Physical Activity Recognition from Acceleration Data under Semi-Naturalistic Conditions

by

Ling Bao

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of Master of Engineering in Electrical Engineering and Computer Science at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2003

© Ling Bao, MMIII. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.

Department of Electrical Engineering and Computer Science

August 21, 2003

Certified by...

Stephen S. Intille

Research Scientist, Changing Places Consortium

Thesis Supervisor

Accepted by ....

Arthur C. Smith

Chairman, Department Committee on Graduate Students
Physical Activity Recognition from Acceleration Data under Semi-Naturalistic Conditions

by

Ling Bao

Submitted to the Department of Electrical Engineering and Computer Science on August 21, 2003, in partial fulfillment of the requirements for the degree of Master of Engineering in Electrical Engineering and Computer Science

Abstract

Achieving context-aware computer systems requires that computers can automatically recognize what people are doing. In this work, algorithms are developed and evaluated to detect physical activities from data acquired using five small accelerometers worn simultaneously on different parts of the body. Acceleration data was collected from twenty subjects in both laboratory and semi-naturalistic environments. For semi-naturalistic data, subjects were asked to perform a sequence of everyday tasks outside of the laboratory. Mean, energy, frequency-domain entropy, and correlation of acceleration data was calculated over 6.71 s sliding windows. Decision table, nearest neighbor, decision tree, and Naive Bayesian classifiers were tested on these features. Classification results using individual training and leave-one-subject-out validation were compared. Leave-one-subject-out validation with decision tree classifiers showed the best performance recognizing everyday activities such as walking, watching TV, and vacuuming with an overall accuracy rate of 89%. The classifier captures conjunctions in acceleration feature values that effectively discriminate activities. This is the first work to investigate performance of recognition algorithms on 20 activities using semi-naturalistic data. These results demonstrate that some activities of daily living can be accurately detected from acceleration data without the need for individual training. However, preliminary findings also indicate that individual training can lead to higher recognition rates given sufficiently large training sets.

Thesis Supervisor: Stephen S. Intille
Title: Research Scientist, Changing Places Consortium
Acknowledgments

This thesis would not have been possible without my advisor, Stephen Intille, who kindly prodded me along throughout the entire project; my parents, who supported me during my arduous years at this Institute; my brother, who entertained me tirelessly, knowingly or otherwise; and my friends, who distracted me from working too hard. I am deeply grateful to all these people.

Thanks also to the National Science Foundation for funding this project (#0112900 ITR/PE: Using context recognition for preventive medicine in the home) and my research over the past year.
List of Figures

4-1  Top view of a hoarder board with accelerometer (a). Hoarders secured to a subject (b). ................................................................. 26

4-2  Acceleration from five biaxial accelerometers for walking, running, and tooth brushing. Signals from both axes of each biaxial accelerometer are shown. ................................................................. 28

4-3  A room with a television, vacuum, sofa, and reading materials. ........ 31

4-4  Five minutes of acceleration data annotated with subject self-report activity labels. Data within 10 s of self-report labels is discarded as indicated by masking. ................................................................. 34

4-5  Differences in feature values aide in discriminating different activities. 36

5-1  Aggregate confusion matrix for C4.5 classifier based on leave-one-subject-out validation for 20 subjects using laboratory and obstacle course data. 46

C-1  Confusion matrix for laboratory data results in [19]. Confusion matrix is not available for naturalistic results. ................................................................. 62

D-1  Obstacle course worksheet, page 1 of 3. ................................................................. 68

D-2  Obstacle course worksheet, page 2 of 3. ................................................................. 69

D-3  Obstacle course worksheet, page 3 of 3. ................................................................. 70

D-4  Activity sequence worksheet. ................................................................. 70

E-1  Visually aligned synchronization signals across five accelerometers. .... 72
| F-1 | Aggregate confusion matrix for C4.5 classifier based on leave hip accelerometer in validation for 20 subjects using laboratory and obstacle course data. | 74 |
| F-2 | Aggregate confusion matrix for C4.5 classifier based on leave wrist accelerometer in validation for 20 subjects using laboratory and obstacle course data. | 75 |
| F-3 | Aggregate confusion matrix for C4.5 classifier based on leave arm accelerometer in validation for 20 subjects using laboratory and obstacle course data. | 76 |
| F-4 | Aggregate confusion matrix for C4.5 classifier based on leave ankle accelerometer in validation for 20 subjects using laboratory and obstacle course data. | 77 |
| F-5 | Aggregate confusion matrix for C4.5 classifier based on leave thigh accelerometer in validation for 20 subjects using laboratory and obstacle course data. | 78 |
| F-6 | Aggregate confusion matrix for C4.5 classifier based on thigh and wrist accelerometer data for 20 subjects using laboratory and obstacle course data. | 79 |
| F-7 | Aggregate confusion matrix for C4.5 classifier based on hip and wrist accelerometer data for 20 subjects using laboratory and obstacle course data. | 80 |
### List of Tables

3.1 Summary of past work on activity recognition using acceleration. The “No. Subj.” column specifies the number of subjects who participated in each study, and the “Data Type” column specifies whether data was collected under laboratory (L) or naturalistic (N) settings. The “No. Sensors” column specifies the number of uniaxial accelerometers used per subject. ........................................... 22

5.1 Summary of classifier results (mean ± standard deviation) using individual training and leave-one-subject-out training. Classifiers were trained on laboratory data and tested on obstacle course data. ....... 41

5.2 Summary of preliminary classifier results (mean ± standard deviation) using individual training and leave-one-subject-out training where both training data sets are equivalent to five laboratory data sessions. .... 42

5.3 Aggregate recognition rates (%) for activities studied using leave-one-subject-out validation over 20 subjects. .......................... 44

5.4 Difference in recognition accuracy (mean ± standard deviation) due to leaving only one or two accelerometers in. Accuracy rates are aggregated for 20 subjects using leave-one-subject-out validation. ....... 45

C.1 Recognition ratios (%) for walking (Level), ascending stairs (Up), descending stairs (Down) in [30]. .......................... 60

C.2 Recognition ratios (%) for standing, walking (Level), ascending stairs (Up), descending stairs (Down) in [35]. .......................... 61

C.3 Confusion matrix for three martial arts moves recognized in [12]. .... 64
C.4 Recognition ratios (%) for walking (Level), ascending stairs (Up), descending stairs (Down) in [29].

C.5 Recognition rates (%) for activities studied in [52].

G.1 Summary of classifier results (mean ± standard deviation) using individual training and leave-one-subject-out training for 2G data. Classifiers were trained on laboratory data and tested on semi-naturalistic data.

G.2 Summary of classifier results (mean ± standard deviation) using individual training and leave-one-subject-out training. Classifiers were trained on laboratory data and tested on naturalistic data.
Chapter 1

Introduction

Context-aware computing can improve human-computer interactions by leveraging situational information about the user [1]. One facet of the user’s context is his physical activity. This work focuses on recognition of everyday activities using acceleration data that could be collected by a mobile computer.

Although existing works discuss physical activity recognition using acceleration [30, 12, 10, 33, 35, 52, 19, 5, 22, 51, 6, 32] or a fusion of acceleration and other data modalities [31, 7, 34], it is unclear how these systems will perform under real-world conditions.

Specifically, most of these works rely on data collected from subjects under artificially constrained laboratory settings to validate recognition results [30, 12, 35, 52, 5, 22, 6]. Some works also evaluate recognition performance on data collected in natural, out-of-lab settings but only use limited data sets collected from one individual [51, 13]. A number of works use naturalistic data but do not quantify recognition accuracy [31, 7, 33, 34, 32]. Lastly, research using naturalistic data collected from multiple subjects has focused on recognition of a limited subset of nine or fewer everyday activities consisting largely of ambulatory motions and basic postures such as sitting and standing [19, 10]. Consequently, it is uncertain how any of these systems will perform in recognizing a variety of everyday activities for a diverse sample population under real-world conditions.

In particular, the possibility that large-scale real-world implementation of physical
activity recognition systems may require training on individuals’ specific patterns must be explored. Training on individual patterns to maximize recognition accuracy is required for effective speech dictation recognition systems [46]. If the same individual training is required for accurate activity recognition, then current systems trained purely on data collected from supervised laboratory settings [30, 12, 35, 52, 5, 22, 6] are impractical for real-world application. Given such a scenario, researchers must explore activity recognition systems using unsupervised data collection and training techniques for individual customization.

Although individual training is needed in speech dictation recognition to customize conditional inter-word probabilities based on a user’s vocabulary and dictation style, accurate word recognition for automated telephone systems does not require individual training. These systems limit the context of the user’s speech and train on large example sets of specific words, simplifying word recognition. For instance, when using a movie theater’s automated telephone system, a user will utter a very limited subset of words such as “times,” “purchase tickets,” or the name of a movie that is playing. Context is exploited to resolve recognition ambiguities. Similarly, the need for individual training of physical activity recognition systems may vary depending upon the activities or “words” to be recognized, and the difficulty of recognizing activities may increase substantially in less constrained, non-laboratory settings.

This paper explores the need for individual training of activity recognition systems and examines the feasibility of an unsupervised approach to such training. The performance of activity recognition algorithms under conditions akin to those found in real-world settings is assessed. Activity recognition results are based on acceleration data collected from five biaxial accelerometers placed on 20 subjects under laboratory and semi-naturalistic conditions. These results show that an overall recognition accuracy of 84.26% can be achieved for 20 everyday activities using five accelerometers without training on individual data. Acceleration of the thigh is the most useful single sensor location for discriminating activities, and acceleration of the thigh and wrist can be used to recognize activities with 80.99% overall accuracy.
Chapter 2

Background

Technological obstacles to large-scale real-world activity recognition systems are diminishing as mobile computing power improves and becomes more prevalent. Mobile computing devices such as personal digital assistants (PDAs) are increasingly widespread with estimates of over 12 million handheld units shipped in 2002 [37]. Watches with GPS receivers are currently available, and researchers have prototyped watches that run Linux [41]. As these technologies continue to develop, it may be possible to create inconspicuous mobile activity recognition systems in the near future. Researchers have already prototyped wearable computer systems that use acceleration, audio, video, and other sensors to recognize user activity [40, 14, 36, 47, 17]. With further advances in miniaturization, it may be possible to develop a mobile activity recognition system in the convenient form factor of a watch. The system could include a built-in computer and radio transceiver. Accelerometers embedded within wrist bands [12], bracelets, adhesive patches, or belts could be placed on various parts of the user’s body, relaying acceleration data using integrated transceivers. The system’s onboard computer could collect acceleration data from its sensors, enabling it to recognize user activities and run real-time context-aware applications.

However, one major unresolved issue with mobile activity recognition systems is customization. High accuracy recognition of some activities may require an individual training phase as is the case with speech dictation recognition [46]. In dictation recognition, intra-word and inter-word level relations across individuals exhibit strong
similarities due to the words and rules shared through spoken language [46]. Nonetheless, accurate recognition relies upon training on individual speech to account for individual variation in word pronunciation and to reduce uncertainty in sentence-level recognition [46]. A comparable parallel may apply to activity recognition.

For instance, ambulation may exhibit strong commonalities in body acceleration between many individuals. However, differences in gait from person to person may still necessitate customization on an individual basis. These differences may be analogous to differences in word pronunciation between individuals in speech recognition. Extending the analogy, if activities such as standing, walking, running, jumping, and sitting are words, then sequences of these activities can be thought of as sentences. Playing basketball could be a sentence level activity involving sequences of running and jumping with certain transition probabilities between the word level activities. Similarly, a sequence of walking, standing, and sitting with certain transition probabilities could be a sentence level representation for commuting on a bus. Paragraph level activities such as working a job could be considered sequences of sentence level activities such as commuting on a bus, working at a desk, and sitting through a meeting with appropriate transition probabilities. Activity recognition systems could train for individual variation in sentence level activities by modelling the individual’s transition probabilities for the sequence of word level activities composing that sentence. The same could be done for the transition probabilities between sentence level activities in paragraph level activities.

To determine the feasibility of wide-scale mobile activity recognition systems, this work assesses the need for individual training of recognition algorithms. The goal is to determine whether individual training is required for accurate recognition of activities such as walking, climbing stairs, eating, and vacuuming. Should the results support individual training of activity algorithms, this work will explore unsupervised training techniques for recognition systems.

Note that recognition of sentence and paragraph level activities consisting of sequences of more basic activities such as those listed above is beyond the scope of this work. Although algorithms such as hidden Markov models may work well for recog-
nizing higher level activities [46], collecting sufficient data to train the Markov models is time consuming. Whereas sentences in speech take seconds to utter, sentence-level activities may take minutes or hours. Furthermore, understanding word-level activity recognition is a prerequisite to sentence- and paragraph-level activity recognition. Thus, this work must focuses on lower level activity recognition.

A data set collected from a diverse sample population is critical in assessing the need for individual training and evaluating the robustness of any proposed recognition algorithms. For example, a small sample set of tall male test subjects exhibit fairly identical walking gaits. Consequently, such a limited data set would not be helpful in assessing the need for individual training of recognition algorithms.

Moreover, if individual variations in activity are significant, an activity recognition algorithm must be tested on data from many subjects to ensure it performs well for a range of individual patterns.

Beyond the size of the data set, the quality of the data is also critical. Specifically, it is essential to verify activity recognition systems on data collected under naturalistic circumstances because laboratory environments may artificially constrict, simplify, or influence subject activity patterns. For instance, laboratory acceleration data of walking displays distinct phases of a consistent gait cycle which can aide recognition of pace and incline [6]. However, acceleration data from the same subject outside of the laboratory may display marked fluctuation in the relation of gait phases and total gait length. In a naturalistic environment, the subject is less aware of the fact that his walking is monitored and may assume a more relaxed, unconventional gait. Furthermore, traffic patterns and mood may frequently alter the subject’s walking pace and style. Consequently, a highly accurate activity recognition algorithm trained on laboratory data may rely too heavily on distinct phases and periodicity of the gait cycle for detecting walking. The accuracy of such a system may suffer when tested on naturalistic data, where there is greater variation in gait pattern.

Due to the above factors, this work compares recognition results for laboratory and semi-naturalistic data collected from twenty test subjects who were aged 17 to 48 and lived in the Boston-Cambridge area.
Chapter 3

Theory

Many past works have demonstrated 85% to 95% recognition rates for ambulation, posture, and other activities using acceleration data [30, 10, 33, 35, 19, 5, 22, 51, 11, 6, 32]. Activity recognition has been performed on acceleration data collected from the hip [30, 35, 51, 11, 6] and from multiple locations on the body [10, 33, 19, 5, 22]. Table 3 summarizes activity recognition results using acceleration from recent works. Additional details about these works are available in Appendix C.

Related work in the field of activity counts and computer vision also support the potential for activity recognition using acceleration. The energy of a subject’s acceleration can discriminate sedentary activities such as sitting or sleeping from moderate intensity activities such as walking or typing and vigorous activities such as running [44, 53, 21, 42, 38, 43]. Body trajectory, which can be derived from acceleration, is used in computer vision to recognize gestures and gait [9, 26, 25].

Although the literature supports the use of acceleration for physical activity recognition, little work has been done to validate the idea under real-world circumstances. Most works on activity recognition using acceleration rely on data collected in controlled laboratory settings [30, 12, 35, 52, 5, 22, 6]. Additionally, all of the literature focuses on recognizing a special subset of physical activities such as ambulation with the exception of [19] which examines nine everyday activities. Interestingly, [19] demonstrated 95.8% recognition rates for data collected in the laboratory as opposed to 66.7% recognition rates for data collected outside the laboratory in naturalistic set-
<table>
<thead>
<tr>
<th>Ref.</th>
<th>Recognition Accuracy</th>
<th>Activities Recognized</th>
<th>No. Subj.</th>
<th>Data Type</th>
<th>No. Sensors</th>
<th>Sensor Placement</th>
</tr>
</thead>
<tbody>
<tr>
<td>[30]</td>
<td>92.85% to 95.91%</td>
<td>ambulation</td>
<td>8</td>
<td>L</td>
<td>2</td>
<td>2 thigh</td>
</tr>
<tr>
<td>[35]</td>
<td>83% to 90%</td>
<td>ambulation, posture</td>
<td>6</td>
<td>L</td>
<td>6</td>
<td>3 left hip, 3 right hip</td>
</tr>
<tr>
<td>[19]</td>
<td>95.8%</td>
<td>ambulation, posture, typing, talking, bicycling</td>
<td>24</td>
<td>L</td>
<td>4</td>
<td>chest, thigh, wrist, forearm</td>
</tr>
<tr>
<td>[19]</td>
<td>66.7%</td>
<td>ambulation, posture, typing, talking, bicycling</td>
<td>24</td>
<td>N</td>
<td>4</td>
<td>chest, thigh, wrist, forearm</td>
</tr>
<tr>
<td>[5]</td>
<td>89.30%</td>
<td>ambulation, posture</td>
<td>5</td>
<td>L</td>
<td>2</td>
<td>chest, thigh</td>
</tr>
<tr>
<td>[22]</td>
<td>N/A</td>
<td>walking speed, incline</td>
<td>20</td>
<td>L</td>
<td>4</td>
<td>3 lower back, 1 ankle</td>
</tr>
<tr>
<td>[51]</td>
<td>86% to 93%</td>
<td>ambulation, posture, play</td>
<td>1</td>
<td>N</td>
<td>3</td>
<td>2 waist, 1 thigh</td>
</tr>
<tr>
<td>[12]</td>
<td>96.67%</td>
<td>Kung Fu arm movements</td>
<td>1</td>
<td>L</td>
<td>2</td>
<td>2 wrist</td>
</tr>
<tr>
<td>[52]</td>
<td>42% to 96%</td>
<td>ambulation, posture, bicycling</td>
<td>1</td>
<td>L</td>
<td>2</td>
<td>2 lower back</td>
</tr>
<tr>
<td>[47]</td>
<td>85% to 90%</td>
<td>ambulation, posture</td>
<td>10</td>
<td>L</td>
<td>2</td>
<td>2 knee</td>
</tr>
</tbody>
</table>

Table 3.1: Summary of past work on activity recognition using acceleration. The “No. Subj.” column specifies the number of subjects who participated in each study, and the “Data Type” column specifies whether data was collected under laboratory (L) or naturalistic (N) settings. The “No. Sensors” column specifies the number of uniaxial accelerometers used per subject.
tings. These results underline uncertainty in the performance of current algorithms for recognizing a broad array of everyday activities in naturalistic, uncontrolled environments.

To address these issues, this work will test activity recognition algorithms on acceleration data collected from 20 subjects performing 20 everyday activities under both laboratory and semi-naturalistic conditions.

Another concern with physical activity recognition is the need for individual training. Recognition accuracy rates of 80% to 95% can be achieved for postures and ambulatory activities using accelerometer data without individual training [30, 35, 19, 5, 51, 47]. However, this may not be the case for all everyday activities. Although comparisons of activity recognition performance with and without the use of individual training is lacking, recognition accuracy for certain activities such as ascending stairs has been shown to improve through use of individual training [29]. Additionally, for an activity such as tooth brushing, individuals may display significant variations in brushing vigor, duration, and posture. In such cases, individual training would improve discrimination of tooth brushing from other similar motions such as window scrubbing. Thus, this work will compare recognition accuracy for a variety of activities with and without the use of individualized training.

Prior literature demonstrates that forms of locomotion such as walking, running, and climbing stairs and postures such as sitting, standing, and lying down can be recognized at 83% to 95% accuracy rates using hip, thigh, and ankle acceleration [30, 35, 19, 5, 22, 51]. Acceleration data of the wrist and arm improve recognition rates of upper body activities [12, 19] such as typing and martial arts movements. All past works have used accelerometers with wires, which may restrict subject movement. Based on these results, this work will use data collected from five wire-free biaxial accelerometers placed on each subject's right hip, dominant wrist, non-dominant upper arm, dominant ankle, and non-dominant thigh to recognize ambulation, posture, and other everyday activities. Although each of the above five locations have been used for sensor placement in past work, no work addresses which of the accelerometer locations provide the best data for recognizing activities. Consequently, this work will
also determine which sensor locations result in the best recognition accuracy. Note that five biaxial accelerometers is equivalent to ten uniaxial accelerometers. This exceeds the maximum number of simultaneous sensors used in past work, which is six uniaxial accelerometers [35]. The increased number of simultaneous sensors should aide in determining which locations are most useful in activity recognition.
Chapter 4

Design

4.1 Overview

Subjects wore 5 accelerometers as they performed a variety of activities under two different protocols for semi-naturalistic and laboratory data collection. Subjects first participated in a semi-naturalistic data collection session and then participated in a laboratory data collection session. Details about the accelerometers, activities, and protocols are provided below.

4.2 Accelerometers

4.2.1 Hardware

Subject acceleration was collected using ADXL210E accelerometers from Analog Devices. These two-axis accelerometers are accurate to ±10 G with tolerances within 2%. Accelerometers were mounted to hoarder boards [20], which sampled at 76.25 Hz with minor variations based on onboard clock accuracy and stored acceleration data on compact flash memory. This sampling frequency is more than sufficient compared to the 20 Hz frequency required to assess daily physical activity [8]. The hoarder board time stamped one out of every 100 acceleration samples, or one every 1.31 seconds. The time stamp consisted of a year, month, date, hour, minute, and second...
field. Continuous running time of the hoarder board was limited by battery life. Four AAA batteries can power the hoarder board for roughly 24 hours. This is more than sufficient for the 90 minute data collection sessions used in this study. A hoarder board is shown in Figure 4-1a.

Previous work shows promising activity recognition results from ±2 G acceleration data [15] even though typical body acceleration amplitude can range up to 12 G [8]. Ideally, this study would only use ±12 G accelerometers to minimize signal saturation. However, due to limitations in availability of ±12 G accelerometers, ±10 G acceleration data was used. Moreover, although body limbs and extremities can exhibit a 12 G range in acceleration, points near the torso and hip experience a 6 G range in acceleration [8].

The hoarder boards were not electronically synchronized to each other and relied on independent quartz clocks to time stamp data. Electronic synchronization would have required wiring between the boards which would restrict subject movements, especially during whole body activities such as bicycling or running. Additionally, a
wire-free sensor system more closely models the long term vision of a mobile activity recognition system that collects data from wireless "bracelet" sensors that do not encumber the wearer.

To achieve synchronization without wires, hoarder board clocks were synchronized with subjects’ watch times at the beginning of each data collection session. Due to clock skew, hoarder clocks and the watch clock drifted between 1 and 3 seconds every 24 hours. To minimize the effects of clock skew, hoarders were shaken together in a fixed sinusoidal pattern in two axes of acceleration at the beginning and end of each data collection session. Watch times were manually recorded for the periods of shaking. The peaks of the distinct sinusoidal patterns at the beginning and end of each acceleration signal were visually aligned between the hoarder boards. Time stamps during the shaking period were also shifted to be consistent with the recorded watch times for shaking. Acceleration time stamps were linearly scaled between these manually aligned start and end points. Further details on the synchronization process is available in Appendix E.

To characterize the accuracy of the synchronization process, three hoarder boards were synchronized with each other and a digital watch using the above protocol. The boards were then shaken together several times during a full day to produce matching sinusoidal patterns on all boards. Visually comparing the peaks of these matching sinusoids across the three boards showed mean skew of 4.3 samples with a standard deviation of 1.8 samples between the boards. At a sampling frequency of 76.25 Hz, the skew between boards is equivalent to $0.0564 \pm 0.0236$ s.

4.2.2 Placement

A T-Mobile Sidekick phone pouch was used as a carrying case for all hoarders. The carrying case was light, durable, and provided protection for the electronics. A carrying case was secured to the subject’s belt on the right hip. Elastic medical bandages were used to wrap and secure carrying cases at sites other than the hip. Typical placement of hoarders is shown in Figure 4-1b. Figure 4-2 shows acceleration data collected for walking, running, and tooth brushing from the five accelerometers.
Figure 4-2: Acceleration from five biaxial accelerometers for walking, running, and tooth brushing. Signals from both axes of each biaxial accelerometer are shown.
No wires were used to connect the hoarders to each other or any other devices. Each hoarder in its carrying case weighed less than 120 g. Lack of wiring and the lightness of accelerometers minimized restrictions in subject movement.

4.3 Activity Labels

Twenty activities were studied. These activities were walking, walking while carrying items, sitting and relaxing, working on computer, standing still, eating or drinking, watching tv, reading, running, bicycling, stretching, strength-training, scrubbing, vacuuming, folding laundry, lying down and relaxing, brushing teeth, climbing stairs, riding an elevator, and riding an escalator. To address ambiguities in activity labels, subjects were provided short sentence descriptions of each activity (see Appendix B). For example, walking was described as “walking without carrying any items in your hand or on your back heavier than a pound” and scrubbing is described as “using a sponge, towel, or paper towel to wipe a window.”

The 20 activities were selected from a comprehensive list of everyday activities [4] and designed to include a range of common everyday activities that involve different parts of the body and range in level of intensity. Whole body activities such as walking, predominantly arm-based activities such as brushing of teeth, and predominantly leg-based activities such as bicycling were included. Sedentary activities such as sitting, light intensity activities such as eating, moderate intensity activities such as window scrubbing, and vigorous activities such as running were recognized.

In selecting activity labels, one major concern was distinguishing style from content [50]. For instance, an algorithm may recognize running but not the exact speed of the running action or window scrubbing but not the vigor of the scrubbing action. In this work, activity labels were chosen to reflect the content of the actions but do not specify the style. Consideration of the style of each of the 20 activities would multiply the number of activity labels, requiring greater amounts of training data and possibly specialized algorithms for distinguishing style. Thus, recognizing style is beyond the scope of this paper.
However, even limiting activity labels to reflect content is challenging. In speech recognition, words are defined by a specific sequence of letters. However, physical activities are not so clearly defined. The activity “scrubbing” can be interpreted as window scrubbing, dish scrubbing, or car scrubbing. To mitigate ambiguity in performance of activities, subjects were given short definitions of the 20 activity labels (see Appendix B). These short definitions resolved major ambiguities in the activity labels while leaving room for interpretation so that subjects could show natural, individual variations in how they performed activities.

4.4 Semi-Naturalistic Protocol

Semi-naturalistic data collection is an intermediate between laboratory and naturalistic data collection. For semi-naturalistic data collection, subjects ran an obstacle course consisting of a series of activities listed on a worksheet (see Appendix D.2). These activities were disguised as goals in an obstacle course to minimize subject awareness of data collection. For instance, subjects were asked to “use the web to find out what the world’s largest city in terms of population is.” Subjects recorded the time they began each obstacle and the time they completed each obstacle. Subjects completed each obstacle on the course ensuring capture of all 20 activities being studied. There was no researcher supervision of subjects while they collected data under the semi-naturalistic collection protocol. Many activities were performed outside of the lab, but some activities such as watching TV, vacuuming, lying down and relaxing, and reading were performed in a “common room” within the lab equipped with a television, vacuum, sofa, and reading materials. This room is pictured in Figure 4-3. No researchers or cameras monitored this room.

Although naturalistic data is most desirable, it requires direct observation of subjects by researchers, subject self-report of activities, or use of the experience sampling method [54] to label subject activities for algorithm training and testing. However, direct observation can be costly and scales poorly for the study of large subject populations, subject self-report is prone to recall errors [28], and the experience sampling
method requires frequent interruption of subject activity, which agitates subjects over an extended period of time. Furthermore, some activities such as folding laundry, riding escalators, and scrubbing windows may not occur on a daily basis. A purely naturalistic protocol would not capture sufficient samples of these activities for thorough testing of recognition systems.

Due to these deficiencies, a semi-naturalistic collection protocol was used in place of a purely naturalistic protocol to overcome its shortcomings while minimizing laboratory control on subjects.

4.5 Activity Protocol

Activity recognition results on laboratory data is designed to evaluate algorithm performance under ideal circumstances and to act as a baseline for comparison with results on semi-naturalistic data. Subjects were requested to perform random sequences of activities defined on a worksheet (see Appendix D.1) during laboratory data collection. Subjects performed the sequence of activities given at their own pace and labelled the start and end times of each activity in this case. Although this data collected under this protocol is considered laboratory data, the protocol allowed sub-
jects to perform their activities anywhere including outside of the laboratory. Also, there was no researcher supervision during the data collection session. Requiring subjects to perform certain activities under this protocol ensured that activities that may appear rarely in everyday life could be captured. These activities include folding laundry, riding escalators, and scrubbing windows. In assessing the robustness of a recognition system, it is important to test on these types of everyday activities since they may not appear frequently enough during naturalistic collection for sufficient sample sizes to be gathered. Subjects were free to perform activities anywhere, but they often performed activities such as watching TV, vacuuming, lying down and relaxing, and reading in the “common room” shown in Figure 4-3.

4.6 Activity Annotation

Under the laboratory data collection protocol, subjects were asked to perform a randomized sequence of activities on a worksheet (see Appendix D.1). For example, the first 3 activities listed on the worksheet might be “bicycling,” “riding elevator,” and “standing still.” As subjects performed each of these activity in the order given on their worksheet, they labelled the start and stop times for that activity and made any relevant notes about that activity such as “I climbed the stairs instead of using the elevator since the elevator was out of service.” Acceleration data collected between the start and stop times were labelled with the name of that activity.

Under the semi-naturalistic protocol, subjects were asked to perform an obstacle course in the form of a worksheet (see Appendix D.2). As subjects performed each of these obstacles in the order given on their worksheet, they labelled the start and stop times for that activity and made any relevant notes about that activity such as “I couldn’t eat any of the snacks provided because I am on a strict diet.” Acceleration data collected between the start and stop times were labelled with the name of that activity. Subjects were free to rest between obstacles and proceed through the worksheet at their own pace as long as they performed obstacles in the order given. Furthermore, subjects had freedom in how they performed each obstacle. For exam-
ple, one obstacle was to “read the newspaper in the House-n common room. Read the entirety of at least one non-frontpage article.” The subject could choose which and exactly how many articles to read.

Start and stop times for labels were hand annotated by the subject, making them imprecise. To minimize mislabelling, data within 10 s of the start and stop times was discarded. For example, if a subject reported that he started running at 11:05:10 and stopped running at 11:07:13, data from 11:05:10 to 11:05:20 and data from 11:07:03 to 11:07:13 was discarded. Only data from 11:05:20 to 11:07:03 was labelled as running. Since the subject is probably standing still or sitting while he records the start and stop times, the data immediately around these times may not correspond to the activity label. Figure 4-4 shows acceleration data annotated with subject self-report labels.

4.7 Feature Extraction

Features were computed on 512 sample windows of acceleration data with 256 samples overlapping between consecutive windows. At a sampling frequency of 76.25 Hz, each window represents 6.7 seconds. Mean, energy, frequency-domain entropy, and correlation features were extracted from the sliding windows signals for activity recognition. Feature extraction on sliding windows with 50% overlap has demonstrated success in past works [15, 52]. A window of several seconds was used to sufficiently capture cycles in activities such as walking, window scrubbing, or vacuuming. The 512 sample window size enabled fast computation of FFTs used for some of the features.

The DC feature is the mean acceleration value of the signal over the window. The energy feature was calculated as the sum of the squared discrete FFT component magnitudes of the signal. The sum was divided by the window length for normalization. Additionally, the DC component of the FFT was excluded in this sum since the DC characteristic of the signal is already measured by another feature. Note that the FFT algorithm used produced 512 components for each 512 sample window. Use of mean [19, 5] and energy [49] of acceleration features has been shown to result in
Figure 4-4: Five minutes of acceleration data annotated with subject self-report activity labels. Data within 10 s of self-report labels is discarded as indicated by masking.
accurate recognition of certain postures and activities (see Table 3).

Frequency-domain entropy is calculated as the normalized information entropy of the discrete FFT component magnitudes of the signal. Again, the DC component of the FFT was excluded in this calculation. This feature may support discrimination of activities with similar energy values. For instance, biking and running may result in roughly the same amounts of energy in the hip acceleration data. However, because biking involves a nearly uniform circular movement of the legs, a discrete FFT of hip acceleration in the vertical direction may show a single dominant frequency component at 1 Hz and very low magnitude for all other frequencies. This would result in a low frequency-domain entropy. Running on the other hand may result in complex hip acceleration and many major FFT frequency components between 0.5 Hz and 2 Hz. This would result in a higher frequency-domain entropy.

Correlation is a feature computed between two different acceleration axes. It is calculated as the dot product of two acceleration axes divided by the window length of 512. Features that measure correlation or acceleration between axes can improve recognition of activities involving movements of multiple body parts [22, 6]. Correlation is calculated between the two axes of each accelerometer hoarder board and between all pairwise combinations of axes on different hoarder boards.

Figure 4-5 shows the use of the above features in discriminating different activities. It is anticipated that certain activities will be difficult to discriminate using these features. For example, “watching TV” and “sitting” should exhibit very similar if not identical body acceleration. In fact, it may be impossible to distinguish these two activities purely from acceleration data. Additionally, activities such as “stretching” may show marked variation from person to person and for the same person at different times. Stretching could involve light or moderate energy acceleration in the upper body, torso, or lower body. The number of possible stretches a subject could engage in makes recognition difficult. Nonetheless, recognition will be performed on these types of activities to test the limitations of algorithms presented.
standing: mean vertical acceleration = 1.05 G

walking: vertical acceleration energy = 68.4

walking: forward acceleration FFT entropy = 0.7

running: vertical acceleration energy = 685.2

bicycling: forward acceleration FFT entropy = 0.9

sitting: mean vertical acceleration = 0.54 G

Figure 4-5: Differences in feature values aide in discriminating different activities.
4.8 Recognition Algorithms

Conventional decision table [24], nearest neighbor [3], decision tree [45], and Naive Bayesian [23] classifiers were tested for activity recognition using the feature vector. Decision based [5] and nearest neighbors [19, 29] algorithms have been used in past work to recognize activities. Naive Bayes is a computationally efficient algorithm that has been used for pattern classification in a variety of applications. The 20 activities we are recognizing can be considered “word level” activities since they do not involve sequences of other more atomic activities. Using “word level” activity recognition, systems could detect “sentence level” activities by evaluating transition probabilities between “word level” activities. Transition modelling could be achieved using hidden Markov models, which have been used to recognize sign language words from video by modelling transition probabilities between signed letters [48].
Chapter 5

Evaluation

5.1 Collection Protocol

Subjects were recruited through flyers seeking research study participants for compensation. Flyers were posted around the MIT campus and were also emailed to the MIT student population. Twenty subjects from the MIT community volunteered. Details on subject recruitment and instructions are given in Appendix A.

Each subject participated in two sessions of study. In the first session, subjects wore five accelerometers and a digital watch. Subjects completed an obstacle course worksheet, noting the start and end times of each obstacle on the worksheet.

In the second session, subjects wore the same set of sensors. Subjects performed the sequence of activities listed on their activity worksheet, noting the start and end times of these activities. This session lasted around 90 minutes.

5.2 Data Collected

Data was collected from 13 males and 7 females. Subjects ranged in age from 17 to 48 with a mean of 21.8 and standard deviation of 6.59. For laboratory data, each subject collected between 54 and 131 minutes of data with a mean of 96 minutes and a standard deviation of 16.7. Eight subjects skipped between one to four activities during laboratory data collection due to factors such as inclement weather, time
constraints, or problems with equipment including the television, vacuum, computer, and bicycle. Each subject collected between 82 and 160 minutes of semi-naturalistic data with a mean of 104 minutes and a standard deviation of 13.4. Subjects performed each activity on their obstacle course for an average of 156 seconds with a standard deviation of 50 seconds. Six subjects skipped between one to two obstacles during semi-naturalistic data collection due to factors listed earlier.

5.3 Results

Mean, energy, entropy, and correlation features were extracted from acceleration data. Activity recognition on these features was performed using decision table [39], instance-based learning (IBL or nearest neighbor) [18, 2], C4.5 decision tree [45], and naive Bayes [16, 27] classifiers found in the Weka Machine Learning Algorithms Toolkit [55].

Classifiers were trained and tested using two protocols. Under the first protocol, classifiers were trained on each subject’s activity sequence data and tested on that subject’s obstacle course data. This individual training protocol was repeated for all twenty subjects. Under the second protocol, classifiers were trained on activity sequence and obstacle course data for all subjects except one. The classifiers were then tested on obstacle course data for the only subject left out of the training data set. This leave-one-subject-out validation process was repeated for all twenty subjects. Mean and standard deviation for classification accuracy under both protocols is summarized in Table G.

Overall, recognition accuracy was highest for decision tree classifiers whereas Bayesian approaches performed poorly. The strong performance of decision tree classifiers is consistent with past work where decision based algorithms recognized lying, sitting, standing and locomotion with 89.30% accuracy [5]. Nearest neighbor is the second most accurate algorithm and its strong relative performance is also supported by past works where nearest neighbor algorithms recognized ambulation and postures with over 90% accuracy [29, 19]. Rule-based activity recognition captures conjunc-
<table>
<thead>
<tr>
<th>Classifier</th>
<th>Individual Training</th>
<th>Leave-one-subject-out Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Table</td>
<td>36.32 ± 14.501</td>
<td>46.75 ± 9.296</td>
</tr>
<tr>
<td>IBL</td>
<td>69.21 ± 6.822</td>
<td>82.70 ± 6.416</td>
</tr>
<tr>
<td>C4.5</td>
<td>71.58 ± 7.438</td>
<td>84.26 ± 5.178</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>34.94 ± 5.818</td>
<td>52.35 ± 1.690</td>
</tr>
</tbody>
</table>

Table 5.1: Summary of classifier results (mean ± standard deviation) using individual training and leave-one-subject-out training. Classifiers were trained on laboratory data and tested on obstacle course data.

...tions in feature values that may lead to higher recognition accuracy. For instance, the C4.5 decision tree classified sitting as an activity having nearly 1 G downward acceleration and low energy at both hip and arm. The tree classified bicycling as an activity involving moderate energy levels and low frequency-domain entropy at the hip and low energy levels at the arm. The tree distinguishes “window scrubbing” from “brushing teeth” because the first activity involves more energy in hip acceleration even though both activities show high energy in arm acceleration. The fitting of probability distributions to acceleration features under a Naive Bayesian approach may be unable to adequately model such rules due to the assumptions of conditional independence between features and normal distribution of feature values. Furthermore, Bayesian algorithms may require more data to accurately model feature value distributions.

Recognition accuracy was significantly higher for all algorithms under the leave-one-subject-out validation process. This indicates that the effects of individual variation in body acceleration may be dominated by strong commonalities between people in activity pattern. Additionally, because leave-one-subject-out validation resulted in larger training sets consisting of data from 19 subjects, this protocol may have resulted in more generalized and robust activity classifiers. The markedly smaller training sets used for the individual training protocol may have limited the accuracy of classifiers.

To control for the effects of sample size in comparing leave-one-subject-out and individual training, preliminary results were gathered using a larger training data
set collected for three subjects. These subjects were researcher affiliates and are not included in the official 20 subject count. Each of these subjects participated in one semi-naturalistic and five laboratory data collection sessions. The C4.5 decision tree algorithm was trained for each individual using data collected from all five of his laboratory sessions and tested on the semi-naturalistic data. The algorithm was also trained on five laboratory data sets from five random subjects other than the individual and tested on the individual’s semi-naturalistic data. The results are compared in Table 5.3. In this case, individual training resulted in an increase in recognition accuracy of above 4.32% over recognition rates for leave-one-subject-out-training. This difference shows that given equal amounts of training data, individual training can recognize activities more accurately than leave-one-subject-out training. However, the certainty of these conclusions is limited by the low number of subjects used for this comparison and the fact that the three individuals studied were all researcher affiliates. Nonetheless, these initial results support the need for further study of individual training for physical activity recognition.

The above results suggest that real-world activity recognition systems can rely on classifiers that are pre-trained on large activity data sets. Although preliminary results show that individual training can lead to more accurate activity recognition given large training sets, pre-trained systems offer greater convenience. Pre-trained systems could recognize many activities accurately without requiring training on data from their user, simplifying the deployment of these systems. Furthermore, since the activity recognition system needs to be trained only once before deployment, the slow running time for decision tree training is not an obstacle. Nonetheless, there may be limitations to a pre-trained algorithm. Although “word-level” activities were accu-

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Individual Training</th>
<th>Leave-one-subject-out Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4.5</td>
<td>77.31 ± 4.328</td>
<td>72.99 ± 8.482</td>
</tr>
</tbody>
</table>

Table 5.2: Summary of preliminary classifier results (mean ± standard deviation) using individual training and leave-one-subject-out training where both training data sets are equivalent to five laboratory data sessions.
rately recognized, higher level “sentence-level” activities may show greater variation between individuals. Ambulation and posture are similar across individuals due to shared physiology, but higher level activities such as playing football or walking the dog are more subject to personal behavioral patterns. For instance, playing football may involve sequences of running and dominant arm acceleration for a quarterback and crouching and sudden bursts of body acceleration due to collision for a lineman.

However, leave-one-subject-out validation still shows deficiencies in recognizing certain activities. Figure 5-1 shows an aggregate confusion matrix for the C4.5 classifier based on all 20 trials of leave-one-subject-out validation. Recognition accuracies for stretching and riding an elevator were below 50%. Recognition accuracies for “watching TV” and “riding escalator” were 77.29% and 70.56%, respectively. These activities do not have simple characteristics and are easily confused with other activities. For instance, “stretching” is often misclassified as “folding laundry” because both may involve the subject moving the arms at a moderate rate. Similarly, “riding elevator” is misclassified as “riding escalator” since both involve the subject standing still. “Watching TV” is confused with “sitting and relaxing” and “reading” because all the activities involve sitting. “Riding escalator” is confused with “riding elevator” since the subject may experience similar vertical acceleration in both cases. “Riding escalator” is also confused with “climbing stairs” since the subject sometimes climbs the escalator stairs.

Recognition accuracy using the decision tree classifier was also computed under a leave-one-accelerometer-in protocol. Specifically, recognition results were computed five times, each time using data from only one of the five accelerometers for the training and testing of the algorithm. The differences in recognition accuracy rates using this protocol from accuracy rates obtained from all five accelerometers are summarized in Table 5.3. These results show the accelerometer placed on the subject’s thigh is the most powerful for recognizing activities. Acceleration of the dominant wrist is more useful in discriminating activities than acceleration of the non-dominant arm. Acceleration of the hip is the second best for activity discrimination. This suggests that an accelerometer attached to a subject’s cell phone, which is often
<table>
<thead>
<tr>
<th>Activity</th>
<th>Recognition Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking</td>
<td>89.71</td>
</tr>
<tr>
<td>Walking while carrying items</td>
<td>82.10</td>
</tr>
<tr>
<td>Sitting and relaxing</td>
<td>94.78</td>
</tr>
<tr>
<td>Working on computer</td>
<td>97.49</td>
</tr>
<tr>
<td>Standing still</td>
<td>95.67</td>
</tr>
<tr>
<td>Eating or drinking</td>
<td>88.67</td>
</tr>
<tr>
<td>Watching TV</td>
<td>77.29</td>
</tr>
<tr>
<td>Reading</td>
<td>91.79</td>
</tr>
<tr>
<td>Running</td>
<td>87.68</td>
</tr>
<tr>
<td>Bicycling</td>
<td>96.29</td>
</tr>
<tr>
<td>Stretching</td>
<td>41.42</td>
</tr>
<tr>
<td>Strength-training</td>
<td>82.51</td>
</tr>
<tr>
<td>Scrubbing</td>
<td>81.09</td>
</tr>
<tr>
<td>Vacuuming</td>
<td>96.41</td>
</tr>
<tr>
<td>Folding laundry</td>
<td>95.14</td>
</tr>
<tr>
<td>Lying down and relaxing</td>
<td>94.96</td>
</tr>
<tr>
<td>Brushing teeth</td>
<td>85.27</td>
</tr>
<tr>
<td>Climbing stairs</td>
<td>85.61</td>
</tr>
<tr>
<td>Riding elevator</td>
<td>43.58</td>
</tr>
<tr>
<td>Riding escalator</td>
<td>70.56</td>
</tr>
</tbody>
</table>

Table 5.3: Aggregate recognition rates (%) for activities studied using leave-one-subject-out validation over 20 subjects.
<table>
<thead>
<tr>
<th>Accelerometer(s) Left In</th>
<th>Difference in Recognition Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hip</td>
<td>$-34.12 \pm 7.115$</td>
</tr>
<tr>
<td>Wrist</td>
<td>$-51.99 \pm 12.194$</td>
</tr>
<tr>
<td>Arm</td>
<td>$-63.65 \pm 13.143$</td>
</tr>
<tr>
<td>Ankle</td>
<td>$-37.08 \pm 7.601$</td>
</tr>
<tr>
<td>Thigh</td>
<td>$-29.47 \pm 4.855$</td>
</tr>
<tr>
<td>Thigh and Wrist</td>
<td>$-3.27 \pm 1.062$</td>
</tr>
<tr>
<td>Hip and Wrist</td>
<td>$-4.78 \pm 1.331$</td>
</tr>
</tbody>
</table>

Table 5.4: Difference in recognition accuracy (mean ± standard deviation) due to leaving only one or two accelerometers in. Accuracy rates are aggregated for 20 subjects using leave-one-subject-out validation.

placed at a fixed location such as on a belt clip, may enable recognition of certain activities.

Confusion matrices resulting from leave-one-accelerometer-in testing (see Appendix F) show that data collected from lower body accelerometers placed on the thigh, hip, and ankle is generally best at recognizing forms of ambulation and posture. Accelerometer data collected from the wrist and arm is better at discriminating activities involving characteristic upper body movements such as reading from watching TV or sitting and strength-training (push ups) from stretching. To explore the power of combining upper and lower body accelerometer data, a leave-two-accelerometers-in protocol was used. This protocol is analogous to the leave-one-accelerometer-in protocol, except that data from two accelerometers is used. Data from thigh and wrist accelerometers and hip and wrist accelerometers were used. These results are also listed in Table 5.3. Note that recognition rates improved over 25% for the leave-two-accelerometers-in results as compared to the best leave-one-accelerometer-in results. Of the two pairs tested, thigh and wrist acceleration data resulted in the highest recognition accuracy. However, both thigh and wrist and hip and wrist pairs showed less than a 5% decrease in recognition rate from results using all five accelerometer signals. This suggests that effective recognition of certain everyday activities can be achieved using two accelerometers placed on the wrist and thigh or wrist and hip.
Figure 5-1: Aggregate confusion matrix for C4.5 classifier based on leave-one-subject-out validation for 20 subjects using laboratory and obstacle course data.
5.4 Analysis

This work shows that individual training is not necessary to achieve recognition rates of over 80% for 20 everyday activities. Using features from previous literature such as mean [19, 5], energy [49], and correlation [22, 6] of acceleration along with frequency-domain entropy of acceleration, this work demonstrates successful recognition of activities under laboratory and semi-naturalistic conditions.

Classification accuracy rates of between 80% to 95% for walking, running, climbing stairs, standing still, sitting, lying down, working on a computer, bicycling, and vacuuming are on par with recognition results using laboratory data from previous works [30, 35, 29, 52, 47, 19, 5, 51]. However, all cited works used data collected under controlled laboratory conditions to achieve their recognition accuracy rates. Considering that subjects could move about freely outside the lab without researcher supervision while collecting laboratory and semi-naturalistic data, the 84.26% overall recognition rate achieved in this work is significant in understanding physical activity recognition under more realistic conditions.

The C4.5 classifier used mean acceleration to recognize postures such as sitting, standing still, and lying down. Ambulatory activities and bicycling were recognized by the level of hip acceleration energy. Frequency-domain entropy and correlation between arm and hip acceleration strongly distinguished bicycling, which showed low entropy hip acceleration and low arm-hip correlation, from running, which displayed higher entropy in hip acceleration and higher arm-hip movement correlation. Both activities showed similar levels of hip acceleration mean and energy. Working on a computer, eating or drinking, reading, strength-training as defined by a combination of situps and pushups, window scrubbing, vacuuming, and brushing teeth were recognized by arm posture and movement as measured by mean acceleration and energy.

“Watching TV,” “stretching,” “riding elevator,” and “riding escalator” were recognized at between 40% and 80% accuracy. The C4.5 decision tree has difficulty in distinguishing the activities from others with similar feature values such as sitting,
folding laundry, standing, and climbing stairs. Higher accuracy rates for recognition of these types of activities may require “sentence level” activity recognition systems. For instance, stretching is usually followed by vigorous exercise. Riding an escalator is usually preceded by standing still while riding an escalator is preceded by walking. Accurate discrimination of activities such as watching TV from others such as sitting may be impossible without audio or visual sensory data.

The stronger performance of leave-one-subject-out training over individual training demonstrated in this work does not invalidate individual training of activity recognition algorithms. Because leave-one-subject-out validation produced training sets 19 times larger than those encountered with individual training, it may be that the individual training data sets collected were just too small or homogeneous. Preliminary experiments with individual training using more training data per individual shows improved performance. However, more data from a greater number of subjects is required to substantiate these initial findings.

Furthermore, it may be possible to augment the power of individual training data by requesting subjects to perform activities in a variety of manners to decrease homogeneity of the training set. This could reduce overfitting and improve recognition rates achieved with individual training. For instance, subjects could be requested to scrub windows, slowly, vigorously, and carefully to collect training data for a number of possible scrubbing motions encountered in real-life situations. However, increasing training data requires more training time, which is a barrier to widespread real-world adoption of activity recognition systems. Nonetheless, combining individual training data with population training data may improve activity recognition rates further without greatly increasing training time.

5.5 Future Work

Although this work demonstrates that training classifiers exclusively on an individual basis may lead to poor activity recognition accuracy, combining individual training with a pre-trained classifier may yield improved activity recognition. In the case of
speech, recognition systems integrate pre-compiled general information about language with individual training to maximize recognition accuracy [46]. If the same paradigm applies for activity recognition, researchers must determine how best to integrate individual training with pre-trained activity models. An effective scheme for weighting and incorporating the individual training data into a pre-trained classifier must be determined.

If real-world systems will train on individual activity patterns, the details of unsupervised training for a mobile recognition system must be addressed. A PDA connected wirelessly to sensors could collect acceleration and other types of data. The PDA could ask the user to perform certain activities to initially train the system for the individual. Once more, this paradigm is akin to that of speech recognition, where users are requested to utter specific words for the system to train on [46].

Lower recognition accuracies for activities such as stretching, scrubbing, riding an elevator, and riding an escalator suggest that higher level analysis is required to improve classification of these activities. Temporal modelling may enable better detection of these and other more complex activities such as “walking a pet” or “playing basketball.” These activities are characterized by a sequence of activities such as walking, standing still, walking, or running, jumping, running. Hidden Markov models exploit the temporal relation of words in sentences to improve speech recognition [46]. Similarly, low-level activities such as walking, sitting, or standing still can be thought of as words and higher-level activities can be thought of as sentences. Then hidden Markov models can exploit the temporal relation of low-level activities in higher-level activities to improve activity recognition.

Temporal information in the form of duration and time and day of activities could also be used to detect activities. For instance, standing still and riding an elevator are similar in terms of body posture. However, riding an elevator usually lasts for a minute or less whereas standing still can last for a much longer duration. By considering the duration of a particular posture or type of body acceleration, these activities could be distinguished from each other with greater accuracy. Similarly, adults may be more likely to watch TV at night than at other times on a weekday.
Thus, date and time can be used to improve discrimination of watching TV from sitting and relaxing. However, because daily activity patterns may vary dramatically across individuals, individual training may be required to effectively use date and time information for activity recognition.

The decision tree algorithm used in this work can recognize the content of activities, but may not readily recognize activity style. Although a decision tree algorithm could recognize activity style using a greater number of labels such as “walking slowly,” “walking briskly,” “scrubbing softly,” or “scrubbing vigorously,” the extensibility of this technique is limited. For example, the exact pace of walking cannot be recognized using any number of labels. Other techniques such as neural nets or regression could be used to recognize parameterize activity style.

Use of other sensor data modalities may further improve activity recognition. Heart rate data could be used to augment acceleration data to detect intensity of physical activities. GPS location data could be used to infer whether an individual is at home or at work and affect the probability of activities such as working on the computer or lying down and relaxing. Decision trees have been used to recognize activities such as working, shopping, attending class from GPS location data with over 90% accuracy [13]. However, use of GPS data may require greater individual training since individuals can work, reside, and shop in totally different locations.
Chapter 6

Conclusion

Using decision tree classifiers, recognition accuracy of over 80% on a variety of 20
everyday activities was achieved using leave-one-subject-out-validation on both labo-
ratory and semi-naturalistic data from 20 subjects. These results are competitive with
activity recognition that only used laboratory data [30, 35, 29, 52, 47, 19, 5, 51]. The
success of leave-one-subject-out validation across 20 subjects suggests that accurate
mobile activity recognition systems using acceleration data do not require training on
individuals. This simplifies the wide deployment of these systems in the real-world.

Furthermore, this work shows acceleration can be used to recognize activities such
as vacuuming, window scrubbing, and working on the computer. This extends pre-
vious work on recognizing ambulation and posture using acceleration [30, 35, 29, 52,
47, 5, 51]. This work indicates that a mobile system consisting of a mobile computer
and small wireless accelerometers placed on an individual’s thigh and wrist may be
able to detect everyday activities in naturalistic settings. The computer could com-
pute mean, acceleration, frequency-domain entropy, and correlation of acceleration
features and run a pre-trained decision tree algorithm on these features. The algo-

rithm could be pre-trained on a large data set of everyday activities. Because decision
trees are slow to train but quick to run, a pre-trained decision tree should be able to
classify user activities in real-time.

Such a mobile activity recognition system would be applicable to many innova-
tive computer applications for mobile computing devices. In the field of preventive
health, such systems could recognize changes in physical activity patterns of the sick or elderly, helping detect the onset of diseases or physical conditions. In the entertainment industry, activity recognition systems could enable portable video game units to solicit user interaction while the user is idle. PDAs using activity recognition could encourage users to engage in healthier behavior by suggesting that they take the stairs the next time they are about to ride an elevator or that they walk to lunch instead of driving.

To encourage further research in the area of physical activity recognition, all data collected for this study is made publicly available. See Appendix H for details on acquiring this data.
Appendix A

Subject Protocol

The following outlines the subject recruitment and data collection process. This protocol was approved by the MIT Committee On the Use of Humans as Experimental Subjects.

1. Flyers publicized the need for research subjects. These flyers were distributed around MIT. Additionally, emails were sent to the MIT community soliciting volunteer research subjects. Gift incentives of $5 certificates to a local ice cream shop were used to attract potential subjects.

2. Potential subject volunteers were informed that any data collected from them during the study would be disassociated with their actual identities and would be used strictly for training and testing pattern recognition algorithms. Participating subjects signed an informed consent form to document their agreement to participate in the study.

3. Subjects participated in two data collection sessions lasting between 60 minutes and 2 hours. During each session, subjects were given instructions found in Appendices A.1 and A.2 and a definition of activity labels shown in Appendix B.

4. Subjects followed given instructions and wore five accelerometers. Acceleration data was collected during these sessions.
5. Subjects who experienced difficulty with their devices or had questions or concerns during the experiment could talk to the research investigator.

6. At the conclusion of the experiment, subjects could submit informal written or oral feedback about the study. Feedback was not required.

7. Each subject received a $5 gift certificate to a local ice cream parlor at the end of the study. Subjects were free to decline this offer.

The following sections detail instructions given to subjects for each of the two data collection sessions.

A.1 Semi-Naturalistic Session Instructions

General Instructions - Session 1

We would like you to complete a series of obstacles in the form of an obstacle course. Completing all the obstacles should take you around 90 minutes and involves walking to certain locations and reaching objectives at those locations. Here are detailed instructions for this session.

1. You will meet with a lab person to receive a digital watch and 5 accelerometers. Wear the watch on your non-dominant wrist. A lab person will secure accelerometers to your hip, dominant wrist, non-dominant upper arm, dominant ankle, and non-dominant thigh. The lab person will give you a worksheet titled "Obstacle Course Worksheet" (See Appendix D.1). The worksheet will consist of a table with numbered "obstacles" for you to perform. These obstacles will usually involve going to a certain destination or performing a specific task.

2. Complete each item on the obstacle course on the worksheet in the order that they are numbered. Once you are ready to complete an obstacle, write down the time on your watch (to the second) next to the activity entry on the worksheet under the "Start Time" column.
3. When you are done with an obstacle on the worksheet, write down the time on your watch (to the second) next to the obstacle entry on the worksheet under the "End Time" column. You can also make notes about any unusual or special circumstances under the notes column.

4. Repeat Steps 2 and 3 until you have completed all obstacles on the worksheet to the best of your abilities. This may require you to go to several different locations and sometimes to move back and forth between locations.

5. When you have completed all obstacles on the worksheet, return to the lab person. He will retrieve your watch and remove any accelerometers from your body. You may ask him any questions or give him comments during this time.

If you have any problems or questions during the study, ask the researcher investigator Ling Bao. You can also call him if you are outside of the building at 617-452-5642.

A.2 Laboratory Session Instructions

General Instructions - Session 2

We would like you to complete a sequence of activities in a particular order given on an activity worksheet. Completing all the activities on the worksheet should take you around 90 minutes and involves such tasks as walking, running, and eating. Here are detailed instructions for this session.

1. You will meet with the lab person to receive a digital watch and 5 accelerometers. Wear the watch on your non-dominant wrist. A lab person will secure accelerometers to your hip, dominant wrist, non-dominant upper arm, dominant ankle, and non-dominant thigh. The lab person will give you a worksheet titled "Activity Worksheet" (see Appendix D.2). The worksheet will consist of a table with numbered activities for you to perform. These activities will usually involving doing such things as watching TV, doing push ups, and standing still.
2. Complete each activity on the activity worksheet in the order that they are numbered. Once you are ready to perform an activity, write down the time on your watch (to the second) next to the activity entry on the worksheet under the "Start Time" column.

3. Perform each activity for at least 1 minute or to the best of your abilities unless otherwise specified on the worksheet. When you are done with an activity on the worksheet, write down the time on your watch (to the second) next to the obstacle entry on the worksheet under the "End Time" column and the location at which you performed the activity under the "Location" column. You can also make notes about any unusual or special circumstances under the notes column.

4. Repeat Steps 2 and 3 until you have completed all activities on the worksheet. This may require you to go to certain locations. For instance, the "riding escalator" activity may require you to find a building with an escalator in it. You are free to choose exactly where you perform each of the activities.

5. When you have completed all obstacles on the worksheet to the best of your ability, return to the lab person. He will retrieve your watch and remove any accelerometers from your body. You may ask him any questions or give him comments during this time.

If you have any problems or questions during the study, ask the researcher investigator Ling Bao. You can also call him if you are outside of the building at 617-452-5642.
Appendix B

Activity List

The following list describes the 20 activities studied. The same list was given to subjects so they were clear on the definition of each activity type.

Walking Walking without carrying any items in your hand or on your back heavier than a pound.
Walking while carrying items Walking while carrying a backpack.
Sitting and relaxing Sitting down inactively. Does not include working on computer or reading while sitting.
Working on computer Using a computer while seated.
Standing still Standing without moving of legs.
Eating or drinking Consuming foods or beverages.
Watching TV Watching television from a seated position.
Reading Reading a book or texts while seated.
Running Jogging at a moderate clip or higher. Only one foot is contacting the ground at a time.
Bicycling Riding a bicycle or using a cycling machine in a gym.
Stretching Doing stretching exercises for the arms and legs.
Strength-training Doing push-ups and sit-ups.
Scrubbing Using a sponge, towel, or paper towel to wipe a window.
Vacuuming Vacuuming a carpet.
Folding laundry Folding clothes neatly.
Lying down and relaxing Lying on your back and being inactive.
Brushing teeth Brushing your teeth with a tooth brush.
Climbing stairs Ascending or descending stairs.
Riding elevator Ascending or descending floors via an elevator.
Riding escalator Ascending or descending floors via an escalator.
Appendix C

Prior Work Details

Details for prior literature relating to physical activity recognition is provided below.

Reference [30]

Activities Recognized walking, ascending stairs, descending stairs
Recognition Accuracy 92.85% to 95.91%

Features standard deviation over 50 samples of forward acceleration, upward acceleration, and thigh angle along with integral of thigh angular velocity between the last 4 zero-crossings of angular velocity; acceleration low-pass filtered with 2.5 Hz cutoff

Algorithm created fuzzy sets of Gaussian distributions based on features for each activity; samples classified as activity with highest likelihood given activity distributions

Subjects 8 (6 males, 2 females)

Data Collection Methodology subjects performed a scripted sequence of walking motions; data collected with microcontroller worn by subject

Sensor Placement 1 biaxial accelerometer and 1 angular velocity sensor placed on outside of thigh; accelerometer axes faced forward and upward relative to subject; axis of angular velocity sensor points forward relative to subject

Sensor Equipment ±2 G accelerometer and angular velocity sensor for detecting thigh angle sampled at 50 Hz; sensors wired to microcontroller and PDA for data storage

Figures and Tables Table C.1
<table>
<thead>
<tr>
<th></th>
<th>Level</th>
<th>Up</th>
<th>Down</th>
<th>Unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>95.91</td>
<td>0.51</td>
<td>0.67</td>
<td>2.92</td>
</tr>
<tr>
<td>Up</td>
<td>0</td>
<td>94.35</td>
<td>0</td>
<td>5.65</td>
</tr>
<tr>
<td>Down</td>
<td>0.51</td>
<td>0</td>
<td>92.85</td>
<td>6.63</td>
</tr>
</tbody>
</table>

Table C.1: Recognition ratios (%) for walking (Level), ascending stairs (Up), descending stairs (Down) in [30].

Reference [35]

Activities Recognized standing, walking, climbing stairs

Recognition Accuracy 83% to 90%

Features 6 axes of acceleration data decimated by 2 with anti-aliasing filtering; ICA is used on filtered data to generate 6 components; components are normalized to have mean of 0 and variance of 1; wavelet coefficients for levels 5 to 8 composed from 256 sample sliding windows with 64 sample shifts to generate 24 channels; power of each channel calculated for 24 feature vector

Algorithm 3 multilayer perception neural networks trained for 3 activities using back propagation; 10-fold cross validation used to test classifiers

Subjects 6

Data Collection Methodology subjects walked a predefined route through an office environment; data collected with laptop computer carried by researcher with subject

Sensor Placement 2 triaxial accelerators placed on left and right side of hip

Sensor Equipment ±2 G accelerometers sampled at 256 Hz; accelerometers wired to laptop PC, which collected data

Figures and Tables Table C.2

Reference [19]

Activities Recognized walking, ascending stairs, descending stairs, sitting, standing, lying supine, talking, bicycling, typing

Recognition Accuracy 95.8% for data collected from subjects performing scripted sequence of activities, 66.7% for naturalistic data
<table>
<thead>
<tr>
<th></th>
<th>Standing</th>
<th>Level</th>
<th>Down</th>
<th>Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standing</td>
<td>90</td>
<td>7</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Level</td>
<td>1</td>
<td>85</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Down</td>
<td>0</td>
<td>16</td>
<td>84</td>
<td>0</td>
</tr>
<tr>
<td>Up</td>
<td>3</td>
<td>14</td>
<td>0</td>
<td>83</td>
</tr>
</tbody>
</table>

Table C.2: Recognition ratios (%) for standing, walking (Level), ascending stairs (Up), descending stairs (Down) in [35].

**Features**  DC and AC components of raw signals separated using simulated resistance-capacitance circuit; mean of DC and AC components for signals over 20 second durations used as features

**Algorithm** classification used weighted sum of absolute differences between feature values and activity reference patterns; samples classified as nearest activity reference using this distance metric

**Subjects** 24

**Data Collection Methodology** each subject performed sequence of activities for fixed durations around the laboratory; 50 minutes of naturalistic data collected from each subject as the subject actly freely; researchers following subjects annotated naturalistic data; signals recorded with portable Vitaport recorder

**Sensor Placement** 4 uniaxial accelerometers placed perpendicular to body surface on sternum, dorsum of the wrist distal from m. extensor carpi ulnaris, frontal aspect of thigh distal from m. rectus femoris, frontal aspect of lower leg, 1 uniaxial accelerometer placed between left ear and mastoid to measure vertical head movement, microphone attached to throat, electrocardiogram with modified Nehb anterior leads

**Sensor Equipment** ±2 G accelerometers, microphone, electrocardiogram; sensors wired to microcontroller for data storage

**Figures and Tables** Figure C-1

**Reference** [5]

**Activities Recognized** lying, sitting, standing, locomotion

**Recognition Accuracy** 89.30%
<table>
<thead>
<tr>
<th>Lying</th>
<th>Sitting</th>
<th>Talking</th>
<th>Typing</th>
<th>Standing</th>
<th>Walking</th>
<th>Descending Stairs</th>
<th>Ascending Stairs</th>
<th>Cycling</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>16</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>114</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>107</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>80</td>
<td>26</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>20</td>
<td>1</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
</tbody>
</table>

Figure C-1: Confusion matrix for laboratory data results in [19]. Confusion matrix is not available for naturalistic results.

**Features**  forward and vertical acceleration, lowpass filtered acceleration with 0.5 Hz cutoff, median of acceleration, mean absolute deviation of acceleration, mean absolute deviation of sum of acceleration channels all over 1 second windows formed features

**Algorithm** decision tree using feature set

**Subjects** 5 (4 males, 1 female)

**Data Collection Methodology** subjects acted freely for 1 hour in studio room except they were requested to perform each type of activity studied at least once for several minutes; video footage used to annotate data; Physilog system recorded signals

**Sensor Placement** uniaxial accelerometer placed at chest to measure vertical acceleration, uniaxial accelerometer perpendicular to front of thigh measured forward acceleration

**Sensor Equipment** accelerometers with unknown sensitivity wired to Physilog recording device

**Reference** [22]

**Activities Recognized** walking speed and incline

**Recognition Accuracy** root mean squared error (rmse) of .12 m/s for speed, 0.014 rad rmse for incline
**Features** 4 acceleration axes, mean, median, variance, skewness, and kurtosis of acceleration channels, covariance between pairs of channels, gait cycle time, and maximal negative acceleration of the heel used as feature vector

**Algorithm** stepwise regression analysis determines features that are most associated with variability of subject speed, incline; these features are used to train 2 neural networks for predicting speed and incline

**Subjects** 20 (10 m, 10 f); subjects were volunteers aged between 19 and 29

**Data Collection Methodology** each subject ran 18 laps around an 80-100 m track at varying speeds; 12 laps of data per subject was used for training, the rest of data was used for testing; data collected with an electronic device carried around waist

**Sensor Placement** triaxial accelerometer placed at lower back, uniaxial accelerometer measuring frontal acceleration placed on Achilles tendon above the ankle

**Sensor Equipment** ±5 G accelerometers with exception of one ±10 G accelerometer placed at heel; accelerometers wired to microcontroller and sampled at 200 Hz

**Reference** [51]

**Activities Recognized** locomotion, standing, sitting, lying down, playing

**Recognition Accuracy** 86% to 93%

**Features** unknown due to proprietary nature of DynaPort Monitor

**Algorithm** proprietary DynaPort ADL Monitor software

**Subjects** 1, a maintenance worker at city hall

**Data Collection Methodology** the subject worked regularly while the DynaPort Monitor recorded acceleration data on 2 days; video footage was used for activity annotation

**Sensor Placement** 2 uniaxial accelerometers at the waist measure forward and upward movement; 1 uniaxial accelerometer on the left thigh measure frontal acceleration; all accelerometers are attached to DynaPort Monitor

**Sensor Equipment** ±14 G accelerometers wired to microcontroller

**Reference** [12]

**Activities Recognized** 3 types of Kung Fu martial arts movements (cuts, elbows, punch blocks)

**Recognition Accuracy** 96.67%
<table>
<thead>
<tr>
<th></th>
<th>Cuts</th>
<th>Elbows</th>
<th>Punch Blocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cuts</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Elbows</td>
<td>1</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Punch Blocks</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

Table C.3: Confusion matrix for three martial arts moves recognized in [12].

**Features** zero crossing rate of first and second derivatives, root mean square, and mean of acceleration channels were computed over sliding windows of 32 samples with 8 samples overlap

**Algorithm** 9 hidden markov models were used for 9 sub-gestures such as "wood cut", "grass cut", "throat cut", "side cut"; 3 markov model trained on sequence of sub-gestures were used to for each of the 3 major gestures

**Subjects** 1, a martial arts instructor

**Data Collection Methodology** martial arts instructor performed 10 examples of gestures for training and 30 sequences of moves, 10 of each type of gesture, for testing; gestures were performed with only the arm that was monitored with accelerometers

**Sensor Placement** 2 orthogonal uniaxial accelerometers secured at wrist measuring acceleration in the direction of the extended thumb and index finger; accelerometer data recorded via computer

**Sensor Equipment** ±2 G accelerometer sampled at 96 Hz; accelerometer wired via PC serial port

**Figures and Tables** Table C.3

**Reference** [29]

**Activities Recognized** walking, ascending stairs, descending stairs

**Recognition Accuracy** 83.3% to 96.3%

**Features** AC components of forward and vertical acceleration lowpass filtered with 5 Hz cutoff extracted; time of last positive and negative peaks in filtered acceleration signals over 0.5 s sliding windows comprise 4 features; lag in cross-correlation between forward and vertical signals is fifth feature
Table C.4: Recognition ratios (%) for walking (Level), ascending stairs (Up), descending stairs (Down) in [29].

<table>
<thead>
<tr>
<th></th>
<th>Level</th>
<th>Up</th>
<th>Down</th>
<th>Unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>96.3</td>
<td>1.7</td>
<td>0</td>
<td>2.0</td>
</tr>
<tr>
<td>Up</td>
<td>11.1</td>
<td>83.3</td>
<td>0</td>
<td>5.6</td>
</tr>
<tr>
<td>Down</td>
<td>0</td>
<td>2.8</td>
<td>95.8</td>
<td>1.4</td>
</tr>
</tbody>
</table>

**Algorithm** nearest neighbor classification using Euclidean distance between feature vector and personalized reference feature vectors for each activity; reference vectors calculated as average of training sample feature vectors for the individual

**Subjects** 6

**Data Collection Methodology** subjects performed a scripted sequence of walking motions; data collected with microcontroller worn by subject

**Sensor Placement** 2 orthogonal uniaxial accelerometers secured at lower back measuring acceleration in the forward and vertical directions

**Sensor Equipment** ±2 G accelerometer sampled at 50 Hz, digital compass, infrared light detector; sensors wired to microcontroller and PDA for data storage

**Figures and Tables** Table C.4

**Reference** [52]

**Activities Recognized** walking, running, ascending stairs, descending stairs, sitting, standing, bicycling

**Recognition Accuracy** 42% to 96%

**Features** maximum acceleration, standard deviation of acceleration, number of acceleration zero crossings, mean of standard deviation of acceleration over 50 samples composed the feature set

**Algorithm** 2-dimensional Kohonen map of 20 by 20 neurons trained on feature set

**Subjects** 1, researcher

**Data Collection Methodology** subject carried a laptop which requested the subject to perform certain activities; the subject was asked to press certain buttons depending on what activity they were doing; this information was used to annotate the data
### Table C.5: Recognition rates (%) for activities studied in [52].

<table>
<thead>
<tr>
<th>Activity</th>
<th>Recognition Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting</td>
<td>96</td>
</tr>
<tr>
<td>Standing</td>
<td>94</td>
</tr>
<tr>
<td>Walking</td>
<td>75</td>
</tr>
<tr>
<td>Running</td>
<td>78</td>
</tr>
<tr>
<td>Climbing stairs</td>
<td>42</td>
</tr>
<tr>
<td>Descending stairs</td>
<td>64</td>
</tr>
<tr>
<td>Riding bicycle</td>
<td>91</td>
</tr>
</tbody>
</table>

Sensor Placement 2 orthogonal uniaxial accelerometers secured above the knee measuring acceleration in the forward and vertical directions

Sensor Equipment ±5 G accelerometers wired to laptop PC

Figures and Tables Table C.5

Reference [47]

Activities Recognized walking, running, ascending stairs, descending stairs, sitting, standing

Recognition Accuracy 85% to 90%

Features 4 features: root mean square and integration of 2 acceleration axes over 2 s windows

Algorithm single layer neural network trained on set of 4 feature

Subjects 10

Data Collection Methodology data collection protocol not described in paper

Sensor Placement 2 orthogonal uniaxial accelerometers placed in the pocket measuring acceleration in the forward and vertical directions

Sensor Equipment ±2 G accelerometers wired to onHandPC
Appendix D

Data Collection Worksheets

Sample obstacle course and activity sequence worksheets are shown below.

D.1 Obstacle Course Worksheet

The obstacle course worksheet lists a fixed order of everyday activities for subjects to perform during the semi-naturalistic session. A scanned obstacle course worksheet is shown below.

D.2 Activity Worksheet

The activity sequence worksheet lists a random order of 20 activities subjects to perform during the laboratory session. Each of the 20 activities is repeated twice on the worksheet. A scanned activity sequence worksheet is shown below.
Everyday Tasks Worksheet

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Start Time (24 hr time HH:MM:SS)</th>
<th>End Time (24 hr time HH:MM:SS)</th>
<th>Special notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn on the House-n TV and VCR and watch the videotape. Watch the tape until Raymond starts kissing his wife, then turn off the TV and VCR.</td>
<td>2:14:42</td>
<td>2:16:44</td>
<td></td>
</tr>
<tr>
<td>Read the newspaper in the House-n common room. Read the entirety of at least one non-frontpage article.</td>
<td>2:17:00</td>
<td>2:22:18</td>
<td></td>
</tr>
<tr>
<td>Sit in the House-n common room and relax. Breathe in deeply. Inhale and exhale as slowly as you can 12 times.</td>
<td>2:22:20</td>
<td>2:23:50</td>
<td></td>
</tr>
<tr>
<td>Look at the painting of Boston in the House-n common room. Count the number of graves in the cemetery on the left side of the painting. Write that number down in the notes column.</td>
<td>2:24:58</td>
<td>2:25:57</td>
<td>54 graves</td>
</tr>
<tr>
<td>Neatly fold the House-n laundry in the common room. Don’t worry, these are clean.</td>
<td>2:26:34</td>
<td>2:28:00</td>
<td></td>
</tr>
<tr>
<td>Vacuum the House-n common room. Try to get all the paper scraps littering the floor. Turn the vacuum off when you're done.</td>
<td>2:28:26</td>
<td>2:31:47</td>
<td>1460 sq ft 7th floor</td>
</tr>
<tr>
<td>Use the Windex and paper towels on the House-n shelves to clean the windows near the entrance of the lab. Be sure to clean all the stains off the windows. Wash your hands in the kitchen afterwards.</td>
<td>2:36:08</td>
<td>2:38:46</td>
<td></td>
</tr>
<tr>
<td>Browse the web on the House-n computer. Use the web to find out what the world’s largest city in terms of population is. Also find a major landmark or tourist attraction of this city. Write down the city’s name and the name of the landmark under the notes column. Try your best; if it takes too long, move on.</td>
<td>2:39:20</td>
<td>2:46:16</td>
<td></td>
</tr>
<tr>
<td>Reward yourself by eating some snacks and drinking some water in the common room. Snacks, cups, and water from the sink can be found in the House-n kitchen.</td>
<td>2:46:57</td>
<td>2:49:11</td>
<td></td>
</tr>
<tr>
<td>Go to the bathroom and brush your teeth after that meal. Remember to brush for at least a few minutes to clean your gums thoroughly.</td>
<td>2:50:40</td>
<td>2:54:09</td>
<td></td>
</tr>
</tbody>
</table>

Figure D-1: Obstacle course worksheet, page 1 of 3.
Put the House_n backpack on and walk from the House_n lab to the elevator banks.

Take the elevator to the first floor and stop in the lobby next to the couches.

Take the stairs down the Kendall T station and browse the T map. Find the northernmost stop on the blue line and write down its name under the notes column.

Ascend the stairs out of the T station.

Walk to Dewey Library and enter the library.

Find any book that looks interesting. Skim the book jacket and write a one sentence synopsis of it in the notes column.

Return the book to its shelf and go back to the 1 Cambridge Center lobby.

Take the elevator to House_n and put down your backpack in the common room.

Do 3 sets of 10 sit ups or as many as you can.

Go to the elevator banks.

Ride the elevator to the lobby.

Go outside and unlock the House_n bike (6-12-6). Bike on the side walks. Staying on the side walks, bike to the quantum bookstore and back. Lock up the bike when you're done.

Head down the stairs of the Kendall T. This time, write down the name of the westernmost T stop on the MBTA map in the notes column.

Ride the escalator of the T back to ground level.

Walk to the abstract statue in front of the green building. Write down the name of the statue (its on a plaque) in the notes column.

Jog at a comfortable pace back to 1 Cambridge Center.

Take the elevator up to the House_n common room.

Do 3 sets of 10 push ups or as many as you can and then drink some water from the fountain on the floor.

---

Figure D-2: Obstacle course worksheet, page 2 of 3.
Stretch your arms, legs, and back in the common room to warm down

Lie down on the sofa and relax for at least a minute after all that exercise

You’re done! Talk to the researcher to conclude this session.

Figure D-3: Obstacle course worksheet, page 3 of 3.

Activity Worksheet

<table>
<thead>
<tr>
<th>Activity</th>
<th>Start Time (ipaq 24 hr time HH:MM:SS)</th>
<th>End Time (ipaq 24 hr time HH:MM:SS)</th>
<th>Location</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brushing teeth</td>
<td>1:48:12</td>
<td>2:00:48</td>
<td>Common</td>
<td></td>
</tr>
<tr>
<td>Walking while carrying items</td>
<td>1:51:26</td>
<td>2:57:57</td>
<td>Hall</td>
<td></td>
</tr>
<tr>
<td>Sitting and relaxing</td>
<td>1:53:49</td>
<td>1:57:37</td>
<td>Common</td>
<td></td>
</tr>
<tr>
<td>Standing still</td>
<td>1:54:49</td>
<td>1:55:46</td>
<td>Common</td>
<td></td>
</tr>
<tr>
<td>Riding escalator (best effort)</td>
<td>1:59:22</td>
<td>1:59:41</td>
<td>SJ</td>
<td></td>
</tr>
<tr>
<td>Brushing teeth</td>
<td>2:04:04</td>
<td>2:05:33</td>
<td>Common</td>
<td></td>
</tr>
<tr>
<td>Vacuuming</td>
<td>2:10:11:16</td>
<td>2:10:19:17</td>
<td>Common</td>
<td></td>
</tr>
<tr>
<td>Walking</td>
<td>2:10:19:19</td>
<td>2:11:29</td>
<td>Office</td>
<td></td>
</tr>
<tr>
<td>Bicycling</td>
<td>2:11:12:12</td>
<td>2:15:16</td>
<td>Kitchen</td>
<td></td>
</tr>
<tr>
<td>Folding laundry (2 minutes)</td>
<td>2:32:57</td>
<td>2:35:04</td>
<td>Office</td>
<td></td>
</tr>
<tr>
<td>Watching TV (2 minutes)</td>
<td>2:43:44</td>
<td>2:45:16</td>
<td>Office</td>
<td></td>
</tr>
<tr>
<td>Climbing stairs (best effort)</td>
<td>2:43:54</td>
<td>2:45:21</td>
<td>Office</td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>2:44:58</td>
<td>2:46:29</td>
<td>Common</td>
<td></td>
</tr>
<tr>
<td>Strength-training</td>
<td>2:45:08</td>
<td>2:46:29</td>
<td>Common</td>
<td></td>
</tr>
<tr>
<td>Bicycling</td>
<td>2:45:00</td>
<td>2:51:37</td>
<td>Office</td>
<td></td>
</tr>
<tr>
<td>Folding laundry (2 minutes)</td>
<td>2:55:21</td>
<td>2:57:33</td>
<td>Office</td>
<td></td>
</tr>
<tr>
<td>Watching TV (2 minutes)</td>
<td>2:56:36:36</td>
<td>3:00:56</td>
<td>Office</td>
<td></td>
</tr>
<tr>
<td>Stretching</td>
<td>3:02:29:29</td>
<td>3:05:56</td>
<td>Common</td>
<td></td>
</tr>
<tr>
<td>Riding escalator (best effort)</td>
<td>3:03:26:26</td>
<td>3:06:54</td>
<td>Hall</td>
<td></td>
</tr>
<tr>
<td>Eating or drinking (1.5 minutes)</td>
<td>3:05:02</td>
<td>3:08:33</td>
<td>Hall</td>
<td></td>
</tr>
<tr>
<td>Scrubbing</td>
<td>3:15:59</td>
<td>3:15:59</td>
<td>Hall</td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>3:19:23:24</td>
<td>3:21:59</td>
<td>Hall</td>
<td></td>
</tr>
<tr>
<td>Running (best effort)</td>
<td>3:23:00:00</td>
<td>3:24:02:00</td>
<td>SJ</td>
<td></td>
</tr>
<tr>
<td>Lying down and relaxing</td>
<td>3:24:10:10</td>
<td>3:25:17:17</td>
<td>SJ</td>
<td></td>
</tr>
<tr>
<td>Eating or drinking (1.5 minutes)</td>
<td>3:24:17:17</td>
<td>3:25:17:17</td>
<td>SJ</td>
<td></td>
</tr>
</tbody>
</table>

Figure D-4: Activity sequence worksheet.
Appendix E

Hoarder Clock Synchronization

Five hoarder boards were shaken together in a sinusoidal fashion along their two axes before and after a data collection session. The sinusoidal signals at the beginning and end of each set of collected data were visually aligned. Time stamps for the aligned sinusoidal signals were adjusted to match across the five hoarder boards. This was done at the beginning and end of the signal. Time stamps in between the beginning and end of the signal were linearly adjusted to be consistent with the aligned time stamps at the beginning and end. This process mitigates the effects of independent clock skew between the hoarder boards. Figure E-1 shows the sinusoidal signal across the five hoarder boards after visual alignment.
Figure E-1: Visually aligned synchronization signals across five accelerometers.
Appendix F

Result Details

Confusion matrices for leave-one-accelerometer-in and leave-two-accelerometers-in results are shown below.
Figure F-1: Aggregate confusion matrix for C4.5 classifier based on leave hip accelerometer in validation for 20 subjects using laboratory and obstacle course data.
Figure F-2: Aggregate confusion matrix for C4.5 classifier based on leave wrist accelerometer in validation for 20 subjects using laboratory and obstacle course data.
Figure F-3: Aggregate confusion matrix for C4.5 classifier based on leave arm accelerometer in validation for 20 subjects using laboratory and obstacle course data.
Figure F-4: Aggregate confusion matrix for C4.5 classifier based on leave ankle accelerometer in validation for 20 subjects using laboratory and obstacle course data.
Figure F-5: Aggregate confusion matrix for C4.5 classifier based on leave thigh accelerometer in validation for 20 subjects using laboratory and obstacle course data.
Figure F-6: Aggregate confusion matrix for C4.5 classifier based on thigh and wrist accelerometer data for 20 subjects using laboratory and obstacle course data.
Figure F-7: Aggregate confusion matrix for C4.5 classifier based on hip and wrist accelerometer data for 20 subjects using laboratory and obstacle course data.
Appendix G

Related Results

A preliminary experiment was conducted prior to this work using ADXL202JE accelerometers from Analog Devices. Two hoarder boarders mounted with the accelerometers were placed in each subject’s pockets and around the subject’s arm. Laboratory and semi-naturalistic data was collected for twenty subjects using protocol similar to those used for the 10 G accelerometer study. Additionally, five subjects participated in naturalistic data collection. For naturalistic data collection, subjects carried a PDA equipped with experience sampling software that interrupted them periodically and asked them for their activity. During naturalistic sessions, subjects acted freely outside the laboratory. For instance, subjects went to work wearing the accelerometers and carrying the PDA.

There were several problems with this preliminary experiment. Because typical body acceleration amplitude can range up to 12 G [8], data collected with the 2 G accelerometers was often saturated. Accelerometers placed in the pocket were not secured firmly and may have collected dampened acceleration of the hip due to loose pockets. Furthermore, many subjects participating in the preliminary experiment were affiliated with the researcher or with the research group. Nonetheless, results for this preliminary study are shown below.
Table G.1: Summary of classifier results (mean ± standard deviation) using individual training and leave-one-subject-out training for 2 G data. Classifiers were trained on laboratory data and tested on semi-naturalistic data.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Individual Training</th>
<th>Leave-one-subject-out Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Table</td>
<td>49.95 ± 11.858</td>
<td>61.93 ± 7.492</td>
</tr>
<tr>
<td>IBL</td>
<td>70.33 ± 7.333</td>
<td>77.59 ± 6.913</td>
</tr>
<tr>
<td>C4.5</td>
<td>73.69 ± 4.082</td>
<td>88.26 ± 3.941</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>32.22 ± 8.593</td>
<td>45.71 ± 6.188</td>
</tr>
</tbody>
</table>

Table G.2: Summary of classifier results (mean ± standard deviation) using individual training and leave-one-subject-out training. Classifiers were trained on laboratory data and tested on naturalistic data.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Individual Training</th>
<th>Leave-one-subject-out Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Table</td>
<td>33.18 ± 13.607</td>
<td>56.48 ± 9.523</td>
</tr>
<tr>
<td>IBL</td>
<td>62.99 ± 7.579</td>
<td>72.79 ± 7.828</td>
</tr>
<tr>
<td>C4.5</td>
<td>68.42 ± 5.997</td>
<td>83.31 ± 4.093</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>22.58 ± 10.366</td>
<td>37.54 ± 8.571</td>
</tr>
</tbody>
</table>
Appendix H

Data Availability

All data collected for this work is freely available to the public for research purposes. Contact Stephen Intille at intille@mit.edu for information on acquiring data. The entire data set is 3.2 GB.
Bibliography


