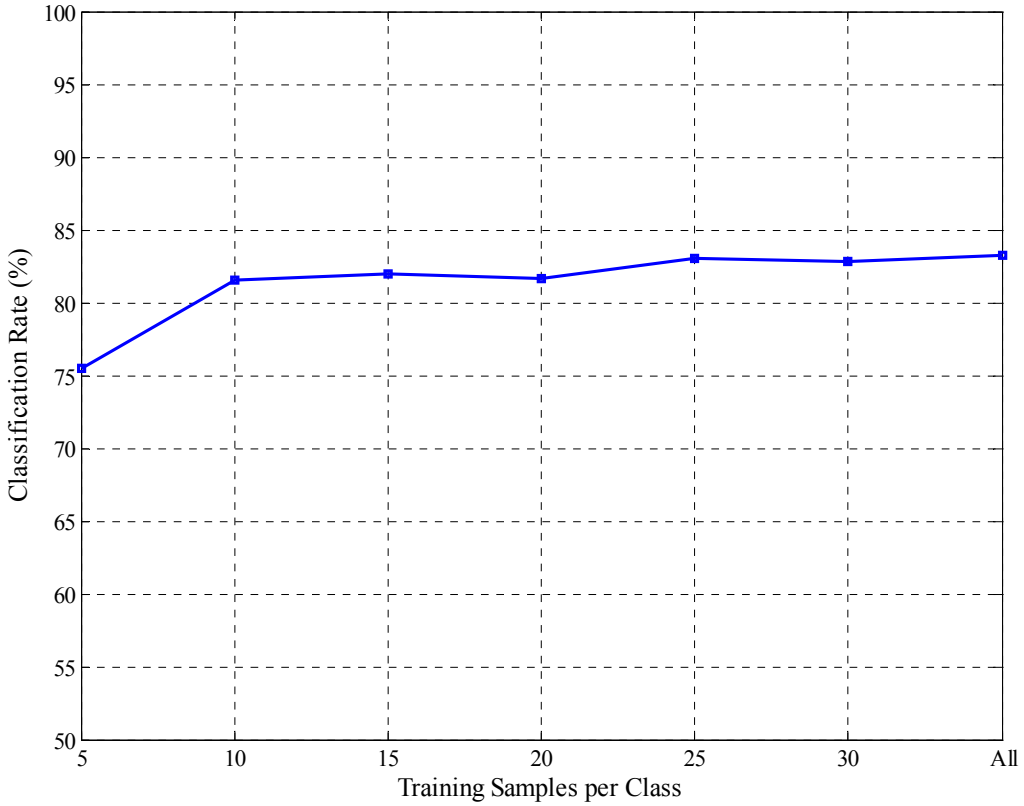


Figure 7.8: Effect of the training set size: classification rate as a function of the number of training samples per class.

Subject-Independent Models

Subject-independent models are always preferable and especially important in applications where training is difficult or impossible to conduct. The disadvantage of the subject-independent models is a loss in performance due to the additional inter-subject

variability. In order to investigate the decline in performance, we use the same pattern recognition approach and representation as in the subject-dependent model, but train the model with data from a number of subjects.

The subject-independent model is built using a leave-one-out approach, i.e., using training data from 14 subjects to build the model, and use the data from the 15th subject left out from the group to test the model. The same procedure is repeated for all subjects, and the results of each tested subject are combined to obtain the overall performance. We tested with different number of mixture components, and models with as few as 4 components give reasonable performance. The difference in performance for 2 and 4 components is statistically significant with a significance level at 0.01. The difference in performance for 4 and 6 components is not statistically significant. We chose to use 4 mixtures.

The classification rate on test data for a model with 4 mixtures is 80.3%. The difference in performance between the subject-independent and the subject-dependent (performance of 84.6% with 2 mixtures) models is statistically significant, with a significance level at 0.01. Although the subject-dependent model yields better results, the relative improvement of 5.35% over the subject-independent model is fairly small. We note that the subjects' weight varied by a factor of 3:1 (maximum weight is 275 lbs, and minimum is 85 lbs). This suggests that the selected feature representation is relatively invariant across subjects.

Large fluctuations in the individual classification rate are expected due to possible mismatches in the training data between existing subjects and the new ones [90]. The individual performances for both methods are shown in Figure 7.9. While for some subjects (subjects 1, 5, 6, 7, 10, 13, and 15) the performance does not degrade considerably with the subject-independent model, there are differences for others. For example, for subject 11, for example, the classification rate drops from 95.3% to 69.4% when using the subject-independent model.

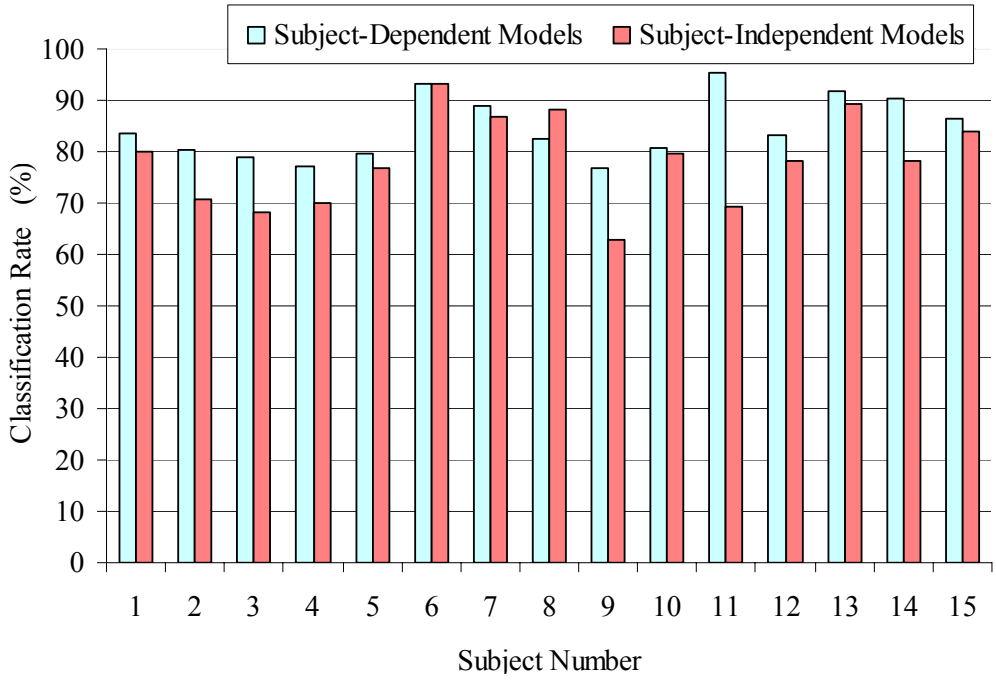


Figure 7.9: Individual performances for subject-dependent and subject-independent models.

7.5 Using a Time-Frequency Representation

Although the feature representation described in Section 7.3.2 was selected considering the physical model of human body, it is possible that a representation based on temporal decomposition of the signals would be more complete and therefore yield better results. Our cursory analysis of the nonlinearity of the underlying system suggested that movements of the small body parts would be nearly linear. This notion provided motivation for investigating an additive time-frequency representation with features extracted using wavelet-based multiresolution analysis (MRA). Through a zooming procedure across scales, the multiresolution analysis may provide a better characterization of the transients that represent different types of movements [91].

Many successful applications of the wavelet transform in pattern classification problems have been reported [92-94]. Englehart [92] uses a wavelet packet based feature set, subject to dimensionality reduction using principal components analysis (PCA), for discrimination of four classes of upper limb motions from two-channel myoelectric signals. Kundu et al. [93] use wavelet coefficients selected using a thresholding operation to discriminate between 6 classes of underwater acoustic signals using a Hidden Markov Model (HMM) as classifier. Yen et al. [94] use the energy values of selected wavelet

packet nodes to classify 8 classes of vibration signals that constitute the signatures of faulty components in industrial machinery equipments.

Details about the feature extraction and feature subset selection steps are described in details next.

7.5.1 Feature Extraction

The classification of each movement is concentrated around a given time interval $[t_0, t_1]$ that defines the start and end of the movement. In the feature extraction step, the discrete wavelet transformation is applied only over the interval containing most of the intensity of the movement. The intensity of a movement refers here to the effort made to make a certain movement.

Due to the fact that 1) each detected movement signal has a specific duration, 2) the wavelet coefficients are used as features, and 3) the number of wavelet coefficients varies with the length of a signal, the wavelet transformation cannot be applied directly on the interval $[t_0, t_1]$. For each movement, the algorithm first finds a segment S of length T, in seconds, which corresponds to the interval where the intensity of the movement is the largest. The wavelet transformation is then applied to the signals $x_{CM}(t)$ and $y_{CM}(t)$ only over the interval defined by S. With this operation, only the segment containing the most concentration of the intensity of the movement is kept for feature extraction.

Let $w_i(t)$ be the function that represents the load cells signal for a given movement defined over a time interval $[t_0, t_1]$, for $i = 1, 2, 3, 4$. The segmentation consists of the following steps:

1. Compute the square differences for each load cell signal, at time t, as follows:

$$SD_i(t) = \left(w_i(t) - w_i(t-1)\right)^2, \qquad t_0 \le t \le t_1,$$

where $w_i(t-1)$ is the load cell signal measured at previous sampling time. The intensity of a movement is captured through the computation of the square differences for each load cell signal because the magnitude of the changes in the signals are related to the force applied at each time t. That is, the force is proportional to the amount of weight that is being moved and, therefore, directly related to the intensity.

2. Compute the sum of the square differences across all load cell signals, at time t, as follows:

$$SSD(t) = \sum_{i=1}^{4} SD_i\,(t), \quad t_0 \le t \le t_1 .$$

3. Estimate a segment of length T that is centered at time instant t_C, where t_C corresponds to the t-coordinate of the centroid of the region below *SSD(t)* on the interval $[t_0, t_1]$. The x-coordinate of the centroid of a region bounded by two curves $f(x)$ and $g(x)$, on the interval $[a,b]$ [75], is given by

$$\bar{x} = \frac{\int_a^b x\big(f(x) - g(x)\big)dx}{\int_a^b \big(f(x) - g(x)\big)dx},$$

In this case, the t-coordinate of the centroid of the region below *SSD(t)* corresponds to

$$t_C = \frac{\sum_{t=t_0}^{t_1} tSSD(t)}{\sum_{t=t_0}^{t_1} SSD(t)},\ t_0 \le t_C \le t_1 .$$

The segment S is thus defined over the interval $\left[t_C - \frac{T}{2}, t_C + \frac{T}{2}\right]$. However, the segment S can be long enough that includes data from the previous or next movement. In order to avoid including data from movements besides the movement being analyzed, the segment S is defined as

$$\left[max\left(t_C - \frac{T}{2}, t_{1,Previous}\right), min\left(t_C + \frac{T}{2}, t_{0,Next}\right)\right] \tag{7.2}$$

where $t_{1,Previous}$ is the offset of the previous movement and $t_{0,Next}$ is the onset of the next movement. Since the segment defined in Equation 7.2 can be smaller than T seconds due to the overlap of the segment and the previous/next movement, the length of the segment S is corrected by performing a symmetric boundary value replication. Figure 7.10 shows an example of the computation of the segment S for a class-1 movement. The load cell signals $w_i(t)$, for i = 1, 2, 3, 4, are shown in the top plot, and the correspondent square differences $SD_i(t)$ are shown in the middle plot. The sum of the square differences across

all load cell signals $SSD(t)$ is shown in the bottom plot, where the vertical dotted lines represent $\left[t_C - \frac{T}{2}, t_C + \frac{T}{2}\right]$ with length T = 3 seconds, and the solid vertical line shows the position of the center of the segment, t_C.

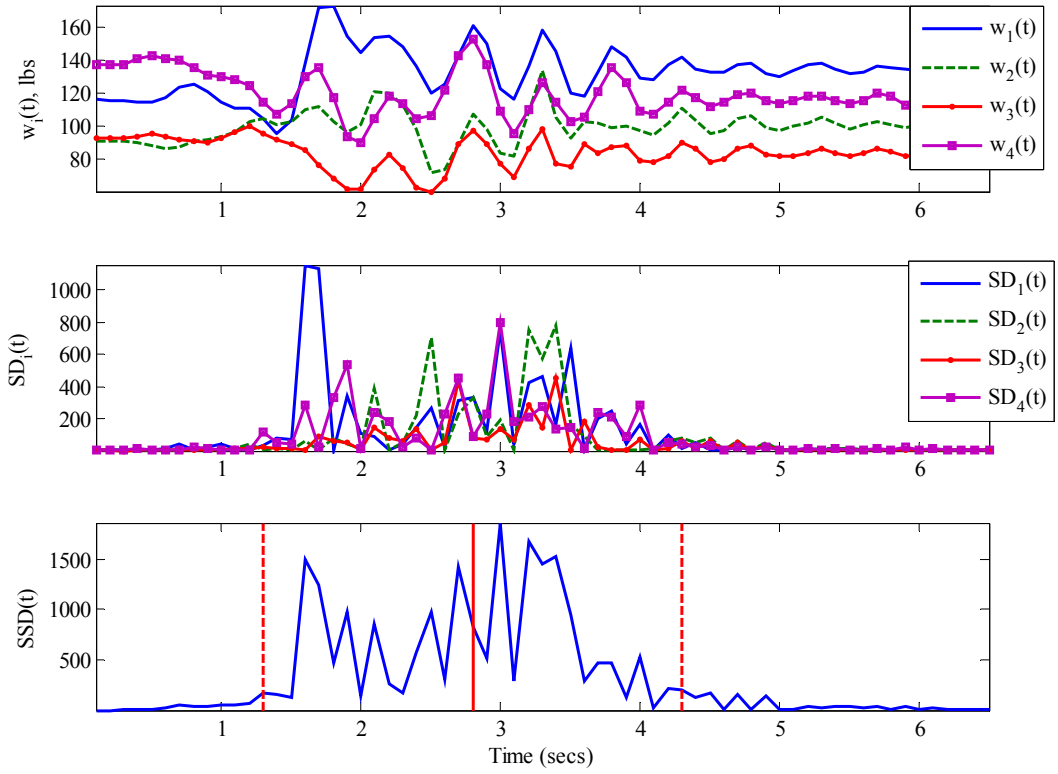


Figure 7.10: Load cell signals $w_i(t)$ in pounds (top) and square differences $SD_i(t)$ (middle), for i = 1, 2, 3, 4, during a class-1 movement. Correspondent SSD(t) is shown in the bottom plot. Vertical dotted lines in the bottom plot show the boundaries of the segment S with length T = 3 seconds, and the solid vertical line shows the center of the segment located at t_C.

The signals $x_{CM}(t)$ and $y_{CM}(t)$ from the segment S are then decomposed using the Discrete Wavelet Transform (DWT) with J_0 levels of decomposition. A review of the DWT is given in Appendix F. From each decomposed signal, the following set of features is extracted:

- The energy of the details coefficients at the first decomposition level E_1: this feature has been chosen to capture the energy distribution of the high frequency components of the movements. The energy of the detail coefficients at the first level is estimated by

$$E_1 = \sum_{n=1}^{2^J/2} D_{1,n}^2 , \qquad (7.3)$$

where $2^J/2$ is the number of elements in the vector D_1 for a signal with length 2^J. In this case, 2^J corresponds to the segment length T.

- The wavelet detail coefficients vectors for the remaining levels D_j, for $j = 2, 3, \ldots, J_0$: the coefficients corresponding to the slower varying components of the movements.

Given that the trajectory of the center of mass signal is sampled at 10 Hz, the detail coefficients in the first level, D_1, represent the signal in the approximately 2.5-5 Hz range, D_2 in the range 1.25-2.5 Hz, D_3 in the 0.625-1.25 Hz range, and so on. In the last step, features from each signal are concatenated into a single feature vector to form a high-dimensional vector. The dimension of the feature vector thus depends on the segment length T and on the choice of the number of decomposition levels J_0.

7.5.2 Feature Subset Selection

In feature selection step, the *sequential forward selection* method [82] is used to reduce the dimensionality of the original feature set. Concerning both the computational efficiency and performance of a recognition system, one is usually interested in a feature space of low dimensionality. The relatively high dimensionality is due to 1) the fact that features are extracted from two signals and concatenated into one feature vector, and 2) the fact that time-frequency representations have high dimension. Therefore, the goal of the feature selection step is to reduce the number of features by selecting the best subset of the original feature set, according to some criterion.

In general, the problem of feature selection can be described as selecting the best subset x of d features from the set y of $D > d$ available features. Feature selection consists of 2 components:

1. **Feature Selection Criterion $J(\cdot)$:** a criterion must be established by which it is possible to judge whether one subset of features is better than another. The feature set can be selected in 2 ways as follows:
   - Design a classifier on the reduced feature set, and choose the feature sets for which the classifier performs well on a separate test/validation set. The criterion is the error rate, and the feature set is chosen to match the classifier. A different feature set may result with a different choice of

classifier [83]. This method is also known as the wrapper method because it conducts a search for a good subset using the classifier itself as part of the selection criterion [95].

- Estimate the overlap between the distributions from which the data are drawn and favor those feature sets for which this overlap is minimal (that is, maximize separabiliy). This is independent of the classifier, but has the disadvantage that the assumptions made in determining the overlap are often crude and may result in a poor estimate of the discriminability [83].

2. **Search Method:** a systematic method must be used for searching through candidate subsets of features. There are 2 basic methods for feature subset selection:
   - Optimal methods: these generally require exhaustive search methods (that evaluate all possible subsets), which are feasible for only very small problems [83]. The branch and bound algorithm avoids an exhaustive search by using intermediate results for obtaining bounds on the final criterion value [82].
   - Suboptimal methods: the optimality of the above methods is traded for computational efficiency [83]. Although they are not capable of examining every feature combination, they assess a set of potentially useful feature combinations [83]. Some of the algorithms include the best individual N features, the sequential forward selection, the sequential backward selection, and "plus l – take away r" selection [83].

Our approach uses the classification rate calculated on the training data as the selection criterion. The training data are split into n approximately equally sized partitions and, for a given subset of features, the statistical model is estimated using n -1 partitions, and the remaining partition is used as test set. The classification results from each of the n runs (n is equal to 3 in this case) are summed to produce the estimated classification rate.

The Sequential Forward Selection (SFS) is the algorithm used as the search method [83]. The SFS method (or the method of set addition) is a bottom-up search procedure that adds new features to a feature set one at a time until the final feature set is

reached. Suppose we have a set of d_1 features, $x = (x_1, x_2, \ldots, x_{d_1})$. Let the feature selection criterion function for the set x be represented by $J(x)$. For each of the features ξ_k not yet selected (i.e., in y - x, where y is the original feature set), the criterion function $J_k = J(x + \xi_k)$ is evaluated. The feature that yields the maximum value of J_k is chosen as the one that is added to the set x. Thus, at each stage, a variable is chosen that, when added to the current set, maximizes the selection criterion. The feature set is initialized to the null set [83]. The algorithm may be terminated when the maximum allowable number of features is reached. The reasons for choosing the SFS algorithm instead of the Sequential Backward Selection (SBS), which starts with a full set of features, are: 1) there are not sufficient data to accurately estimate the large number of parameters resulting from a high-dimensional feature vector, and 2) it is computationally more demanding to start with the full set of features and delete one at a time. Although, in theory, going backward from the full set of features may capture interacting features more easily [95].

7.5.3 Parameters Optimization

The selection of the feature subset depends on the choice of the segment length T, the wavelet mother ψ, and the number of decomposition levels J_0. Since the feature selection is performed in conjunction with the classifier, the choice of the final subset also depends on the number of mixture components M and the feature subset dimension. The dimension of the feature subset is defined last, based on the performances resulting from the optimized parameter values of T, ψ, J_0 and M. The parameters optimization performs a restricted search of the parameter space according to the parameter values defined by data analysis and by restrictions of the classifier as follows:

- **Segment Length:** the length of the segment where most of the intensity of a movement is concentrated, T, needs to be long enough to include sufficient information to discriminate among different classes. The search for the segment length spans from 2 to 10 seconds (the upper limit is approximately the mean duration plus one standard deviation of the class-1 movements – the longest class).

- **Wavelet Mother:** there are no uniformly optimal ways of determining which wavelet is more sui for a given classification problem. Therefore, we selected the most commonly used wavelets, namely Haar, Daubechies, Symmlets, and Coiflets (the mother wavelet properties are described in Appendix F), with the following filter lengths: Haar (filter length is 2), Daubechies of order $n = 2$ to $n = 5$ (filter length is $2n$), Symmlets of order 2 to 5 (filter length is $2n$), and Coiflets of order 2 (filter length is $6n$).
- **Level J_0 of Partial DWT:** although the decomposition can be carried out to its maximum depth $J = log_2 N$, where N represents the length of the decomposed signal, most applications require only a subset of the coefficients. In order to avoid the decomposition of levels that contain only one coefficient, the values for J_0 vary from 2 to the largest integer corresponding to $log_2 T$.
- **Number of Mixture Components:** the selected values for M vary from 2 to 5 mixtures. Given the limited number of training samples per class, we restricted the maximum number of mixtures to 5 in order to have sufficient data for the model estimation.

The optimization is based on the overall classification rate on the training data, as described in Section 7.5.2, and consists of 4 steps. In the first step, the performances of all combinations of parameters are computed. In the second step, for each feature subset dimension, the combinations corresponding to the best performances, which are statistically significant from the remaining performances, are selected. In the third step, only the combinations that are common in all feature set dimensions are kept. In the last step, the combination of parameters with the least complexity is chosen. The optimal values were found to be: segment length T is equal to 8 seconds, the wavelet mother ψ is db6, the number of decomposition levels J_0 is 5, and the number of mixture components M is 2. There are no statistical differences in performance for $T = 8, 9$, $J_0 = 5, 6$, and for $M = 2, 3$.

The remaining parameter to be selected is the dimension of the feature subset. A rank plot of the classification rate versus the feature subset dimension, for the optimal parameter values and for the first 10 features, is shown in Figure 7.11. The performance on the training data for a 3-dimensional feature subset is 90.3%, and it is 92.9% for a 6-

dimensional feature subset. The addition of features beyond 6 dimensions does not improve the performance, as shown in Figure 7.11. The difference in performance for a feature subset with 3 and 6 dimensions is statistically significant with a significance level of 0.01. Therefore, for the wavelet-based features, the best result on the training data is obtaining using a feature subset with 6 dimensions.

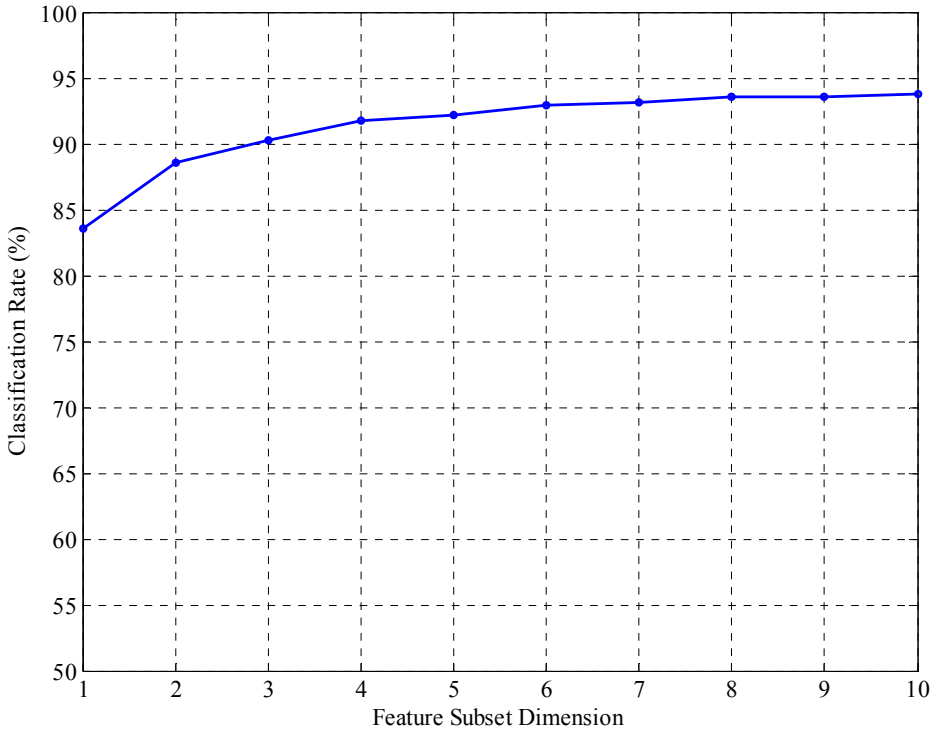


Figure 7.11: Rank plot of the classification rate versus the feature subset dimension, for the optimized parameter values.

7.5.4 Results

Using the optimal parameter values, the overall classification rate on the test data for a 3-dimensional feature vector is 82.2%. The correspondent confusion matrix is presented in Table 7.3. The overall classification rate on the test data for a 6-dimensional feature vector is 83.5%, and the difference in performance between the system with a 6-dimensional and a 3-dimensional wavelet-based feature vector is not statistically significant. Because there was no difference in performance on the test data, and because the number of degrees of freedom for a GMM depends on the dimension of the feature set and on the number of mixtures, we compare the performance of the 2 representations (time domain and time-frequency domain) on the test data based on the same number of

degrees of freedom for a fair comparison. Thus, unless otherwise indicated, the results are presented based on a 3-dimensional feature vector.

Table 7.3: Confusion matrix for the wavelet-based representation for the 3-class classification problem: large, medium and leg movements.

		Estimated		
		Class 1: Large	**Class 2: Medium**	**Class 3: Legs**
True	**Class 1: Large**	321	8	7
	Class 2: Medium	28	386	91
	Class 3: Legs	10	53	203

These results suggest that despite the nonlinearity of the bed-human system discussed in Section 3.2, for the purposes of movement classification, the bed-human system can be viewed in terms of an additive representation. As noted above, this is consistent with the fact that for movements of small body parts, the system approximates a linear system response. Therefore, an additive representation is sufficient to represent fast and slow portions of a movement.

An analysis of the selected features for each subject shows that the energy of the details coefficients at the first decomposition level of $y_{CM}(t)$ appears at the top 3 best features for 60% of the subjects. Such feature represents the high-frequency components of the movement along the *y*-direction (side to side of the bed). Note that this feature conveys similar information to the feature used in the time-domain – the sample variance of the trajectory in the *y*-direction. Thus, how the movement is performed along the *y*-axis is important for discriminating the type of movement.

Another similarity observed across subjects is that, besides the energy of the details coefficients at the first level, the coefficients that correspond to the extremities or the middle of the decomposed segment are selected for classification. Given that a long segment (8 seconds) was chosen and class 2 and 3 are on average 4 seconds long, this implies that those coefficients are being selected because they can discriminate class 1 from class 2-3 movements just by using the information at the extremities of the segments. However, the long segment does not contain sufficient information in the extremities to better discriminate between class-2 and class-3 movements. Therefore, we

can speculate that the coefficients from the middle of the segment are selected to minimize the confusion between class 2 and class 3.

The detection errors in the onset and offset of the movements have no effect on the performance of the classifier. The performance of the classification using the estimated boundaries is 81.5%. Such performance is not statistically different from the performance using the ground truth boundaries. It shows that the wavelet-based features are robust to small differences in the detected boundaries. The errors in the onset and offset of the movements have less or no affect on the performance because the feature extraction is centered on the time interval containing most of the intensity of the movement.

Comparison between Feature Representations

Despite their simplicity, the time-domain features contain enough information to perform as well as the wavelet-based features on the classification task. The difference in performance between the two representations is not statistically significant. A comparison of the confusion matrices of the 2 representations (Table 7.2 and Table 7.3) shows that the degree of feature overlap that causes the confusion between classes 2 and 3 has not been solved with this new representation. Figure 7.12 shows the individual performances for the two feature presentations used.

In addition, both representations are robust to differences in bed type. No differences in the classifier's performance were observed when data from different bed types were combined or tested separately for the subjects with data from 2 different bed types.

7.5.5 Alternative Representations Investigated

This section presents the classification results obtained from alternative signal representations. We experimented with the following representations: raw load cell signals, subband energy features, wavelet packet-based features, fusion of feature types, principal component analysis (PCA) and linear discriminant analysis (LDA) for dimensionality reduction.

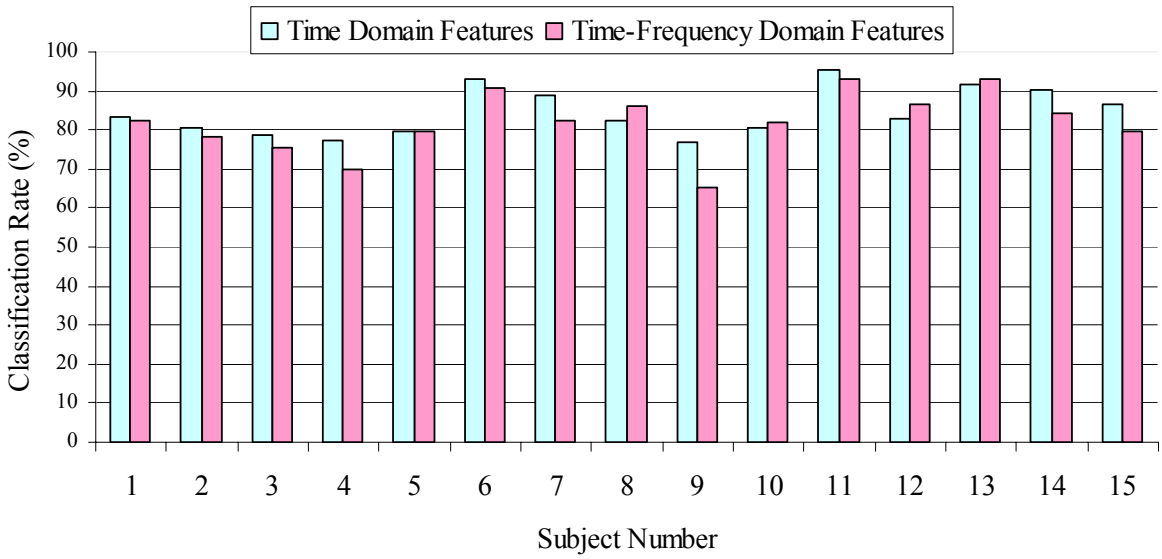


Figure 7.12: Individual classification performances for the time-domain and time-frequency domain feature representations.

One advantage of applying the wavelet transformation on the signals $x_{CM}(t)$ and $y_{CM}(t)$ instead of applying directly on the raw load cell signals $w_i(t)$ is that the original feature set is smaller when using the trajectory of the center of mass because the number of signals to be decomposed is reduced by half. The originally selected feature set based on the trajectory of the center of mass reduces the number of signals to be decomposed. However, because additional information may be present in the raw load cell signals that is lost when estimating the trajectory of the center of mass, it is important to consider the performance of the raw load cell signals. The performance of the movement classification approach based on the wavelet decomposition of $w_i(t)$ is 80.7%, and the difference between the performances is not statistically significant. The results thus show that there is not additional information in the raw load cell signals.

Because the energy of the detail coefficients at the first decomposition level was one of the top 3 best features for a majority of subjects, we also examined the performance of the system using the subband energy at all 5 decomposition levels. The performance of the movement classification approach with a subband energy feature representation is better than the performance with the proposed wavelet-based features. The subband energies for each of the 5 decomposition levels are computed with Equation 7.3. The same method for feature subset selection was applied on the subband energies to select the top 3 features from the 10 (5 levels x 2 decomposed signals) original features.

The classification performance is 85.4% for the subband energy feature representation. The difference in performance between the subband energy and the proposed wavelet-based representations is statistically significant with a significance level of 0.01, and the relative improvement in performance is 3.89%. This representation discards the temporal information when replacing the wavelet detail coefficients vectors by the correspondent subband energies. One reason for such difference in performance can be that, due the lack of shift invariance of the DWT, differences in the temporal alignment of the patterns in a class introduce an additional source of intra-class variance into the coefficients. Another reason can be that, while the temporal information is discarded, more information about the class is conveyed by the subband energies because it takes into account all the coefficients.

Given that the frequency resolution in the WT is poor in the high-frequency region, we investigated the use of the wavelet packet transform (WPT) to determine if additional discriminatory information is obtained when the high-frequency portion of the signals is further decomposed. The filtering process used for extraction of the wavelet-based features is applied recursively only to the most recently lowpass filtered component of the signal. To allow for the decomposition of the low-frequency and the high-frequency components of the signal, it is necessary to use wavelet packets [96]. The iterative process of decomposing the low-frequency and the high-frequency components generates a binary wavelet packet tree structure, where the nodes of the tree represent subspaces with different frequency localization characteristics. In classification problems, the best wavelet packet basis is estimated from the wavelet packet tree using the local discriminant basis algorithm proposed by Saito [97]. The algorithm is a modified version of the best-basis selection algorithm used in signal compression [98]. It selects an orthonormal basis from the wavelet packet tree that maximizes a discriminant measure calculated on the time-frequency energy maps of the classes. Common choices for the discriminant measure are the symmetric relative entropy and the Euclidean distance. The classification performance for the WPT, using the symmetric relative entropy as the discriminant measure and applying the feature selection method on the coefficients corresponding to the best basis, is 80%. Similar results are obtained using the Euclidean

distance. This result shows that increasing the frequency resolution in the high-frequency portion of the signals does not improve discrimination.

Because of the differences between the feature representations (time-domain versus time-frequency domain), it is important to consider the performance when the representation are fused. We tested a representation that is composed by the time-domain features and the subband energy features to verify if the combination of these two representations provide complementary information for discrimination. The performance of the movement classification approach based on the fusion of the features is 85.6%. The differences in the performances based on each type of feature alone and the combined set are not statistically significant. The result shows that, despite the differences between the feature representations, both types of features do not provide complementary information when they are fused.

One common approach for selecting features is to find a transformation that reduces the dimensionality of the original feature vector. We investigated the use of PCA and LDA as alternative methods for feature selection. PCA seeks directions in feature space that best represent the data in a sum-squared error sense [74]. It produces an orthogonal coordinate system in which the axes are ordered in terms of the amount of variance in the data for which the corresponding principal components account. If the few principal components account for most of the variation, then these may be used to describe the data, thus leading to a reduced-dimension representation [83]. Since PCA does not take into account the class labels of the feature vectors, there is no guarantee that the directions of maximum variance are good for discrimination. Using the first 2 and 3 principal components, the classification performance with PCA is 76.2% and 79.5%, respectively. LDA seeks a transformation that maximizes the between-class separability and minimizes the within-class variability through the Fisher's criterion [83]. That is, it is aimed at finding a projection where samples from the same class are projected very close to each other and, at the same time, the projected class means are as far apart as possible. For a classification problem with N classes, a transformation is produced to a space of dimension at most $N-1$. The classification performance with LDA is 68.2% for a two-dimensional feature vector (in this case, $N = 3$). The performance of the wrapper method for feature selection (82% for a two-dimensional feature vector) is superior to the

performances of the PCA and LDA methods for this classification problem. However, the relative improvement of the wrapper method (82.2%) over PCA (79.5%) for a three-dimensional feature vector is only 3.4%, and it comes at a higher computational cost, and it is dependent on the classifier. Also, PCA outperforms LDA. Generally speaking, LDA is optimal when the distributions of the features for each class are unimodal and separated by the scatter of means [80]. Our classification results thus far suggest that the distributions are multimodal and there is overlap between classes 2 and 3, what could explain why LDA is not effective.

7.6 Performance of the Integrated Detection and Classification System

In this section, the performance of the system is presented by integrating the results of the detection and classification components described in Sections 6.2 and 7.3, respectively. We show the performance of the system in detecting and correctly classifying movements, as well as correctly detecting periods with no movement.

The system is evaluated in the following manner. For each subject, movement data from a fixed and pre-defined number of fixed-protocol trials randomly selected are used for training, and the remaining fixed-movement trials and the free-protocol trial are used for testing. The performance of the system is evaluated on the test set, based on the correctly detected and classified movement time periods, and based on the correctly detected non-movement time periods. Therefore, errors due to 1) falsely detected and missed movements, 2) false alarms and miss detections at the onsets and offsets, and 3) misclassifications are taken into account as the system output is continuously compared with the ground truth along the total time period corresponding to the testing time. Table 7.4 shows the confusion matrix for the continuous evaluation of the system for all subjects that is adapted from Ward [99]. The second column in Table 7.4 shows the total time, in seconds, corresponding to movement periods (classified according to the 3 classes suggested in this work) and non-movement periods in the ground truth. The shaded values show the time periods correctly recognized by the system as being a movement period corresponding to a certain class of movement or non-movement period. The remaining values in the third to sixth columns represent the time periods wrongly recognized according to the type of errors listed above. Errors due to false alarms in the

movement boundaries and falsely detected movements, as well as errors due to misses in the movement boundaries and missed movements are not specified by class, and only the total time periods corresponding to these errors are reported (1336.9 and 692.8 seconds, respectively). The last column shows the percentage of time correctly recognized for each type of action.

Table 7.4: Confusion matrix of the integrated movement detection and classification system.

		Total Time (secs)	Estimated (secs)				% Time Correct
			Class 1	Class 2	Class 3	Not Moving	
True	**Class 1**	2923.7	2659	40.5	27	692.8	90.9
	Class 2	2314.9	155.2	1296.8	500.3		56.0
	Class 3	1298.3	5.7	145.8	1013.8		78.1
	Not Moving	29098.1	1336.9			27761.2	95.4

Based on the percentage of time that each type of action was correctly recognized by the system, the system correctly identified class-2 movements only 56% of the time, and showed percentages close to 80% or higher for all the other types of action. Because of the errors in the classification for class-2 movements, and because all the missed movements correspond to class-2 movements, the overall performance (detection and classification) for this class deteriorates in comparison to the classification performance.

The accuracy of the system is defined as

$$accuracy = \frac{cTP + TN}{Total\ Time},$$

where *cTP* stands for correct true positive, and corresponds to the correctly detected and classified movement time periods. *TN* stands for true negative, and corresponds to the correctly detected non-movement time periods. Total time is the total testing time. Based on this definition, the accuracy of the system is 91.9%, for a total testing time of 35,635 seconds (approximately 9.9 hours of data). The accuracy of the system is significantly better than the performance on just the classification because the data contain long non-movement time periods that are correctly detected by the system.

7.7 Summary

In this chapter, we presented an approach for subject-dependent and subject-independent classification of movements in bed from load cell signals. Given a time interval where a movement has been detected, the goal is to classify the movement into one of the following classes: posture shifts, medium amplitude movements, and isolated leg movements. We presented two feature representations: a time-domain representation that is based on simple descriptors of the trajectory of the body center of mass, and a time-frequency domain representation of the trajectory signals using wavelet-based MRA. The approach was evaluated on data from 15 subjects.

The time-domain features used in this work are: the distance between initial and end points of the trajectory of the body center of mass during movement, the length of the trajectory, and the variance of the trajectory in the *y*-direction (side to side of bed). The distribution of the feature vectors for each class of movements is modeled by a Gaussian mixture density. We showed that increasing the number of mixture components beyond 3 increases the error, and models with as few as 2 components give reasonable performance. The overall classification rate on the test data for 2 mixtures is 84.6%. The most frequent classification errors are between classes 2 and 3. The classifier mistakenly classified movements consisting of leg movements and very small adjustments of head or torso (class 2) as leg movements (class 3). One reason for such mistakes can be that the small movements in the upper body do not substantially affect the overall trajectory of the center of mass.

We showed that this approach is applicable in real settings because it does not require a large amount of training data. Given the present dataset, a minimum of 10 samples per class is necessary to achieve comparable classification results. We also showed that the time-domain representation is relatively invariant across subjects. In particular, it is relatively invariant to their weight and height, because similar results were found when comparing subject-dependent models against a subject-independent model in a dataset where the subjects' weight varied by a factor of 3:1.

The fact that a movement can be decomposable into a series of movements of different intensity provided motivation for investigating a time-frequency representation using wavelet-based multiresolution analysis (MRA). In the feature extraction step, the

discrete wavelet transformation is applied only over the interval containing most of the intensity of the movement. Following feature extraction, feature selection is performed using the sequential forward selection method to reduce the relatively high dimensionality of the original feature set.

The classification performance of the wavelet-based features is 82.2%, and the time-domain features perform as well as the wavelet-based features on the classification task. Also, the classification performance of the subband energy features is better than the performance of the proposed wavelet-based features, with a performance of 85.4%. One reason for such difference in performance can be that differences in the temporal alignment of the patterns in a class introduce an additional source of intra-class variance into the coefficients. This is due to the lack of shift invariance of the DWT. Another reason can be that, while the temporal information is discarded, more information about the class is conveyed by the subband energies because it takes into account all the coefficients.

We also showed that there is no benefit in combining these feature representations. Despite the differences between them (time-domain versus time-frequency domain), both types of features do not provide complementary information when they are fused.

The performance of the integrated detection and classification system was presented. The performance of the system is evaluated on the test set, based on the correctly detected and classified movement time periods, and based on the correctly detected non-movement time periods. The system correctly identified class-2 movements only 56% of the time, and showed percentages close to 80% or higher for all the other classes and for the non-movement periods. The overall accuracy of the system is 91.9%, for a total testing time of approximately 9.9 hours of data. Because the data contain long non-movement time periods that are correctly detected by the system, the accuracy of the system is significantly better than the performance on just the classification.

Chapter 8

Conclusions

This thesis presented a system for assessment and classification of movement in bed with load cells. The main goal was to examine the extent to which the load cell signals can be used to infer clinically meaningful aspects of movement and sleep behaviors. Mobility in bed was characterized by the periods of postural immobility, movement times, and the types of movement performed. We focused on the assessment of major posture shifts and smaller movements. This thesis also presented a methodology for determining typical rest-activity patterns from sleep-related parameters estimated from the load cells.

In Chapter 3, we discussed the issues involved in the inference of movement information from load cell signals by inverting the bed-human system response. The inadequate signal dimensionality of the load cell signals and the nonlinearity of the bed-human system are the two major problems encountered in the inversion of the system response. Such problems motivated the use of pattern recognition as the approach for inference of movement information.

Chapter 4 detailed the protocol for load cell data collection. Data were collected when the subjects were awake using two different protocols, free movement and fixed movement (pre-determined set of movements), to allow both diversity and uniformity of movements. The protocol for data collection in a realistic setting was also described, which included the collection of load cell, actigraph, and light data, and the completion of sleep diaries.

The methodology for determining typical rest-activity patterns was described in Chapter 5. The algorithm for estimation of sleep-related parameters incorporated arbitrary rules, as well as information about a person's in-bed profile and about the use of the bedroom's light to estimate bedtimes and get up times. The methodology was developed using load cell and light data collected for a period of 2 weeks from 2 residents of an assisted-living facility (Elite Care study).

We compared the algorithm estimates of bedtime and get up time with the sleep diaries reports completed by caregivers. The discrepancies found between the objective measures from the sensors and the subjective measures from diaries are due to the fact that there is always some difficulty in recollecting exact bedtimes or the number and length of periods out of bed during the night. Thus, our method is particularly valuable for providing information that is not available all the times (it complements the work of the caregivers) and that may be missing in the sleep diaries.

One limitation of our method is that it assumes that the bed is only used for sleeping. This is not true when a person is in bed reading or watching television. Thus, the total time in bed is overestimated when such situations occur. In the future, this limitation may be mitigated using additional information, e.g., status of various devices, in conjunction with Bayesian estimation techniques.

Also, another limitation of the algorithm is the use of thresholds to determine bedtimes and get up times. The threshold values were based on the data from two people only, and based on a rule used in actigraphy for determining sleep onsets. There may be errors in the algorithm estimates because of these imposed thresholds. However, the use of different threshold values would still result in differences between the algorithm estimates and the sleep diary estimates.

The notion of an in-bed profile for characterization of a person's typical patterns of sleep behavior can be extended to other applications. It can be used in other applications such as localization of a person in a house, where the knowledge about a person's habits throughout the day is important. For example, by placing motion sensors in each room of a house, a profile of the occupancy of each room can be estimated.

In Chapter 6, we presented an approach for detection of movement in bed. The general idea of the approach is to assess the energy in each load cell signal in short

segments, and then form a weighted combination with scaling coefficients that are inversely proportional to the distance of each load cell to the center of mass of the body. The approach reliably detects movements: the average equal error rate (EER) across 15 subjects was 3.22% (± 0.54). Most of the errors were due to misses and false alarms at the onsets and offsets of the movements, and not for falsely detecting or missing a complete movement. We showed some evidence of the generalizability of the proposed detector for different bed sizes. We compared the performance results on subjects tested in two different beds, and showed that the bed type has no effect on the performance of the detector.

The person's weight has no effect on the performance of the proposed detector. In particular, the data suggested that there is no correlation between the subject's weight and the EER. A linear relationship between the EERs and the subjects weight is weak ($r = 0.047$). This shows that the approach does not impose any constraints on its use with very light people.

We also showed that, in the absence of training data to estimate the likelihood functions in the likelihood ratio test, least square regression can be used for approximating the decision threshold. When a new subject is presented to the system and training data are not available, it is possible to find an approximation for the threshold for the new subject by linear regression of the thresholds of a known set of subjects. Although this method only provides an approximation for the threshold, we have not found a statistically significant difference in the performance of the detector between the EER thresholds (thresholds derived from the EER measure) and the approximated thresholds.

We demonstrated how the detection approach derived in a laboratory setting can be applied in a real setting to estimate clinically relevant measures. The load cell setup was installed in the beds of 2 residents of an assisted-living facility, and data were collected for a period of 2 weeks. The clinical measures estimated from the detected movements are: the number of postural immobility periods per night, the number of movements per night, and the frequency of movements for each third of the night.

The study at the assisted-living facility helped us to better understand some of the issues involved in the application of the integrated (detection and classification) system in

a home. The issues are: installation, calibration, and collection of training data. Regarding installation of the load cells under the bed, the load cells were easily installed in a queen and a king beds. The subjects did not complaint about the sensors during the study. However, during recruitment, one of the prospective participants declined to participate in the study because the load cells add approximately 1.5" to the bed height. This modification can cause a problem for elderly people with difficulties in getting in and out of bed, depending on the height of the bed. The load cell sensor model used in the study is large, and a smaller model can be used to solve this issue. The calibration of the load cells can be done in the same way as described in Appendix A, by applying a range of known test weights.

Although no labeled data were collected at the assisted-living facility, the protocol for laboratory data collection used in this thesis could be adapted to a home setting. The experimental protocol can be identical as the one used, except for the number of trials performed and the use of video as ground truth. As showed in Chapter 7, the system does not require extensive training. To achieve comparable results for the data examined in this work, at least 10 samples per class are necessary to train the model for each person. Thus, the number of trials could be reduced to a minimum of two. Recordings can be used to instruct the subject to perform pre-defined movements at a beep sound. Instead of video, an observer in the room can record movement times and also the times and types of extra movements performed (not included in the recordings). Approximate movement times estimated from the detection algorithm can be compared to the observer's annotations to identify misses, eliminate false alarms and assign labels to the data.

Given that actigraphs measure the amount of movement based on movement of the body location of placement of the sensor, we showed that the actigraph under-reports movements as compared to load cells. Therefore, for the detection of a wider range of movements in the bed rather than the movement of a specific limb, the load cells provide more accurate results than actigraphs.

In Chapter 7, we described a framework for classification of movement in bed. Movements in bed were classified according to the following classes: major posture shifts, medium amplitude movements, and leg movements. Two feature representations for movement classification were introduced: a time-domain representation of the

trajectory of the body center of mass, and a time-frequency domain representation of the trajectory based on wavelets. The time-domain features explored spatial and (indirectly) temporal aspects of the movements. The distribution of the feature vectors extracted from a person's movement data was modeled by Gaussian Mixture Models.

Despite their simplicity, the time-domain features contain enough information to perform well on the classification task: the overall classification rate on the test data was 84.6%. We showed that this feature representation is relatively invariant across subjects. In particular, this representation is relatively invariant to their weight and height because the improvement of the subject-dependent models over the subject-independent model is relatively small.

The features extracted using wavelet-based multiresolution analysis explored the fact that a movement can be decomposable into a series of smaller movements with different intensity. Thus, through a zooming procedure across scales, the multiresolution analysis can provide a characterization of the transients that represent different types of movements.

Since the number of wavelet coefficients varies with the length of a signal and each detected movement signal has a specific duration, the discrete wavelet transformation is applied only over a fixed-length interval containing most of the intensity of the movement. Because the segment can be long enough that includes data from the previous or next movement, a symmetric boundary value replication is performed to correct the length of the segment. This becomes a problem when two or more movements happen very close to each other because there may be not enough samples to perform the boundary value replication.

Time-frequency representations have relatively high dimension, which required the use of a feature subset selection technique to reduce the dimensionality of the original feature vector. The technique used in this thesis was the sequential forward selection algorithm. The classification rate calculated on the training data was used as the selection criterion. The classification performance for the wavelet-based features was 82.2%. The relative success of the additive representation suggests that classification of movements of the smaller body parts may be related to small-signal linearity.

Although the feature subset selection method used to reduce the dimensionality of the wavelet-based original feature vector was superior to alternative methods such as PCA and LDA, one of the drawbacks is that it requires a classifier to select the features. That is, the features are only optimal with respect to the chosen classifier. The feature selection process has to be repeated every time the classifier changes.

The difference in performance between the 2 feature representations is not statistically significant. In addition, both representations are robust to the onset and offset errors in the detected boundaries, and to the type of bed.

The performance of the movement classification approach with a subband energy feature representation is better than the performance with the proposed wavelet-based features. The classification performance was 85.4% for the subband energy feature representation. The difference in performance between the subband energy and the proposed wavelet-based representations is statistically significant with a significance level of 0.01, and the relative improvement in performance was 3.89%. One reason for such difference in performance can be that, due the lack of shift invariance of the DWT, differences in the temporal alignment of the patterns in a class introduce an additional source of intra-class variance into the coefficients. Another reason can be that, while the temporal information is discarded, more information about the class is conveyed by the subband energies because it takes into account all the coefficients.

Despite the differences between the feature representations (time-domain versus time-frequency domain), they did not provide complementary information when they were fused. The performance of the movement classification approach based on the fusion of the time-domain features and the subband energy features was 85.6%. The differences in the performances based on each type of feature alone and the combined set are not statistically significant.

Finally, our conclusions are based on experimental results obtained using data from a small set of healthy subjects under controlled laboratory conditions. To examine if these results hold with data collected in realistic settings would require testing with data collected from a wider range of subjects in a sleep laboratory, for example, where video data are available for ground truth.

8.1 Future Work

This work lays out a number of directions in which the assessment of movement in bed can be pursued further.

A natural extension of this work should include the long-term monitoring of bedtime and getup times. Changes in sleep habits due to seasonal variations in the duration of daylight or due to changes in circadian patterns of rest and activity could be examined by extending the method to not only estimate but also monitor changes in bedtime and getup times.

Despite the simplicity and the good classification performance of the time-domain features, there is still difficulty in discriminating between class 2 and class-3 movements. New feature representations must be investigated to improve the classification performance.

Since the evaluation of the system was based on voluntary movements that were performed during wake periods, another extension of this work is to study how the system can be improved to detect abnormal movements during sleep periods. One of the aspects that could most benefit from further study is to differentiate restless legs syndrome (RLS) and periodic limb movements during sleep (PLMS) patients from controls. The system could be evaluated against bi-lateral anterior tibialis electromyography from overnight recordings of controls and patients. Since RLS patients constantly move their legs to relieve the tingling sensations caused by this disorder, we speculate that our system could be used as an aid for diagnosis or treatment of RLS because it can differentiate leg movements from other movements. Our system can be employed in such cases to monitor the frequency of leg movements. For PLMS, the criteria used in activity sensors to differentiate normal leg movements from pathological movements are: leg movements (single events) with duration varying from 0.5 to 5 seconds, with an inter-movement interval of 4 to 90 seconds, in a sequence of at least 4 consecutive movements. It will be interesting to investigate the use of such criteria in our approach to detect PLMS.

Since mobility in bed is impaired for Parkinson's disease patients, and turning in bed becomes difficult, the system could be used to assess slowness of motion and long periods of immobility to improve sleep quality.

Another important extension of this work would be to study the use of load cells under the bed to measure other physiological functions listed in the framework for assessment of sleep presented in Chapter 1, such as breathing. The regularity of breathing peaks at certain periods during the night. Breathing patterns show the highest level of regularity during NREM sleep stages III and IV [64]. One way to test the feasibility of the system to detect respiratory movements and monitor breathing patterns is by comparing the data from load cell and PSG (coupled with a nasal thermistor [5]) during NREM sleep stages III and IV.

Bibliography

[1] M. Irwin, J. Mcclintick, C. Costlow, M. Fortner, J. White, and J. C. Gillin, Partial Night Sleep Deprivation Reduces Natural Killer and Cellular Immune Responses in Humans, *FASEB Journal*, 10, (1996), 643-653.

[2] Talk about Sleep, Inc., *An Intro to Sleep: What is Sleep?*, http://www.talkaboutsleep.com/sleep-disorders/archives/intro.htm. Date viewed: 12/01/2004.

[3] M. H. Kryger, T. Roth, and W. C. Dement, *Principles and Practice of Sleep Medicine*, Fourth ed. Philadelphia, MA: Elsevier Saunders, 2005.

[4] A. Muzet, Dynamics of Body Movements in Normal Sleep, in *Proceedings of Eighth European Congress on Sleep Research '1986* (Szeged, Hungary, 1986), pp. 232-234.

[5] A. Culebras, *Clinical Handbook of Sleep Disorders*, First ed. Newton, MA: Butterworth-Heinemann, 1996.

[6] M. S. Aldrich and M. W. Naylor, Approach to the Patient with Disordered Sleep, in *Principles and Practice of Sleep Medicine*, *Impact, Presentation, and Diagnosis*, M. H. Kryger, T. Roth, and W. C. Dement, Eds., Third ed. Philadelphia, MA: W. B. Saunders Company, 2000, pp. 521-525.

[7] S. Chokroverty, W. A. Hening, and A. S. Walters, *Sleep and Movement Disorders*, First ed. Philadelphia, PA: Elsevier Science, 2003.

[8] M. A. Carskadon, W. C. Dement, M. M. Mitler, C. G. Guilleminault, V. P. Zarcone, and R. S. Spiegel, Self-reports versus Sleep Laboratory Findings in 122 Drug-free Subjects with Complaints of Chronic Insomnia, *American Journal of Psychiatry*, 133, (1976), 1382-1388.

[9] K. Tsuchiyama, H. Nagayama, K. Kudo, K. Kojima, and K. Yamada, Discrepancy between Subjective and Objective Sleep in Patients with Depression, *Psychiatry and Clinical Neurosciences*, 57, (2003), 259-264.

[10] A. E. Rogers, C. C. Caruso, and M. S. Aldrich, Reliability of Sleep Diaries for Assessment of Sleep/Wake Patterns, *Nursing Research*, 42, 6, (1993), 368-372.

[11] R. E.Tractenberg, C. M. Singer, and J. A. Kaye, Symptoms of Sleep Disturbance in Persons with Alzheimer's Disease and Normal Elderly, *Journal of Sleep Research*, 14, (2005), 177-185.

[12] S. Ancoli-Israel, R. Cole, C. Alessi, M. Chambers, W. Moorcroft, and C. P. Pollak, The Role of Actigraphy in the Study of Sleep and Circadian Rhythms, *Sleep*, 26, 3, (2003), 342-392.

[13] A. Sadeh and C. Acebo, The Role of Actigraphy in Sleep Medicine, *Sleep Medicine Reviews*, 6, 2, (2002), 113-124.

[14] I. Verbeek, E. C. Klip, and A. C. Declerck, The Use of Actigraphy Revised: The Value for Clinical Practice in Insomnia, *Perceptual and Motor Skills*, 92, (2001), 852-856.

[15] G. Jean-Louis, D. F. Kripke, W. J. Mason, J. A. Elliott, and S. D. Youngstedt, Sleep Estimation from Wrist Movement Quantified by Different Actigraphic Modalities, *Journal of Neuroscience Methods*, 105, (2001), 185-191.

[16] S. T. Aaronson, S. Rashed, M. P. Biber, and A. Hobson, Brain State and Body Position, *Archives of General Psychiatry*, 39, (1982), 330-335.

[17] J. J. van Hilten, H. A. M. Middelkoop, E. A. M. Braat, E. A. v. d. Velde, G. A. Kerkhof, G. J. Lighart, A. Wauquier, and H. A. C. Kamphuisen, Nocturnal Activity and Immobility across Aging (50-98 Years) in Healthy Persons, *Journal of the American Geriatrics Society*, 41, (1993), 837-841.

[18] M. T. Hyyppa and E. Kronholm, Sleep Movements and Poor Sleep in Patients with Non-Specific Somatic Complaints - II. Affective Disorders and Sleep Quality, *Journal of Psychosomatic Research*, 31, 5, (1987), 631-637.

[19] M. R. Lemke, P. Puhl, and A. Broderick, Motor Activity and Perception of Sleep in Depressed Patients, *Journal of Psychiatric Research*, 33, (1999), 215-224.

[20] G. Aubert-Tulkens, C. Culee, and K. H. Rijckevorsel, Ambulatory Evaluation of Sleep Disturbance and Therapeutic Effects in Sleep Apnea Syndrome by Wrist Activity Monitoring, *The American Review of Respiratory Disease*, 136, (1995), 851-856.

[21] L. S. Bennett, B. A. Langford, J. R. Stradling, and R. J. O. Davies, Sleep Fragmentation Indices as Predictors of Daytime Sleepiness and nCPAP Response in Obstructive Sleep Apnea, *American Journal of Respiratory and Critical Care Medicine*, 158, (1998), 778-786.

[22] B. Phillips, Movement Disorders: A Sleep Specialist's Perspective, *Neurology*, 62, (2004), S9-S16.

[23] A. S. Walters, Toward a Better Definition of the Restless Legs Syndrome, *Movement Disorders*, 10, 5, (1995), 634-642.

[24] National Sleep Foundation, *2005 Sleep in America Poll*, http://www.sleepfoundation.org/hottopics/index.php?secid=16&id=245. Date viewed: 05/02/2006.

[25] J. Wilde-Frenz and H. Schulz, Rate and Distribution of Body Movements during Sleep in Humans, *Perceptual and Motor Skills*, 56, (1983), 275-283.

[26] S. Gori, G. Ficca, F. Giganti, I. D. Nasso, L. Murri, and P. Salzarulo, Body Movements during Night Sleep in Healthy Elderly Subjects and Their Relationships with Sleep Stages, *Brain Research Bulletin*, 63, (2004), 393-397.

[27] J. Alihanka and K. Vaahtoranta, A Static Charge Sensitive Bed - A New Method for Recording Body Movements during Sleep, *Electroencephalography and Clinical Neurophysiology*, 46, (1979), 731-734.

[28] J. Alihanka, Sleep Movements and Associated Autonomic Nervous Activities in Young Male Adults, *Acta Physiologica Scandinavica*, Supplementum 511, (1982), 1-85.

[29] A. Culebras, Who Should be Tested in the Sleep Laboratory?, *Reviews in Neurological Diseases*, 1, 3, (2004), 124-132.

[30] W. W. Tryon, Issues of Validity in Actigraphic Sleep Assessment, *Sleep*, 27, 1, (2004), 158-165.

[31] S. Gorny, R. P. Allen, D. T. Krausman, and C. J. Earley, Evaluation of the PAM-RL System for the Detection of Periodic Leg Movements during Sleep in the Lab and Home Environments, *Sleep*, Suppl 21, (1998), 183.

[32] K. Tuisku, M. M. Holi, K. Wahlbeck, A. J. Ahlgren, and H. Lauerma, Quantitative Rest Activity in Ambulatory Monitoring as a Physiological Marker of Restless Legs Syndrome: A Controlled Study, *Movement Disorders*, 18, 4, (2002), 442-448.

[33] M. A. King, M.-O. Jaffre, E. Morrish, J. M. Shneerson, and I. E. Smith, The Validation of a New Actigraphy System for the Measurement of Period Leg Movements in Sleep, *Sleep Medicine*, 6, 6, (2005), 507-513.

[34] J. Kazenwadel, T. Pollmacher, C. Trenkwalder, W. H. Oertel, R. Kohnen, M. Kunzel, and H. P. Kruger, New Actigraphic Assessment Method for Periodic Leg Movements (PLM), *Sleep*, 18, 8, (1995), 689-697.

[35] W. Poewe and B. Hogl, Akathisia, Restless Legs and Periodic Limb Movements in Sleep in Parkinson's Disease, *Neurology*, 63, Suppl 3, (2004), S12-S16.

[36] P. Gehrman, C. Stepnowsky, M. Cohen-Zion, M. Marler, D. F. Kripke, and S. Ancoli-Israel, Long-Term Follow-Up of Periodic Limb Movements in Sleep in Older Adults, *Sleep*, 25, 3, (2002), 340-343.

[37] H. R. Colten and B. M. Altevogt, "Sleep Disorders and Sleep Deprivation - An Unmet Public Health Problem," National Academies Press, Washington, DC 2006.

[38] M. Chan, E. Campo, E. Laval, and D. Estève, Validation of a Remote Monitoring System for the Elderly: Application to Mobility Measurements, *Technology and Health Care*, 10, (2002), 391-399.

[39] M. Chan, E. Campo, and D. Estève, Assessment of Activity of Elderly People Using a Home Monitoring System, *International Journal of Rehabilitation Research*, 28, 1, (2005), 69-76.

[40] T. Rachwalski, S. Irvine, J. W. T. Steeper, D. R. Inkster, and C. Wells, Poor Quality of Sleep is a Precursor of Mobility-Related Adverse Events, in *Proceedings of Evidence-based strategies for Patient Falls and Wandering '2005* (Clearwater, FL, 2005), pp. NA.

[41] Y.-L. Huang, R.-Y. Liu, Q.-S. Wang, E. J. W. Someren, H. Xu, and J.-N. Zhou, Age-Associated Difference in Circadian Sleep-Wake and Rest-Activity Rhythms, *Physiology & Behavior*, 76, (2002), 597-603.

[42] S. W. Lockley, D. J. Skene, and J. Arendt, Comparison between Subjective and Actigraphic Measurement of Sleep and Sleep Rhythms, *Journal of Sleep Research*, 8, (1999), 175-183.

[43] K. Kayed, P. E. Hesla, and O. Rosjo, The Actioculographic Monitor of Sleep, *Sleep*, 2, 2, (1979), 253-260.

[44] S. Feldman and J. Abernathy, Management of Sleep Disorders in the Elderly, *Journal of the American Society of Consultant Pharmacists*, 15, Supplement 2, (2000), 1-14.

[45] N. Kleitman, *Sleep and Wakefulness*, Third ed. Chicago, 1967.

[46] E. J. W. Someren, Actigraphic Monitoring of Movement and Rest-Activity Rhythms in Aging, Alzheimer's Disease, and Parkinson's Disease, *IEEE Transactions on Rehabilitation Engineering*, 5, 4, (1997), 394-398.

[47] R. J. W. Dunnewold, J. I. Hoff, H. C. J. v. Pelt, P. Q. Fredrikze, E. A. H. Wagemans, and B. J. J. v. Hilten, Ambulatory Quantitative Assessment of Body Position , Bradykinesia, and Hypokinesia in Parkinson's Disease, *Journal of Clinical Neurophysiology*, 15, 3, (1998), 235-242.

[48] K. R. Chaudhuri and P. Martinez-Martin, Clinical Assessment of Nocturnal Disability in Parkinson's Disease: The Parkinson's Disease Sleep Scale, *Neurology*, 63, Suppl 3, (2004), S17-S20.

[49] National Sleep Foundation, *Restless Leg Syndrome (RLS)*, http://www.sleepfoundation.org/sleeptionary/index.php?id=23. Date viewed: 05/02/2006.

[50] L. Lu, T. Tamura, and T. Togawa, Detection of Body Movements during Sleep by Monitoring of Bed Temperature, *Physiological Measurements*, 20, (1999), 137-148.

[51] T. Tamura, J. Zhou, H. Mizukami, and T. Togawa, A System for Monitoring Temperature Distribution in Bed and Its Application to the Assessment of Body Movement, *Physiological Measurements*, 14, (1993), 33-41.

[52] H. F. M. Van der Loos, N. Ullrich, and H. Kobayashi, Development of Sensate and Robotic Bed Technologies for Vital Signs Monitoring and Sleep Quality Improvement, *Autonomous Robots*, 15, 1, (2003), 67-79.

[53] E. Rauhala, M. Erkinjuntti, and O. Polo, Detection of Periodic Leg Movements with a Static-Charge-Sensitive Bed, *Journal of Sleep Research*, 5, (1996), 246-250.

[54] T. Harada, T. Mori, Y. Nishida, T. Yoshimi, and T. Sato, Body Parts Positions and Posture Estimation System Based on Pressure Distribution Image, in *Proceedings of 1999 IEEE International Conference on Robotics and Automation '1999* (Detroit, Michigan, 1999), pp. 968-975.

[55] Tactex Controls Inc., *What is Kinotex?*, http://www.tactex.com/kinotex.php. Date viewed: 06/23/06.

[56] J. Kaartinen, I. Kuhlman, and P. Peura, Long-term Monitoring of Movements in Bed and Their Relation to Subjective Sleep Quality, *Sleep and Hypnosis*, 5, 3, (2003), 145-153.

[57] J. Kaartinen, M. Erkinjutti, and E. Rauhala, Automatic SCBS Analysis of Motor and Automatic Nervous Functions Compared with Sleep Stages, *NeuroReport*, 7, (1996), 1102-1106.

[58] Canadian Institute for Health Information, "Home Care Roadmap Indicators Data Standard," Canadian Institute for Health Information, Ottawa, Ontario 04/2004 2004.

[59] J. A. Hobson, T. Spagna, and R. Malenka, Ethology of Sleep Studied with Time-Lapse Photography: Postural Immobility and Sleep-Cycle Phase in Humans, *Science*, 201, 4362, (1978), 1251-1253.

[60] Scaime, *Load Cells*, http://www.scaime.com/english/pages_e/loadcells.html. Date viewed: 12/06/2004.

[61] R. S. Figliola and D. E. Beasley, *Theory and Design for Mechanical Measurements*, First ed. New York, NY: John Wiley & Sons, Inc., 1991.

[62] The UK's National Measurement Laboratory, *Glossary of Terms Used in Force Metrology*, http://www.npl.co.uk/force/faqs/glossary.html. Date viewed: 02/21/05.

[63] C. P. Pollack and P. Stokes, Circadian Rest-Activity Rhythms in Demented and Nondemented Older Community Residents and Their Caregivers, *Journal of the American Geriatrics Society*, 45, (1997), 446-452.

[64] B. W. Carlson, V. J. Neelon, and H. Hsiao, Evaluation of a Non-Invasive Respiratory Monitoring System for Sleeping Subjects, *Physiological Measurements*, 20, (1999), 53-63.

[65] A. Sadeh, P. J. Hauri, D. F. Kripke, and P. Lavie, The Role of Actigraphy in the Evaluation of Sleep Disorders, *Sleep*, 18, 4, (1995), 288-302.

[66] Mini Mitter Co., Inc., *Actiwatch*, http://www.minimitter.com/Products/Actiwatch/index.html. Date viewed: 03/09/05.

[67] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Reading, Massachusetts: Addison-Wesley Publishing Company, Inc., 1992.

[68] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "Mini-Mental State": A Practical Method for Grading the Cognitive State of Patients for The Clinician, *Journal of Psychiatric Research*, 12, (1975), 189-198.

[69] G. S. Alexopoulos, R. C. Abrams, R. C. Young, and C. A. Shamoian, Cornell Scale for Depression in Dementia, *Biological Psychiatry*, 23, (1988), 271-284.

[70] R. E. Tractenberg, C. S. Singer, J. L. Cummings, and L. J. Thal, The Sleep Disorders Inventory (SDI): An Instrument for Studying Sleep Disturbance in Persons with Alzheimer's Disease, *Journal of Sleep Research*, 12, (2003), 331-337.

[71] C. Acebo, A. Sadeh, R. Seifer, O. Tzischinsky, A. R. Wolfson, A. Hafer, and M. A. Carskadon, Estimating Sleep Patterns with Activity Monitoring in Children and Adolescents: How Many Nights are Necessary for Reliable Measures?, *Sleep*, 22, 1, (1999), 95-103.

[72] R. J. Cole and D. F. Kripke, Automatic Sleep/Wake Identification from Wrist Activity, *Sleep*, 15, 5, (1992), 461-469.

[73] R. D. Hippenstiel, *Detection Theory - Applications and Digital Signal Processing*, First ed. Boca Raton, FL: CRC Press LLC, 2001.

[74] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Second ed. New York, NY: John Wiley & Sons, Inc., 2001.

[75] G. Strang, *Calculus*. Wellesley, MA: Wellesley-Cambridge Press, 1991.

[76] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, First ed. New York, NY: Chapman and Hall Ltd., 1986.

[77] R. A. Johnson, *Miller and Freund's Probability and Statistics for Engineers*, Sixth ed. Upper Saddle River, NJ: Prentice Hall, Inc., 2000.

[78] G. Staude, C. Flachenecker, M. Daumer, and W. Wolf, Onset Detection in Surface Electromyographic Signals: A Systematic Comparison of Methods, *EURASIP Journal on Applied Signal Processing*, 2, (2001), 67-81.

[79] G. K. Kanji, *100 Statistical Tests*, Second ed. Thousand Oaks, California: SAGE Publications Ltd., 1999.

[80] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Second ed: Academic Press, 1990.

[81] E. Kronholm, E. Alanen, and M. T. Hyyppa, Sleep Movements and Poor Sleep in Patients with Non-Specific Somatic Complaints - I. No First-Night Effect in Poor and Good Sleepers, *Journal of Psychosomatic Research*, 31, 5, (1987), 623-629.

[82] A. K. Jain, R. P. W. Duin, and J. Mao, Statistical Pattern Recognition: A Review, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, 1, (2000), 4-37.

[83] A. Webb, *Statistical Pattern Recognition*, First ed. New York, NY: Oxford University Press Inc., 1999.

[84] D. H. Kil and F. B. Shin, *Pattern Recognition and Prediction with Applications to Signal Characterization*, First ed. Woodbury, NY: American Institute of Physics, 1996.

[85] I. Kingma, H. M. Toussaint, D. A. C. M. Comissaris, M. J. M. Hoozemans, and M. J. Ober, Optimizing the Determination of the Body Center of Mass, *Journal of Biomechanics*, 28, 9, (1995), 1137-1142.

[86] L. Lee, *Gait Analysis for Classification*, (Ph.D. thesis, Massachusetts Institute of Technology, 2002), pp. 110.

[87] Y. Huang, K. B. Englehart, B. Hudgins, and A. Chan, A Gaussian Mixture Model Based Classification Scheme for Myoelectric Control of Powered Upper Limb Prostheses, *IEEE Transactions on Biomedical Engineering*, 52, 11, (2005), 1801-1811.

[88] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, First ed. Norwell, MA: Kluwer Academic Publishers, 1991.

[89] N. Stergiou, *Innovative Analyses of Human Movement*, First ed. Champaign, IL: Human Kinetics Publishers, 2004.

[90] X. Huang, A. Acero, and H.-W. Hon, *Spoken Language Processing*, First ed. Upper Saddle River, NJ: Prentice Hall PTR, 2001.

[91] S. Mallat, *A Wavelet Tour of Signal Processing*, Second ed. San Diego, CA: Academic Press, 1999.

[92] K. Englehart, *Signal Representation for Classification of the Transient Myoelectric Signal*, (Ph.D. thesis, University of New Brunswick, 1998), pp. 361.

[93] A. Kundu, G. C. Chen, and C. E. Persons, Transient Sonar Signal Classification Using Hidden Markov Models and Neural Nets, *IEEE Journal of Oceanic Engineering*, 19, 1, (1994), 87-99.

[94] G. G. Yen and K.-C. Lin, Wavelet Packet Feature Extraction for Vibration Monitoring, *IEEE Transactions on Industrial Electronics*, 47, 3, (2000), 650-667.

[95] R. Kohavi and G. H. John, Wrappers for Feature Subset Selection, *Artificial Intelligence Journal*, 97, 1, (1997), 273-324.

[96] D. B. Percival and A. T. Walden, *Wavelets Methods for Time Series Analysis*, First ed. Cambridge, UK: Cambridge University Press, 2000.

[97] N. Saito, *Local Feature Extraction and Its Applications Using a Library of Bases*, (Ph.D thesis, Yale University, 1994), pp. 227.

[98] K. Englehart, B. Hudgins, and P. A. Parker, A Wavelet-Based Continuous Classification Scheme for Multifunction Myoelectric Control, *IEEE Transactions on Biomedical Engineering*, 48, 3, (2001), 302-311.

[99] J. A. Ward, P. Lukowicz, and G. Troster, Gesture Spotting Using Wrist Worn Microphone and 3-Axis Accelerometer, in *Proceedings of Joint Smart Objects and Ambient Intelligence-EUSAI'2005 Conference '2005* (Grenoble, France, 2005), pp. 99-104.

[100] R. W. Hamming, *Digital Filters*, Third ed. Mineola, New York: Dover Publications, Inc., 1998.

[101] R. Agarwal and J. Gotman, Digital Tools in Polysomnography, *Journal of Clinical Neurophysiology*, 19, 2, (2002), 136-143.

[102] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, First ed. Wellesley, MA: Wellesley-Cambridge Press, 1996.

[103] Robi Polikar, *Multiresolution Analysis: The Discrete Wavelet Transform*, 01/12/2001, http://users.rowan.edu/~polikar/WAVELETS/WTpart4.html. Date viewed: 02/20/2006.

Appendix A

Load Cells Calibration

The relationship between the force applied on a load cell and the load cell output value is established during calibration. Although data can be reported using the observed digitized outputs, calibration is convenient because each load cell has a different gain, i.e., resistance values can differ between load cells slightly. The variation is due to the following [60]:

1. The difference between each load cell's temperature (metal flexes differently at different temperatures).

2. How straight the strain gauges were applied to the surface of the cell during manufacturing. While great care is given to making sure the gauges are applied absolutely straight, they are very small, and this work has to be done by hand.

Therefore, we can compare the output of different load cells by converting the raw individual outputs to weight values.

We perform calibration by applying a range of known test weights, and obtaining a calibration curve by observation of the outputs. The observed output for each load cell i, for $i = 1, 2, 3, 4$, is a digitized value o_i and, after calibration, we have an estimation of $o_i = f(w)$, where w is the weight of the known test weight applied on the load cell. Even though we are dealing with force, what is usually expressed in lbf (pound-force), we report the values in kg. This choice is solely based on the fact that people are more familiar with such unit, which facilitates comprehension.

For each tested weight *w* and for each load cell *i*, data is collected at 200 Hz, for ten seconds. Then, the mean value of the digitized output is calculated and used as the correspondent ordinate o_i of the calibration curve. We apply the sequence of test weights once with increasing load (upscale direction), and once with decreasing load (downscale direction). The test weights used are 25-lb disc plates, and we use 8 plates. An aluminum platform is placed on the load cell to provide a more s support when many test weights are used during calibration. This platform measures 10.2" x 10.2" and weights 1.4 kg (3 lbs). Figure A.1 shows the setup used for calibration, with the platform supporting two 25-lb plates. The weight of the platform is added to the known weight placed on the load cells. After verification of the weight of each plate used in the calibration, the input range extends from 0 to 88.3 kg (194.7 lbs). Figure A.2 shows the plots of the average digitized output versus the weight for load cells 1-4 for upscale and downscale runs. From the plots, we can see that there is no hysteresis in the system.

Figure A.1: Setup for calibration: a platform is used to provide a more s support for the test weights.

The form of the calibration curve $o_i = f(w)$ is determined by a curve fitting technique. The outputs o_i are assumed to be linearly proportional to the force, and the relationship is determined by the least squares method [77]. From $o_i = f(w)$, it is straightforward to find $w = f(o_i)$. Therefore, for each load cell *i*, we have

$$w_i = c_i o_i + d_i + \varepsilon_i$$

where o_i is the digitized value from load cell i, c_i is the slope of the least-square approximation for load cell *i*, d_i is the intercept, ε_i is the error due to electrical noise and vibration on the floor, and w_i is the weight, in kg, of the load sensed by load cell *i*. The values for c_i and d_i, for load cells i = 1, 2, 3, 4 are given in Table A.1.

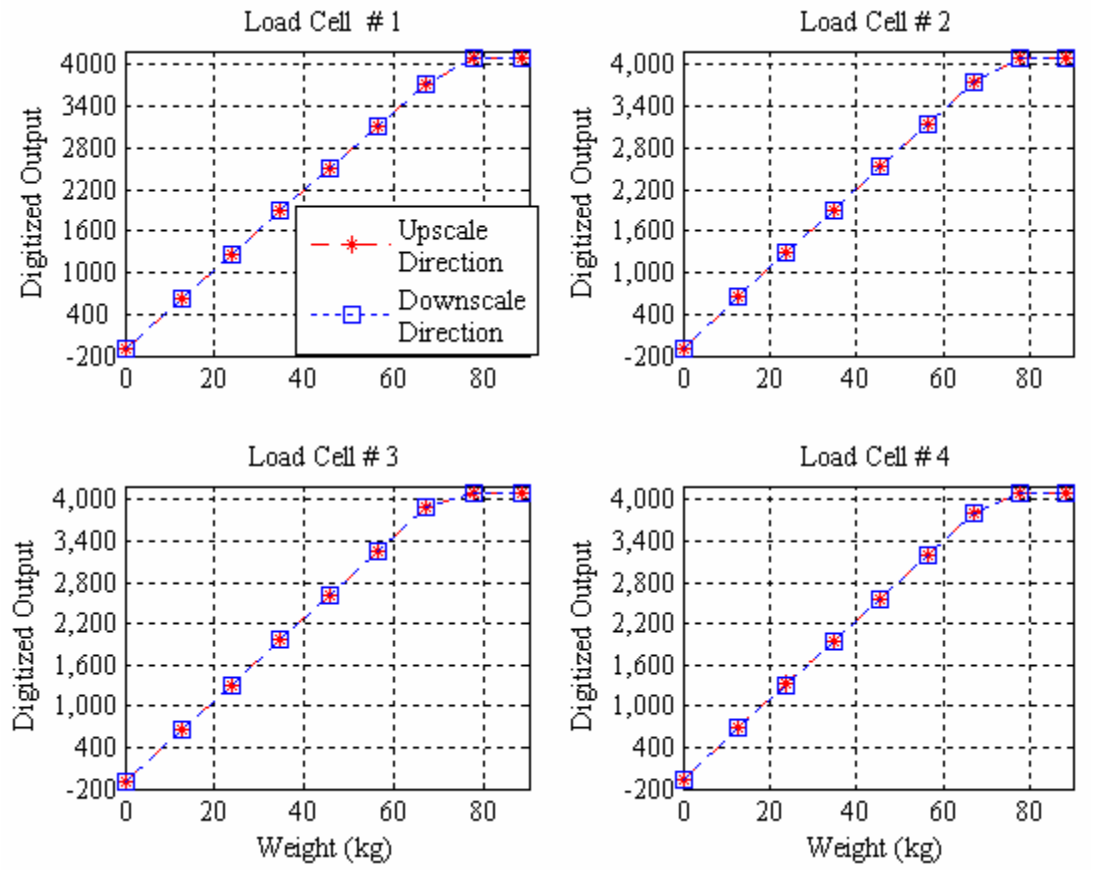


Figure A.2: Mean digitized output versus weight, in kilograms, for load cells 1-4 during upscale and downscale runs.

Table A.1: Calibration parameters (c_i and d_i) for each load cell.

Load Cell #	c_i	d_i
1	0.0197	-1.3770
2	0.0197	-1.5716
3	0.0195	-2.0294
4	0.0198	-2.4493

Appendix B

Analysis of Video Data

In this appendix, we describe in details the analysis of the video data for the purpose of obtaining the ground truth for the experiments described in Sections 4.1 and 4.2.

Cloth Bands Tracking

From the video, we want to determine the correct movement intervals, i.e., to determine where each movement starts and ends by tracking the trajectories of the cloth bands worn by the subjects.

The location of every cloth band, consequently the respective part of the body, is estimated using template matching [67]. Such technique uses the normalized cross-correlation to find the closest match between a given template and an image. The region of the image with highest correlation is defined as the closest match. First, templates from each cloth band were extracted from the first frames of the video, when the subject is lying on his/her back with straight legs and arms. Second, each frame is divided into regions of interest prior to template matching. For example, to search for a match for the cloth band located in the head, the algorithm just searches in a region of the image around the bed top. Third, the template matching is performed for each cloth band. The location of each cloth band is determined by the location in the image with highest cross-correlation according to the following steps:

- Compute the correlation coefficient [67] between a gray-level image *f(x,y)* of size *M* x *N* that represents the region of interest of each for each cloth band, and a template *m(x,y)* of size *J* x *K* that represents cloth band template at point (*r*,*s*) :

$$c(r,s)=\frac{\sum_x\sum_y [f(x,y)-\bar{f}(x,y)][m(x-r,y-s)-\overline{m}]}{\left\{\sum_x\sum_y [f(x,y)-\bar{f}(x,y)]^2 \sum_x\sum_y [m(x-r,y-s)-\overline{m}]^2\right\}^{1/2}},$$

where $r = 0, 1, 2,\ldots, M-1$, $s = 0, 1, 2,\ldots, N-1$, $\overline{m}$ is the average value of the pixels in $m(x, y)$, $\bar{f}(x,y)$ is the average value of $f(x, y)$ in the region coincident with the current location of the template *m*, and the summations are taken over the coordinates to both *f* and *m*.

- Select the location candidates from the first 10 highest correlation coefficients.
- Compute the Euclidean distance between all location candidates, and select only the locations of the candidates whose distance fall below a given threshold.
- The location of each cloth band is given by the median value of the selected candidates' locations.

Errors during tracking of the cloth bands occur, and they are mainly due to occlusion of the cloth bands and the fast movements that are aliased in the video acquisition. To reduce the occlusion errors, we used a method to automatically correct trajectories that were wrong based on distance constraints. If a cloth band is occluded, the closest match will probably be on the location of another cloth band of similar color, or some location in the image far from the last time the cloth band was correctly located. If the distance between the last detected location and the next is larger than a certain threshold, we assume the last seen location as the correct one.

We tested different cloth band color combinations to minimize the confusion of the color in the cloth bands on the limbs, because these are typically closer to each other. Also, a different color combination may need to be used if the person is wearing clothes with colors that match one of the cloth band colors.

Camera Calibration

Calibration of the camera is necessary to convert the (*x*, *y*) coordinates of the image converted into real world coordinates, i.e., coordinates of the scene. Under the planar scene assumption, camera calibration is accomplished by recognition of four landmarks located at known coordinates of the scene. The four pink circles on the bed posts, as seen in Figure 4.2, are used as the landmarks for camera calibration. The coordinates of these landmarks are used to generate an affine transformation between the 3-dimensional scene and the image of the bed. The general transformation takes the form

$$\begin{bmatrix} x_s \\ y_s \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_4 & \alpha_5 & \alpha_6 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix},$$

where (x_i, y_i) are the image coordinates and (x_s, y_s) are the correspondent coordinates in the scene (world coordinates), which are measured in centimeters. Therefore, the approximate location of each cloth band can be expressed in terms of the coordinates of the scene.

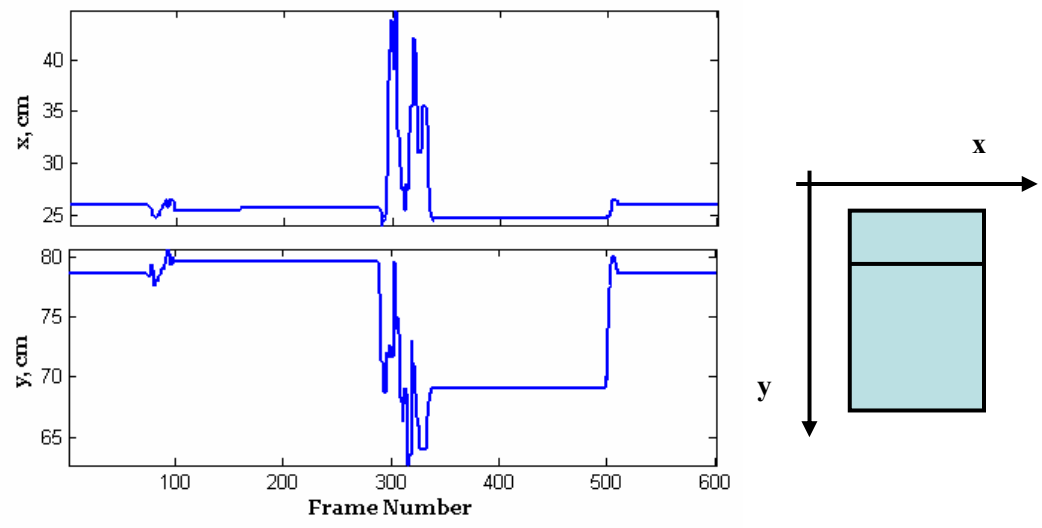
During calibration, we compute the values of α_i (i = 1, 2,…,6) by solving two systems of equations, one for α_1, α_2, α_3, and another for α_4, α_5, and α_6. Since certain movements include all 3 dimensions to be performed (e.g., moving an arm by rising it from left to right), the transformation between the 3-dimensional scene and the image of the bed in 2-dimensional results in some errors caused by the projection of the 3-dimensional space into 2-dimensional space. However, since that occurs only during certain movements, we believe that the approximation is good enough for this application.

Labeling the Data

The approximate location of each cloth band is estimated from the procedure described above. Since the goal is to determine the correct movement intervals, i.e., to determine where each movement starts and ends, we need to obtain this information from the cloth band trajectories. As we explained next, this is done in a semi-automatic way.

Let's assume that *x(t)* and *y(t)* are two functions of time representing the *x* and *y* coordinates of each cloth band trajectory along a certain trial. Figure B.1 (a) shows an example of the trajectory of the head during a sequence composed by an arm movement,

a posture shift, and a leg movement. Horizontal axis is the frame number, and the vertical axis is the value of *x* and *y* coordinates in centimeters. Figure B.1 (b) shows the reference directions chosen for *x* and *y*, so that for the head trajectory in this example, *x* and *y* do not change much during an arm movement (approximately 1 cm or 0.4"each), and y changes more (approximately 10 cm or 4") that *x* (approximately 1 cm or 0.4") during a leg movement.



(a) (b)

Figure B.1: (a) *x* and *y* coordinates, in centimeters, of the trajectory of the head during a sequence composed by an arm movement, a posture shift, and a leg movement. Time, in the horizontal axis, is represented by the frame number. (b) *x* varies along the bed width, and *y* varies along the bed length.

From the trajectories *x(t)* and *y(t)* of the six points (i.e., head, arms, legs and torso), we need to determine the movement intervals for each body part and, from those intervals, estimate the interval that corresponds to the entire movement. The algorithm for estimation of the intervals of each body part $[t_{begin}, t_{end}]_{x_i}$ and $[t_{begin}, t_{end}]_{y_i}$, for $i = 1, 2, \ldots, 6$, consists of the following steps:

1. Use time stamps from beep sounds to provide a coarse segmentation: during each trial, every subject hears a beep sound that indicates the time to start a movement. The algorithm uses the time stamps of these beeps to create a coarse segmentation of each movement. The segmentation is performed by selecting a window centered at the time of the beep, and analyzing the 5 seconds preceding it, and the 15 seconds following it. The choice for 5

seconds at the beginning is motivated by the fact that, in many occasions, subjects started moving before hearing the beep. We chose 15 seconds because, approximately every 15 seconds, a recording is played to instruct the subject to perform a movement.

2. Perform edge detection: the changes in the trajectories *x(t)* and *y(t)* can be seen as sharp transitions in the signal. The algorithm employs a filter that approximates the second derivative, using the estimate $u''(t) = u(t+1) - 2u(t) + u(t-1)$ [100], followed by hard thresholding to obtain a first approximation of the intervals of each body part $[t_{begin}, t_{end}]_{x_i}$ and $[t_{begin}, t_{end}]_{y_i}$.
3. Post-processing: after estimating the intervals of each body part $[t_{begin}, t_{end}]_{x_i}$ and $[t_{begin}, t_{end}]_{y_i}$, for i = 1, 2,…, 6, the interval for the movement is given by taking the longest interval in $[t_{begin}, t_{end}]_{x_i}$ and $[t_{begin}, t_{end}]_{y_i}$, for i = 1, 2,…, 6, i.e., the beginning of a movement B_{begin} and the end of a movement B_{end} are given by
$$B_{begin} = \min[t_{begin}, \forall x_i, \forall y_i]$$
$$B_{end} = \max[t_{end}, \forall x_i, \forall y_i]$$
.
4. Visual confirmation: it is rare that results of automatic methods like this one can be used without human validation [101]. Therefore, the final step requires a visual review of the automatically generated boundaries, during which the reviewer can change them. The most common errors are due to small movements at the very beginning (or end) of movements that are difficult to track by the proposed method, and sometimes can only be detected by visual inspection. Therefore, we visually inspect the video only at the frames closer to the values of B_{begin} and B_{end}. However, because of possible interscorer differences, different results are obtained by different scorers.

After validating the boundaries, it is also necessary to visually inspect the video to assign labels for each movement, according to the classes defined in Section 7.1. Even though we should know the labels beforehand (because the sequence of movements to be performed is pre-defined), there are cases when we must confirm them. For example, in

the case of isolated leg movements, we must verify if the subject has not moved any other part of the body such as the arms. We also examine the video and the results of the automatic detection of the cloth band movements to determine if a subject has performed a movement not included in the protocol during the periods when he/she should stay still. If a movement occurs during these intervals, it is also included and labeled accordingly to its type. Movements that do not belong to the classification system defined in Section 7.1, such as coughing and deep breathing are labeled as a separate class.

The capability of this video technique to track the location of the cloth bands could be further extended to be used as a method for assessment of movement. Movement could be assessed by designing a mapping from the cloth bands locations to the load cells outputs, and incorporating constraints imposed by the structure of the human body and the kinematics of the human body movements to simplify the problem. Since the problem of tracking human movement with video has been extensively addressed in the literature, we do not address such problem in this work.

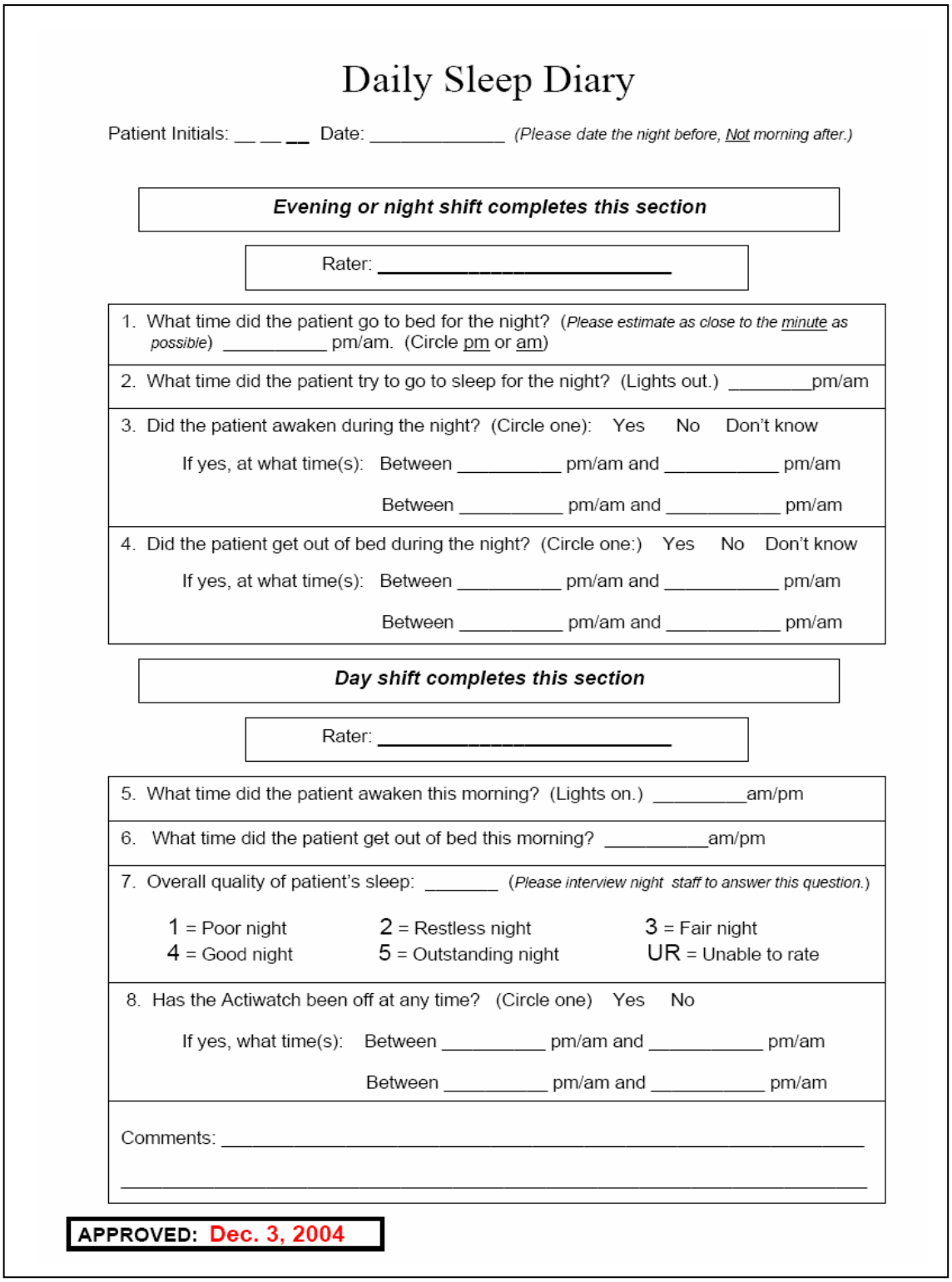
Appendix C

Sleep Diary

According to the study described in Section 4.3, caregivers were asked to complete a sleep diary for each subject, for a period of two weeks each. Figure C.1 shows the sleep diary used in the study, which was approved by the Review Board of the university. The daily sleep diary includes observations of bedtime, lights out, lights on and morning arising, in addition to comments regarding nighttime awakening and sleep quality. The caregivers received detailed instructions about the completion of the sleep diary from a nurse, prior to the beginning of the study. Question #7 includes a scale of sleep quality from 1 to 5 that refers to:

1. Poor night: resident was pacing all or most of night, and had less than 2 hours of sleep or rest.
2. Restless night: resident slept 2 to 4 hrs, and was getting up frequently.
3. Fair night: slightly restless; resident slept less than 6 hours, but did get some restful sleep.
4. Good night: resident slept 6 to 8 hours with few awakenings.
5. Outstanding night: resident slept more than 8 hrs with few or no awakenings.

UR: Unable to rate: resident spent night off campus; household crisis.



Daily Sleep Diary

Patient Initials: __ __ __ Date: _____ *(Please date the night before, Not morning after.)*

Evening or night shift completes this section

Rater: _____

1. What time did the patient go to bed for the night? (*Please estimate as close to the minute as possible*) _____ pm/am. (Circle pm or am)
2. What time did the patient try to go to sleep for the night? (Lights out.) _____pm/am
3. Did the patient awaken during the night? (Circle one): Yes No Don't know
   If yes, at what time(s): Between _____ pm/am and _____ pm/am
   Between _____ pm/am and _____ pm/am
4. Did the patient get out of bed during the night? (Circle one:) Yes No Don't know
   If yes, at what time(s): Between _____ pm/am and _____ pm/am
   Between _____ pm/am and _____ pm/am

Day shift completes this section

Rater: _____

5. What time did the patient awaken this morning? (Lights on.) _____am/pm
6. What time did the patient get out of bed this morning? _____am/pm
7. Overall quality of patient's sleep: _____ (*Please interview night staff to answer this question.*)

   1 = Poor night 2 = Restless night 3 = Fair night
   4 = Good night 5 = Outstanding night UR = Unable to rate

8. Has the Actiwatch been off at any time? (Circle one) Yes No
   If yes, what time(s): Between _____ pm/am and _____ pm/am
   Between _____ pm/am and _____ pm/am

Comments: _____

_____

APPROVED: Dec. 3, 2004

Figure C.1: Sleep diary for the study at Elite Care.

Appendix D

ROC Curves

In this appendix, we show the ROC (receiver operating characteristic) curves for the detection approach described in Section 6.2. The ROC curves indicate the relationship between hit rate and false alarm rate for each subject. The ROC curve is estimated by continuously changing the value of the decision threshold, i.e., a given threshold defines an operating point on the ROC curve. The EER is the point where the false alarms and miss detection rates are equal, and both errors have the same cost. The EER is reported for each tested value of the analysis window length L, and the circles show the locations of the EER.

The ROC curves for subjects 1 to 9, who participated in the LAB1_TWINSIZE experiment described in Section 4.1, are shown in Figure D.1. Figure D.2 shows the ROC curves for subjects 10 to 15, who participated in the LAB2_FULLSIZE experiment described in Section 4.2.

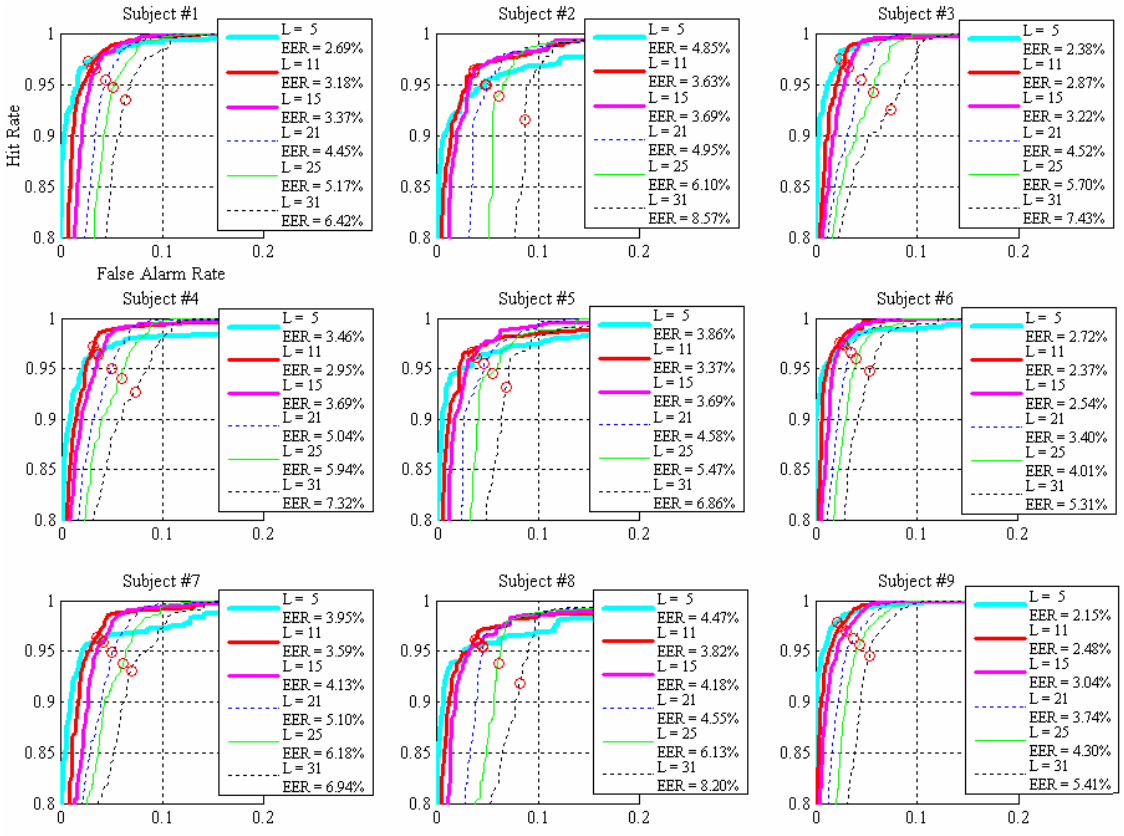


Figure D.1: ROC curves for subjects 1 to 9. The EER is reported for each tested value of *L*. The red circle corresponds to the EER point.

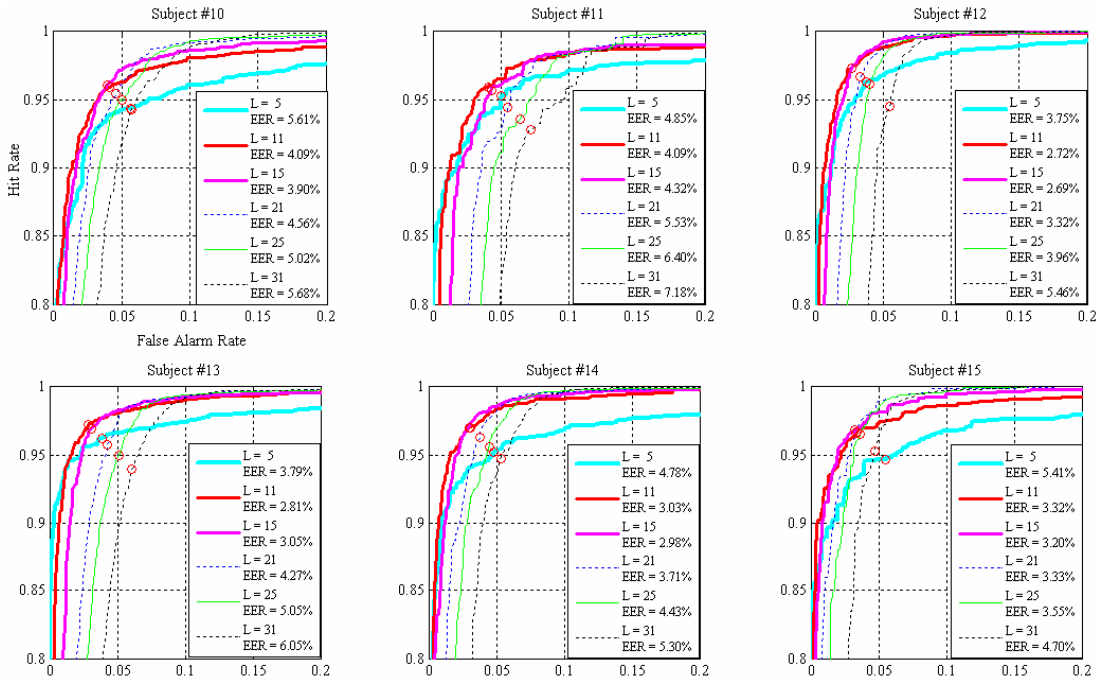


Figure D.2: ROC curves for subjects 10 to 15. The EER is reported for each tested value of *L*. The red circle corresponds to the EER point.

Appendix E

Gaussian Mixture Model Parameters Estimation

The algorithms used to initialize and estimate the parameters $\boldsymbol{\Theta} = \{\varpi_1, \ldots, \varpi_M, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_M, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_M\}$ of the Gaussian Mixture Models (GMMs) in Section 7.3.3, respectively the Linde, Buzo and Gray (LBG) and the expectation maximization (EM) [74] algorithms, are described in this appendix.

The Linde, Buzo and Gray Algorithm

The means of the Gaussians densities are initialized using centroids estimated by the Linde, Buzo and Gray (LBG) algorithm [88]. The centroids are obtained by successive splitting starting from a single global centroid. The algorithm can be summarized in the following steps:

1. Given a set of *N* training vectors, let the mean vector, which is denoted by z, be the centroid of the initial cluster generated by all training vectors.
2. Split the initial cluster into 2 new clusters: create 2 centroids $z + \epsilon$ and $z - \epsilon$, where ϵ is a fixed perturbation vector. For each training vector, compute its Euclidean distance from each centroid $z + \epsilon$ and $z - \epsilon$, and assign each vector to the closest cluster.
3. Repeat step 2 by splitting and clustering the current clusters until reaching the desired number of clusters, which in this case correspond to the number of mixture components *M*. Because the original LBG algorithm only allows the number of clusters to be a power of 2, we used a modified version to allow for any number of clusters.

The estimated centroids for each mixture component *i* are denoted by μ_i^{LBG}.

The Expectation Maximization Algorithm

The expectation maximization (EM) is an iterative algorithm for finding the maximum-likelihood parameter estimates for the case of incomplete data, where in the mixture of Gaussians the probability of assigning a sample x_n to the i^{th} Gaussian component is unknown [74]. It is assumed that the samples in $X = (x_1, x_2, \ldots, x_N)$ are independent and identically distributed (i.i.d.). In a generative model of the data in the training process, each sample $x_n \in X$ is generated by only one of the Gaussian components. The goal is to obtain the parameter values $\hat{\Theta}$ which maximize the likelihood of *X* given the data, as follows

$$\hat{\Theta} = \underset{\Theta}{argmax}\, p(\mathrm{X}|\Theta) = \underset{\Theta}{argmax} \prod_{n=1}^{N} p(\mathrm{x}_n|\Theta).$$

The maximum-likelihood estimate $\hat{\Theta}$ is the value of Θ that maximizes $p(X|\Theta)$.

To implement the EM algorithm it is necessary to have initial estimates for the mixing weights, means, and covariances of the Gaussians densities. In this work, the parameters of each mixture component *i* are initialized to

$$\begin{aligned} \varpi_i^{(0)} &= \frac{1}{M} \\ \mu_i^{(0)} &= \mu_i^{LBG} \\ \Sigma_i^{(0)} &= \Sigma_i^{LBG} \end{aligned},$$

where the initial estimate for the covariance matrix for each mixture component *i* is calculated based on the training vectors assigned to each cluster *i* by the LBG algorithm.

Given an initial estimate of the mixture density parameters, the EM algorithm iterates the following two steps:

E-step: estimate the probability that $\mathbf{x}_n$ is assigned to the i^{th} mixture component given the current parameter estimate $\boldsymbol{\Theta}^{(p)}$ and the data X using

$$p(\mathrm{x}_n, i|\Theta^{(p)}, \mathrm{X}) = \frac{\varpi_i^{(p)} p(\mathrm{x}_n|\mu_i^{(p)}, \Sigma_i^{(p)})}{\sum_{j=1}^{M} \varpi_j^{(p)} p(\mathrm{x}_n|\mu_j^{(p)}, \Sigma_j^{(p)})},$$

where $\boldsymbol{\Theta}^{(p)} = \left\{\boldsymbol{\varpi}^{(p)}, \boldsymbol{\mu}^{(p)}, \boldsymbol{\Sigma}^{(p)}\right\}$ represent the mixture density parameters after the p^{th} iteration.

M-step: re-estimate the mixing weights, means, and covariances of the Gaussians densities using the data set weighted by $p\left(\mathrm{x}_n, i \middle| \Theta^{(p)}, \mathrm{X}\right)$ to maximize the likelihood of the data, according to the following equations:

$$\varpi_i^{(p+1)} = \frac{1}{N}\sum_{n=1}^{N} p\left(\mathrm{x}_n, i \middle| \Theta^{(p)}, \mathrm{X}\right)$$

$$\mu_i^{(p+1)} = \frac{\sum_{n=1}^{N} p\left(\mathrm{x}_n, i \middle| \Theta^{(p)}, \mathrm{X}\right)\mathrm{x}_n}{\sum_{n=1}^{N} p\left(\mathrm{x}_n, i \middle| \Theta^{(p)}, \mathrm{X}\right)}$$

$$\Sigma_i^{(p+1)} = \frac{\sum_{n=1}^{N} p\left(\mathrm{x}_n, i \middle| \Theta^{(p)}, \mathrm{X}\right)\left(\mathrm{x}_n - \mu_i^{(p+1)}\right)\left(\mathrm{x}_n - \mu_i^{(p+1)}\right)^T}{\sum_{n=1}^{N} p\left(\mathrm{x}_n, i \middle| \Theta^{(p)}, \mathrm{X}\right)}$$

Appendix F

The Discrete Wavelet Transform

The wavelet theory involves representing general functions in terms of simpler fixed building blocks at different scales and positions [102]. The continuous wavelet transform (CWT) is defined as follows:

$$\Psi(a,\tau) = \int x(t)\psi_{a,\tau}(t)dt \qquad \text{(F.1)}$$

As seen in Equation F.1, the CWT decomposes a signal in the time domain into a two-dimensional function in the time-scale plane (a,τ). The wavelet coefficient $\Psi(a,\tau)$ measures the time-frequency content in a signal, indexed by the scale parameter a and the translation parameter τ that indicates the translation in time. The definition of the CWT shows that the wavelet analysis is a measure of similarity between the basis function $\psi_{a,\tau}(t)$ and the signal itself. Here the similarity is in the sense of similar frequency content. The wavelet basis function,$\psi_{a,\tau}(t)$, is also known as the mother wavelet.

The wavelet basis function is a family of short-duration high-frequency and long-duration low-frequency functions defined as

$$\psi_{a,\tau}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-\tau}{a}\right),\ a > 0,\ \tau \in \Re.$$

The term mother implies that the functions with different region of support that are used in the transformation process are derived from one main function. In other words, the

mother wavelet is a prototype for generating other window functions. As an example, the Haar wavelet is shown in Figure F.1 (a), and the Daubechies mother wavelet of order 2 ('db2') is shown in Figure F.1 (b).

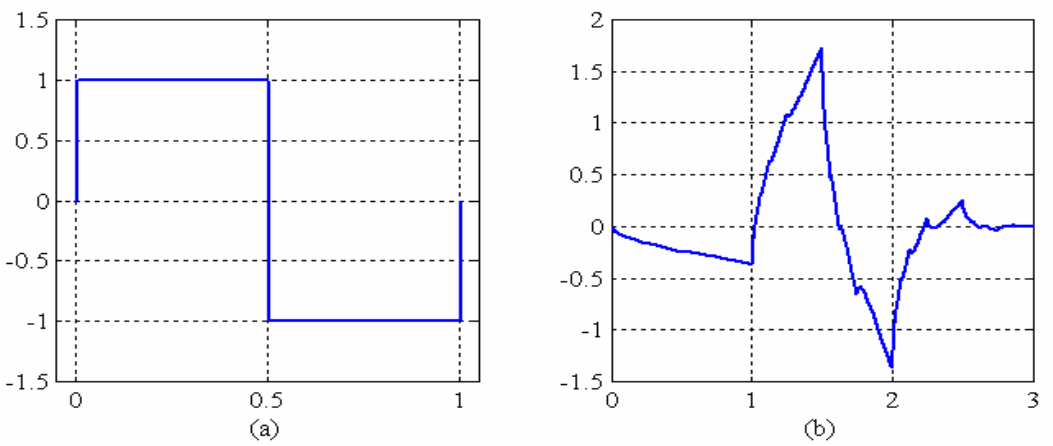


Figure F.1: (a) Haar wavelet. (b) Daubechies wavelet of order 2.

Depending on the parameter a, the wavelet function dilates or contracts in time causing the corresponding contraction or dilation in the frequency domain. When a is large ($a > 1$), the basis function becomes a stretched version of the mother wavelet ($a = 1$) and demonstrates a low-frequency characteristic. When a is small ($a < 1$), this basis function is a contracted version of the mother wavelet function and demonstrates a high-frequency characteristic [96]. Some of the characteristics of the mother wavelet are summarized in the following properties [96]:

- Support: wavelets can be divided based on their duration or support into infinite and finite support wavelets. In practice, finite support wavelets are more popular due to their use in multiresolution filter banks.
- Vanishing moments: the higher the number of vanishing moments of a wavelet function, the better it models the smooth part of a signal.
- Regularity: smooth basis functions are desired in applications where derivatives are involved. Smoothness also corresponds to better frequency localization of the filters.

The most commonly used wavelet mother functions that are subject to empirical evaluation in this work (as described in Section 7.5.3) are:

- **Daubechies**: compactly supported wavelets that are extremely asymmetric (introducing large phase distortion) and have the highest number of vanishing moments for a given support width. The Daubechies wavelet of order n (denoted by *dbn*) has a filter length of $2n$, a support width of $2n - 1$, and n vanishing moments. The Haar wavelet is the Daubechies wavelet of order one.
- **Symmlets**: compactly supported wavelets with minimal asymmetry and the highest number of vanishing moments for a given support width. The near symmetry introduces minimal phase distortion into the transform. A Symmlet of order n (denoted by *symn*) has a filter of length $2n$, a support width of $2n - 1$ and n vanishing moments.
- **Coiflets**: compactly supported wavelets designed to yield the highest number of vanishing moments for both mother wavelet and the scaling function for a given width. A Coiflet of order n (denoted by *coifn*) has a filter length of $6n$, a support width of $6n - 1$, and $2n$ vanishing moments for the mother wavelet and the scaling function.

The discrete wavelet transform (DWT) can be thought of as a judicious subsampling of the continuous wavelet transform in which we deal with just dyadic scales 2^j, $j = 1, 2, 3,\ldots$[96]. In the discrete case, filters of different cutoff frequencies are used to analyze the signal at different scales. The DWT employs two sets of functions, called scaling functions and wavelet functions, which are associated with low pass and highpass filters, respectively. As illustrated in Figure F.2, the original signal $x[n]$ is first passed through a highpass filter $g[n]$ and a lowpass filter $h[n]$. After the filtering, half of the samples can be eliminated according to the Nyquist's rule, since the signal now has a highest frequency of $\pi/2$ radians instead of π. The signal can therefore be downsampled by 2, and this constitutes one level of decomposition [103].

The multiresolution analysis (MRA) is then achieved by repeatedly decomposing the signal into approximation and added detail using a series of successive lowpass and highpass filters. This subband filtering process is applied recursively to the most recently lowpass filtered portion of the signal. Thus the coarse or approximation space of the DWT corresponds to a lowpass filtered version of the signal, while the detail space of the DWT corresponds to the highpass filtered version of the signal. The outputs of the

lowpass filters are referred as approximation coefficients, and the outputs of the highpass filters are referred as detail coefficients. As illustrated in Figure F.2, D_j denotes the vector of wavelet detail coefficients and A_j denotes the approximation coefficients of a signal at a given level j. Each vector D_j contains $2^J/2^j$ coefficients. At each level, the filtering and subsampling result in half the number of samples (and hence half the time resolution) and half the frequency band (and hence double the frequency resolution). Figure F.2 demonstrates the wavelet decomposition procedure for a partial DWT of level $J_0 = 3$. If a signal has 2^J samples, the maximum decomposition level is J. If we stop after $J_0 < J$ repetitions, we obtain a level J_0 partial DWT of $x[n]$. Partial DWTs are commonly used in practice when a scale beyond J_0 is no longer of interest [96].

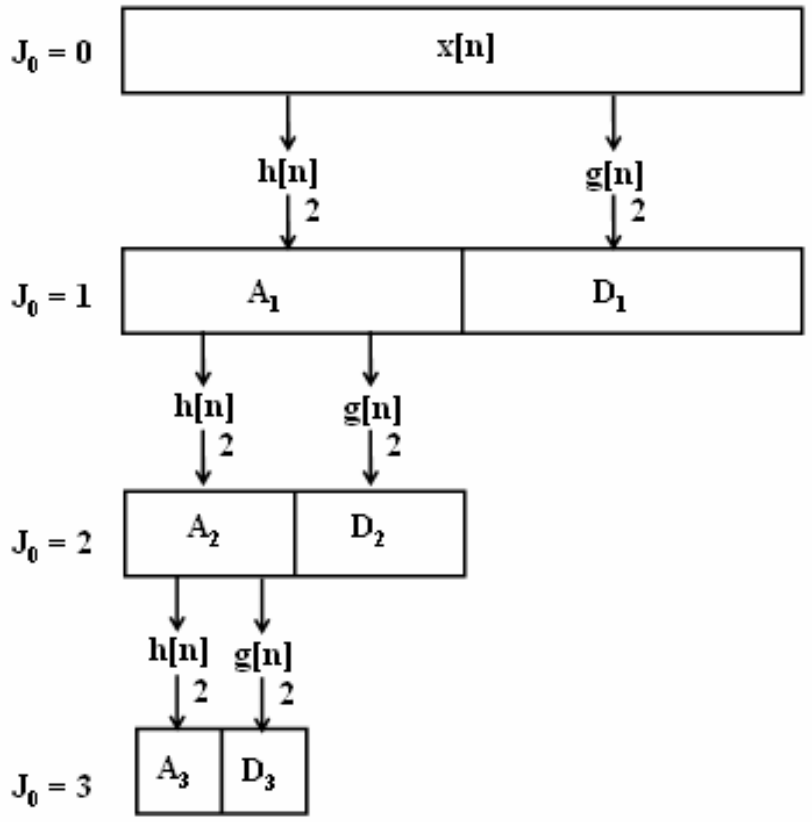


Figure F.2: Flow diagram illustrating the wavelet decomposition procedure for a partial DWT of level $J_0 = 3$.