UNIVERSITY OF SOUTHAMPTON

Detecting and Monitoring Behavioural Change Through Personalised Ambient Monitoring

by

James D. Amor

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the
Faculty of Engineering and the Environment
Institute of Sound and Vibration Research

October 2011

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND THE ENVIRONMENT
INSTITUTE OF SOUND AND VIBRATION RESEARCH

Doctor of Philosophy

DETECTING AND MONITORING BEHAVIOURAL CHANGE THROUGH PERSONALISED AMBIENT MONITORING

by James D. Amor

Bipolar disorder (BD) is one form of mental illness and is estimated to affect around 0.4–1.6% of the population. The disorder is characterised by recurrent episodes of mania and depression and is estimated to cost the UK economy £5.21 billion a year. Many people with BD self-monitor their behaviour to help them identify the early warning signs of an affective episode. The Personalised Ambient Monitoring (PAM) project has been conceived take ideas from existing telehealth systems and apply them to BD. By using a distributed network of discreet, unobtrusive sensors, the user's behavioural patterns can be monitored and deviations in their behaviour can be detected. In doing so it is hoped that the early warning signs can be detected and that this can be used to assist them in their self-monitoring.

The PAM system is being developed by a multi-disciplinary team based at the ISVR and the School of Management at the University of Southampton, the School of Electrical and Electronic Engineering at the University of Nottingham and the Department of Computing Science and Mathematics at the University of Stirling.

This thesis presents the background and motivations for the PAM project, the approach the project will take, a review of appropriate data analysis techniques and the experimental work that has been undertaken in the investigation of accelerometry for activity monitoring, the use of a wireless camera to monitor a complex environment and the use of multiple sensors to capture behaviour patterns in a technical trial.

Results from the technical trial show that it is possible to process information from a variety of sensors to identity activity signatures and behavioural patterns in normal controls. When these activity patterns are trained on week-days, the results presented show that it is possible to identify weekend days as being behaviourally different.

Contents

List of Figures

List of Tables

Nomenclature

General

x, y, z	Variables
$\mathbf{X}$	Data matrix
$\mathbf{x}$, $\mathbf{y}$	Data vector
t	Time
$P(x = y)$	Probability
$p(x)$	Probability density function
N	Number of data-points
n, m	Number of dimensions
i, j	(Subscripts) Indexes
$f(x)$	Function

Chapter 4

x, y, z	Axes of acceleration
A	Total acceleration
k	Number of neighbours
f	Feature
F	Partial feature set
$\mathcal{F}$	Set of all features
c	Class variable
$I(x, y)$	Mutual information function
$\mathbf{\Sigma}$	Covariance matrix
$\boldsymbol{\alpha}$	Eigenvector
λ	Eigenvalue
$\mathbf{Y}$	Projection matrix
$\mathbf{S}$	Subjective dissimilarity matrix
s_{ij}	Point in $\mathbf{S}$
E	Neuroscale error parameter
α	Neuroscale supervision parameter
d_{ij}, d^*_{ij}	Neuroscale Distance measures
$\mathbf{U}$	Partition matrix
μ_{ij}	Point in $\mathbf{U}$
c	Number of clusters
$\mathbf{V}$	Cluster centres matrix
$\mathbf{A}$	Norm inducing matrix
$D^2_{ijA_j}$	Clustering distance measure
tp	True positive
fp	False positive
tn	True negative
fn	False negative

Chapter 5

p	Pixel
H	Change mask
$h(p)$	Pixel in change mask
I	Image
$\mathbf{r}$	Region of support
$\mathcal{I}$	Pixel illuminance ratio
$W(x/y)$	Wronskian function
$y(p)$	Pixel intensity
$i(p)$	Pixel illumination component
$r(p)$	Pixel reflectance component

Chapter 6

T	Number of days
s	(Subscript) Data-stream index
L^p_s	Total log-likelihood
$L\alpha_s$	Training log-likelihood
$L\beta_s$	Testing log-likelihood
ρ_s	Consistency measure
D_s	Behavioural difference score

W_{avg}	Average WBD score
S	Set of HMM states
q	HMM state
M	Number of observable values
V	Set of HMM output values
v	HMM output value
A	Set of HMM transition probabilities
a_{ij}	HMM state transition probability
B	Set of HMM emission probabilities
$b_j(k)$	HMM emission probability
π	Set of HMM initial state probabilities
π_i	HMM initial state probability
Λ	HMM parameter set
O	HMM Observation sequence
Q	HMM state sequence
$\alpha_t(.)$	Forwards step
$\beta_j(.)$	Backwards step
$\delta_t(.)$	Viterbi algorithm step
$\psi_t(.)$	Viterbi algorithm tracking array
$\xi_t(.,.)$	Compound probability in Baum-Welch
$\gamma_t(.,.)$	Compound probability in Baum-Welch
$\mathcal{X}$	Data set drawn from a probability density function
$\mathcal{Y}$	EM hidden parameters
$\mathcal{Z}$	EM complete data set
$\mathcal{L}(.)$	Likelihood function
Θ	EM parameter set
$Q(.)$	EM E-step function
k	(Superscript) Number of time series
$\mathbf{z}$	CPM Latent trace
G	Length of latent trace
ϵ	Adjustment factor in G
$\boldsymbol{\iota}$	Sequence of hidden states
ι	CPM Hidden state
τ	CPM Time state
ϕ	CPM Scale state
R	Number of scale states
ζ	Noise parameter
u	Scale parameter
$\mathcal{N}(.)$	Normal distribution
T	CPM State transition probability
$D(x,y)$	CPM Distance measure
L^p	CPM likelihood
χ	Smoothing penalty
$\mathcal{P}$	Regularising term

Declaration of Authorship

I, James David Amor declare that the thesis entitled Detecting and Monitoring Behavioural Change Through Personal Ambient Monitoring and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as:

  J.D. Amor and C.J. James. Personalised ambient monitoring: wireless cameras for activity detection. In *Proceedings of the 5th UKRI PG Conference in Biomedical Engineering and Medical Physics*, 2009.

  J.D. Amor and C.J. James. Personalized ambient monitoring: Accelerometry for activity level classification. In *4th European Conference of the International Federation for Medical and Biological Engineering*, 2009.

  J.D. Amor and C.J. James. Behavioral pattern detection from personalized ambient monitoring. In *Proceedings of the 32nd Annual International Conference of the IEEE EMBS*, 2010.

Signed: Date:

Acknowledgements

Thanks to Prof. C. James for his input, the members of the PAM project team for their parts of the PAM project and for the project funding from EPSRC grant number EP/F005091/1.

Chapter 1

Overview

1.1 Introduction

A person's health is a subjective measure of their overall condition in both body and mind. It is important that a state of good health be maintained in order that the person experience a good quality of life. In the pursuit of good health, attention must be paid to both physical health and mental health and whilst the provision for physical healthcare in the UK is sufficient, mental illnesses are still not well understood [68]. There is still significant room for improvement in the provision of care for mental health and in maintaining a state of wellbeing in people with mental illness. It is for this reason that the provision and improvement of care for mental illness is an ongoing priority for the UK government [40].

Mental illnesses are medical conditions that can cause direct problems with a person's mood, emotion, thoughts and relationships and can interfere with a person's day-to-day functioning [80]. There are a large number of mental illnesses such as depression, dementia and schizophrenia. It is estimated that approximately 1 in 4 people will suffer from some form of mental illness over the course of their lives [40] and this has an incredibly high social and economic impact.

The cost to the economy of mental health illnesses in England was estimated to be £48.6 billion a year in 2008[1] [66]. This cost includes both service costs (£22.5 billion) and lost earnings (£26.1 billion). In relation to these costs, it is estimated that more than 58 million working days are lost a year due to mental illness [30].

Bipolar disorder (BD) is one such form of mental illness and is estimated to affect around 0.4–1.6% of the population [5] and affects a variety of age groups. The disorder is characterised by recurrent episodes of mania and depression, which are interspersed with

[1] The study focused on "specific conditions covering major mental health problems" and included depression, anxiety, schizophrenia, BD, eating disorders, personality disorder, child/adolescent disorders and dementia. Less common disorders for which data were not available were not included in the study.

periods of remission. Manic symptoms are varied and can include an increase in energy, a decreased need of sleep and an increase in goal directed activity. Depressive symptoms are equally varied and represent the opposite side of the coin, typically including decreased energy, increased need of sleep and a general lack of enjoyment in nearly all activities. Episodes can last for between 2 weeks and 5 months for mania and slightly longer for depression. Episodes that are severe enough can often result in hospitalisation [5, 81].

The potentially severe symptoms of BD and the increased risk of hospitalisation mean that BD is very disruptive to a person's life. Affective episodes can result in undesired activity, such as excessive spending on credit cards whilst manic, or a neglect of personal hygiene and care whilst depressed, to pick two examples. Hospitalisation is also a disrupting event in itself. It can be very difficult because of this disruption for someone with BD to successfully hold down long term employment and this, coupled with the expense of hospitalisation, is the driver behind the economic cost of BD, which was estimated to be £5.21 billion a year in 2008. This figure breaks down into £1.64 billion in service costs and £3.57 billion in lost earnings [66].

There are a number of treatments available for BD, both pharmaceutical and psychological and both types of treatment are often used together. Pharmaceutical treatment focuses on the biological aspects of BD and aims to stabilise the patient's mood with drug based treatment. The gold standard is lithium, but other treatments can be used instead if lithium is not effective or produces severe side effects [51]. The general focus of the psychological approaches is to counteract the non-biological aspects of BD by recognising prodromes and symptoms and taking appropriate action when necessary [7, 44]. Many people with BD self-monitor their behaviour and mood to help them identify not only prodromes and symptoms but also any events that could trigger an affective episode. By doing this, the early warning signs can be detected and acted upon early and the severity of an affective episode can be greatly diminished [94].

The high economic cost of mental illness, in conjunction with an ageing population in developed countries and the need to provide better care for mental illness is creating a strong socioeconomic driver for the provision of low-cost, effective healthcare solutions that can be delivered outside of the hospital environment. This has been realised in the development of telehealth and telecare systems for providing exactly this sort of healthcare for a number of conditions, such as diabetes [71], dementia [27], chronic obstructive pulmonary disorder (COPD) [18] and for assisted living in general [107]. Compared to the advances in telecare for dementia and other conditions, there has been relatively little work performed that examines the development of such a system for BD. Such systems as do exist often focus on providing a technological implementation of existing mood reporting systems, such as ChronoRecord [11] and iMonitor [62], or are more broad in scope and deal with the diagnosis of mental illness [104]. The Personalised Ambient Monitoring (PAM) project [43, 91] has been conceived to take ideas from these systems and from telehealth systems that exist for other illnesses and apply them to BD.

The project is aiming to develop a self-management system for people with BD that can automate some of the self-management that they perform as a way to help control their condition. By using a distributed network of discreet, unobtrusive sensors, the user's behavioural patterns can be monitored and an activity signature, which shows the normal behaviour patterns of the user, can be identified. From this established pattern changes in the user's behaviour can be detected and alerts can be generated to be sent to the user and potentially to their carers if the user so desires. Building onto this system, the PAM project is aiming to develop some computationally intelligent data analysis that can combine behavioural change information, expert knowledge about BD and knowledge about a user's specific illness pathways to determine the user's health trajectory.

The PAM system is being developed by a multi-disciplinary team based at the ISVR and the School of Management at the University of Southampton, the School of Electrical and Electronic Engineering at the University of Nottingham and the Department of Computing Science and Mathematics at the University of Stirling. The research questions being addressed by the project as a whole are:

1. Is it possible to obtain, in an automatic, ambient and unobtrusive manner, activity signatures from the mentally ill that provide information about the trajectory of their health status?
2. If this is so, can this information be used to assist their healthcare?

Currently, the PAM project has just completed the feasibility study stage, which was concentrating on the development and testing of the technology platform upon which the higher level functionality can be built. A prototype system has been developed and has undergone a technical trial in order to test the system and to gather data that can be used to develop appropriate data analysis techniques.

As part of the PAM project, the author's work focused on the data analysis for the PAM project, in particular on the detection of activity signatures and changing behavioural patterns. As such, the principal research questions that are being addressed in this thesis are:

1. Is it possible to obtain activity signatures from normal controls?
2. Is it possible to detect changes in behavioural patterns in normal controls?

In answering these questions the author's work took in both high-level and low-level data analysis required for the PAM system, including the analysis of the data from a technical trial of the entire PAM system. Three sensors were investigated by the author; cameras, accelerometers and passive infra-red (PIR) sensors. The hardware for the camera and the PIR sensors were integrated into the system by the author and the hardware for the accelerometer was implemented and integrated into the system by the team at the University of Nottingham. Development of data analysis routines for all three was undertaken by the author. Data gathered from the technical trial were also analysed by the author, which required the fusion of the sensor data from the camera, accelerometer, PIRs and other sensors developed in Nottingham and Stirling, and for the use of a variety of data analysis methods.

The remainder of this thesis presents the work that the author has performed on the PAM project. Chapter 2 presents background information on BD, including its diagnostic criteria, symptoms and treatment; on telehealth systems for acute, sub-acute and chronic care in general and for BD in particular; and on the technology used in telecare system. Chapter 3 presents the system requirements, system design and data processing architecture of the PAM system and some background information on the data analysis techniques that have been used. Chapters 4 and 5 present work that has been carried out on the development of two of the sensors for the PAM system. The use of an accelerometer and the ways in which accelerometry data can be processed are discussed in Chapter 4. A controlled study was performed to gather a body of data that was used to develop appropriate analysis routines to allow the inclusion of accelerometers in the PAM system. The use of a wireless camera as a single sensor for complex environments is discussed in Chapter 5. Four image change detection algorithms were evaluated on several image sequences in order to select an appropriate image processing algorithm. Chapter 6 presents the technical trial of the prototype system, the data analysis that is used to detect activity signatures and behavioural change and the results of running this data analysis on the data from the technical trial. Finally, Chapter 7 presents some general discussion on the current state of the PAM system, some ideas for further work and the conclusions that can be drawn both from the work on the PAM project as a whole and from the work that the author has performed on the detection of activity signatures and behavioural change.

Chapter 2

Background

2.1 Introduction

The PAM project is a multi-disciplinary project and as such draws on several different fields, including psychiatric health; telecare and telemedicine; complex data analysis; and hardware design and communication protocols. This chapter provides the background information relevant to the development of the PAM system. As such BD and telehealth systems are covered in this chapter whilst Chapter 3 covers the design of the PAM system and some background on data analysis. This chapter is split into two parts. Firstly, a discussion of BD is presented, including the diagnostic criteria, symptoms and principal treatments. Secondly, an examination of existing telehealth and telecare systems that provide similar functionality to what the PAM project is aiming towards is presented.

2.2 Bipolar Disorder

BD is a mental disorder characterised by recurring episodes of mania and depression, termed affective episodes. The disorder can be very disruptive, as both manic and depressive episodes can severely affect a person's daily life, and relapses often result in hospitalisation.

2.2.1 Symptoms

There are three distinct types of episode in BD: manic, depressive and mixed. Whilst the typical manic and depressive symptoms are experienced by most people with BD, there is also room for considerable variability in the specific symptoms a person may experience. Symptoms of a particular episode type can also vary from episode to episode in the same person [7].

A. Mania

Manic episodes are classified as either hypomania or mania and both share similar symptoms, with hypomania being a less severe case of mania. The major symptoms of a manic episode as listed by the ICD-10 [81] and DSM-IV [5] are a persistent, elevated, expansive or irritable mood; increased energy and activity; and feelings of well being. These can be described as euphoric, cheerful, unusually good or high. A person often experiences some symptoms of increased sociability, talkativeness, flights of ideas, distractibility, overfamiliarity, increased sexual energy and a decreased need of sleep, although not all of these will be present. For a hypomanic episode the symptoms are not severe enough that they significantly interfere with occupational or social functioning [5, 44, 51, 67, 81].

During a full manic episode a person's mood is significantly altered, irrespective of their circumstances, and it is not uncommon for a person to develop grandiose ideas and an inflated self esteem. The change in mood is accompanied by more of the additional symptoms described above, all to a greater extent than for hypomania [5, 44, 51, 81].

Mania (but not hypomania) may also present with psychotic symptoms, in which a person may develop delusions based on their inflated self esteem and grandiose ideas. In extreme circumstances these may develop into delusions of identity, which can also have a religious aspect. The flights of ideas coupled with fast speech may make the person incomprehensible and the person's increased activity may result in neglect of their own well being, such as not eating or drinking enough [81].

B. Depression

Depressive episodes are classified as either mild, moderate or severe. The typical symptom of depression is depressed mood, often described as depressed, sad or hopeless, which leads to a loss of interest and enjoyment in nearly all activities. The person's depressed mood is present most days and for most of the day. Energy is also reduced, which leads to increased fatigue and less activity. Other symptoms that may be present are reduced concentration, reduced self-esteem and self-confidence, ideas of guilt, bleak views of the future, ideas or acts of self harm, disturbed sleep and loss of appetite [5, 81].

In mild depressive episodes the typical symptoms are present, but not severe, as are some of the additional symptoms. A person will normally have some difficulty with normal occupational and social functioning. In some individuals the observable occupational and social functioning will appear normal but will require increased mental effort on the part of the person [5].

A person with a moderate depressive episode will present with more severe symptoms and will have more of the additional symptoms to a greater degree. The person will have considerable trouble continuing with normal occupational and social functioning.

In severe depressive episodes, the person's normal social and occupational functioning will cease completely [5, 81].

C. Mixed Episodes and Rapid Cycling

A mixed episode is one in which a person exhibits both manic and depressive symptoms nearly every day. Alternatively, the person may present with a rapidly alternating mood, from depressive to manic or hypomanic on a daily or hourly basis. The DSM-IV also includes agitation, insomnia, appetite dysregulation, psychotic features and suicidal thinking as symptoms of a mixed episode. The social and occupational functioning of the person is often significantly affected [5, 81].

Any diagnosis of BD can be augmented by 'with rapid cycling'. In order for this to be used, a person must experience a minimum of four affective episodes in a twelve month period. This is distinct from the rapid switching of symptoms in a single mixed episode [82].

2.2.2 Diagnosis

The World Health Organisation's ICD-10 [81] and the American Psychiatric Association's DSM-IV [5] provide the diagnostic criteria that are used for the diagnosis of BD. The diagnosis can be made when a person is experiencing an affective episode, either manic or depressive, or if the person is in remission.

There are two principal types of BD; bipolar I disorder and bipolar II disorder. In order to be diagnosed with bipolar I disorder, a person has to have had (including the current episode if the patent is not in remission) at least one manic-type episode (mania, hypomania or mixed) and one other affective episode (mania, hypomania, mixed or depressive). If the patent is currently manic or hypomanic then the they must have had at least one other affective episode. If the patient is currently depressed the requirement is strictly for a previous manic-type episode.

The diagnosis of bipolar II disorder can be made if a person suffers from major depressive episodes with episodes of hypomania [5], however, the presence of any full manic episodes prohibit this diagnosis from being made. If a person has only ever had a single manic episode they can be diagnosed with 'bipolar disorder, single manic episode' [5]. If a person has not had any episodes of mania then the diagnosis of recurrent depressive disorder is more appropriate [81].

The above diagnostic criteria mean that mania, rather than depression, is the dominant factor in diagnosis of BD. A person has to have had one or more manic-type episodes in the past; episodes of depression are not strictly necessary for diagnosis, but are a common occurrence.

2.2.3 Predictors and Triggers of Relapse

As with many mental disorders, there are several factors that can be either predictors or triggers of relapse in BD. Predictors are factors that are indicative of shorter times to relapse, both for manic, depressive and mixed episodes. Triggers however, are factors that can directly bring about, or help to bring about, episodes of relapse. Many triggers contribute to either depression or mania, but not both.

Predictors generally relate to a person's environment and mental state. The home environment, in terms of social support networks and close family and friends, is important. A lack of social support has been linked to an increased risk of relapse; mainly of depressive but also of manic episodes [44]. The emotional attitudes of close family and friends, if hostile or overly interfering, are strong predictors for both types of episode. A person's mental state during remission can also be an important factor as any subsyndromal symptoms that are present during remission, or partial remission, increase the risk of future relapse [44, 86].

One of the most robust predictors is medication non-adherence [7, 44]. People who discontinue their medication, without clinical advice to do so, relapse in a shorter time than people who maintain their medication. Current substance abuse has also been shown to be related to an increased risk of manic, but not depressive, relapse [51, 86].

Triggers in BD generally relate to disrupting and stressful events in a person's life. Disruptions to the sleep-wake cycle and the circadian rhythm seem to be particularly important [7, 31, 94]. Disturbances that cause high levels of social disruption have been linked with episodes of mania but not to depression [31]. Particularly stressful life events are linked to both manic and depressive recurrence [31, 44]. Excessive activity, overstimulation and sleep deprivation can also contribute to the onset of an affective episode [7, 44, 94].

2.2.4 Prodromes and Early Warning Signs

Prodromes are the pre-symptomatic signs that a person with BD may suffer a relapse. The prodromal stage of an episode can last for anywhere between 1–84 days (mean 20.5 days) for mania and for between 2–31 days (mean 11 days) for depression [51]. Prodromes are very similar to the symptoms of their respective episodes, albeit to a lesser degree. Prodromes of mania include increased activity, decreased need for sleep, talkativeness and an increase in goal directed behaviour. Prodromes for depression include interrupted sleep, loss of interest in activities and feelings of sadness [51].

It has been shown that people with BD can recognise and act upon their prodromes [51, 94]. This early action can help to prevent the prodromes developing into full symptoms. There are a great many things that people with BD do to intervene in the prodromal stages. These include cancelling work or social plans, increasing medication, sleeping more and making an appointment with a clinician [94].

2.2.5 Treatment

There are two different treatment types for BD, pharmacology and psychotherapy. Pharmacology is treatment with pharmaceutical drugs, whereas psychotherapy is treatment with psychological techniques. Both can be, and often are, used together in the same treatment plan.

A. Pharmacology

The first choice treatment in pharmacology for BD is lithium, which acts as a mood stabiliser and helps to reduce the swings between mania and depression. Lithium is also the only drug that has been shown to reduce suicidality [44]. Lithium is used both for the acute phases of BD and for prophylaxis, where it is taken regularly and in low doses [51].

Lithium is not a totally effective treatment; 20–40% of people with BD do not respond to lithium therapy [51] and three quarters of people who use it report side effects. These include thirst, frequent urination, tremor, weight gain and memory problems [44]. The side effects can be lessened by reducing the dosage, but this also reduces the efficiency of the treatment. This has led to the use of other drugs that have fewer side effects, such as anti-seizures, anti-depressants and anti-psychotics.

B. Psychotherapy

There are several different approaches taken in psychotherapy for BD, such as cognitive therapy, interpersonal and social rhythm therapy, family therapy and collaborative practice management. Psychotherapy does not involve pharmaceutical drugs but instead seeks to counteract the non-biological factors of the disorder and reduce the severity of the mood swings through cognitive and behavioural changes [7, 44, 51].

Cognitive therapy aims to counter a person's disruptive thought processes and assumptions, and provide the person with a set of cognitive skills that can help them better cope with the disorder. There are normally three areas that cognitive therapy focuses on: psychoeducation, cognitive and behavioural skills, and self monitoring and maintenance. Psychoeducation teaches the person about the disorder, how it is treated and how certain stressors can exacerbate symptoms and bring on episodes. Developing cognitive and behavioural skills enables the person to recognise and react to maladaptive thoughts, assumptions and overly positive or negative thinking. This helps the person to better cope with the prodromes of an episode and prevent the episode from worsening. Finally the self-management and maintenance aspect sets the person up to monitor their own symptoms and take control of their illness [44, 51].

Interpersonal and social rhythm therapy (IPSRT) [31, 44] is a behavioural technique that aims to regularise a person's circadian rhythm and sleep-wake cycles. This directly counters many of the triggers of episode relapse, such as sleep deprivation, overstimulation and social disruption. IPSRT focuses on five main areas: the link between mood and life events, keeping a regular daily routine, identification of early rhythm deviation, acceptance of the diagnosis and symptom management [31].

Family therapy is a therapy aimed at a person, their family and their close friends. This form of therapy deals with negative emotion that the family and friends may express toward the person, encouraging more emotionally sensitive communication and problem solving; and promoting greater tolerance and acceptance of the disorder. Ultimately, family therapy provides the person with a more supportive and constructive environment and contributes to their long term emotional stability [44].

Collaborative practice management aims to allow a person to be better able to manage their own illness and care. It is similar to other chronic disease management methodologies and focuses on three areas: illness self management skills, healthcare provider support, and improved access to, and care from, primary nurses [44].

2.3 Telehealth and Healthcare Technology

The rapid development in information and communication technology in the last few decades, coupled with an increasing population size and an ageing population has led to the development of systems to deliver healthcare outside of the traditional hospital environment. The delivery of healthcare in this way significantly reduces costs, as the cost for hospital care is quite high, and allows the patient to be treated in a comfortable and familiar home setting. The term telehealth has been coined to describe healthcare delivered remotely with the use of information and communications technology [77].

2.3.1 Telehealth Terms

The term telehealth is an over-arching term that covers the provision of healthcare for acute, subacute and chronic situations. The terms telemedicine and telecare appear in the literature as more specific terms for these two sides of telehealth. The former refers mainly to the provision of medical care for acute conditions, whereas the latter refers to the provision of home based care for a chronically ill patient.

The term telemonitoring also occurs in the literature and refers to the capture and transfer of medical data relating to a patient and their condition; the capture and transmission of an acutely ill patient's electrocardiogram (ECG) and blood oxygen saturation (SpO_2) from a care home to an external doctor for example [20].

2.3.2 Acute Treatment

Acute treatment is treatment for an immediate need and over a very short time-span following the initial incident. For example, the immediate care of a patient following a stroke or a heart attack. The focus is on the treatment of the patient's condition and the capture and transmission of any relevant medical data. Telecare in the acute phase is often synchronous, with parties at both ends of the telecommunications link being present and communicating with each other and transferring medical information in real time. A good everyday example of synchronous communication is a telephone call. Continuing the example of stroke care, a combination of video-conferencing and the acquisition and transmission of a patient's X-ray computed tomography (CT) scan can be used to improve their care [52].

2.3.3 Sub-Acute Treatment

Sub-acute treatment covers treatment in the medium term; further past the initial incident than the acute phase, but not far enough to be considered chronic care. For example, the capture and transmission of images of burn injuries as part of the patient referral process [108]. Telecare in the sub-acute phase can either be synchronous or asynchronous. Asynchronous communication occurs when both parties are communicating, but not at the same time. A good everyday example of this would be email, where one party sends a message that the other party reads at a later time and then responds to. This means that the parties at either end of the system do not have to be present at the same time. For telemedicine, asynchronous communication often takes the form of a store-and-forward system, where medical data are captured by one party and transferred to the other to be analysed at a later time. Teleradiology [103] and teledermatology [61] are both examples of this type of system.

2.3.4 Chronic Treatment

Telecare for the treatment of chronic conditions focuses on improving the patients's quality of care and on reducing the number of hospital admissions due to relapse or worsening of the patient's condition. It has also been reported that people receiving telecare have a higher compliance with their treatment programme than those without telecare [56]. The telecare systems for chronic care often incorporate telemonitoring to gather data that can be used by the patient's clinician (or by the patient themself) to better inform their care. These data generally fall into one of three categories: medical, behavioural or contextual.

The medical data gathered in a such a telemonitoring system is directly related to the patient's illness. For example, in diabetes it is the blood glucose level that is important [71], whereas in COPD measures such as spirometry (a measurement of lung function) are used [18]. The data will be gathered with the appropriate medical device and it is likely, depending on the condition being monitored, that several different medical parameters will be measured. The data gathering can be patient-led, where it is the patient who takes medical readings and transfers them to the clinician, or carer-led, where a nurse or other suitably trained carer makes a visit to the patient to record and transfer medical data. The transmission of data from home visits for COPD care [25] is a good example of carer-led data gathering, whereas the self-measurement and transmission of blood glucose levels by diabetic people [35] is a good example of patient-led data gathering.

The behavioural data gathered in a telemonitoring system is usually related to the patient's physical activity and behaviour. The limitation of physical activity in patients with asthma [29] and the gait and mobility in patients with COPD [64] are good examples of behavioural data. The data can be captured through movement sensors, such as accelerometers or tilt-switches, or through a questionnaire. If a questionnaire is used, very little, if any, processing is required to make the data useful. However, if an accelerometer is used for example, the raw data must be processed to extract the desired features, such as orientation, mobility and energy expenditure [64]. Accelerometers are considered in more detail in Section 2.3.6.

The contextual data gathered in a telemonitoring system consists of all the supplementary information that is gathered by the system, such as a patient's medical history and logs of calls to call-centres [25]. Contextual information is not necessarily directly relevant to the condition being monitored but serves to provide additional background information to the patient's carers which can be useful in the long term care of the patient.

A. Data Use

The data collected in chronic disease management systems is primarily used by the clinician to manage the patient's care programme [18, 56] and may be made available to nurses who are involved in the patient's care [25, 71]. In addition to the clinician and nurses, the data can also be made available to the patient so that they can review their condition and progress. This is particularly important in diabetic control, for example, where the person needs to know their blood glucose levels so that they can keep them to within the range set by their clinician [35].

Since the data collected in most chronic disease management systems are simple medical measurements and used a direct feedback to inform the care of the person there is very little analysis or further processing performed on the data, other than forwarding it to the interested parties. However some telecare systems do go further than data forwarding and include some level of data analysis. One example of very basic data analysis is the generation of alerts for important events, such as having a particularly high or low blood pressure measurement for hypertension control [56], or not adhering to the self testing schedule for asthma monitoring [29].

2.3.5 Health and Wellbeing

Telecare for health and wellbeing focuses on the maintenance of a good state of health and wellbeing in a person. Various facets of the person's behaviour and status are monitored so that changes in these can be acted upon to improve their condition. Such systems are used in a wide variety of settings, to monitor the elderly for example, so as to detect deteriorating health conditions and act upon them, or to monitor the amount of exercise a person is doing, as an aide to a healthy lifestyle. The care of people with dementia is also a strong driving force behind these types of system. Whilst also being a chronic illness dementia is strongly associated with ageing and as such dementia care systems are discussed in this section.

Regardless of the application, a telecare system in this sort of context will aim to get an overall view of the user's activity and behavioural patterns, rather than concentrating on monitoring of specific symptoms, although this can be a feature.

In order to gather the required data the user's activity is monitored in a variety of ways. The approaches taken use a variety of different types of sensor and vary in the number of sensors used. Some approaches look at monitoring a single device, such as the television [75], to give an indication of the user's behaviour patterns whilst other approaches use a large variety and number of sensors to monitor the user's activity and behaviour in the home in a detailed manner [107].

The simplest approach is the use of a single device to give an idea of the user's activity and behaviour. The device may monitor the user directly or may monitor some aspect of their environment. Accelerometers have been used to monitor the elderly —a wrist worn unit which incorporated an alarm button was used to measure the user's movement and allow the user to call for assistance if need, in the case of a fall for example [49]. For monitoring of the environment television (TV) usage has been used to gauge the user's behaviour patterns [75]. The TV fits in with the user's behaviour patterns as they tend to watch it at roughly the same times each day. If the user's health status changes, their behaviour changes also and with that their TV usage, which becomes more random. By analysing the amount of randomness in the user's TV usage, their health condition can be estimated.

The approach of monitoring the environment can be scaled up to include more sensors. These can be used to monitor other electrical devices [73, 74] or can be used to monitor the behaviour of the user more directly by employing sensors such as magnetic door switches and PIR detectors [23, 55, 107]. These types of system can include sensors specific to the situation, the use of fall detection sensors [23, 78, 17] for the elderly for example. The addition of multiple sensors to a system improves its ability to detect behaviour patterns but will also increase the amount of data to be transmitted as well as the computational complexity of the data processing. Furthermore, the increase in system complexity, by adding sensors or providing more powerful processing capability, comes with an associated cost. This results in a trade-off between cost and complexity. The system cost will ultimately restrict the sensors and processing power available and these will act as constraints on the system design.

A. Data Use

The data in telecare systems for health and wellbeing is used to provide an up-to-date picture of the patient's health and wellbeing. These data can be provided directly to the patient, in the case of monitoring for healthy living, or to the patient's carer or family, in the case of monitoring the elderly [55, 73]. The data collected is often transformed into a simple measure so that it is easy to interpret for non-expert users. More innovative solutions are occasionally seen, the use of a Sony AIBO (a small robotic dog) for example [55], which was used to provide a carer with a real-time sound and video link to an elderly person's home.

Several systems use the data they acquire to generate alerts. For simple systems, with only a few sensors, a simple threshold trigger can be used. For example, with a wearable accelerometer, an alarm can be triggered if the user is inactive for a long period of time [49]. In a more complex system, with several sensors, a data-fusion approach can be taken in which data from many sensors is fused to determine if an alert needs to be raised. For example a smart fall-sensor has been developed [78], which fuses accelerometer, tilt-switch and skin-vibration information to determine if the user has fallen.

B. Dementia Care

Dementia is an umbrella term used to cover several brain disorders that all present with a increasing deterioration in brain function over time. The typical symptoms of dementia are confusion, loss of memory and problems with speech [99]. As a result of the confusion and loss of memory people with dementia often wander and become lost.

This forms one of the key requirements for dementia care systems, which have to provide an overall picture of the elderly person's behaviour as well as help with wandering and memory loss.

There are a number of approaches that have been taken with these kinds of systems. In general the approach is similar to the more generic elder care systems and uses a number of ambient sensors in the person's home to monitor their activity over time [27]. The Just Checking [88] system for example uses a set of wireless movement sensors in the home to track the location of the elder and allow users to make inferences about the elder's activity. Bluetooth based localisation in the home has also been used for this purpose [47]. The general framework of these systems can be extended with specific technology targeted at the memory loss and wandering aspects of dementia. Intelligent data analysis can be used to determine the state of the patient and telephone interventions can be used if necessary [27]. It is also possible to track the location of the elder outside of the home, using GPS for example [69].

2.3.6 Healthcare Technology

Whilst there are several telecare systems described in the literature, there are also a number of stand-alone, wearable devices that have been developed for use in a telecare environment. Telecare devices are specifically designed to measure one or two parameters, either physical, such as posture [22, 60] and acceleration [100], or physiological, such as blood glucose [35] and Electrocardiograph (ECG) [109], and as such each device is different. However, they do share a common set of characteristics that is important for any device in a telehealth system.

Primarily, any wearable device has to be small enough to be wearable and unobtrusive. Whilst this is not always possible, especially in the development stages, it is of primary concern for the end product as a device which is too obtrusive will not be used. A second but equally important concern is the power consumption of the device. If it is to be useful, a device must have a reasonably long battery life, which means a low power consumption. A low power consumption is usually accomplished through the use of components with low power needs, but can also be achieved through the use of processing algorithms that minimise the amount of computation that has to be performed and the amount of data that needs to be written to memory or transmitted [50].

A. Accelerometers

Accelerometers have been used in several clinical research fields, such as sleep assessment [39] and circadian rhythm analysis [100]. The use of accelerometers to record and monitor human activity, specifically rest-activity cycles and circadian rhythm, is known as actigraphy (the word actimetry is also used but is less common). Accelerometers can be used in single or double configurations. If only one accelerometer is being used it is often placed on the wrist [39], the sternum [22] or the leg [50]. If two accelerometers are used, they are placed on the body either side of a joint, such as placement above and below the waist; on the trunk and the thigh [60]. Placements either side of the knee are also possible.

The actigraphy data captured by the accelerometer shows the acceleration of the accelerometer over time, in one, two or three axis. The acceleration is directly linked to the speed that the accelerometer is moving at and, because the accelerometer is attached to a subject, is a measure of the movement of the subject. At coarse resolutions, over a long time-scale, it is possible to observe the circadian rhythm of the subject but for many applications the data has to be processed to extract specific information on a much finer time-scale.

A number of approaches have been taken to the processing of actigraphy data to extract different pieces of information and each approach is different. Typically, the signal will be filtered to remove background noise and gravity induced components. The frequencies that are used for filtering vary across the approaches with some examples being 0.25 Hz high pass [65], 3 Hz low pass [60], 0.5–11 Hz bandpass [100] and 2–2.5 Hz bandpass [39]. Further processing on the data also varies across the different approaches but the data are usually abstracted into various features, such as mean and standard deviation over a window [60], before being specifically processed.

B. Patient Diaries

Patient diaries are tools that have been used extensively to collect subjective information from patients. Information that cannot be collected with a monitoring device, such as experience of pain and quality of life, can be collected with a diary [97]. The subjective information gained with a diary can be used by clinicians, to inform a patient's treatment plan, and by researchers, to evaluate potential treatments. The information contained in a diary is very valuable but comes with problems of reliability and compliance.

A patient's compliance with a diary routine is a measure of how often they make accurate entries at the times they are supposed to. Someone that makes accurate diary entries at the right times has high compliance and someone that does not make accurate entries and/or makes them at the wrong time has poor compliance. With paper based diaries compliance is often very poor [102], resulting in missing entries, entries made after events or faked entries. Of the three possible forms of poor compliance, faked entries are the worst as the patient is giving a false representation of what has been happening. Diary entries made after events can also be misrepresentative of what has actually been happening as the patient's recollection may be unreliable and biased [97]. A person's recall is also affected by their mood; a negative mood at the time of recollection can cause the recall of more negative things for example [37].

In addition to the problems of compliance there are problems with data reliability. In a paper based diary, the patient is free to write down anything and may not comply with the expected entry format. This may lead to numeric answers that are out of range or textual answers that are hard to interpret, both of which lessen the value of the diary data.

As an alternative to a paper based diary approach, an electronic diary can be used instead. Electronic diaries have an advantage in that they can be augmented with features to improve compliance such as prompting for input, not accepting input outside of certain time windows and constraining input to valid ranges [97]. These measures improve compliance and also improve the quality of the data. Care must be taken with the use of an electronic diary that any prompting functions do not cause irritation to the subject. Ecological momentary assessment [96] is an assessment protocol that can be used in electronic diaries to prompt for subject input at randomly determined times within set parameters. This has the advantage of fully assessing the subject over a period of time, without the need for excessive prompting.

2.3.7 Technology for Bipolar Disorder

There is an increasing body of work that examines the use of technology in the treatment of BD. Existing systems mostly focus on automating existing treatment methodologies to make them easier to use. The ChronoRecord [11] project is a good example of this type of system. In contrast to this, newer approaches for monitoring BD use a range of different sensors to gather behavioural information. The MONitoring, treAtment and pRediCtion of bipolAr Disorder Episodes (MONARCA) project is one such example [41].

ChronoRecord is an automated mood diary system which provides patients an electronic way to fill in what would otherwise be a paper-based diary [11, 10]. Information is collected on mood, medication usage, sleep, life events and menstrual cycle [9, 8] once per day and weight information is collected once per week. The data collected by the ChronoRecord system is used for longitudinal studies of BD, such as the Maudsley Bipolar eMonitoring Project [63], but is not actively used as a self-management tool.

MONARCA is a new European project aiming to develop a multiparametric system to perform “Quantitative analysis of every-day life activities and state” for people with BD [33, 41]. The system will collect a variety of data using a mobile phone, wrist worn device, Electroencephalography (EEG) sensor and a sock sensor, which will measure galvanic skin response (GSR) and pulse rate [46]. The system will perform complex behavioural analysis and emotion recognition and feed these results back to the patient, their caregivers and doctors.

There are other projects, such as the P-cube platform [104] that are investigating the use of specific technologies to monitor people. In the P-cube architecture, the person is monitored with an accelerometer and a microphone, both of which are connected to a mobile device that serves as a gateway. Positional data are also acquired from a mobile phone. The data gathered is used to detect different activities and emotions and this information is then fed back to the person.

2.4 Contributions

Through the work that will be presented in this thesis, the author has made several contributions. These are:

- The development of the activity signature detection. The idea of extracting activity signatures in itself is a novel idea as is the application of the CPM algorithm to achieve this.
- The fusion of several time-series in an intelligent and automatically weighted way to achieve a data analysis approach that can be applied to a system independently of the sensors that are used.
- The development and integration of the camera and PIR sensors to the system.
- The data analysis for the accelerometer.
- The incorporation of the data from the PAM sensors into the high-level data analysis.

2.5 Summary

This chapter presented the background information on BD and telecare systems. BD is an illness that can exhibit significant variation in the symptoms that any particular person will exhibit, both in comparison to other people and in comparison to that person's previous episodes. Symptoms typically present in a manic episode include decreased need of sleep, increased energy and talkativeness and an elevated mood. Conversely, the symptoms typical of a depressive episode are increased sleep, lack of energy and depressed mood. It is evident from an examination of the literature on BD that behaviour and mood-state are closely linked; mood-state can effect changes in behaviour, but changes in behaviour, especially very disruptive or stressful changes, can also effect a change in mood-state. The capture of a broad range of behaviour has therefore got to be a primary consideration in the development of any self-management system and in order to successfully capture this behaviour the person must be monitored both in their home, and also in an ambulatory setting.

For the capture of behavioural data, the literature on telehealth and telemonitoring must be turned to. This reveals that there are a number of systems that look at a person's behaviour patterns as a means of assessing their health. These are mostly limited to care of the elderly in an assisted living context and dementia, and tend to focus on the home environment. Monitoring in the home can be as simple as a single sensor or expanded to use a network of sensors to give a comprehensive picture of the behaviour in the home. There are fewer systems that make use of an ambulatory component, but in those that do accelerometers are widely used as a measure of activity.

In developing a self-management system for BD, the PAM project must use both environmental and ambulatory monitoring techniques to gather the data that is needed. Attention must be paid to sensors used to ensure that the correct behaviour is monitored. Care must also be taken that the system function correctly from a technical and a usability standpoint as a system that fails in these areas will also fail to capture behaviour patterns correctly. The design requirements, sensors and system architecture used in the PAM system are discussed in Chapter 3, along with some background information on the data analysis techniques that have been used.

Chapter 3

The PAM Approach

3.1 Introduction

It is evident from the review of the literature presented in Chapter 2 that in developing a self-management system for BD a wide variety of elements must be taken into consideration. These include the various facets of BD; end-user usability concerns; hardware design and communication considerations; and algorithm and data processing requirements. This chapter presents the PAM approach to developing a BD self-management system. The project organisation is discussed briefly, followed by a discussion of the system requirements and system design. The data processing architecture that is being used is presented and followed by an overview of the data analysis techniques that have been used.

3.2 Project Organisation

The PAM project is a collaborative project between three universities: Southampton (both ISVR and the School of Management), Nottingham and Stirling. The project has been split into four general areas: sensors and hardware, communication and networking, operational modelling and data analysis. Each area is being developed by one group; Nottingham are looking at sensors and hardware, Stirling are developing the communication and networking, the School of Management at Southampton are dealing with operational modelling for the project and ISVR, including the author, are investigating the data analysis.

The development of the system has been partially a collaborative affair and partially an individual one. The entire team collaborated on developing the requirements for the PAM system, the high-level systems design and developing the list of sensors to be used in the system. Once this preliminary work was accomplished, the team diverged to work on their own particular areas.

The team at the University of Stirling developed the communication infrastructure and rules engine for the PAM system and, in particular, the software that ran on the mobile phone [15, 16]. The Stirling team were also responsible for the back-end data base that was used in the technical trial.

The team at the University of Nottingham developed the majority of the sensors for the PAM system, including the wearable node and the environmental node. A range of other sensors, including door switches, pressure mats, light and sound were also developed by the Nottingham team and either integrated directly into the environmental or wearable node, or made to interface with the PAM system [90]. The communication links between the mobile phone and the wearable unit were developed between the Nottingham and Stirling teams.

The work that the author took responsibility for, and accomplished, focused primarily on the development of the camera and PIR sensors; the data processing chain for the camera, PIR and accelerometer; and the high-level data analysis. The camera sensor development included identification of suitable hardware, design and identification of a suitable image processing algorithm and the implementation of these things for the technical trial. The work for the PIR sensor included the selection of appropriate hardware, development of the mid-level data processing and the implementation in the technical trial. The development work for the accelerometer included an initial data gathering trial to determine suitable processing methodology and the application of this methodology to the data gathered in the technical trial. The high-level data analysis was performed on the data gathered from the technical trial and takes the output from the sensors and uses this to detect activity signatures and differences in behavioural patterns.

3.3 System Requirements

The overall goal of the PAM project is to develop a self-management system that can identify a person's normal behaviour patterns, identify when they change and predict the person's mood state based on those changes. Because the behaviour of a person with BD can be affected by their mood state and, in some circumstances, also effect an affective episode a broad a spectrum of behaviour as possible needs to be captured. It is important that the self-management system be designed with this goal in mind but that other considerations that are common to telehealth systems, such as power consumption, unobtrusiveness and usability, are kept in mind. The requirements for the system break down into four categories. The first three of these relate to the PAM system in general and are grouped in this section. The final requirements set relates directly to the work the author has performed on the data analysis for the PAM system and is presented in the following section. It is also prudent to keep in mind that currently the PAM system is in a prototype stage and that there are some design considerations that differ between an end-product solution and a prototype solution and these are also discussed.

3.3.1 Monitoring for Bipolar Disorder

The PAM system is aimed at people with BD and must be designed with this principally in mind. It is the aim of the PAM system to capture behaviour patterns, which necessitates a broad variety of sensors to capture the correct data. This means data capture in both home and ambulatory settings is needed. In addition, the time granularity of the captured data must be fine enough that small changes in behaviour can be captured whilst not overloading the system's capacity to process data.

It is also important that the system be modular in design. Since BD is a very variable illness, as discussed in Chapter 2, any self-management system must be able to cope with this variability. A modular system in which sensors can be swapped in and out depending on exactly what needs to be monitored is an effective way to achieve this. The requirements on the PAM system stemming from monitoring for BD are:

1. The system must capture a broad variety of behaviour.
2. The system must be able to capture data in both home and ambulatory settings.
3. Data must be captured with a sufficiently fine time granularity that patterns of activity are visible over the course of a day.
4. The system must be modular and able to target different aspects of behaviour.

3.3.2 Usability Requirements

For any self-management system to be useful, and more importantly used, attention must be paid to the end-user and their requirements. Any system designed must be unobtrusive and require minimal user input (this can range from uploading information to changing batteries) as the less the user has to do, the easier the system is to use; and the easier it is to use, the more likely it is to *be* used. It is also of paramount concern that the user not be frustrated or annoyed by the system as this can lead to the system being abandoned by the user. The usability requirements for the PAM system are:

1. The system must be as unobtrusive as possible.
2. The battery life of any device must be sufficiently long that the user does not have to replace batteries or charge the device excessively often.

3. The user-interface must be simple and easy to use.
4. The user must be minimally involved with maintaining the system.
5. The system must be modular and configurable to the needs of the individual user.

3.3.3 Hardware and Communication Requirements

For the PAM system to be able to perform the tasks required it must have the correct hardware and communications infrastructure. A framework must be provided for sensors to be connected together and for data to be passed around and processed. The hardware and communications requirements are:

1. The system must provide appropriate sensors to capture a person's behaviour patterns.
2. The system must provide an appropriate communications infrastructure.
3. The system must be able to cope with any ambulatory components dropping in and out of communication range.
4. The system must provide sufficient high-level computing power that data can be processed to extract behavioural patterns.
5. The system should allow for data to be stored long-term for further analysis and academic interest.

3.3.4 End-Product Design Requirements

In addition to the requirements listed above there are a number of additional considerations that must be kept in mind for an end-product solution. These are focused on making a product saleable and suitable for the purpose for which it is designed. At this stage of the project an end-product is not being developed, but these considerations are worth keeping in mind. The end-product requirements are:

1. The system must be able to process all the data in real time.
2. The system must provide some way of generating alerts based on predictions of a person's mental state.
3. The system should be expandable.

3.4 Data Processing Requirements

The PAM system will generate a large volume of data which will need to be processed in a sensible and efficient manner in order to satisfy the goals of the project. For a prototype system this processing can be done in a mostly off-line way, but for an end-product solution this must all be performed in real-time. There is also a need for data to be pre-processed to reduce the data communication volume, which requires some low-level processing capability. The data processing requirements for the PAM system are directly related to the author's work, and to the research questions posed for this thesis in Chapter 1, and serve to drive a number of the decisions that have been made. The data processing requirements for the system are:

1. The data processing algorithms must be able to handle a modular system and allow for different configurations of sensors.
2. The low-level data processing algorithms must be implementable on the low-level system components.
3. The high-level algorithms must be able to cope with missing sections of data.
4. There should be ideally zero, or at worst absolute minimal, user involvement with the data processing.
5. The data processing should enable activity signatures and behavioural patterns to be detected.
6. The data processing should enable changes in behavioural patterns to be detected.
7. Both the high-level and intermediate levels of data processing should be easily communicable to the users.

3.5 System Design

In order to capture a person's behaviour patterns and to meet the system requirements discussed above the PAM system is abstracted as a set of sensors, nodes and smart-nodes to gather behavioural data and perform the required data-processing and storage tasks[1].

The broad approach of the PAM system is to have two smart-nodes as controllers for the system, one for the home setting and one for the ambulatory setting. These smart-nodes can be connected to nodes, or directly to sensor units, and node units can be connected

[1] The use of the word abstracted is intended to mean that the PAM system can be described in terms of sensors, nodes and smart-nodes, but that the actual implementation is slightly more involved. Several of the units in the actual system perform two, sometimes arguably three, of the conceptual roles. It is easiest however to describe the system in abstract terms.

to sensors. The PAM system also employs a remote server that is used as a long-term data-sink. In this way the PAM system builds up a sensing and data-processing network that fulfills the system requirements.

The system architecture for the PAM project is shown in Figure 3.1. The components outlined above are used to create two sub-systems. One for ambulatory monitoring and one for environmental monitoring. For the environmental sub-system, the smart-node takes the form of a laptop or PC and there is some flexibility in the choice of component used here. The smart-node for the ambulatory sub-system is restricted to being a smart-phone.

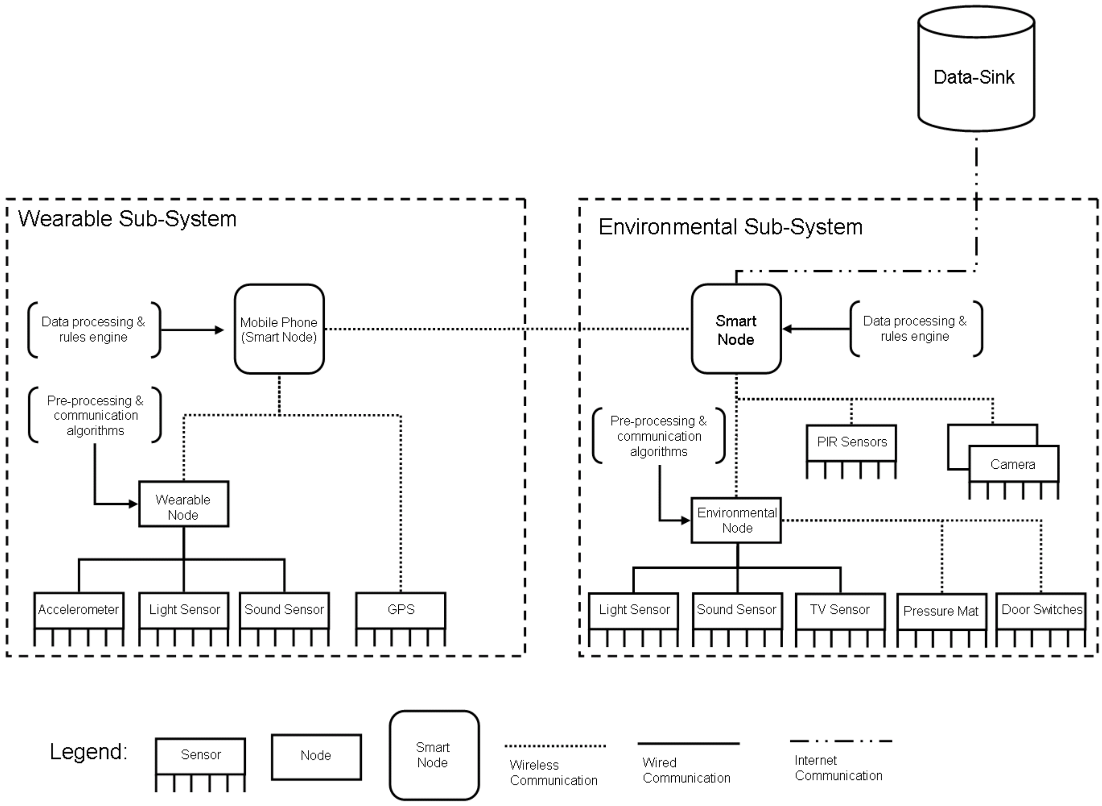


FIGURE 3.1: System architecture for the PAM System showing the sensors, nodes, smart-nodes and the communication links between them. The system is divided into a wearable sub-system and an environmental sub-system, which are each controlled by a smart-node..

3.5.1 Sensors

Sensor units are the most basic unit in the system. A sensor is a single unit which captures a single type of data. Examples of sensor units include light sensors and sound sensors. Each sensor has to connect to either a node, or a smart node. A sensor node can have some form of data processing capability, in either hardware of firmware, but this capability is limited to what is provided with the sensor to do the very basic pre-processing. The PAM project does not alter any of this data processing and it is therefore transparent for all practical purposes.

The sensors that have been used attempt to cover a wide variety of behavioural aspects that are likely to have a consistent pattern when the user is in remission but that could change when the user starts to have an affective episode. A person's behaviour consists largely of where they are and what they are physically doing at any given time and the sensors that have been chosen attempt to capture both aspects of behaviour[2]. The list shown in Table 3.1 shows the sensor set that has been used in, or proposed for use in, the PAM system. A photo of the PAM kit that was used in the technical trial is shown in Figure 3.2.

Table 3.1: The sensor set developed for the PAM system showing the sensors, the system component that they belong to and the data captured for each sensor.

Sensor	Sub-System	Data Captured
Accelerometer	Wearable	Physical activity level
GPS	Wearable	Location
Bluetooth module	Wearable	Encounters with other Bluetooth devices
Light sensor	Both	Natural and artificial light levels
Sound sensor	Both	Ambient noise levels
PIR sensor	Environmental	Room occupation
Door switch	Environmental	Opening and closing of doors
Wireless camera	Environmental	Activity in complex scenes
TV IR Sensor	Environmental	TV remote control usage
Pressure Mat	Environmental	Presence of person on mat
Patient diaries	Wearable	Capture subjective information

Accelerometers have been used successfully in various settings for recording activity and are used as part of the wearable node to capture information on the physical activity of the user. Detecting physical activity is one of the key areas in the PAM project and a change in physical activity patterns would likely be a component of a behavioural pattern change. Work is presented in Chapter 4 that investigates the use of accelerometry for the PAM project.

The GPS is used in the PAM system to gather data on the user's location (the 'where they are' side of behaviour) outside of the home environment. It is likely that the user visits the same locations on a regular basis (home, work, the supermarket etc.) and that

[2]It is important to note that behaviour is very time dependent; something normal at 0900 could be very abnormal at 0300. When behavioural patterns are discussed in this work, a time correlation is implicitly implied. This applies most often to the comparison of a 'normal' behaviour with a different 'abnormal' one. When a particular aspect of behaviour is described, without an explicit time context, as 'normal' and a corresponding 'abnormal' behaviour is given, it is assumed that the two examples are for the same time of day and that time of day is largely unimportant for the given example.

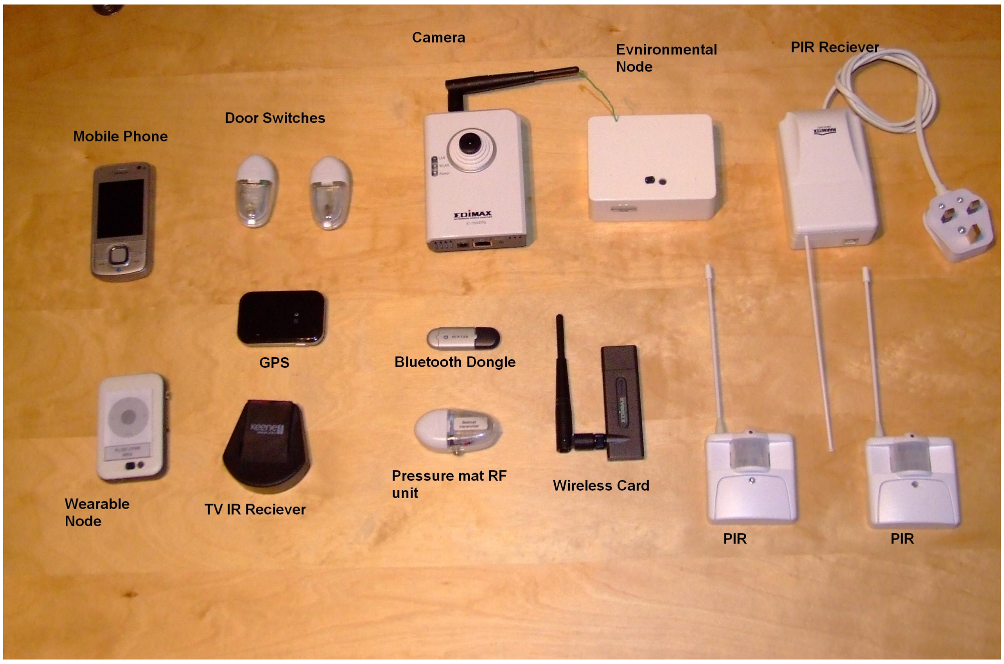


FIGURE 3.2: The sensors for the PAM self-management system that were used in the technical trial. Both sub-systems are shown; the wearable sub-system, comprising the phone, wearable node and GPS, and the environmental sub-system, comprising everything else.

a deviation from the patterns of location, especially events that were not just one-off, would be present if their behaviour changed.

The Bluetooth module on the mobile phone, as well as providing some of the communications infrastructure, can also provide information on other Bluetooth enabled devices that have been encountered. Bluetooth encounters have been included as an experimental idea, with the reasoning being that a person is likely to encounter some of the same devices on a fairly regular basis and that a deviation from this indicates that something has changed, although the other sensors would have to be used to provide context for this change. An example of this might be that the user's phone encounters the blue tooth IDs *Alice's_phone* and *Bob's_phone* every day. If the user's phone stops encountering *Alice's_phone* and *Bob's_phone* and instead encounters *Eve's_phone* for several days then this could indicate that something has changed.

Light sensors and sound level sensors are used to detect ambient conditions in both subsystems of the PAM system. It is reasonable to expect a user to be exposed to similar ambient conditions on a day-to-day basis, especially for natural light (seasonal variations notwithstanding) and being able to detect any change in these, such as might be caused by a change in the time of day the user opens the curtains, is a useful capability to have in the system.

PIR sensors are used to determine the room occupation (the presence or not of a person in a room) in a user's home. The effectiveness of PIR sensors for this task is heavily dependent on the number of occupants in the home. For a single occupant, their location can easily be derived if enough sensors are used [79]. For multiple occupants however, the PIR sensors will not be sufficient to localise the patient but instead will only be able to determine if a room is in use.

A wireless camera is used to monitor activity in the kitchen. Since the kitchen is a complex environment a single camera can be used to monitor a range of things that would otherwise require a number of individual sensors. Change detection is performed on successive images to determine activity in the scene and areas of interest (AOI), such as cupboards, the cooker and the fridge, are identified and where activity in these areas is detected, an event is recorded. Work is presented in Chapter 5 that covers the use of the camera in the PAM system.

Magnetic contact switches are used for the detection of door opening and closing. The cupboard doors in the kitchen are the primary target for these sensors, although they could be applied to any door. These sensors are particularly useful in the kitchen as they can cover some of the blind-spots that the camera may have.

The Television IR sensor is used to detect the usage of the user's IR based remote controllers and will in fact respond to any IR controller, not solely those for the television. It is included in the system in the event that a user has a very rigid pattern of device usage, which could deteriorate during an affective episode. The pressure mat is included for largely the same reason. It is not intended to target any specific activity, but can be used in a variety of areas where it could pick up a consistent behavioural pattern which may be subject to change during an affective episode.

3.5.2 Nodes

Node units are the second unit in the PAM system and are used as controllers for sensors. A node can be connected to one or more sensors and is also connected to a smart-node. The node unit is responsible for handling and communicating the data from each of the sensors it is connected to. A node has a fair amount of processing power which is used to perform the data control tasks specific to that node. In addition the node may be required to perform some data-processing tasks prior to the data being communicated to the smart-node.

The PAM system makes use of three node units, two custom designed and one off-the-shelf. The off-the-shelf node is the camera, which is shown on the system diagram as a merged sensor-node unit. This is because the camera has many of the functions of the nodes in the system (data processing and communication) but is inextricably linked to the camera sensor.

The two custom designed node units in the system, one each for the environmental and wearable sub-systems, are used to control and handle data from a number of sensors, as shown in the system diagram. The node units are designed and built by the PAM team at the University of Nottingham. The wearable node contains an on-board accelerometer, light sensor and sound sensor. The environmental node contains an on-board light sensor and sound sensor, has a wired connection to the IR sensor and wireless connections to the pressure mat and door switches.

3.5.3 Smart-Nodes

Smart-nodes are the third component in the PAM system and are used as the main data-processing, control, storage and communication units. The PAM system employs two smart-nodes. One to control the wearable sub-system and one to control the environmental sub-system. Each smart-node has considerable processing power, which is used to perform high-level data-processing tasks. The wearable smart-node is a mobile phone and configured so that the it can communicate with the environmental smart-node when it is in communication range so that data can be transferred. The environmental smart-node can be either a PC or a laptop and has access to the internet so that it can communicate to the data-sink.

3.6 Data Processing Architecture

In order to process the large volumes of data that the PAM system will generate a multi-layer architecture has been developed, as shown in Figure 3.3. The pre-processing level is implemented at the node level (or smart-node if a sensor connects directly to a smart-node) and will perform feature extraction to transform the raw data from a sensor into a set of characterizing features. The primary advantage in transforming the raw data into features is that less data needs to be transmitted, which helps prolong the battery life of the sensor platform as there is less load placed on the power intensive wireless communication modules. It is also easier, possibly necessary, to perform higher level processing on features rather than on raw data. The sensor level processing may be omitted for sensors that produce very simplistic data, such as the PIR.

The time-series generation layer is intended to process the features passed to it and convert the data into a time-series that can be used to convey some basic meaning. In the case of the accelerometry data, the features might include measures of amplitude and frequency and time-series would show the activity score over time. Each time-series contains enough information that it is understandable in isolation and represents the lowest level of detail where data coming from the sensors can be understood easily.

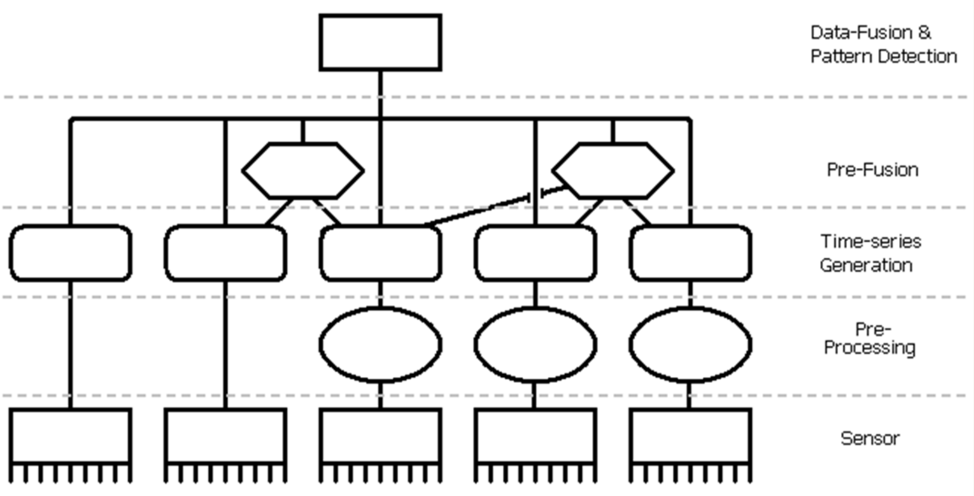


FIGURE 3.3: Data processing architecture used in the PAM system showing the sensor, pre-processing, time-series generation, pre-fusion and data fusion layers.

The pre-fusion layer is intended to allow the combination of processed sensor data to allow for multi-sensor behaviours and interactions to be delivered to the fusion layer. For example, the combination of PIR, light sensor and pressure mat could be used to infer that a person was asleep; the pressure mat placed in the door-way to the bedroom is triggered, followed by some PIR activity and light readings. Subsequently if the PIR and light readings cease but the pressure mat is not triggered again it can be inferred that the person is remaining still in a room with no lights; and sleep is a logical conclusion. This has not been implemented to-date, but is a source of possible expansion in the system.

The data fusion and pattern detection layer is brings together all of the time-series components and provide the behavioural pattern analysis capabilities of the system. This layer makes use of time-series alignment algorithms to detect the behavioural patterns. New data can then be compared against these patterns to detect changes in behaviour. Work is presented in Chapter 6 that covers the behavioural detection in the system.

3.7 Integration with Current Treatment of BD

The system envisaged for the PAM project provides several functions that enable it to be integrated with a user's treatment plan. The primary function of the system is to act as a self-management system to detect behavioural changes and to identify the onset of an affective episode. If the system can identify the onset of an episode before the user does, the user can use this forewarning to take corrective action to lessen the severity of the episode. Further to the detection of episode onsets, the user's clinician or carer can be informed, if the user so desires, so that they can take the appropriate action to provide care for the user.

The long term data storage of the system would also allow the system to fit in with a self-monitoring and control process, whereby the user monitors their own behaviour to ensure that they do not do things that could trigger an affective episode. The system could be set up to automatically check certain conditions, such as time spent asleep, and provide warnings if the user's behaviour deviated from their normal patterns by a specified amount.

3.8 Technical Trial

As part of the development of the PAM self-management system, a technical trial was carried out in order to test the hardware and software components of the system. Three complete sets of equipment were installed in the homes of three of the PhD students on the project. The systems were installed and left to run for three months, with bug-fixes and system maintenance being applied as necessary. This work is presented in Chapter 6.

3.9 Data Analysis Background

The PAM system produces a large volume of complex data, where patterns may be difficult to determine and where the data may be multivariate. It is therefore essential that these data be processed in a sensible and efficient manner. There is a large volume of work in the field of complex data analysis and the PAM project draws on a number of different techniques to solve a range of different problems. The principal problems that can be solved through the use of these techniques, such as regression, optimisation and dimension reduction are discussed and this is followed by a brief, non-mathematical overview of the techniques themselves. The detailed technical and mathematical details of the data analysis techniques that have been used are presented as they are used in Chapters 4, 5 and 6.

3.9.1 Complex Data

Data from the PAM system is drawn from several different sources, be they separate sensors or different input channels from the same sensor. This leads to very different types of data that need to be processed.

The standard data format for many analysis approaches is to represent the data as a vector data-structure where a data-structure, $\mathbf{x}$, with n variables, or features, is represented as the vector $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$. Many variables can be directly represented in this manner, individual measurements from instrumentation for example. However, when dealing with some more complex forms of data, such as time-series measurements, is often necessary for the raw data to be pre-processed and for features to be extracted and placed into an appropriate data structure prior to any further analysis.

3.9.2 Feature Extraction

Feature extraction is the process of extracting features from raw data to facilitate further analysis and to reduce the volume of the data. A distinction must be made between feature extraction, which aims to extract information from raw data, and feature selection, which aims to reduce the number of features in a data-structure and is discussed below. The goal of feature extraction is to extract meaningful information from potentially noisy data and this is used in the initial processing stages for both the accelerometer data and the camera data. Extracted information can range from simple statistical measures, such as mean and standard deviation, to more complicated information such as a measure of change between two images. There are two particular areas that could be considered as feature extraction and are used as such in the PAM system; signal processing and image processing.

A. Signal Processing

Signal processing is the processing of a signal to extract some desired piece or pieces of information, or to perform some operation on the signal. A signal is formally defined as being “a physical quantity that varies with time or space” [89].

In the real world signals are analogue; that is they exist in continuous time. In moving into the digital world, such analogue signals have to be sampled and digitised so that they are only defined at discrete times. This enables them to be manipulated digitally.

The overall goal of signal processing is to extract the desired information from the signal or to perform some operation on the signal to change it in some way. Examples of operations that could be performed include filtering the signal to remove a specific frequency or to boost the power of the signal at certain frequencies. Information to extract could include as examples, information about the dominant frequency in the signal or an estimate of the power in the signal over a specified time window.

There exist a number of methods for the manipulation of signals, such as the z-transform, Fourier-transform and digital filters. In particular, the Fourier-transform is one of the principal methods for identifying the frequency components of the signal and transforms the signal from the time domain into the frequency domain and is used in the accelerometer feature extraction. A good overview of digital signal processing techniques can be found in [42] and [89].

B. Image Processing

Image processing is the processing of an image, or sequence of images, in order to extract some required information from the image or to perform some operation on the image. An image is a collection of pixels and operations on the image can affect the entire image, or be restricted to a specific group of pixels.

Image processing tasks range from the very simple, such as enlargement and rotation, through to the more involved, such as interpolation and edge detection, to the complex, such as image segmentation and depth perception from image pairs. Information to be extracted through image processing could include locations of changes in image sequences, or information about the colour components in the image.

A number of methods that can be applied to signals, such as interpolation, Fourier transforming and a number of different filtering operations can be applied to images. A good overview of image processing techniques can be found in [76].

The principal use of image processing in the PAM system is in image change detection on the images produced by the camera. Image change detection is the task of detecting differences between two images, or between an image and a background estimate for the scene. It is often desirable that these changes be the result of a physical change in the scene, rather than changes caused by differing lighting conditions. Since a camera will produce images with different pixel values under different lighting conditions a simple pixel-by-pixel differencing will not produce good results. This leads to the use of more complicated algorithms and Chapter 5 presents the work that has been undertaken on the selection of an appropriate image change detection algorithm.

3.9.3 Data Analysis Tasks

There are five major tasks that can be accomplished with complex data analysis methods: classification, regression (also termed function estimation), clustering, optimisation and dimension reduction.

Classification is the task of assigning a discrete classification, or class label, to a data-structure. The algorithm is typically given a set of pre-labelled data-structures and is trained to recognise the implicit link between the data-structure and the class label. Once the algorithm has been trained, new data-structures can be fed in and class labels assigned [12, 14, 19]. The task of identifying behavioural changes could be formed as a classification problem with the task being to classify data into one of two categories; ‘normal’ or ‘not normal’.

Regression (also termed **function estimation**) is very similar task to classification but with an output that is continuous, rather than discrete. Regression also makes use of known input-output pairs and the algorithm in this case is trained to predict the output based on the input data-structure [12, 19]. In the PAM system, the identification of behavioural pattern for a particular data-stream is a function estimation problem where the input is the collection of time-series obtained for the data-stream and the desired output is the underlying pattern, which can be viewed as a function of time.

Clustering is the task of identifying similarities in the data, with no other knowledge of the desired output. The data are separated into n clusters where all the data-structures in a particular cluster are more similar to each other than they are to the data-structures not in the cluster. Clustering can be absolute, with each data-structure being a member of only one cluster, or fuzzy, with data-structures having membership values for more than one cluster [12, 19]. A clustering approach is used in the investigation of the accelerometry data in Chapter 4.

Optimisation is the task of finding the global maximum or minimum of an objective function, subject to a set of constraints. The aim of an optimisation algorithm is to find the maxima or minima in a faster time than using a brute force approach (exhaustively trying each combination) [12]. Many different techniques result in some form of optimisation problem and there are several of these throughout the PAM system data analysis.

Dimension reduction (also termed **feature selection**) is the task of reducing the dimensionality of the data-structure, whilst preserving as much differential information as possible. This is often done to reduce the volume of data that needs to be transmitted, or to make the data suitable for visualisation [19]. Dimension reduction is used in the PAM system as part of the accelerometry data processing chain.

3.9.4 Data Analysis Methods

There are a great many methods for analysing complex data and they draw their inspiration from a number of different fields, including machine learning, statistics and computational intelligence. The data analysis performed for the PAM system makes use of several of these techniques and a brief overview of these is presented here, whilst the technical and mathematical detail is presented in the relevant chapters.

A. Machine-Learning

Whilst not a specific method in and of itself, machine-learning (sometimes referred to as computational intelligence) is a term applied to any technique that allows an algorithm to 'learn' and to improve its performance in a given task based on its experience. A formal definition [70] is that:

> "A computer program is said to **learn** from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

Machine learning techniques can be further categorised into supervised and unsupervised learning, where supervised learning algorithms make use of known input-output pairs in the training period. Unsupervised algorithms do not make use of such known information and are instead completely data driven.

B. Artificial Neural Networks

Artificial Neural Networks (ANNs) are a algorithmic tool that uses a network of simple artificial neurons to approximate a real-world function and are used in the PAM project data analysis as part of the Neuroscale algorithm for the analysis of the accelerometry data in Chapter 4. An ANN is typically composed of layers of artificial neurones in which each neurone is connected to all of the neurones in the next level, as shown in Figure 3.4. Each neurone can be made to perform some function on its inputs and the weights on the connections between the neurones can be weighted to alter the outputs of the neurones.

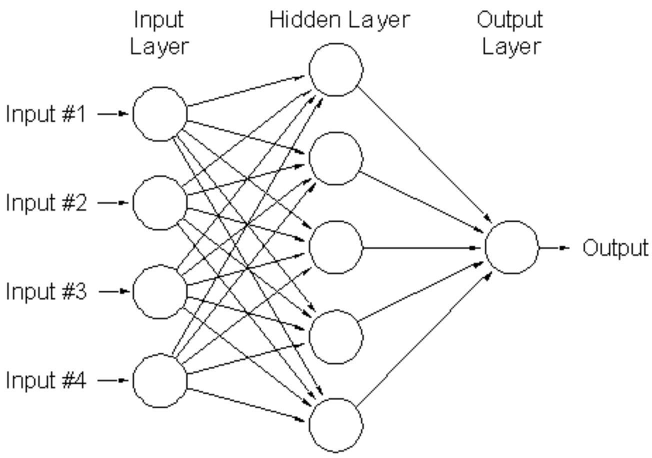


FIGURE 3.4: Diagram of a neural network showing an input layer, a hidden layer and an output layer

ANNs can be adapted by altering the weighting on the connections and the function parameters of the neurones and it is by doing this adaptively that an ANN can learn to approximate a function. In order for the network to learn a cost function has to be defined. For supervised learning, the cost function will be based on a difference between the network's output and the desired output. For unsupervised learning, the cost function will be dependant on the function that is being approximated.

Once the cost function is selected, the network is trained by systematically altering the weightings and function parameters based on the outcome of the cost function. This corresponds to a straight-forward optimisation process to minimise the cost function by varying the network parameters and there are many algorithms that can be employed for this purpose. ANNs have been applied successfully to a number of areas including speech recognition and interpreting visual scenes [70, 83].

C. Hidden Markov Models

Hidden Markov Models (HMMs) are a statistical modelling tool that can be used to model a Markov process where there is a hidden process that has given rise to an observed set of results. A Markov process is one in which the next state of the system is dependent only on the current state of the system. To model the underlying process a series of connected states is used, where each state is connected to one or more other states and there are probabilistic chances to transition from one state to the next. In each time step, an observable output is emitted, according to the emission probability of the possible outputs from the state, and a state transition occurs, according to the transition probabilities. The emitted observable output from the model corresponds to the real-world measurable results and the interior state of the model corresponds to the underlying real-world process that is being modelled [14].

There are three problems that can be posed of an HMM [92]. Firstly the model can be run, starting in an initial state and generating the emitted observations and state transitions, according to the relevant probabilities. Secondly, and perhaps more usefully, a sequence of observed results can be taken and the probability of the observation sequence given the model and the best (normally the most probable) sequence of states that produced the observation sequence can be calculated. Thirdly, in the situation where the model structure is known, the parameters are not known and there is at least one sequence of observed values, the parameters that maximise the probability of generating the observation sequence(s) given the parameters can be calculated. HMMs have been successfully applied to problems such as speech recognition and DNA modelling [12] and are used in the PAM system data analysis in the identification of behavioural patterns from the technical trial in Chapter 6.

D. Clustering

Clustering algorithms are used to analyse the clustering of data-points in the data and are used in the analysis of the accelerometry data in Chapter 4. A cluster is a set of data-points that are more similar to each other than they are to the rest of the data set. A clustering algorithm will, when given the number of clusters in advance, determine the location and data-point memberships for the clusters in the data space. This can be done in a hard partition, where each data-point is a member of at most one cluster, or in a fuzzy partition, where a data-point can be a member of more than one cluster.

Cluster determination is a data driven approach and does not rely on labelled training data [19].

E. Principal Component Analysis

Principal Component Analysis (PCA) [45, 98] is a transformation of the data from its raw feature space onto a space defined by a set of orthogonal combinations of those original features, called principal components. The principal components are ordered so that the maximum variability in the data is accounted for by the first principal component and the next, orthogonal variability accounted for by the second principal component, and so on. This ordering means that PCA can be used as a dimensionality reduction tool and is used in the PAM system as part of the processing of the accelerometry data.

F. Expectation Maximisation

The Expectation Maximisation (EM) algorithm [24] is a method for determining the maximum likelihood estimate of a set of parameters in a statistical model. It is an iterative method and consists of an expectation stage followed by a maximization stage. In the expectation stage, the expected value of the log-likelihood function is evaluated based on the current estimates for the parameters. In the maximization stage, the value of the expectation is maximised by updating the parameters. The two steps are iterated until convergence occurs. EM is used in the PAM system as part of the optimisation for the HMM parameters in the technical trial data analysis in Chapter 6.

3.9.5 Image Change Detection

Change detection is the process of comparing two images and identifying the areas in the two images that are different [93]. More formally, the goal is to produce a change mask H of binary labels $h(p)$ for each pixel p. The change mask shows the differences between a reference image, I_r and the current image, I_c, where $h(p) = 1$ indicates that there has been a change in that pixel, and $h(p) = 0$ indicates no change. It is also desirable, in most situations, that the changes detected be insensitive to changes in the lighting conditions and only relate to changes in the physical scene.

Change detection techniques are normally applied to two images of the same physical scene. In most situations this will mean that the camera has not moved while the two images are taken. The time between the two images, or frame rate, is of less importance and a lot of techniques are applicable regardless as to the frame rate. However, there are some techniques that are used to analyse video sequences that require a high frame rate in order to work.

There are several different things that can be done with image change detection techniques such as detecting the differences in two images; analysing video for motion and anomoly detection; and detecting moving objects in the foreground. The image change detection in the PAM system focuses on detecting the user and their activity and therefore is concerned with the detection of moving foreground components in an image sequence. There are two components in this task, estimating the background and detecting changes between the background and the current image.

A. Background estimation

The task of background estimation is to identify the background in a series of images in which the background may not be wholly visible in any single image. Additionally, the background may change over time, particularly with respect to lighting changes and objects entering and remaining in the scene. There are several approaches to estimating the background ranging from simple averages to complex statistical models.

One of the easiest ways to extract a background reference is through a temporal median over several images [76]. For each pixel, the values from the image set are median filtered to get the background estimate. This has the effect of removing any intruding objects from the scene to produce the background estimate. There are a couple of inherent and related problems with this form of background estimation in that any object that is still enough for long enough will eventually become part of the background due to the median filtering. This can lead to objects that need to be tracked disappearing and not being picked up by the change mask. Increasing the length of the median filter can compensate for this problem. However, increasing the length of the temporal median introduces a similar problem where if an object that was considered to be part of the background is removed, it will take some time for it to be removed from the background estimate. This can lead to ghosting, where the object will be identified in a change mask despite being no longer present. The remedy for this is to decrease the length of the temporal median filter. In reality, the length of the median filter is best set depending on the characteristics of the scene to be observed and the desired operating characteristics of the algorithm.

A slightly more involved technique is to examine a histogram for the red, green and blue components in the RGB colour-space for each pixel over many frames. The background values for each pixel correspond to the highest values from each histogram [110]. This technique can also be applied to a grayscale colour-space. The limitations of this technique are similar to those of the temporal median. Furthermore, neither this approach nor the temporal median cope very well with multi-modal backgrounds where two different values for a pixel should both be considered background. This typically occurs when the background is changing more rapidly than the algorithm is updating.

More complicated techniques extend the histogram approach by using a mixture of Gaussians to model the histogram for each pixel [101]. These types of models no longer provide a background image as such, but provide a way to directly determine if a pixel is in the background or not. In the case of a single Gaussian, the mean and standard deviation can be updated each frame based on new pixel values. To determine the background membership for a particular pixel it can be compared to the Gaussian by looking at where the pixel value falls in relation to the mean and standard deviation. If it is far enough outside the mean then it is deemed to be foreground. This idea can be extended to multiple Gaussians, which are then able to handle multi-modal backgrounds. Other techniques that have been used include kernel density estimation, sequential kernel density approximation and eigenbackgrounds [87].

B. Detecting Changes

As stated above, the goal of change detection is to produce a change mask showing the difference between two images. There are a number of techniques that can be applied to achieve this. Simple, naive techniques such as taking the difference between two images and thresholding them provide very quick ways of detecting changes. These change masks tend to be very noisy and the application of median filters can help to reduce some of this noise. However, due to the simplicity of simple differencing these techniques are not lighting invariant. That is, a change in lighting will register in the change mask, and for a number of applications this is highly undesirable.

To deal with illumination changes it is necessary to separate the physical objects in the scene from the illumination of the scene. By doing this, changes in the physical objects can be detected independent of the lighting. This typically makes use of a shading model which states that the pixel value recorded by the camera, the pixel intensity, is the product of the illumination and the reflectance of the lighting from the surfaces of the objects in the scene. The key focus in many change detection algorithms therefore is to find a way to examine the reflectance component by removing or otherwise accounting for the illumination component [93].

Homomorphic filtering is an approach to this problem that attempts to separate the illumination and reflectance components by transforming the pixel intensities into a logarithmic space. In this space, the multiplicative effect of illumination and reflectance becomes an additive effect and the two can be separated with a low-pass filter [106].

The Wronskian change detector does not separate the two components but rather focuses on the assumption that, for any reasonably small grouping of pixels, a lighting change will produce linear scaling effect from one image to the next but that a physical change will produce a non-linear effect. The Wronskian change detector therefore compares corresponding regions in two images for linear independence in a vector representation of the region. If the two regions are linearly dependent then there has been no physical change in the scene in that region [28].

3.10 Summary

This chapter has presented the requirements and design of the PAM system, the data processing architecture and an overview of the data analysis techniques that have been used. The design of the PAM self-management system requires that several different things be taken into account. Principally these are the symptoms and pathways of BD; system function and usability; and general system design principals. The design presented here draws on the literature presented in Chapter 2 and ties these areas together to provide the design for a system that can perform the data capture and data processing tasks required.

The data processing architecture presented uses a three layer architecture to take data from the sensors, convert it into a time-series and then fuse time-series from multiple sensors to provide the behavioural pattern detection capabilities in the PAM system. There are several data analysis techniques that have been used to accomplish this and the background information on these has been presented in this chapter.

In the next two chapters, two of the sensors that require the most complicated data processing are presented. Chapter 4 covers an exploratory analysis of the data obtainable from an accelerometer and presents an investigation of the ways in which these data can be processed. This analysis is used to inform the selection of an appropriate processing methodology for incorporating accelerometry data into the PAM system. Chapter 5 covers the use of a camera as a single sensor for complex environments and presents a comparison of four image change detection algorithms and the selection of a suitable algorithm to perform image change detection in the PAM system. The data analysis and data fusion for the remaining sensors in the PAM system is covered in Chapter 6, which presents the technical trial of the PAM system, the high-level data analysis and the results obtained.

Chapter 4

Accelerometry

4.1 Introduction

As one of the sensors in the PAM system, triaxial accelerometers are being used as part of the wearable node in order to obtain information about the physical activity of the user. The aim of using the accelerometer is that physical activity can be monitored and that this will be one of the key indicators of behavioural pattern. In the PAM system the accelerometer is integrated into the wearable node and worn on the hip and to facilitate the inclusion of the accelerometer data into the PAM data processing architecture the data must be processed so that a single measure can be obtained for a given length of time. This allows for the 24-hour long time series to be generated and the subsequent integration into the high-level processing.

This chapter presents some preliminary work that was undertaken prior to the development of the wearable node that investigated the use of accelerometers for the PAM system. A controlled data gathering study was conducted to gather data that were used to explore and develop the analysis routines necessary to obtain behavioural information from a triaxial accelerometer. The study involved 19 people without BD (male = 9, female = 10) who were asked to wear a small triaxial accelerometer on the wrist whilst completing some well-defined everyday activities. The data gathered from the study were processed with a variety of algorithms in order to explore the data space and identify and develop appropriate methods for handling the data. Ethical approval for the study was obtained from the ISVR ethics committee (ISVR ethics reference number 895). This chapter presents the work that has been undertaken to investigate the use of accelerometers in the PAM system and the ways in which the data can be processed in order for them to be useful. Some of this work was presented at EMBEC2008 [3] and a copy of the poster and accompanying paper are included in Appendix A.

4.2 The Problem

In order for the accelerometer to be a useful sensor in the PAM system it must fit into the data processing architecture outlined in Chapter 3. This means that there must be some data processing that takes the output from the accelerometer and produces one or more time-series that show behavioural patterns over the course of a single day.

There are a range of processing techniques that can be used to achieve this and the problems being investigated in this chapter are twofold. Firstly, the accelerometry data are to be explored to determine what algorithms are suitable for use and what can be done with the data. Secondly, a method must be chosen to sensibly process the accelerometry data so that it can be included in the PAM system.

4.3 Experimental Protocol

To answer the two questions being posed of the accelerometer data a controlled study was performed. Participants were asked to wear an MSR 145 triaxial accelerometer [36] and complete the tasks shown in Table 4.1. The accelerometer was secured to the outside of the participant's wrist using a length of elasticated sports bandage that had been modified to securely hold the accelerometer. The wrist was chosen as a location due to the ease of attaching the device that location. Figure 4.1 shows the accelerometer unit and the orientation in which it was placed on the wrist. Accelerations were recorded to $\pm 2\ g$ at a rate of 50 Hz. This was chosen so that there was sufficient information captured in the acceleration signal whilst not exceeding the device's on-board memory over the course of running a single participant through the study.

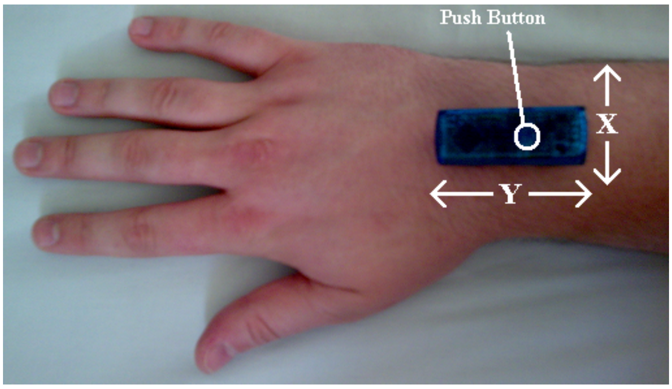


FIGURE 4.1: Photograph of the MSR 145 accelerometer placed on the wrist showing the orientation of the device with respect to the wrist. The z-axis comes out of the page. The accelerometer is secured to the wrist with a piece of elasticated sports bandage (not shown).

Prior to each experiment the accelerometer was reset to the factory settings to ensure consistency. In order to easily distinguish between different activities, the on-board push button was used to start and stop the device. The participant was asked to hold their hand and arm flat and steady at the beginning and end of each activity to allow the start and end of each activity to be clearly identified in the trace.

Table 4.1: Data Gathering Activities.

Activity	Time	Notes
Walking	15 minutes	The participant completes a walk around the university grounds.
Typing	10 minutes	The participant types out a short piece of prose about the history of the University and is stopped after 10 minutes regardless of completion of the task.
Sitting	10 minutes	The participant sits on the chair provided and watches 10 minutes of a DVD played through a laptop at medium volume.
Brisk Walk	10 minutes	The participant completes the same walk around the university as for the first task, but at a brisk pace.
Listening to Music	10 minutes	The participant lies on the bed provided and listens to music played through laptop speakers at medium volume.

4.4 Data Processing Methodology

The data gathered from the study were passed through feature extraction and normalised using z-scores, where each data point is represented by a number that shows how many standard deviations that datum is above or below the mean. This was applied to each feature independently to normalise the data between features. The normalised data were processed with two methods, Neuroscale and PCA, to project the data into a two-dimensional space so that they could be visualised. Clustering techniques were applied to the two-dimensional projections to automatically categorise the data in a purely data-driven way. As a final step, the identified clusters were used as a classifier so that new data could be projected into the same space and categorised by the cluster into which they were projected. The steps in both processing chains are shown below.

PCA Approach

The PCA approach used the following steps.

1. Feature extraction and normalisation.

2. Feature selection with the maximum relevance, minimum redundancy algorithm (mRMR) [85].

3. Projection of the data into two dimensions with PCA.

4. Clustering with the Gustafson-Kessel (GK) fuzzy clustering algorithm [38].

5. Classification with k-Nearest Neighbour (k-NN) classifier [21] and cluster projection.

Neuroscale Approach
The Neuroscale approach used the following steps.

1. Feature extraction and normalisation.

2. Feature selection with mRMR.

3. Projection into two dimensions with the Neuroscale algorithm.

4. Clustering of data-points with the GK algorithm.

5. Classification with k-NN and cluster projection.

4.4.1 Feature Extraction and Selection

Feature extraction was performed on the raw data to extract relevant features for further processing. The data were high-pass filtered at 0.05 Hz to remove the DC signal component and each ten minute activity was split into three five minute segments, with the middle segment having a 50% overlap. Five minute segments were chosen because this represents a reasonably sized amount of activity when analysing behavioural patterns for BD. The behavioural changes that need to be detected are over a few days and it is therefore not necessary to have a very fine time resolution. It is necessary however to ensure that as patterns change over the course of a day they can be determined. The choice of five minute segmentation allows for both of these criteria to be met. The 50% overlap was chosen so that data in the middle of the activity recordings were analysed as a contiguous block rather than being split across the end of one segment and the beginning of the next. This ensures that any significant sections of the recording remain intact. Feature extraction was performed for each activity recording on each of the three axes (X, Y and Z) and on the absolute magnitude of the instantaneous acceleration given by

$$\mathrm{A}(t) = \sqrt{\mathrm{x}^2(t) + \mathrm{y}^2(t) + \mathrm{z}^2(t)}, \tag{4.1}$$

where A(t) is the absolute magnitude, x, y and z are the axis of the accelerometer and t is time [32]. The sections of the trace in which the participants hand and arm were kept still, whilst the device was being started and stopped, were trimmed from the data prior to analysis. The features used are shown in Table 4.2.

TABLE 4.2: The feature list extracted from 5 minute section of the X, Y and Z axis and A during the preprocessing of the accelerometry data and the rationale for using them.

No.	Feature	Rationale
1 2 3 4	Maximum amplitude Minimum amplitude Amplitude range Mean Amplitude	These are basic statistical measures and should give an idea of some of the basic characteristics of the data section.
5 6 7 8	Variance Standard deviation Skewness Kurtosis	These are more advanced statistical measures and give an idea of the variability of the data section, how weighted it is towards positive or negative acceleration and how many large spikes there are in the section
9 10	Variance of variance Variance of mean	These provide a measure of the stability of the data section. The section is split into 30 second segments and the variance and mean for each 30 second segment is calculated. Subsequently, the variance of the 30 second segment variances and means is calculated.
11-15	Power spectrum in five bins	This gives an idea of the frequency components of the signal
16	Double integration of the section	This provides the distance the accelerometer moved over the course of the section.
17	Mean of absolute amplitude	This gives an idea of the average vigorousness of movement.
18	Variance of absolute amplitude	This gives a measure of the variance in the speed of movement.

Feature selection was performed to reduce the size of the feature set being passed to the Neuroscale algorithm. There are some features in the feature list that are closely related to each other, standard deviation and variance for example. These are included since it may be the case that one or the other is more useful in combination with respect to the other features. Feature selection is employed to algorithmically remove features that do not contribute much non-redundant information to the feature set. The mRMR algorithm, detailed in Section 4.5.1, was used to perform the feature selection. At this stage, for both processing chains, the data were divided into a training set and a test set. The training/test split was 2/3 training to 1/3 test and the data were divided with each data-point having a random chance to be assigned to one of the two sets. Feature selection was implemented in the mRMR Matlab toolbox [84].

4.4.2 Data Projection

Data projection into two dimensions was performed to allow for easy visualisation of the data and to facilitate a visual examination of the data. By projecting into two dimensions it is possible to examine the data-space for any groupings of data-points that may exist and to identify broad patterns in the data. A two-dimensional visualisation of the data is also an useful visual tool to communicate information to users and this is a key consideration for the PAM project and one of the requirements listed in Chapter 3. Two methods were used to project the data. PCA, detailed in Section 4.5.2, and Neuroscale, detailed in Section 4.5.3.

PCA reduces the dimensionality of the data to produce principal components (PCs) that are linear combinations of the input variables. It is widely used algorithm for this class of task and was chosen as a simple approach for projecting the accelerometry data. The Neuroscale algorithm performs a similar task but creates outputs that are non-linear and also aims to preserve the shape of the underlying data. Neuroscale also offers a tunable degree of supervised machine learning where a known data structure can be imposed on the mapping of the data points. Other algorithms, such as Sammon Mappings [95] and Kohonen Self-Organising Maps [48], offer a similar functionality, however, these algorithms do not have the supervised component that Neuroscale offers. As a result, Neuroscale was chosen as a more advanced approach to projecting the accelerometry data.

The projection with PCA is a straight forward application of the PCA algorithm. No machine learning steps are used and the resultant projection is purely data-driven. The data are projected onto the two principal components and the maximum amount of variability in the data is preserved.

The projection of the data with the Neuroscale algorithm is slightly more involved as the algorithm allows the degree of supervision to be altered. The algorithm was provided with a set of target projections based on the known class labels of the input data. This allows for the imposition of a known structure on the data and provides a way to increases the separation of the data classes. In both the PCA and Neuroscale cases it was the training data that was used in the initial projection phase. The Neuroscale algorithm was implemented using the Netlab Matlab toolbox [72].

4.4.3 Clustering and Classification

Clustering was performed on the data projections to examine the separability of the activity classes in the two-dimensional projections. By running a clustering algorithm on the data the groupings in the data can be automatically determined and the activity represented by the cluster can be identified. Clustering is a natural step once the data have been projected into two dimensions, especially for the Neuroscale projection, which produces a projection especially amenable to clustering. This leads directly to the use of these identified clusters as a means of classification and the use of the k-NN classifier provides a second means of classification which acts as a comparison for the cluster based classifier.

Clustering was performed with the GK algorithm. This choice was made for three reasons. Firstly, because it is a fuzzy algorithm it allows for data-points to be members of more that one cluster, which allows for better handling of marginal cases. Secondly, the algorithm allows for the shape of the cluster to change, which allows for a better fitting of the data where the clusters are the same volume, but different shapes as is the case with the data from this study. Finally, in comparison to the other algorithm tested that offers similar characteristics, the Gath-Geva (GG) algorithm, the GK provided significantly better results. Clustering was only performed on the training data, and once the data had been clustered, each cluster was assigned a class label corresponding to the majority label of the training points assigned to the cluster. The GK algorithm was implemented using the Matlab Fuzzy Clustering Toolbox [6].

Classification was performed in two ways. The first method was to use the existing clusters; the test data were projected into the projection-space, either through the PCA or Neuroscale mappings, and each point classified according to the cluster into which it fell. The second method was the k-NN classifier, detailed in Section 4.5.4. The results of the classification are presented in terms of precision, recall and the F_1 measure, which are explained in Section 4.5.6.

4.4.4 One-dimensional and Higher-dimensional Projections

As an additional investigation of the data gathered from the accelerometer, some projections of the data into both a one-dimensional space and into higher-dimensional (3D, 4D and 5D) space were produced using both Neuroscale and PCA. These were designed to see how the data fell into other dimensioned spaces and how this affected the classification accuracy of the Clustering and k-NN classifiers.

4.4.5 Algorithm Optimisation

The algorithms used to process the data have several parameters that need to be tuned to produce a good performance. Both PCA and Neuroscale can have a varying number of features passed to them and the neuroscale algorithm can have the number of centres used in the ANN tuned. The k-NN algorithm can have the number of neighbours, k, tuned. As part of this investigation, the data projection into 1–5 dimensions is used, and the tuning was repeated for each combination of projection method and classifier in each dimension space.

For the PCA/k-NN combination a two variable search was performed on the features passed to PCA and on the k nearest neighbours. This created a 2D optimisation space which showed the classification results with the feature and k parameters that generated them. Each classification result was an average over 100 runs with different training and testing splits of the data. Values for the number of features and k were chosen to be the lowest possible values with a high classification accuracy. Table 4.3 shows the results of this optimisation.

Table 4.3: Optimised parameters for the PCA/k-NN combination.

Dimensionality	Features	k
1	72	17
2	72	9
3	72	10
4	72	13
5	72	4

For the Neuroscale/k-NN combination, a two stage search was used. Initially a two variable search was used on the space of features and hidden centres for the Neuroscale projection and k was set to 13 to identify the approximate region of optimality. Subsequently a smaller region (± 10) of feature and hidden centre numbers was optimised whilst varying k in order to identify the best combination. Table 4.4 shows the results of this optimisation.

Table 4.4: Optimised parameters for the Neuroscale/k-NN combination.

Dimensionality	Features	Centres	k
1	17	58	13
2	17	55	13
3	17	61	9
4	17	66	4
5	17	61	3

For the Neuroscale/Clustering combination, the parameters were tuned by running the data analysis whilst changing the number of features and centres. In general, the optimisation space showed a plateau where the accuracy did not rise significantly once feature and centre had reached certain values. Values for features and centres were chosen such that they were as low as they could be whilst still being on the plateau and providing good classification accuracy. Table 4.5 shows the results of the optimization.

Table 4.5: Optimised parameters for the Neuroscale/Clustering combination.

Dimensionality	Features	Centres
1	20	49
2	18	50
3	14	61
4	18	71
5	17	67

4.5 Data Processing Techniques

4.5.1 Maximum-Relevance Minimum-Redundancy Feature Selection

Feature selection using the criteria of mRMR is a feature selection technique that aims to select a subset of features from a list where the selected subset contains as much information as possible whilst containing as little redundancy (duplicate information) as possible [26, 85]. The technique is run on a data-set for which there are a set of features and an optional class label, c. Assuming that there is an existing set of features that has been selected that needs to have one feature added to it, F_{m-1}, the feature that is added must add the most information to the selected feature set whilst minimising the redundancy. The mth feature is the feature that maximises the condition

$$\max_{f_j \in \{\mathcal{F} - F_{m-1}\}} \left[I(f_j; \mathsf{c}) - \frac{1}{m-1} \sum_{f_i \in \{F_{m-1}\}} I(f_j; f_i) \right], \tag{4.2}$$

which, assuming $\{F_{m-1}\}$ is the feature set with $m-1$ features, aims to select the mth feature from the set $\{\mathcal{F} - F_{m-1}\}$, where $\{\mathcal{F}\}$ is the set of all features, that is maximally dependant on the class variable, c, and minimally redundant with respect to the other features in the feature set $\{F_{m-1}\}$, based on the mutual information, $I(.)$. The class variable is essentially, for training purposes, another feature and corresponds to the known classification of the training data-points. Mutual information between two random variables is calculated as

$$(f_j; f_i) = \int\int p(f_j, f_i) \log \frac{p(f_j, f_i)}{p(f_j)p(f_i)}\, df_j\, df_i. \tag{4.3}$$

For discrete variables, the integrals become summations and the joint and marginal probabilities can be calculated by tallying the relevant samples in the data. For the purposes of the work on the accelerometry data, the data passed to the mRMR were truncated to three decimal places for the calculation of mutual information.

The selected feature set is built up recursively, starting by selecting the single best feature and then successively adding one feature to the set. For the initial feature selection equation 4.2 does not require redundancy to be taken into account and so becomes

$$\max_{f_j \in \{\mathcal{F} - F_{m-1}\}} I(f_j; \mathsf{c}). \tag{4.4}$$

The feature list that is generated by this algorithm contains all the features from the feature space, ordered such that the mth feature is the feature that adds the most value to the list with $m - 1$ features. The list therefore needs to be trimmed to obtain the optimal number of features. This is typically performed once a classifier has been built. A classifier is constructed and tested starting with the first feature and sequentially adding features into the classification, working down the list. This continues until an acceptable classification error is reached, or until the error converges to within acceptable limits. The list of features used is therefore the top n features in the list that lead to the highest acceptable classification error.

4.5.2 PCA Analysis

The goal of PCA [45] is to reduce the dimensionality of a data-set whilst retaining as much of the variability in the data-set as possible. Where the data in question is a set of n-dimensional random variables of the form $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$, the dimension reduction can be achieved by transforming the data onto m principal components (PCs). These take the form $\boldsymbol{\alpha}_d^T\mathbf{x}$, where $d = [1, m]$. For the first PC $\boldsymbol{\alpha}_1^T\mathbf{x}$ should have maximum variance. Subsequently, $\boldsymbol{\alpha}_d^T\mathbf{x}$ should have maximum variance subject to being uncorrelated to $\boldsymbol{\alpha}_1^T\mathbf{x}$, $\boldsymbol{\alpha}_1^T\mathbf{x}$, …, $\boldsymbol{\alpha}_{d-1}^T\mathbf{x}$. In general, most of the variance in the data-set should be accounted for by the first m PCs with $m << n$.

Assuming that the data-set has a known covariance matrix $\boldsymbol{\Sigma}$ the dth PC is given by $z_d = \boldsymbol{\alpha}_d^T\mathbf{x}$, where $\boldsymbol{\alpha}_d^T$ is an eigenvector of $\boldsymbol{\Sigma}$ corresponding to the dth largest eigenvalue λ_d. This definition can be derived by examining the first PC $\boldsymbol{\alpha}_1^T\mathbf{x}$ and observing that in order to maximise the variance, $\text{var}\left(\boldsymbol{\alpha}_1^T\mathbf{x}\right) = \boldsymbol{\alpha}_1^T\Sigma\boldsymbol{\alpha}_1$, the term $\boldsymbol{\alpha}_1$ must be maximised. This cannot be achieved without a constraint on $\boldsymbol{\alpha}_1$ so the constraint $\boldsymbol{\alpha}_1^T\boldsymbol{\alpha}_1 = 1$ is used.

To perform this maximization Lagrange multipliers can be used so that the problem becomes to maximise

$$\boldsymbol{\alpha}_1^T \boldsymbol{\Sigma} \boldsymbol{\alpha}_1 - \lambda \left(\boldsymbol{\alpha}_1^T \boldsymbol{\alpha}_1 - 1 \right), \tag{4.5}$$

where λ is a Lagrange multiplier. Differentiation with respect to $\boldsymbol{\alpha}_1$ produces

$$\boldsymbol{\Sigma} \boldsymbol{\alpha}_1 - \lambda \boldsymbol{\alpha}_1 = 0. \tag{4.6}$$

It follows that λ is an eigenvalue of $\boldsymbol{\Sigma}$ and $\boldsymbol{\alpha}_1$ the corresponding eigenvector. In order to obtain the maximum variance for $\boldsymbol{\alpha}_1 \mathbf{x}$ it is λ that must be maximised, which leads to $\boldsymbol{\alpha}_1$ being the eigenvector that corresponds to the largest eigenvalue. In general, as stated above, the dth PC of $\mathbf{x}$ is $\boldsymbol{\alpha}_d^T \mathbf{x}$, where $\boldsymbol{\alpha}_d^T$ is an eigenvector of $\boldsymbol{\Sigma}$ corresponding to the dth largest eigenvalue λ_d.

4.5.3 The Neuroscale Algorithm

The Neuroscale algorithm [57, 58, 59], is designed to project high-dimensional data onto a low-dimensional space whilst, at the same time, preserving the distance ratios between points. Thus two points that are very close together in the high-dimensional space remain close together when projected onto the low-dimensional space and two points that are far apart remain far apart after projection. This allows the relationships and structure in the high-dimensional data to be visualised for data analysis. Unknown data-points can also be projected onto the low-dimensional space using the generative mapping produced by the Neuroscale algorithm. Given an n-dimensional feature space, $\mathbf{X}$, containing N data-points, $\mathbf{x}_i$; and some prior knowledge of the underlying topography, $\mathbf{S}$, containing points $\mathbf{s}_{ij}$, the m-dimensional (where $m < n$) projection-space, $\mathbf{Y}$, containing data-points $\mathbf{y}_i$, is created to minimise the stress measure

$$E = \sum_{j=1}^{N} \sum_{i>j}^{N} \left((1-\alpha) d_{ij}^* + \alpha s_{ij} - d_{ij} \right)^2, \tag{4.7}$$

where d_{ij}^* are the inter-point Euclidean distances in the feature space and d_{ij} are the corresponding distances in the projection-space, defined as

$$d_{ij}^* = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j)}, \tag{4.8}$$

$$d_{ij} = \sqrt{(\mathbf{y}_i - \mathbf{y}_j)^T (\mathbf{y}_i - \mathbf{y}_j)}, \tag{4.9}$$

and $\mathbf{s}_{ij}$ is part of a subjective dissimilarity matrix corresponding to the ideal distances between the points in the projection-space and α is used to control the influence of $\mathbf{s}_{ij}$ in the equation. The subjective dissimilarity matrix $\mathbf{S}$ is constructed as a distance matrix for the points in $\mathbf{X}$ where the distances are the desired distances of the data-points from one another in the projection space $\mathbf{Y}$. Thus, $\mathbf{s}_{ij}$ is the desired distance between points i and j in the projection space and the matrix $\mathbf{S}$ shows the ideal output of the Neuroscale algorithm. The effect this matrix has on the computation is controlled by α and serves to alter the degree of supervision from unsupervised at $\alpha = 0$ to fully supervised at $\alpha = 1$.

In order to map between $\mathbf{X}$ and $\mathbf{Y}$ a radial basis function (RBF) ANN is used. The RBF ANN takes n inputs and produces m outputs. An RBF is a function where the output depends only on the distance of the input from the origin or some other centre. There are many different functions that can be used in an RBF ANN and Neuroscale utilises a thin plate spline function of the form $f(x) = x^2 \ln(x)$. The parameters that need to be optimised in the RBF ANN are the centres and output weightings for each RBF. This is performed with the shadow targets algorithm [105], which uses an EM type approach to iteratively calculate the error measure, E, and update the model parameters based on the error. Figure 4.2 shows the Neuroscale approach going from a multi-dimensional space to a two-dimensional space.

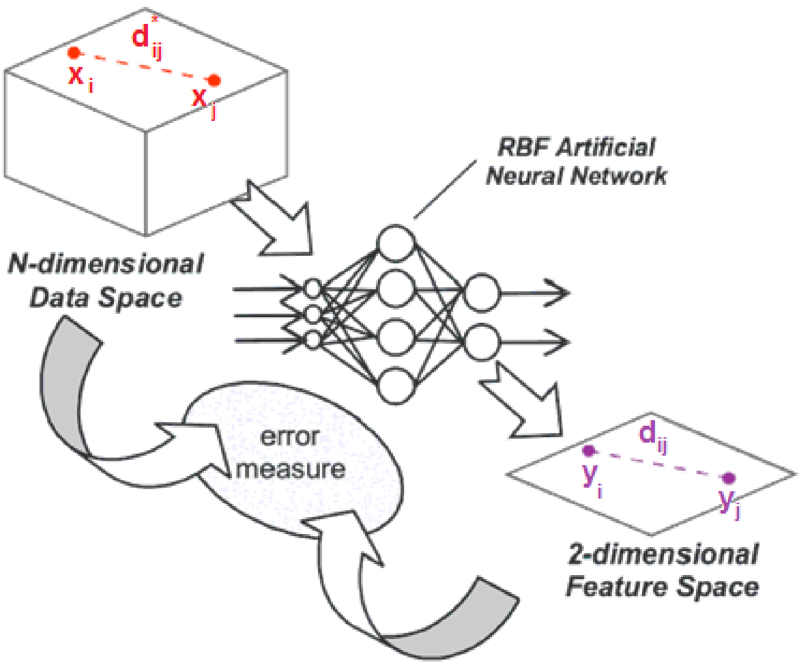


FIGURE 4.2: Graphical diagram of the Neuroscale approach showing the multi-dimensional input, the RBF ANN and the two-dimensional output.

4.5.4 k-Nearest Neighbour

The k-NN [21] algorithm is a very simple classification algorithm. A test point is projected into the training space and the category labels of the k nearest training points are used to determine the classification of the test point. The value of k is normally chosen to be odd and once the nearest neighbours have been identified, the classification corresponds to the most dominant class amongst the k neighbours. The number for k is normally chosen to be odd to minimise the chances of having equal numbers of class labels amongst the neighbours.

4.5.5 Clustering

Clustering takes a matrix of N data vectors, $\mathbf{X}$, where each data vector contains n measured variables for each data-point. A partition matrix $\mathbf{U}$ is used to assign memberships to each of the c clusters for each data-point. The matrix $\mathbf{U}$ has N rows, corresponding to the data-points in $\mathbf{X}$, and c columns, corresponding to the clusters, the number of which must be supplied in advance. The structures of $\mathbf{X}$ and $\mathbf{U}$ are

$$\mathbf{X}^T = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{Nn} \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1c} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{N1} & \mu_{N2} & \cdots & \mu_{Nc} \end{bmatrix}. \tag{4.10}$$

For hard partitions, each data-point in $\mathbf{U}$ can be a member of only one cluster, leading to the following constraints [6]

$$\mu_{ij} \in \{0, 1\}, \quad 1 \leq i \leq N, \; 1 \leq j \leq c, \tag{4.11}$$

$$\sum_{j=1}^{c} \mu_{ij} = 1, \quad 1 \leq i \leq N, \tag{4.12}$$

$$0 < \sum_{i=1}^{N} \mu_{ij} < N, \quad 1 \leq j \leq c, \tag{4.13}$$

where (4.11) constrains the membership values for the data-points to 0 or 1, (4.12) constrains each data-point to be a member of only one class and (4.13) constrains the clusters such than no cluster can be empty or contain all the data-points.

For fuzzy partitions, a data-point can be a member of more than one cluster, which leads to the following constraints for fuzzy partitions [6]

$$\mu_{ij} \in [0,1], \quad 1 \leq i \leq N,\ 1 \leq j \leq c, \tag{4.14}$$

$$\sum_{j=1}^{c} \mu_{ij} = 1, \quad 1 \leq i \leq N, \tag{4.15}$$

$$0 < \sum_{i=1}^{N} \mu_{ij} < N, \quad 1 \leq j \leq c, \tag{4.16}$$

where (4.14) constrains the membership values for the data-points to the range 0–1, (4.15) constrains the total membership of each data-point to sum to one and (4.16) constrains the clusters such than no cluster can be empty or contain all the data-points with their membership at one.

The membership values determined in a clustering algorithm will be derived by minimising an objective function. Typically this function aims to keep the points in each cluster as close to each other as possible. The specific objective function used gives rise to a number of different clustering algorithms such as k-means [1], GK [38] and GG [34]. For the accelerometry work presented here, the GK algorithm has been used and cluster membership is evaluated by minimising the objective function

$$J(\mathbf{X}; \mathbf{U}, \mathbf{V}, \mathbf{A}) = \sum_{i=1}^{N} \sum_{j=1}^{c} (\mu_{ij})^m D^2_{ijA_j}, \tag{4.17}$$

where $\mathbf{V}$ is a vector of cluster centres and $\mathbf{A}$ is the set of norm inducing matrices used in the calculation of $D^2_{ijA_j}$. Both $\mathbf{A}$ and $\mathbf{V}$ must be optimised, along with $\mathbf{X}$ and $\mathbf{U}$. The parameter m is used to control the fuzziness of the partitions. The distance norm (a measure of the distance between the points in a cluster and the cluster centre) $D^2_{ijA_j}$ is calculated using

$$D^2_{ijA_j} = (\mathbf{x}_i - \mathbf{v}_j)^T \mathbf{A}_j (\mathbf{x}_i - \mathbf{v}_j), \quad 1 \leq i \leq N,\ 1 \leq j \leq c, \tag{4.18}$$

where each cluster has a separate norm-inducing matrix, $\mathbf{A}_j$, which is allowed to vary whilst maintaining a constant determinant. This allows the shape of the cluster to change whilst maintaining a constant volume.

4.5.6 Classification Accuracy Measures

In order to determine the accuracy of a classifier, a test set, which the classifier was not trained on, is passed through the classifier to assess its performance on unseen data.

From the classification of the test set it is possible to determine the number of correct and incorrect classifications the classifier has made. In a simple two-class classifier, these can be represented as true positives (tp), false positive (fp), true negative (tn) and false negative (fn). There are two commonly used methods for numerically analysing the performance of a classifier, *sensitivity and specificity*, and *precision and recall.*

Sensitivity and specificity are widely used in the biomedical field and are defined for binary (two class) classifiers. Sensitivity is the proportion of actual positives that are correctly identified. Specificity is the proportion of actual negatives that are correctly identified. Sensitivity and specificity are defined as

$$\text{Sensitivity} = \frac{tp}{tp + fn}, \tag{4.19}$$

$$\text{Specificity} = \frac{tn}{tn + fp}. \tag{4.20}$$

Precision and recall are widely used in the information retrieval and machine learning fields and are similar to sensitivity and specificity. Recall is the same as sensitivity whereas precision is the positive prediction value, that is, the probability that an item given a positive classification is an actual positive. Precision and recall are defined as

$$\text{Precision} = \frac{tp}{tp + fp}, \tag{4.21}$$

$$\text{Recall} = \frac{tp}{tp + fn}. \tag{4.22}$$

It is often common for precision and recall to be combined into a single measure. The most common of these is the F_1 score and is defined as

$$\mathrm{F}_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \tag{4.23}$$

It can be seen from the two sets of equations that the two methods of evaluating classifier accuracy are similar, however, precision and recall are more suited to a multi-class classifier. They do not use the true negative count, which, in a multi-class classifier will likely be large compared to the other numbers. Having a large true negative count will quickly dominate the measure of specificity and lead to less informative measures.

4.6 Results

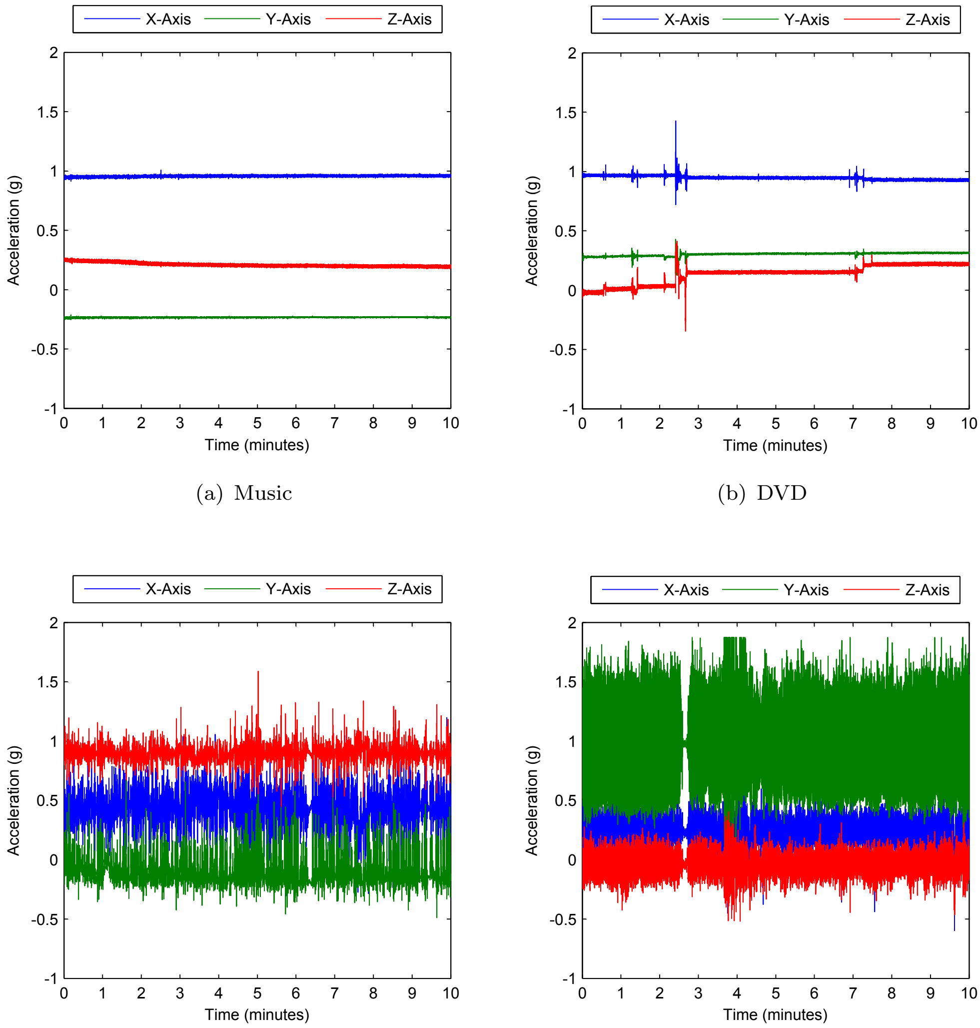
The study was run with 19 participants, 9 male and 10 female. Figure 4.3 shows example traces from four of the activity categories recorded in the study. The traces are visually different in the characteristics they display. The trace for *Walk Slow* for example, shows a very high level of activity across all three axis. This level of activity is evident to some extent in the *Typing* trace, but absent almost entirely from the *Music* and *DVD* traces. There are also some similarities in the traces, *Music* and *DVD* for example both display long periods with little or no movement. The *DVD* trace does however show some posture changes, identifiable by the short periods of activity followed by a change in the level of the trace. The visual differences in the activity traces offer a good indication that different types of activity produce activity traces that display differing characteristics. There is also an element of clipping evident in Figure 4.3(d) and this comes from the data being outside of the $\pm 2\,g$ range of the accelerometer.

4.6.1 Feature Selection

Table 4.6 shows the result of applying the mRMR feature selection to the features extracted from the accelerometry data. Only the top 18 features are shown since it would be impractical to list them all.

TABLE 4.6: Feature list ordered by mRMR showing the top 18 features.

Feature	Axis
Maximum	A
Variance of variance	Z
Variance of variance	X
Variance of variance	Y
Minimum	Y
Variance of variance	A
Maximum	Y
Kurtosis	Z
Minimum	X
Range (Max-Min)	Y
Kurtosis	Y
Variance of mean	Y
Minimum	A
Variance of mean	A
Minimum	Z
Variance of mean	X
Maximum	Z
Maximum	X



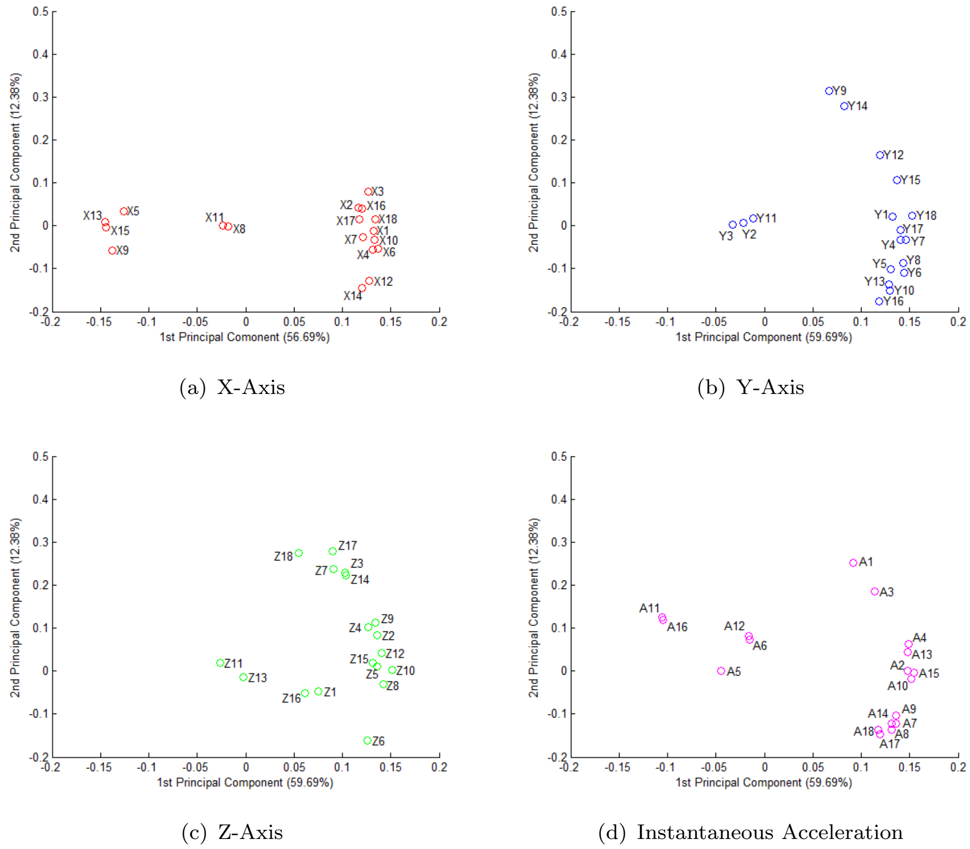
(a) Music (b) DVD

(c) Typing (d) Walk Slow

FIGURE 4.3: Example of four triaxial accelerometry traces obtained in the data gathering study; *Music* (a), *DVD* (b), *Typing* (c) and *Walk Slow* (d). Small posture changes can be observed in (b) as short time duration spikes in the trace. In general, the four traces look visually different in their characteristics.

4.6.2 PCA Processing

Figure 4.4 shows how the various features extracted from the accelerometry data map to the first and second PC. These graphs show that the first PC uses most of the available features and weights most of them with roughly the same loading. The second PC however makes use of significantly fewer features. In particular, features (where the letter corresponds to the x, y, z and instantaneous acceleration (A) axis and the number to the numbers in Table 4.2) Y9 (variance of variance), Y14 (PSD 4th bin), Z17 (mean of absolute amplitude), Z18 (variance of absolute amplitude), A1 (maximum amplitude) and A3 (amplitude range) contribute most to the second PC.



(a) X-Axis

(b) Y-Axis

(c) Z-Axis

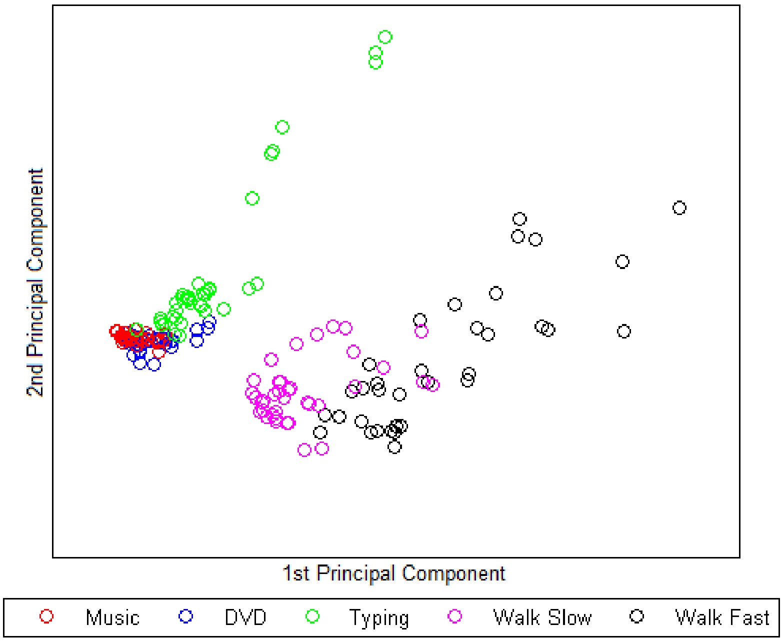
(d) Instantaneous Acceleration

FIGURE 4.4: Scatter plots of the features drawn from the accelerometry data showing the loading scores for each feature on the first and second PC. The three axis of acceleration (X, Y and Z) are shown along with the instantaneous acceleration combination (A). Data-point labels on the scatter plots refer to the axis and feature number from Table 4.2, thus the label X1 refers to the 1st feature from the X-Axis data.

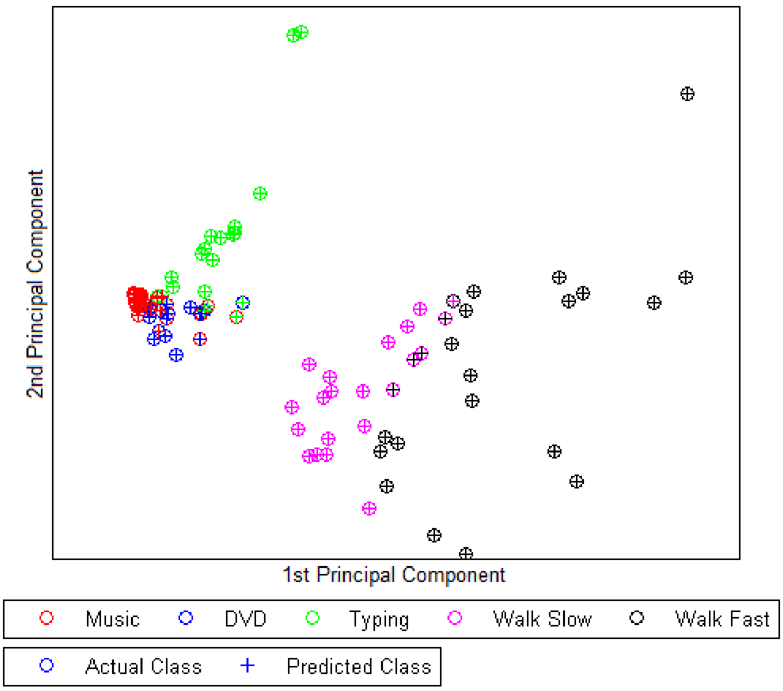
Figure 4.5(a) shows the projection of the training-set from a train/test division of the data-set into two dimensions using PCA. The activity classes have been separated to some degree and there is a general trend for the lower intensity activities to be on the left of the projection (*Music* and *DVD*) and the higher intensity activities on the right (*Walk Slow* and *Walk Fast*). There is however a large degree of overlap between all of the classes.

Figure 4.5(b) shows the result of running the k-NN classifier on the PCA projection of the test-set. It can be seen from this figure that a large number of data-points are correctly classified, but that there are also a substantial number of incorrectly classified data-points. This is particularly evident in the overlap between *Walk Slow* and *Walk Fast*, and between *Music* and *DVD*. This is shown in more detail in Table 4.7 which shows the confusion matrix obtained from averaging 100 runs of this classifier combination with random train/test splits of the data-set.

Figure 4.6 shows box-plots of the precision, recall and F_1 scores obtained from running the PCA projection and k-NN classifier 100 times on different training/testing splits of

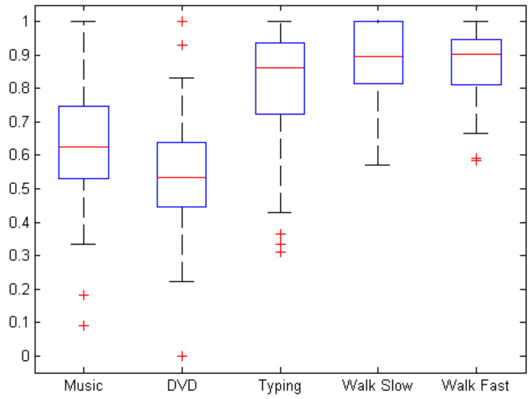


(a) Training Set Projection

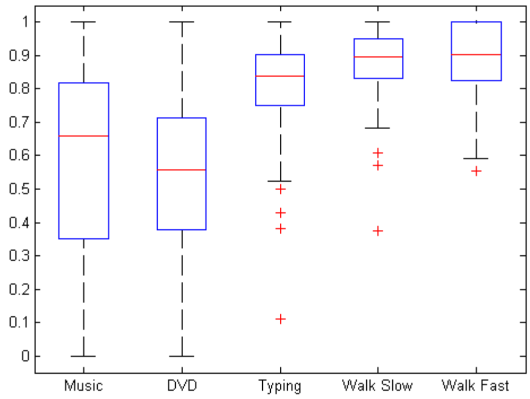


(b) Test Set Projection

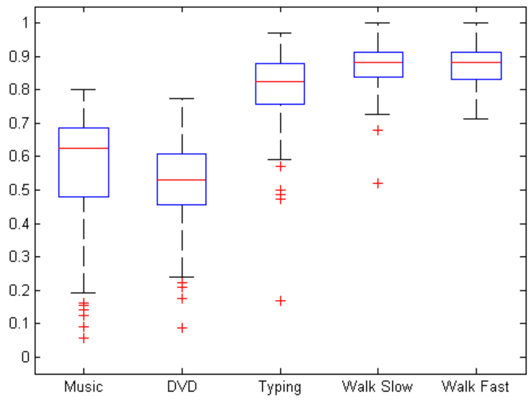
FIGURE 4.5: Typical PCA projection, (a), and k-NN classification, (b), on example training and test data sets using 72 features and setting k to 9. The training projection shows some degree separability in the data classes. The k-NN classification shows a visually good classification performance.



(a) Precision



(b) Recall



(c) F_1

FIGURE 4.6: Box-plots of the precision, (a), recall, (b), and F_1, (c), scores across the five classes using the PCA projection and k-NN classification over 100 train/test splits of the data using 72 features and setting k to 9. All three performance measures are modest. The exception to this are the results for the *Walk Slow* and *Walk Fast* classes.

TABLE 4.7: 2D PCA/k-NN classification confusion matrix obtained from 100 random training/testing splits of the data with 72 features and setting k to 9.

		Predicted Class				
		Music	DVD	Typing	Walk Slow	Walk Fast
Actual Class	Music	11.02	7.54	0.98	0	0
	DVD	5.22	10.5	1.93	0	0
	Typing	1	1.8	14.87	0.1	0.01
	Walk Slow	0	0	0.02	16.21	2.65
	Walk Fast	0	0	0	2.15	16.19

the data. The PCA algorithm was run on 72 features and k was set to 9 for the k-NN algorithm. The results of the classification are reasonable, with the average of the median scores across all five classes being 0.7639 for precision, 0.7706 for recall and 0.7492 for the F_1 score. Furthermore, the spread of scores for both precision and recall across all five categories is very high, especially the recall for the DVD and Music categories. The average misclassification rate for this classifier combination is 0.2538.

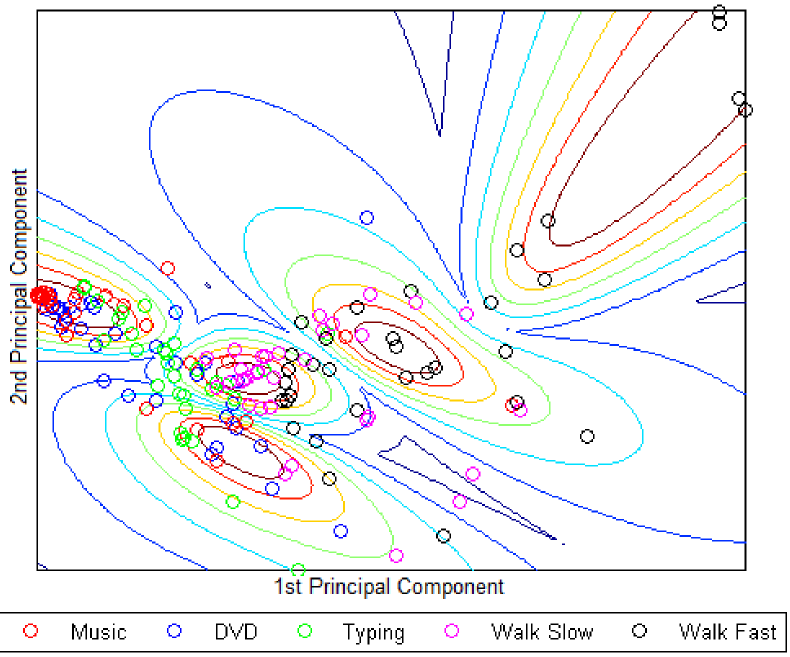


FIGURE 4.7: Example of clustering on the PCA projection of a training data set using 6 features. The contours show cluster membership values. It can be seen from the figure that the clustering performs very poorly on the PCA projection.

Figure 4.7 shows the result of applying the GK clustering algorithm to a typical training-set. The algorithm has completely failed to identify the correct clusters in the data. An attempt was made to run the cluster based classifier on these data, but due to the poor clustering performance (on a large proportion of runs the automatic labelling of the clusters failed to identify a cluster for one or more classes) this did not lead to a usable classifier.

4.6.3 Neuroscale Processing

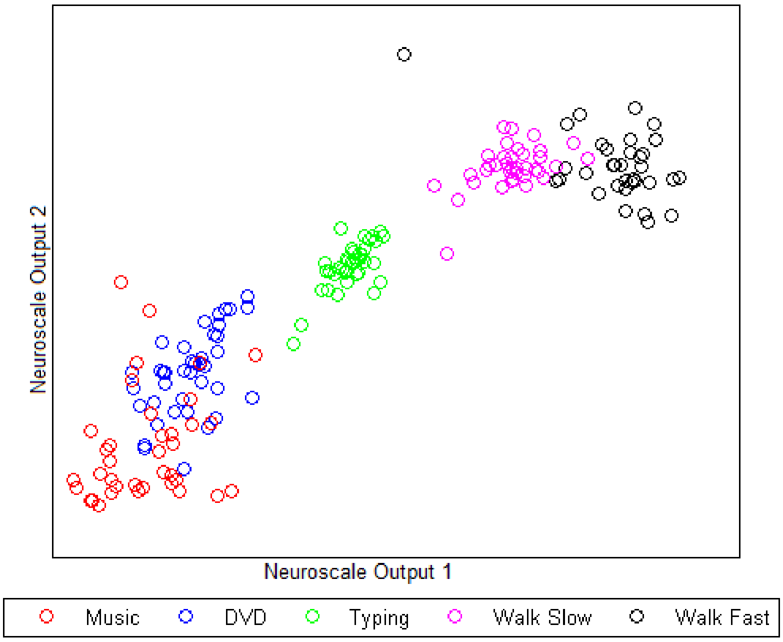
Figure 4.8(a) shows the projection of the training-set from a train/test division of the data-set into two dimensions using the Neuroscale algorithm with 16 features and 57 hidden centres. As with PCA projection, it is the shape and the organisation of the data that is important and as such, the Neuroscale projections have no axis labels. In general the classes have been well separated, with the exception of *Music* and *DVD*, which share significant overlap, and the partial exception of *Walk Fast* and *Walk Slow*, which share a small amount of overlap. As with the PCA projection, there is a trend for low intensity activity to be placed to the left of the plot and high intensity activity to be placed to the right.

Figure 4.8(b) shows the result of the k-NN classifier on the Neuroscale projection of the test-set, setting k to 13. It can be seen from this figure that the majority of the data-points are correctly classified, but that there is significant misclassification between the *Music* and *DVD* classes. This is shown in more detail in Table 4.8 which shows the confusion matrix obtained from averaging 100 runs of this classifier combination with random train/test splits of the data-set.

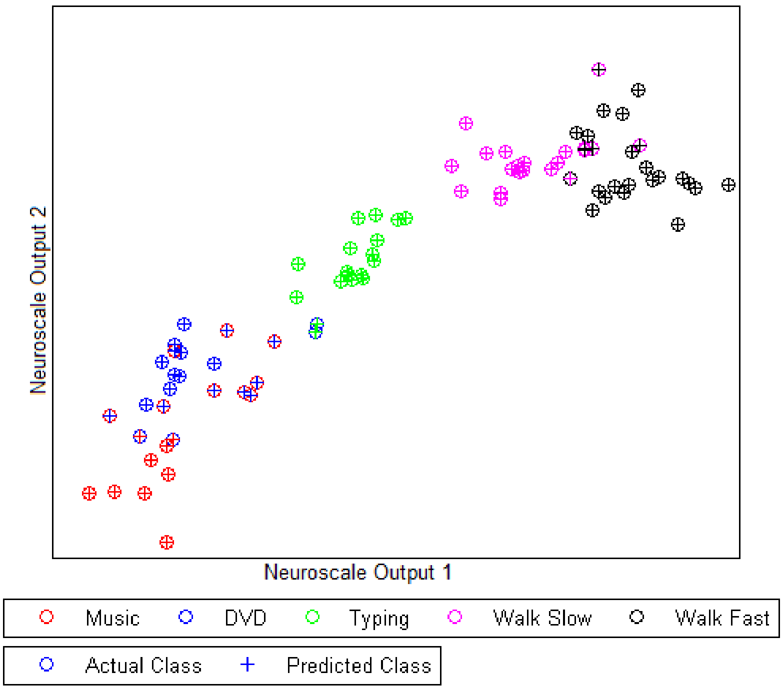
TABLE 4.8: 2D Neuroscale/k-NN classification confusion matrix obtained from 100 random training/testing splits of the data with 16 features and 57 hidden centres.

		Predicted Class				
		Music	DVD	Typing	Walk Slow	Walk Fast
Actual Class	Music	11.02	4.73	0.6	0.01	0
	DVD	4.86	7.41	2.22	0	0.01
	Typing	0.09	1.64	14.04	0.01	0
	Walk Slow	0	0	0.21	15.97	1.08
	Walk Fast	0.01	0	0.02	1.87	14.17

Figure 4.9 shows box-plots of the precision, recall and F_1 scores obtained from running the Neuroscale projection and k-NN classifier 100 different train/test splits of the data, passing 16 features to the Neuroscale algorithm, with 57 hidden centres and setting k to 13. The performance scores for the classification are generally very good with the exception of the *Music* and *DVD* classes, which have lower scores across all three performance measures. The spread of the performance scores across all five classes is reasonably large, although it is smaller for the Neuroscale k-NN results than it is for the PCA. The average of the median scores for these data are 0.7892 for precision, 0.8052 for recall and 0.7879 for the F_1 score. The misclassification rate for this classification combination is 0.2171.

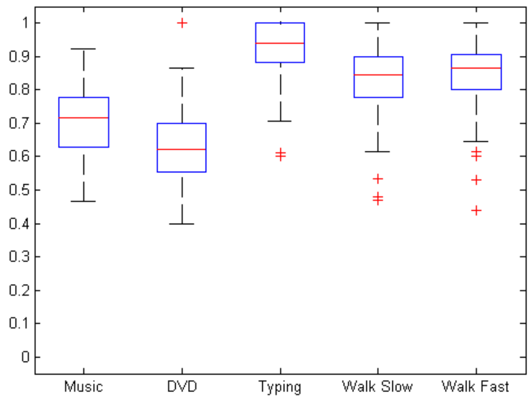


(a) Training Set Projection

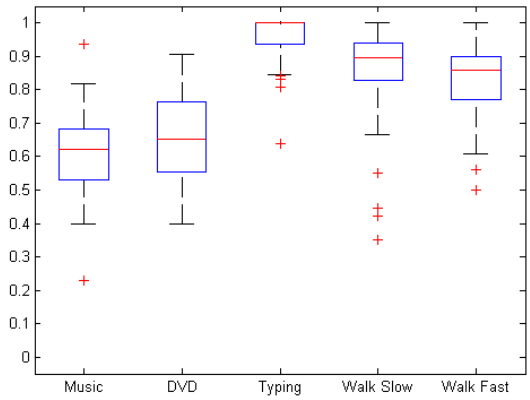


(b) Test Set Projection

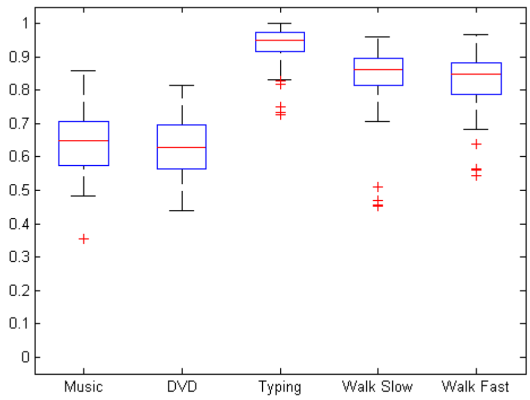
FIGURE 4.8: Typical Neuroscale projection, (a), and k-NN classification, (b), on example training and test sets using 16 features and 57 hidden centres and setting k to 13. The training projection shows a large degree of separability between the data classes. This is reflected in the k-NN classification, which visually produces a very good classification performance.



(a) Precision



(b) Recall



(c) F_1

FIGURE 4.9: Box-plots of the precision, (a), recall, (b) and F_1, (c), scores across the five classes using the Neuroscale projection, with 16 features and 57 hidden centres, and k-NN classification. All three performance measures are reasonably high for this method. In particular, the *Typing* class demonstrates very high scores for both precision and recall.

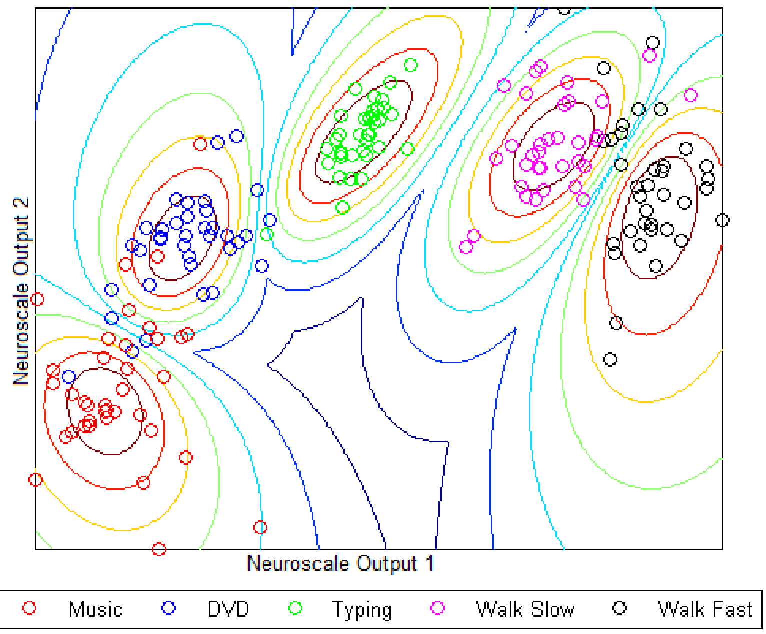
Figure 4.10 shows the projection and clustering of the training and test sets from a train/test division of the data-set into two dimensions using Neuroscale with 18 features and 50 hidden centres. The projection here shares the same characteristics as the previous Neuroscale projection in terms of class separability and overlap. It can be seen from the figure that the good separation of the classes leads to a good performance in the clustering algorithm.

Figure 4.10(b) shows that when the test-set is projected through the Neuroscale mapping the data-points generally fall into the correct cluster. The major exceptions to this are the *Music* and *DVD* clusters where there is once again a good deal of overlap. This is shown in more detail in Table 4.9.

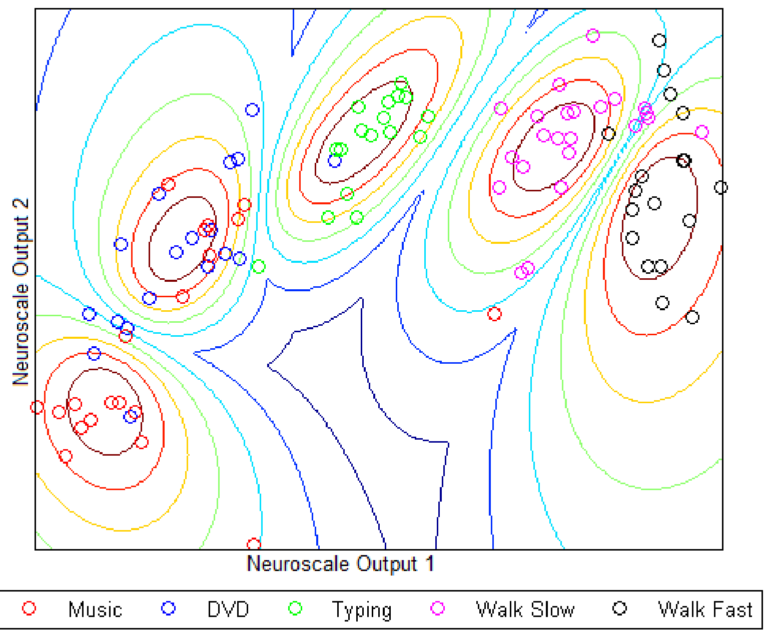
TABLE 4.9: 2D Neuroscale/Clustering classification confusion matrix obtained from 100 random training/testing splits of the data with 18 features and 50 hidden centres.

		Predicted Class				
		Music	DVD	Typing	Walk Slow	Walk Fast
Actual Class	Music	10.69	4.69	1.02	0	0
	DVD	4.85	7.22	1.86	0	0
	Typing	0.13	1.86	14.08	0.02	0
	Walk Slow	0	0.01	0.18	15.79	1.09
	Walk Fast	0.01	0	0.03	1.8	13.91

Figure 4.11 shows box-plots of the precision, recall and F_1 scores obtained from running the Neuroscale projection and cluster based classifier 100 times with different train/test splits of the data, passing 18 features to the Neuroscale algorithm, with 50 hidden centres. As with the k-NN classification, the results of the cluster-based classification are generally very good. The same performance issues exist with the *Music* and *DVD* classes, which again have lower scores across all three performance measures. The spread of the performance scores across all five classes is slightly larger than in the k-NN case, but there are significantly more outliers. The average of the median scores for these data are 0.7910 for precision, 0.7981 for recall and 0.7935 for the F_1 score. The misclassification rate for this classifier combination is 0.2215.

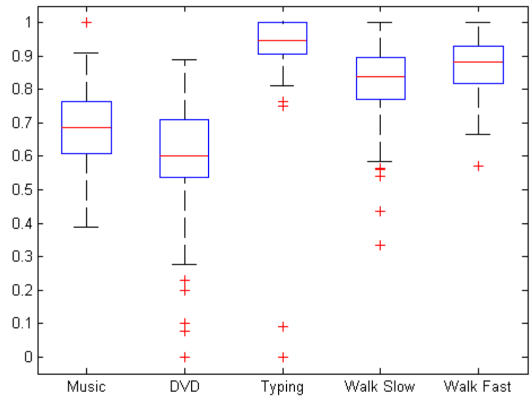


(a) Training Set Projection

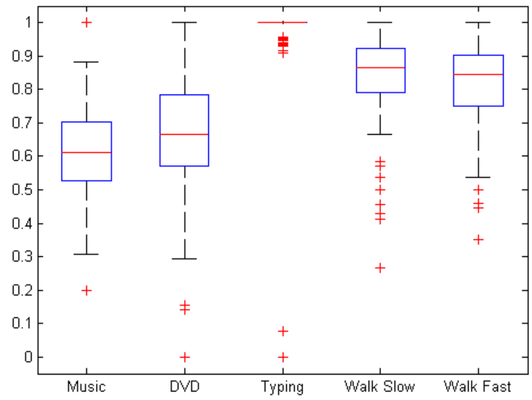


(b) Test Set Projection

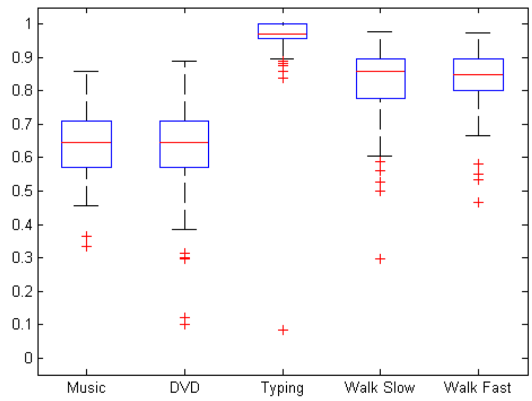
FIGURE 4.10: Typical Neuroscale projection and clustering of example training, (a), and test, (b), data sets using 18 features and 50 hidden centres. The contours show cluster membership values and the clusters determined by the training set are maintained for the test set. It can be seen visually from the figure that the Neuroscale projection allows for a good clustering performance and that when projecting new data into the same space the new points fall into the correct clusters.



(a) Precision



(b) Recall

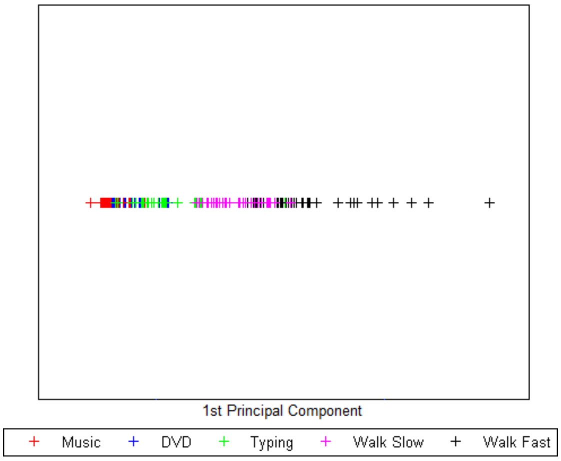


(c) F_1

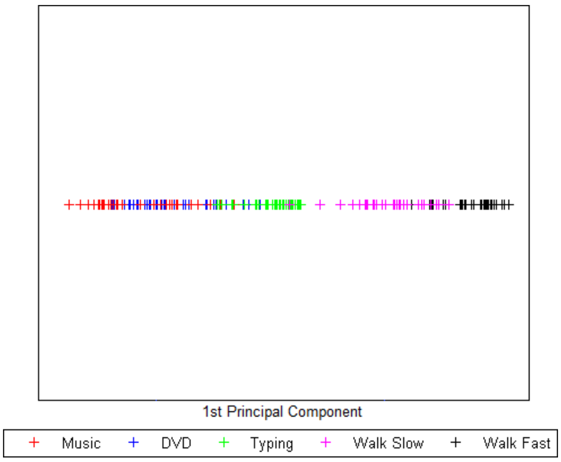
FIGURE 4.11: Box-plots of the precision, (a), recall, (b), and F_1, (c), scores across the five classes using the Neuroscale projection, with 18 features and 50 hidden centres, and the cluster-based classifier. The performance measures are generally high, but the min-max spread is very large for a number of classes, especially *DVD*. There are also a large number of outliers.

4.6.4 One-dimensional and Higher-dimensional Projections

Figure 4.12(a) shows the projection of the training data set with PCA onto the 1st PC. In general, the classes have come out in order of activity intensity, from *Walk Fast* at one end to *Music* at the other. In the higher activity end of the projection (*Walk Slow* and *Walk Fast*) the data are well separated, but becomes more confused towards the lower intensity end of the scale. Figure 4.12(b) shows the projection of the training data set into one dimension with the Neuroscale algorithm. Again, the data categories have been separated according to intensity. It can be seen from this figure that the lower end of the activity spectrum has been separated more than in the PCA projection.



(a) PCA



(b) Neuroscale

FIGURE 4.12: One dimensional projection of the training data through PCA, (a), and Neuroscale, (b). Both projections order the data left to right from least intense to most intense. It can be seen from the figure that the Neuroscale projection separates the low intensity end of the spectrum more than the PCA projection.

Tables 4.10, 4.11 and 4.12 show the confusion matrixes for the three classification approaches after projection into a one-dimensional space.

TABLE 4.10: 1D PCA/k-NN classification confusion matrix obtained from 100 random training/testing splits of the data with 72 features and setting k to 17.

		Predicted Class				
		Music	DVD	Typing	Walk Slow	Walk Fast
Actual Class	Music	7.54	8.97	1.81	0.01	0
	DVD	3.91	9.68	4.59	0.01	0
	Typing	0.55	3.95	10.64	2.64	0.81
	Walk Slow	0	0	0.53	15.19	3.48
	Walk Fast	0	0	0	2.44	15.35

TABLE 4.11: 1D Neuroscale/k-NN classification confusion matrix obtained from 100 random training/testing splits of the data with 17 features, 58 hidden centres and setting k to 13.

		Predicted Class				
		Music	DVD	Typing	Walk Slow	Walk Fast
Actual Class	Music	13.92	3.92	0.31	0	0.03
	DVD	8.25	7.97	0.74	0.02	0.03
	Typing	0.06	7.96	9.02	0.01	0
	Walk Slow	0	0.12	5.07	11.77	1.62
	Walk Fast	0.04	0.08	0.49	3.92	13.65

TABLE 4.12: 1D Neuroscale/Clustering classification confusion matrix obtained from 100 random training/testing splits of the data with 20 features and 49 hidden centres.

		Predicted Class				
		Music	DVD	Typing	Walk Slow	Walk Fast
Actual Class	Music	10.42	5.57	0.81	0.02	0.05
	DVD	4.10	9.88	1.45	0.04	0.01
	Typing	0	1.71	14.78	0.23	0
	Walk Slow	0	0.01	0.92	14.25	1.67
	Walk Fast	0.02	0.02	0.12	2.67	14.15

The following three tables show the misclassification rates from all three classifier combinations obtained from 100 runs using random splits of the data-set over varying dimensionalities. The data for the 2D projections is repeated here for completeness. Table 4.13 shows an decrease in misclassification rate as the dimensionality of projection increases, however both Table 4.14 and Table 4.15 show their lowest misclassification rate at four and three dimensional projection respectively. The loss of classification accuracy as the dimensions are reduced from the optimal down to the two that have been used is 3.65% for the PCA/k-NN combination, 2.28% for the Neuroscale/k-NN combination and 2.88% for the Neuroscale/Clustering combination.

Table 4.13: Misclassification rate for PCA/k-NN classifier projecting onto varying numbers of principal components.

Number of Components	Misclassification Rate
1	0.3615
2	0.2538
3	0.2450
4	0.2321
5	0.2173

Table 4.14: Misclassification rate for Neuroscale/k-NN classifier projecting with a varying number of Neuroscale outputs.

Number of Components	Misclassification Rate
1	0.3671
2	0.2071
3	0.2028
4	0.1943
5	0.2097

Table 4.15: Misclassification rate for Neuroscale/Clustering classifier projecting with a varying number of Neuroscale outputs.

Number of Components	Misclassification Rate
1	0.2340
2	0.2215
3	0.1927
4	0.2401
5	0.2167

4.7 Discussion

The results shown above show that there are a variety of approaches that can be taken in order to process accelerometry data, extract useful information and classify different activity categories. The plots, in Figure 4.3, of the raw data obtained in the study show that there are differences in the accelerometry traces obtained when the participant is performing different activities. These differences, which can be identified visually, are a strong indication that the different activity classes are algorithmically separable and also reinforces the evidence that accelerometers can be used to capture information about a person's activity levels.

4.7.1 Data Projection

The projection of the data, both through PCA, shown in Figure 4.5(a), and through Neuroscale, shown in Figures 4.8(a) and 4.10(a), provide more evidence that the activity classes are separable, but also that there is an underlying structure and ordering to the data.

In looking at the separability, or otherwise, of the activity classes, it can be seen from both projections that in the projection-space the data separate into rough groupings but that there is some overlap. This can be more clearly seen from the confusion matrix tables and is the main reason for the low classification results. This is to be expected to some degree, especially between classes that are similar in the physical activity recorded. The *Music* and *DVD* classes are a good example of this, as are the *Walk Slow* and *Walk Fast* classes. This is particularly evident in the PCA projection and especially from the failure of the clustering algorithm to identify the correct clusters in the data, shown in Figure 4.7.

In examining the projections for the underlying structure to the data, it can be seen from the PCA projections in both one and two dimensions, shown in Figure 4.12(a) and Figure 4.5(a) respectively, that the data-points arrange into a broad scale of activity intensity, with the least active (*Music*) to the left and the most active (*Walk Fast*) to the right. This is a significant result because the PCA projection is completely data-driven, which means that there is no reliance on using pre-labelled data and machine learning to achieve the result. There is good reason therefore to extrapolate that this approach would produce a similar mapping if applied to real-world data, as opposed to carefully controlled data.

4.7.2 Classification

The results shown for the PCA projection and k-NN classification show the activities undertaken in the study can be correctly identified using these methods to around 75% accuracy (based on an F_1 score of 0.7492). This is a modest result as far as the classification of the activity categories is concerned. It is the classes that are most separate in the projection that achieve the best results in the classification. Notably the *Walk Fast* and *Walk Slow* classes, which are the least overlapped groupings and exhibit good scores for precision, recall and F_1. In contrast, the *Music* and *DVD* classes exhibit very poor performance and a very variable performance when run on 100 different train/test splits of the data. It can be seen from Figure 4.5(a) that these two classes are very overlapped and this is the driving factor in the poor results for these two classes. The modest overall performance is not entirely unexpected however, given the large degree of class overlap in the data projection one would not expect a classification algorithm to perform particularly well. This overlap is also the driving reason behind the failure of the GK algorithm to cluster the data.

The results obtained from the Neuroscale clustering and k-NN classification show that the activities undertaken in the study can be correctly identified using these methods to around 79% accuracy (based on an F_1 score of 0.7879). This is an improvement over the PCA/k-NN approach and can be attributed to the superior performance of the Neuroscale algorithm in separating the data. This performance stems from the fact that Neuroscale was used in its machine learning form, with the subjective similarity matrix used to provide the target for the mapping. As one would expect, the increase in separation of the classes of the projected data improves the performance of the classifier. The overall performance of the classifier is hampered somewhat by its relatively poor performance on the *Music* and *DVD* classes. The box-plots shown in Figure 4.9 show that both of these classes have significantly lower scores across all three performance measures. This can be attributed to the similarity in the physical activity observed. Both consist of long periods of inactivity interspersed with brief periods of movement. This leads to the accelerometry traces looking very similar, as shown in Figure 4.3, and to the subsequent misclassification. However, performance with this approach is much better than with the PCA/k-NN approach.

The results obtained from the clustering based classification on the Neuroscale projection of the data show that the activities undertaken in the study can be correctly identified using these methods to around 79% accuracy (based on an F_1 score of 0.7935). This is a very similar result to that obtained with the k-NN classifier. The only real difference between the two classification results being the larger range of values for the cluster based classification performance measures and the higher number of low lying outliers in the performance characteristics. This implies that the cluster based classifier is less robust to differences in the composition of the training and test sets that the k-NN classifier.

There is one further point of interest with regards to the PCA projection of the data and classifier accuracy. It was discovered during the data analysis that, contrary to expectations, the performance of the PCA projection and k-NN classification were not improved by limiting the feature list. For some settings of k, limiting the feature list did improve the classification accuracy, but for the optimal settings the full feature list was used.

4.7.3 Higher-dimensional Projections

It can be seen from the tables in Section 4.6.4 that a better classification performance can be achieved by projecting and classifying in higher dimensions. For example, the PCA/k-NN approach works best in the five-dimensional space and the Neuroscale/Clustering approach works best in the three-dimensional space. This is likely due to a greater degree of separation of the data clusters in a higher dimensional space. However, there is a fundamental limit on the classification performance that stems from the natural overlap in the data classes. Furthermore, whilst a better classification accuracy can be achieved in higher dimensional space, this has little further bearing on the PAM system since, as discussed in the next section, a higher-dimensional space is unsuited for the PAM system.

4.7.4 Accelerometry for PAM

In order for an accelerometer to be successfully used in the PAM system the data that it provides must ultimately fit into the data processing architecture, discussed in Chapter 3, and the device and processing methods must also adhere to the system requirements discussed in the same chapter. In order to fit into the data processing architecture, the data must be processed to a level where a single number can be used to derive information about any given (short) segment of time and therefore allow for a time-series to be generated that will show behavioural patterns. Additionally, in order to meet the system requirements for usability and not being obtrusive to the end user, it is highly desirable that any processing algorithm does not need to be trained by the end user. In essence, the processing must either be trained automatically (unsupervised learning), be trained before being deployed (supervised learning), or have no training element at all. Of the three choices, unsupervised learning is the more favourable option. An unsupervised learning approach allows the PAM system to make use of an automated training phase, which fits in with the goal of determining ‘normal’ behaviour patterns, whilst at the same time not requiring any input from the user.

The choice of unsupervised learning leads to the selection of the PCA projection, since this can be performed entirely on the raw data, whereas Neuroscale, in the form in which it is used here, requires a set of labelled training data. The choice of unsupervised learning also leads to the rejection of both forms of classification; k-NN is inherently a supervised learning algorithm, and the cluster-based approach, whilst being unsupervised does not work on the PCA projected data. Instead, the single dimension projection through PCA can be used to generate a single number for any given time window. This takes advantage of the fact that the PCA projection naturally organises the data according to intensity of activity, and therefore produces an indicative scale onto which new data-points can be projected and an activity score assigned. Higher-dimensionality spaces cannot be used as they would not naturally reduce to a single number and are therefore an unsuitable fit with the overall data processing architecture.

Furthermore, the PCA projection will have to be made using the full feature set, as no classifier will be used with which the feature set can be trimmed. This is not an issue however since it has been shown that on this data PCA performs optimally with the entire feature set. It could also be remarked that this approach could be replaced with simply taking a signal power measurement and further work is presented in Chapter 6 that compares this measure against the PCA approach for data obtained in the technical trial.

There is a further complication which must be addressed at this stage, which is the fact that the device used for this preliminary work and the device used in the technical trial are different and worn on different parts of the body. It is reasonable to assume that the performance of the single dimension PCA projection on the different devices will be similar. The validation of this claim is presented in Chapter 6 along side the comparison of signal power and the PCA projection.

4.8 Summary

The work that has been presented in this chapter has been on the investigation of triaxial accelerometers as sensors in the PAM system and an exploration of the ways in which the data can be manipulated; and on the selection of suitable algorithms to incorporate the accelerometer data into the PAM system. The data gathered through a controlled data gathering study have been analysed with both PCA and Neuroscale and classification of the data into one of five activity categories has been performed with k-NN and a clustering approach based on GK clustering. Accuracy results of around 79% have been achieved in both classifiers from the Neuroscale projection.

Moving forward into using an accelerometer for the PAM system, the choice of using PCA projection onto a single dimension has been made. By using this method, a scale of activity can be automatically created for each end user and subsequent activity can be projected into this space and an activity score assigned. This will enable the accelerometer data processing to fit well within the data processing framework discussed in Chapter 3. The next chapter examines the use of a camera for change detection in complex environments and the integration of both the camera and accelerometer sensors to the PAM system is discussed in Chapter 6 along with further work that validates the choice of the PCA processing approach.

Chapter 5

Image Change Detection

5.1 Introduction

As one of the sensors for the PAM system, a camera is being used, as a single sensor, to monitor complex environments that may otherwise require multiple sensors. The aim of using the camera in this way is that it can be used to monitor various AOIs in the environment instead of the collection of sensors that would otherwise be needed. The camera is set-up in such a way as to see the entire room and image change detection is performed on successive images to derive a change mask, which shows areas of change in the scene. The AOI can be checked for changes and a separate data-stream of changes in each AOI can be generated. The kitchen is a good example of this sort of situation as it contains AOIs such as the cooker, sink and cupboards, which can all be monitored with a single camera.

This chapter discusses the use of a camera in the PAM system, the selection of an appropriate image change detection algorithm and how the choice of image resolution affects the processing speed and results. The image change detection algorithm must be robust to changes in lighting conditions and detect only the changes that are of interest. Seven image sequences were recorded to provide a range of different test conditions and image resolutions and four image change detection algorithms were tested. Results are presented that show the effectiveness of the algorithms under these test conditions. Some of this work was presented at PG Biomed09 [2] and a copy of the accompanying paper is included in Appendix A.

5.2 The Problem

In order for the camera to be a useful sensor in the PAM system it must fit into the data processing architecture outlined in Chapter 3. This means that there must be some data processing that takes the output from the camera and produces one or more time-series that show behavioural patterns over the course of a single day.

Image change detection, as discussed in Chapter 3, is a method that can be used to determine the differences between two successive images and therefore identify areas of activity in the scene. For the PAM system the camera is used in this way and areas of change in predefined AOIs are looked for. Each AOI is then treated as a separate data-stream. By totalling the number of changes in a set time-window it is possible to create a time-series suitable for inclusion in the PAM system's data processing architecture.

The principal problem being addressed in this chapter is therefore the selection of a suitable image change detection algorithm that detects the changes that are important to the PAM system, those relating to the user and their actions, and disregards any others. This breaks down into three issues that have to be overcome. Firstly, the difference in two images will not necessarily reflect a change in the physical scene that makes up the image, but may be caused by a difference in lighting conditions within that scene, be this ambient light or light from a specific source. A secondary concern is that an object that is stationary in a scene will be absorbed into the background estimate over time; this is desirable for inanimate objects, but not desirable if the person(s) being monitored are stationary for some time. Finally the resolution of the image and specifically the effect this has on processing time and resource use is also a concern as a smaller image is faster and less resource intensive to process, but using a smaller image could result in a loss of information. This loss of information could be detrimental either to the definition of AOIs or to the detection of change in the images.

5.3 Method

In order to identify a suitable change detection algorithm a set of four test sequences were recorded to capture a variety of lighting and physical conditions and four different processing algorithms were evaluated against each other using the different image sequences. In addition, a further three test sequences were recorded using different image resolutions to test the algorithm performance on different image sizes.

5.3.1 Camera Specification

The camera used for the image change detection work is an Edimak IC-1520DPg Digital Pan / Tilt Wireless Network Camera. The full technical specifications for the camera are shown in Table 5.1.

Table 5.1: The technical specification for the camera.

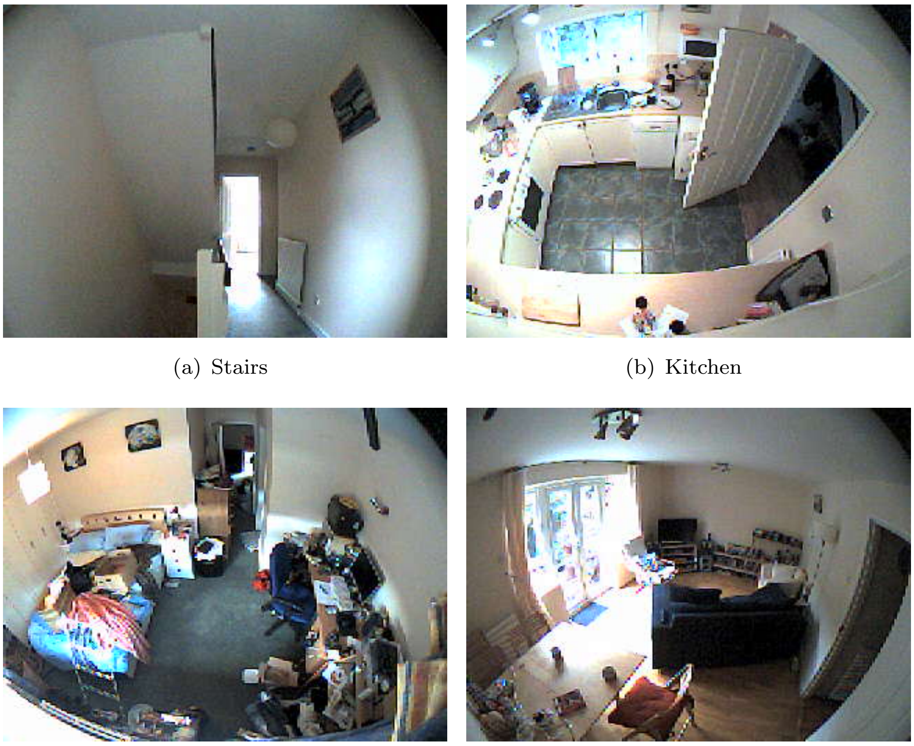
Item	Specification
Model Number	IC-1520DPg
Memory	2 MB Flash and 16 MB SDRAM
Image sensor	1.3M Pixels CMOS sensor
Lens	Manual Focus, F=2.0
Field of view	128 degrees horizontal and 99 degrees vertical
Distortion	¡ 98%
Button	Reset
Transmit power	18dBm $\pm$ 2dBm
Antenna	RP-SMA Detachable 3bNi Antenna
Power	DC12V, 1A Power adapter
LED Indicators	Power, LAN and WLAN
Dimension	$114 \times 79 \times 50$ mm
Humidity	10 ˜90% (Non-Condensing)
Temperature	0 ˜45° Celsius
Certification	CE, FCC

5.3.2 The Test Sequences

The test sequences were recorded to try and cover the key problem areas for any potential algorithm. These include, slow and fast lighting changes, moved objects (i.e. those that have been introduced to the scene and should then be background), stationary subjects (i.e. people who enter the scene and stand still, who should not be identified as background). The test sequences recorded as a test set for lighting and physical changes are *Stairs*, *Kitchen*, *Bedroom* and *Sitting Room*. In addition to these, a set of high, medium and low resolution sequences of the same subject activity were recorded to test the algorithm performance at different resolution. All the image sequences were recorded at 0.1 Hz. This frame rate was selected for two reasons. Firstly, the camera is being used to identify activity and meaningful activity is highly unlikely to be missed with this sampling rate, whilst transient events, which are of less interest, are not likely to be picked up. Secondly, the 0.1 Hz rate allows the images to be processed in real time on a low-cost system where other processes may be competing for system resources, such as the scenario with the PAM system. The following table outlines the major features of each test sequence and Figure 5.1 sows background images from the four test sequences. The test sequences are shown in the figures that follow; *Stairs* in Figure 5.2, *Kitchen* in Figure 5.3, *Bedroom* in Figure 5.4 and *Sitting Room* in Figure 5.5. The reading order for all the image sequences is across (left to right) then down.

TABLE 5.2: The test sequences used in the camera experiments.

Sequence	Description
Stairs	The subject is very mobile in the sequence and appears in isolated frames. There are ambient lighting changes and a light spot on the wall.
Kitchen	There is some object movement in this sequence and the subject is stationary for a period. There is also a sharp lighting change.
Bedroom	The sequence contains lots of objects being moved around, including the subject, and has fairly consistent lighting.
Sitting Room	The subject is still for extended periods and there are changes in ambient lighting.
Low/Med/High Resolutions	This sequence has the subject performing a set series of actions in the kitchen. The three sequences are at different resolutions (640 × 480, 320 × 240 and 160 × 120).



(a) Stairs (b) Kitchen

(c) Bedroom (d) Sitting Room

FIGURE 5.1: Background images from the four test sequences showing *Stairs*, (a), *Kitchen*, (b), *Bedroom*, (c), and *Sitting Room*, (d).

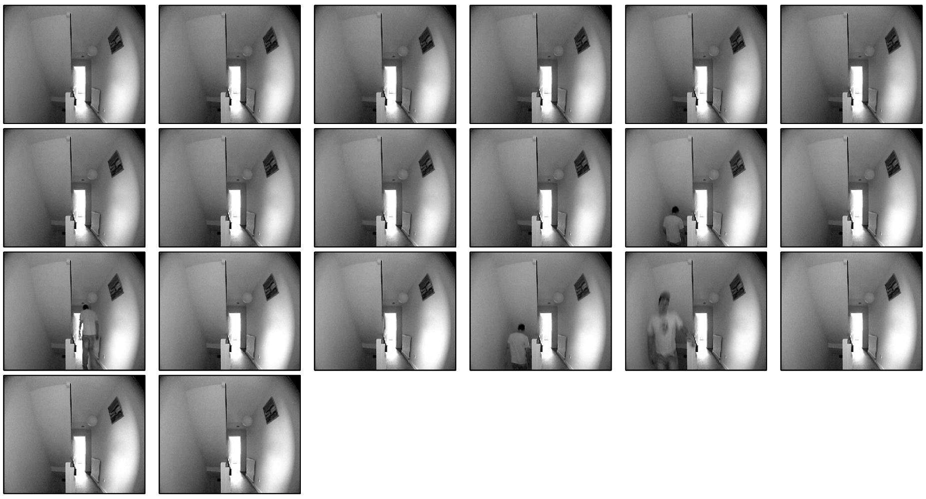


FIGURE 5.2: The stairs image sequence showing the subject moving up and down the stairs. The subject is in isolated frames in this image sequence. There is a light spot on the right hand wall and some ambient lighting changes.

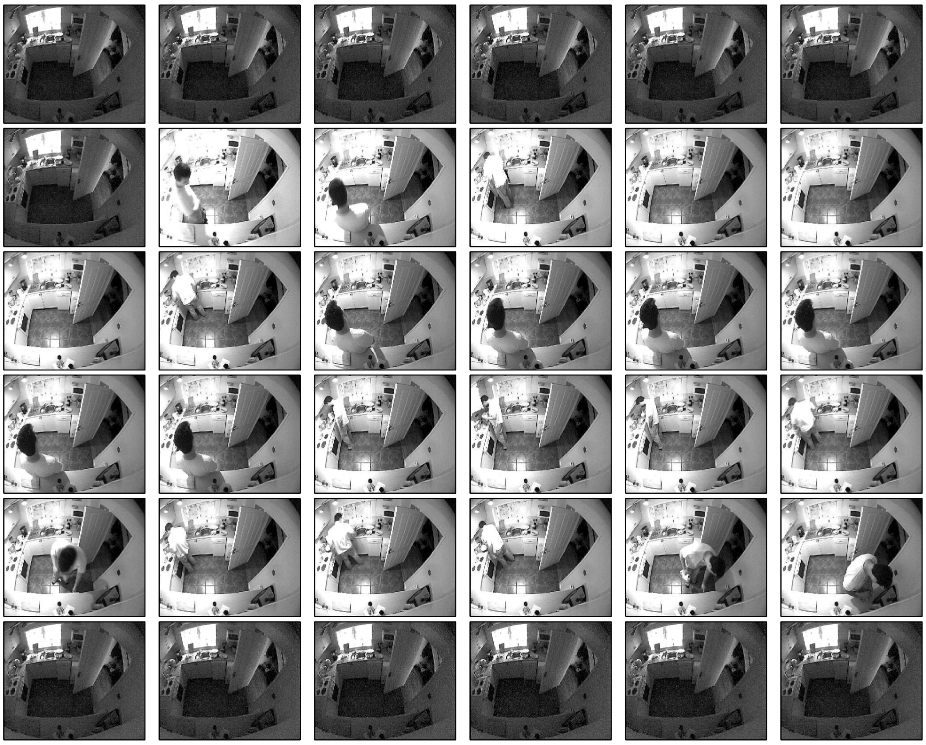


FIGURE 5.3: The kitchen image sequence showing the subject moving around in the kitchen. The subject is present in a large number of images in the sequence and remains still across several images. There are two sharp lighting changes in this image sequence resulting from the kitchen light being switched on and then off.

FIGURE 5.4: The bedroom image sequence showing the subject moving things around the bedroom. There are a large number of items being moved around in this image and consistent lighting levels throughout.

FIGURE 5.5: The sitting room image sequence showing the subject first moving about and then being still in the sitting room. The subject is still for extended periods of time and there are some lighting changes in this image sequence.

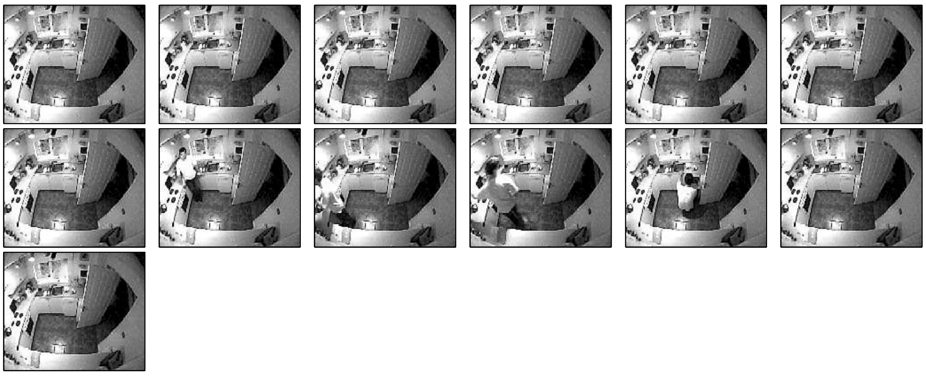


Figure 5.6: The low resolution image sequence recorded at 160×120 pixels showing the subject moving around in the kitchen.

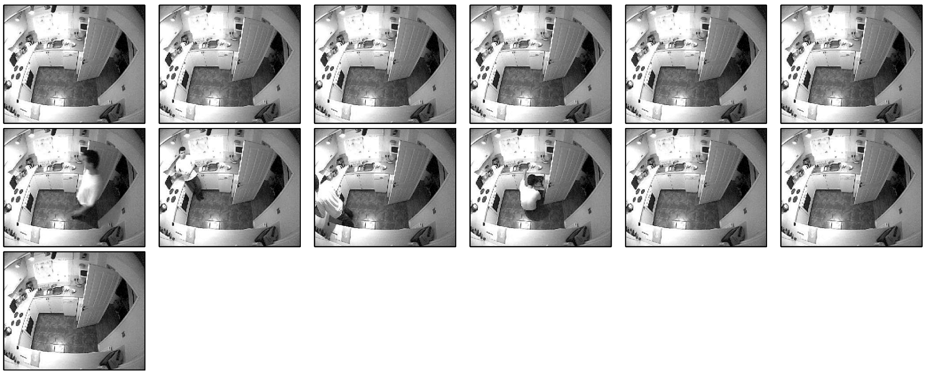


Figure 5.7: The medium resolution image sequence recorded at 320×240 pixels showing the subject moving around in the kitchen.

Figure 5.8: The high resolution image sequence recorded at 640×480 pixels showing the subject moving around in the kitchen.

5.4 Evaluation Criteria

The aim of the change detection algorithm trial is to identify the algorithm that best detects changes in the physical scene, irrespective of lighting changes. A secondary concern is that of computational time; it is preferable that an algorithm execute quickly as this places less strain on the system, especially in an environment such as PAM where there will be several system components competing for the same resources.

Tied into these two concepts is that of image size, which affects both the level of detail available for setting AOIs and the speed at which an algorithm will run. It is necessary that an image be sufficiently detailed that AOIs can be suitably defined, and that a processing algorithm is capable of discerning changes effectively on that image size.

The criteria for algorithm assessment are set out in the following list. They are a subjective set of assessment criteria but cover the key areas for an algorithm to be successful for the PAM system.

1. Invariance to changes in lighting conditions, including both slow and fast changes.
2. Ability to correctly identify changes in the image.
3. Change detection on smaller image sizes.
4. Algorithm run time.

5.5 Image Change Detection Techniques

For this work, the general structure of the algorithm is to take a temporal median and use this as the background. A temporal median is the median value for each pixel across a set of images and will extract the background pixel provided that the background is visible in more than half of the frames in the set. The background image and current image are then differenced, using one of the algorithms described below, and a black and white image (the change mask), which shows up areas of change, is produced. A set of pre-defined AOIs are cross referenced with the change mask to identify any activity in the AOIs. The image processing uses the following steps:

1. Background estimation with a temporal median of 3 images.
2. Background subtraction / change detection using a change detection algorithm.
3. Activity detection within pre-defined AOIs.

5.5.1 Pixel Difference Algorithm

The Pixel Difference (PIXD) algorithm is the simplest processing algorithm and serves as a naive benchmark for the other algorithms to be compared against. The PIXD algorithm works on a gray-scale image representation with the image being transformed from an RGB space with Matlab's inbuilt rgb2gray(.) function. Pixel values are in the range 0–255. The background estimation is not used in PIXD and the change mask is created by differencing the current and previous image, thresholding and median filtering as follows.

1. Difference current and previous images.
2. Threshold the difference image with a cut-off at a pixel value of 20.
3. Apply 5-pixel square median filter to remove speckle noise.

5.5.2 In-House Algorithm

The In-House (IH) algorithm is based on background estimation and pixel differencing to determine the difference between the background image and the current image. The IH algorithm works on a gray-scale image representation with the image being transformed from an RGB space with Matlab's inbuilt rgb2gray(.) function. Pixel values are in the range 0–255. The image is then thresholded and median filtered to clean up the result. This algorithm has been constructed as a slightly more advanced change detector than PIXD. A background estimation is used and the current image and background are differenced. A median filter is then applied to remove speckle noise. The algorithm follows the following steps.

1. Calculate difference image between background and current image.
2. Threshold the difference image with cut-off at a pixel value of 20.
3. Apply 10-pixel square median filter to remove speckle noise.

5.5.3 Wronskian Change Detector

The Wronskian Change Detector (WCD) [28] is based around the idea of linear independence. The image representation used is the Y component (luminance values) of the YCbCr color-space and the image must be converted from an RGB representation to a YCbCr with Matlab's inbuilt function. Pixel values are in the range 0–255.

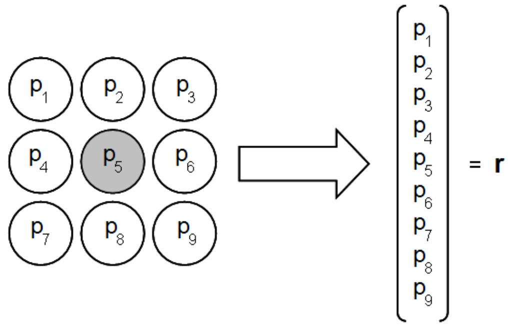


FIGURE 5.9: Computation of a region of support, **r** for a pixel showing the central pixel, its neighbours and how they combine into the region of support vector.

An image I is converted to a vector model where each pixel is represented by a vector of itself and a square of surrounding pixels, known as the region of support, denoted as **r**. The construction of **r** is shown in Figure 5.9.

To compare two images for physical changes in the scene a pixel by pixel assessment of physical change is performed. Regions of support are computed for each pixel in the current image, I_c, and the background image, I_b. Since it is only regions of support for identically located pixels from both images that are compared, no location subscript is necessary and the notation $\mathbf{r_c}$ will be used for a region of support computed from a pixel in the current image, and $\mathbf{r_b}$ for the corresponding region of support from the background image. For each region of support it is necessary to identify the individual vector components and these will be denoted as $x_i \in \mathbf{r}_b$ and $y_i \in \mathbf{r}_c$.

Now, to perform the pixel by pixel assessment of physical change, the regions of support for two corresponding pixels from I_b and I_c are used to compute the Wronskian, W, for a pixel such that

$$W\left(\frac{x_i}{y_i}\right) = \frac{1}{n}\sum_{i=1}^{n}\frac{x_i^2}{y_i^2} - \frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{y_i}, \tag{5.1}$$

where n is the number of pixels in the region of support and the term $1/n$ is used to normalise the vector dimensions. If the result, W, is less than a threshold level then no physical change has taken place in the image at the pixel location from which the regions of support were computed. The computation of W is applied systematically to each pair of pixels in the two images to be compared to produce a change mask which shows the physical changes in the image. The threshold value needs to be tuned for specific situations and circumstances and for the work presented here a value of 0.4 was used.

There is a further addition to the WCD to improve its ability to distinguish between high illuminance intruders (objects that have appeared in the image that have a high illuminance) and global illuminance changes. The pixel illuminance ratio,

$$\mathcal{I}_i = \frac{x_i}{y_i}, \tag{5.2}$$

is calculated for pixels in the reference image, $x_i \in I_r$, and current image, $y_i \in I_c$. By computing the number of $\mathcal{I}_i < 1$ and comparing this to a threshold a global lighting change can be identified. This threshold is set to 1/5 of the total number of pixels in the image. If the global illuminance has changed, the standard WCD, $W(x_i/y_i)$ is used. If the global illuminance has not changed, then both the standard WCD and the inverse, $W^*(y_i/x_i)$, are calculated, since the reverse will then detect high valued ratios, and will detect high illuminance intruders. The value taken is the maximum of W and W^*. The WCD algorithm now becomes

1. Apply WCD with 0.4 threshold to background and current images.
2. Apply 10-pixel square median filter to remove speckle noise.

5.5.4 Homomorphic Filtering Algorithm

The homomorphic filtering algorithm (HM) is based on [106], which uses homomorphic filtering on an image to separate the illuminance and reflectance components of the image. The HM algorithm assumes that all surfaces in the image are Lambertian. That is, that the apparent brightness of the surface is the same regardless of the viewing angle. A Lambertian surface is therefore one where all reflections are diffuse; where a beam of light hitting the surface will scatter in many directions. Dull, non-shiny surfaces, such as unvarnished wood, are a good example of this. Under this assumption a function is formed to relate the pixel intensity to the scene illumination and surface reflectance such that

$$y(p) = i(p) \times r(p), \tag{5.3}$$

where $y(p)$ is the pixel intensity, $i(p)$ the illumination component and $r(p)$ the reflectance component for a pixel p in the image. In order to extract the reflectance component, which provides information about the objects in the scene, a homomorphic filter can be applied to separate the illumination and reflectance components. This is achieved by taking the log of $y(p)$ such that

$$\log(y(p)) = \log(i(p)) + \log(r(p)). \tag{5.4}$$

A low-pass filter can subsequently be applied to $\log(y(p))$ to extract $\log(i(p))$. Subtracting this from $\log(y(p))$ yields the reflectance component $\log r(p))$. Both the illumination component and reflectance component can then be exponentiated back and the filter has been applied. Figure 5.10 shows how the HM algorithm is applied.

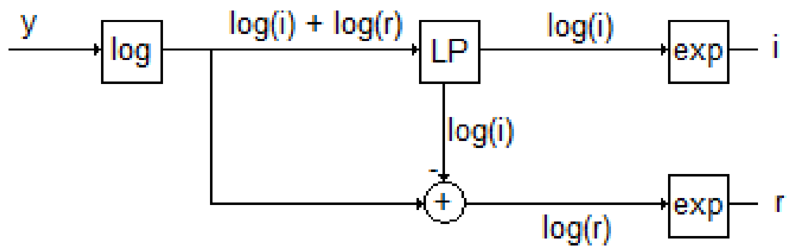


FIGURE 5.10: The application of the HM algorithm to split an image into illumination and reflectance components.

For the processing of the camera images, the HM algorithm is performed on a gray-scale image representation, which is obtained using Matlab's inbuilt rgb2gray(.) function. Pixel values are in the range 0–255. The HM algorithm now becomes

1. Apply homomorphic filter to current and background images.
2. Difference current and background images.
3. Threshold difference image at pixel value of 20.
4. Apply 10-pixel square median filter to remove speckle noise.

5.6 Results

Figures 5.11 to 5.17 show the results of applying the three main processing algorithms to the stairs and kitchen test sequences and the result of applying the PIXD algorithm to the kitchen sequence. The other results sequences can be found in Appendix B. Rather than examine each image, the results for each algorithm are discussed in relation to the assessment criteria stated above. When reference needs to be made to a specific frame in an image sequence, it will be noted first by row going down, then by column going right for example frame (2,3) means the third frame on the second row down for a given image sequence.

The PIXD algorithm performs badly in respect to both the first and second criteria, and was not tested against the third and fourth due to its (predictably) poor performance. Figure 5.11 shows ghosting, where a change in one frame is carried across two frames, which is particularly noticeable in frames (2,6), (3,2) and (3,6). There is also significant light-induced noise in Figure 5.11, which can be seen in frames (3,2) and (3,4) and from (3,4) to (3,6).

The IH algorithm performs significantly better than the PIXD algorithm in terms of identifying the correct changes, in that it does not suffer from ghosting. It does however suffer from object fade, where a stationary object will slowly fade into the background as the background estimation moves forwards. This is an issue for all the algorithms and can be seen very clearly in all of the kitchen sequences (Figures 5.15 to 5.17) in frames (3,3) to (3,6). In terms of light-induced noise, the IH algorithm does no better than the PIXD algorithm and light-induced noise can be seen in several places, notably in Figure 5.12 frame (3,5) and Figure 5.15 frames (2,5) and (2,6).

Both the WCD and the HM algorithms perform similarly in terms of correct change detection and invariance to light-induced noise. The HM performs slightly better in identifying the correct areas, as can be seen in a contrast of Figures 5.14 and 5.13 on frame (3,5), where the HM algorithm identifies more of the subject. The HM algorithm under-performs the WCD in terms of light-induced noise in low light conditions. This can be seen in many of the sequences, but is particularly noticeable in a contrast of Figures 5.17 and 5.16 on frames (1,3) to (1,6) and (6,3) to (6,6), where the WCD records no change, whilst the HM records noise. One final point regarding both algorithms is that where an abrupt lighting change occurs the HM algorithm whites-out on both a dark to light change and a light to dark change, whereas the WCD only whites-out on a dark to light change. The WCD's behaviour on a light to dark change is to identify those areas that have not changed in intensity. This behaviour can be seen in Figures 5.17 and 5.16 on frames (6,1) and (6,2).

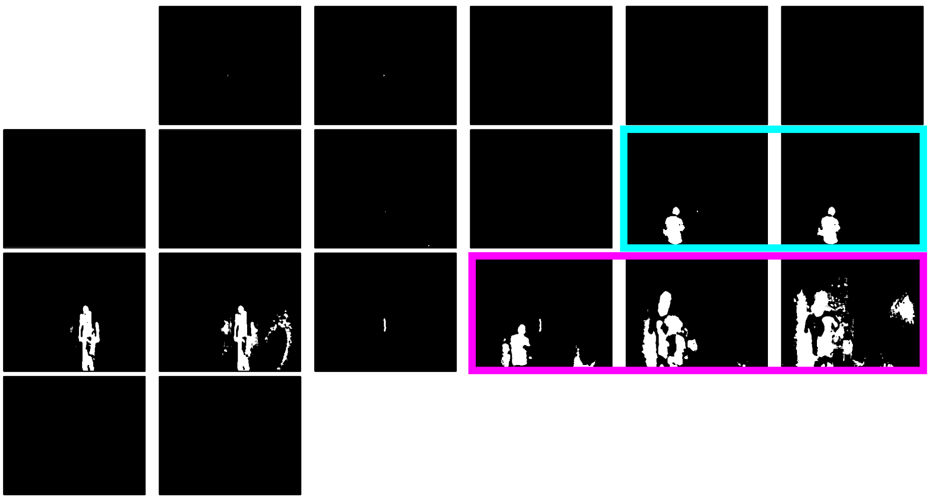


FIGURE 5.11: The stairs image sequence processed with the PIXD algorithm. The image sequence shows poor results; lighting induced changes can be seen, highlighted pink, in frames (3,4) to (3,6) and ghosting can be seen, highlighted blue, in frames(2,5) and (2,6).

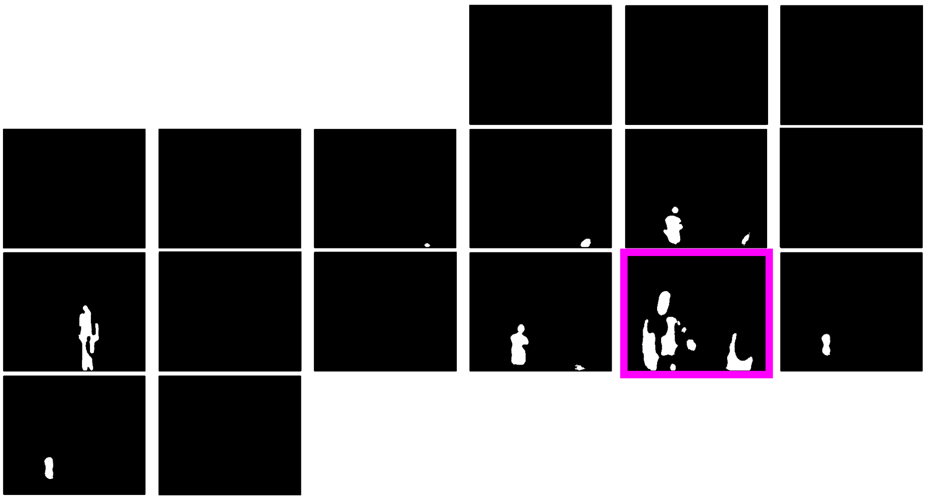


FIGURE 5.12: The stairs image sequence processed with the IH algorithm. The image sequence shows average results; no ghosting occurs, but light induced noise can be seen, highlighted pink, in frame (3,5).

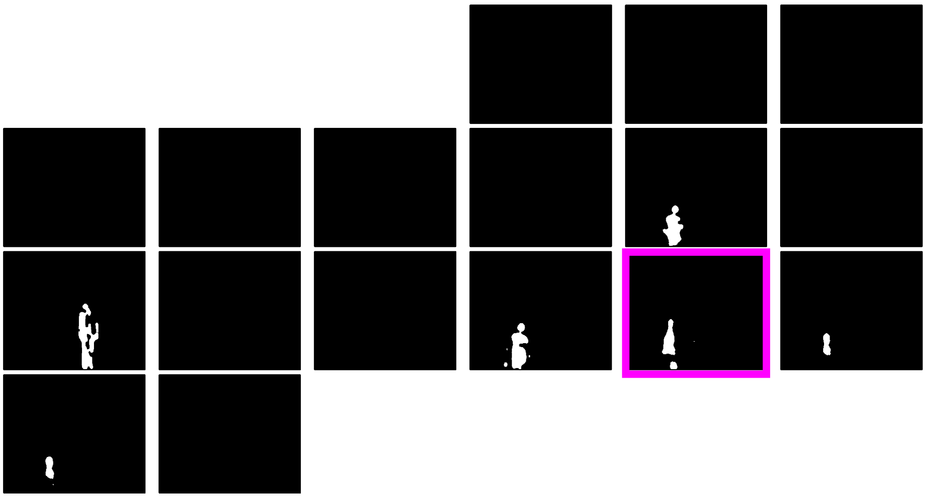


FIGURE 5.13: The stairs image sequence processed with the WCD algorithm. The image sequence shows good results, although only a partial detection of the subject is made in frame (3,5), highlighted pink.

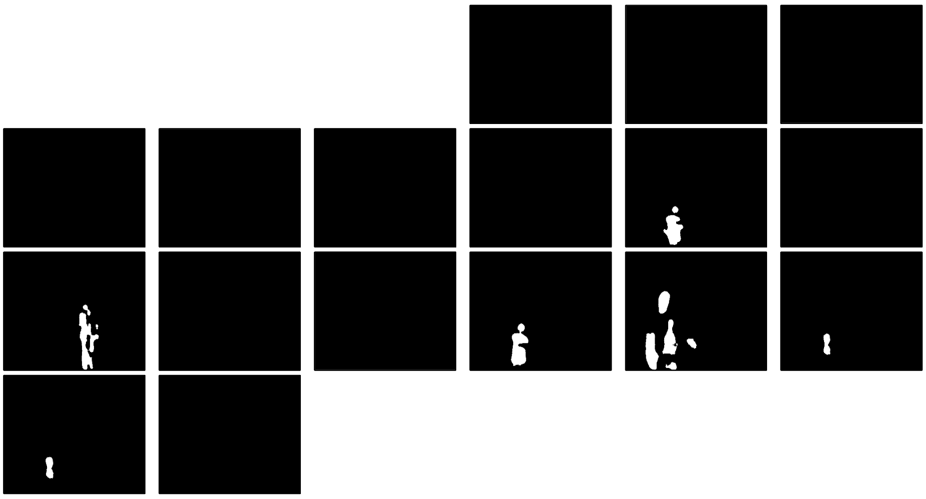


FIGURE 5.14: The stairs image sequence processed with the HM algorithm. The image sequence shows good results; no ghosting or light induced noise is evident.

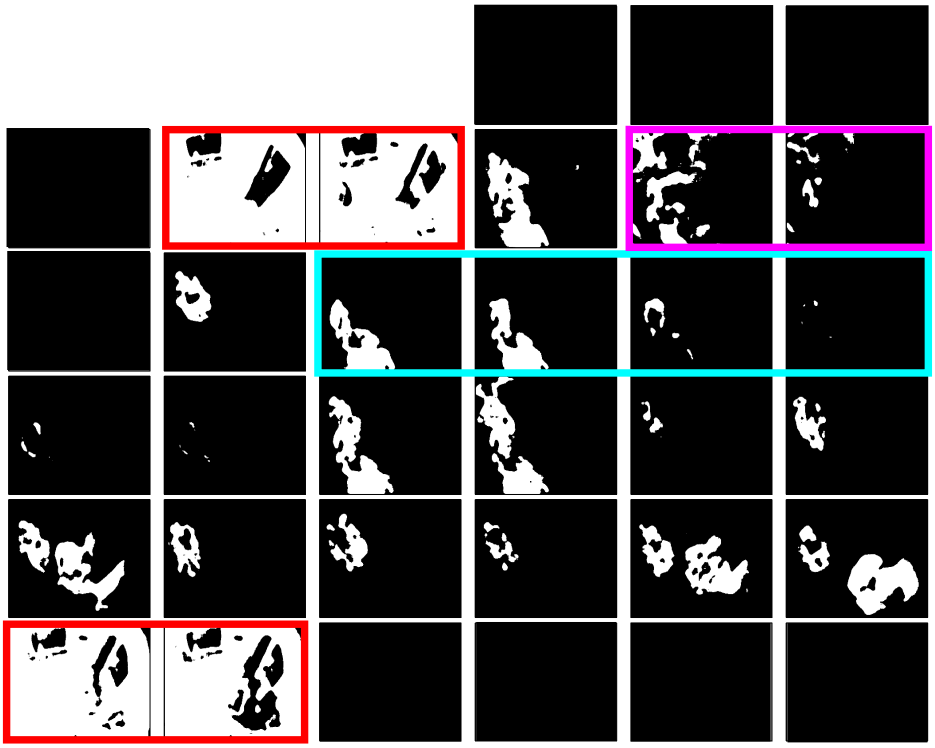


FIGURE 5.15: The kitchen image sequence processed with the IH algorithm. The image sequence shows poor results; object fade can be seen, highlighted blue, in frames (3,3) to (3,6), lighting induced noise can be seen, highlighted pink, in frame (2,5) and (2,6) and white-out can be observed in frames (2,2), (2,3), (6,1) and (6,2), highlighted red, corresponding to the two sharp lighting changes in this image sequence.

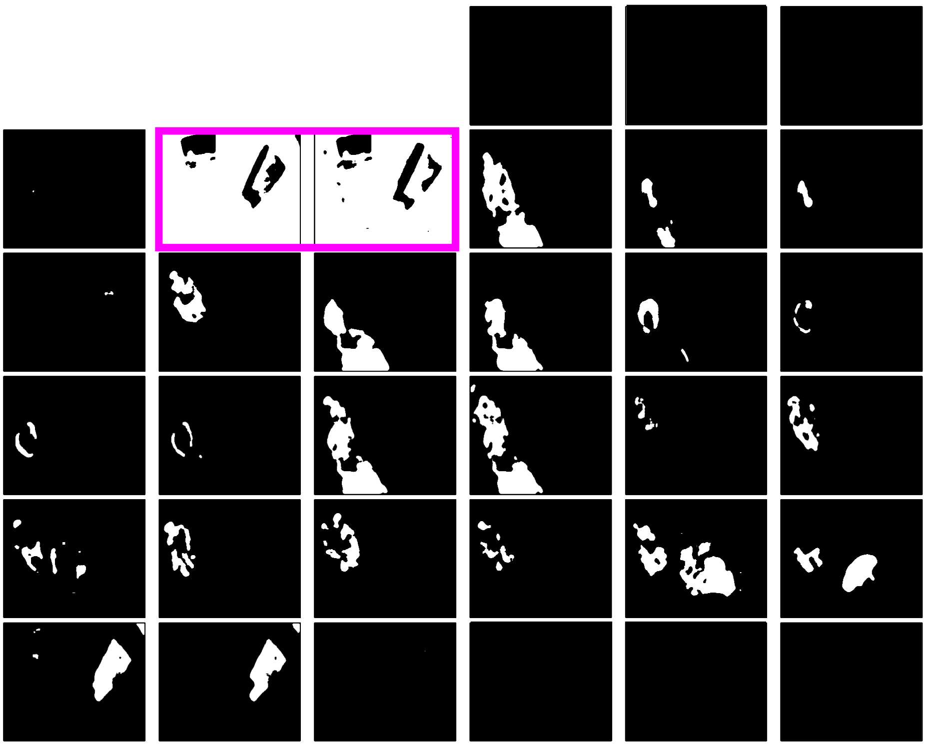


Figure 5.16: The kitchen image sequence processed with the WCD algorithm. The image sequence shows good results; there is little lighting induced noise and white-out can only be observed in frames (2,2) and (2,3), highlighted pink.

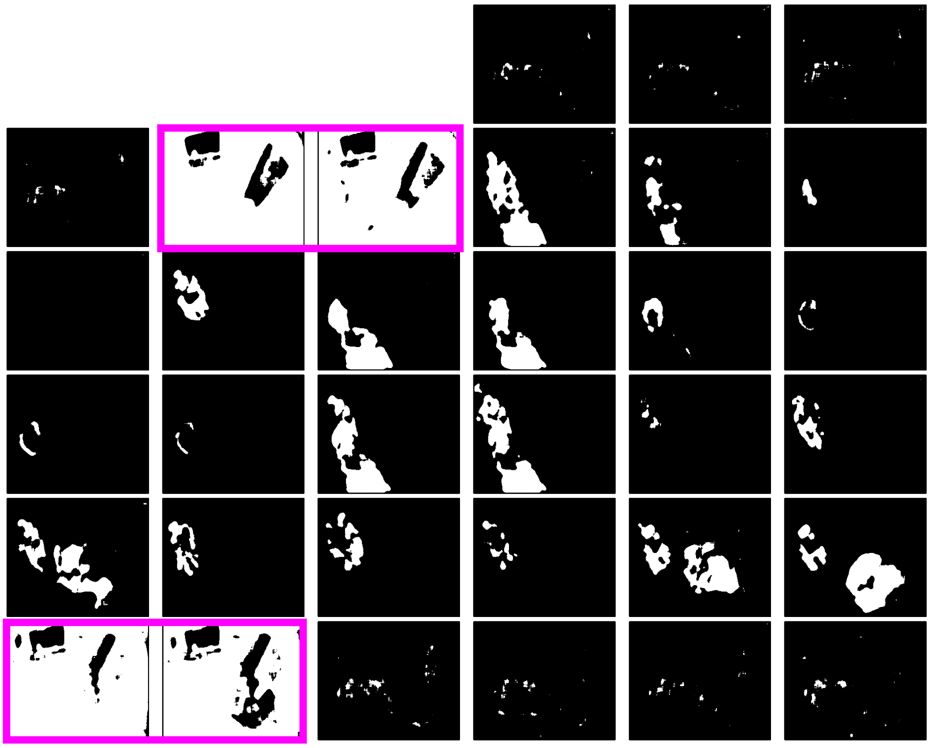


Figure 5.17: The kitchen image sequence processed with the HM algorithm. The image sequence shows good results although there is a low level of light induced noise throughout the image sequence and white-out can be observed in frames (2,2), (2,3), (6,1) and (6,2), highlighted pink.

Figures 5.18 to 5.23 show the result of applying the WCD and the HM to the low, medium and high resolution images. The change detection performance on the medium and high resolution images is adequate and produces good responses. The performance on the low resolution image sequence is inadequate, with the change mask producing vague, ill defined responses. The lack of detail available in the low resolution images, as shown in Figure 5.6 prevents the image processing algorithms performing properly and may cause problems when defining AOIs.

The algorithm run times for processing the different sized images are given in Table 5.3. Both algorithms were coded in Matlab with as efficient an implementation as possible and run on the same PC with no other tasks running. Times shown are an average of five runs. The HM algorithm clearly performs the fastest on all the image sizes. The IH and PIXD algorithms were not run-time tested due to their poor performance in the other areas of the testing. The HM algorithm makes some use of inbuilt Matlab functions, which the WCD does not do. This may contribute to the faster run speed that the HM algorithm exhibits.

TABLE 5.3: Average run times in seconds for the WCD and HM algorithms on different image sizes.

Image Size	HM	WCD
Low (160×120)	0.064	0.232
Medium (320×240)	0.236	1.206
High (640×480)	0.942	4.485

FIGURE 5.18: The low resolution image sequence processed with the WCD showing poor change detection results.

FIGURE 5.19: The low resolution image sequence processed with the HMF showing poor change detection results.

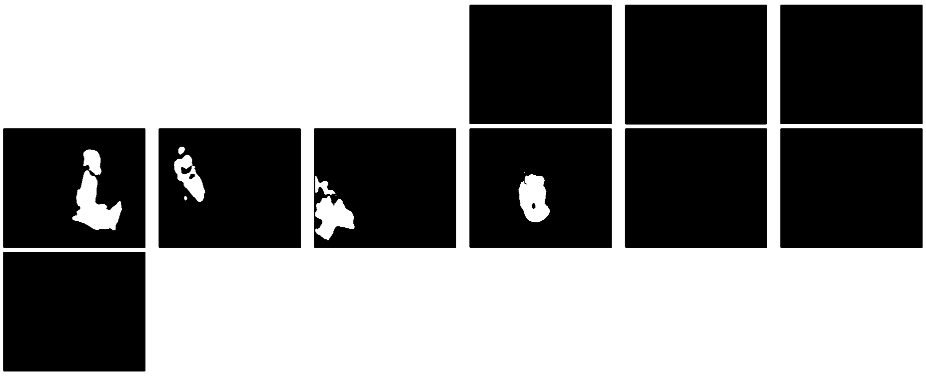


FIGURE 5.20: The medium resolution image sequence processed with the WCD showing adequate change detection results.

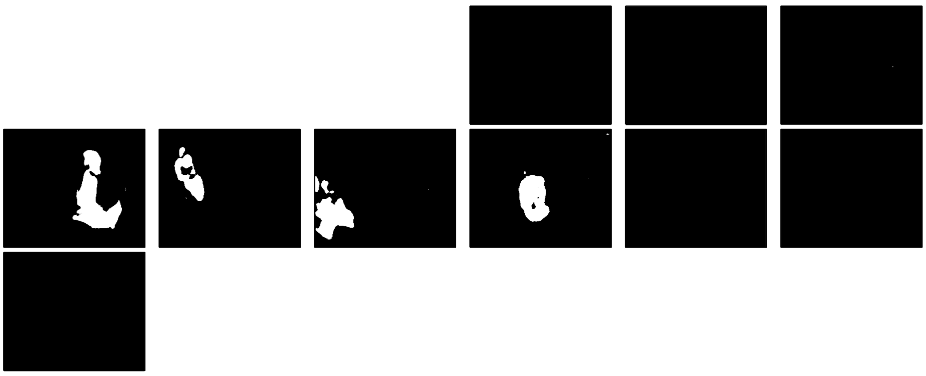


FIGURE 5.21: The medium resolution image sequence processed with the HMF showing adequate change detection results.

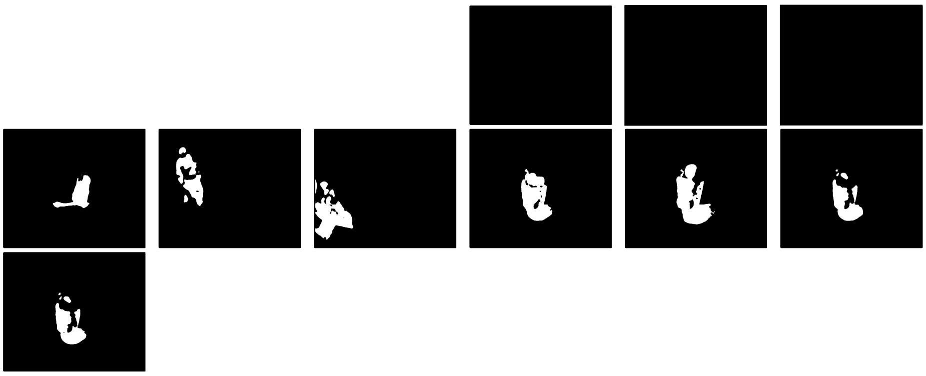


FIGURE 5.22: The high resolution image sequence processed with the WCD showing adequate change detection results.

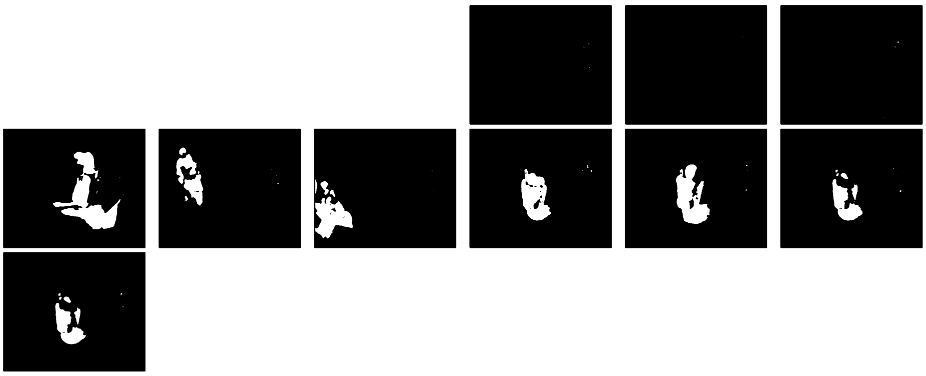


FIGURE 5.23: The high resolution image sequence processed with the HMF showing adequate change detection results.

5.7 Discussion

The results of the algorithm trial show that both the WCD and the HM algorithms perform well when compared to the four assessment criteria. Both algorithms are capable of identifying changes in the physical scene and are quite resistant to light-induced noise. The HM algorithm is slightly better at determining the correct physical changes under bright lighting conditions and the WCD struggles if the illuminance of the background and the illuminance of an intruding object are similar. The WCD on the other hand performs better in low light levels where the HM algorithm becomes prone to light-induced noise. The WCD also has a better performance at abrupt light-to-dark lighting changes.

In terms of computational load and image size, the HM algorithm is significantly faster on all image sizes. The change detection quality on medium resolution images is good enough however, that these can be used without a loss in detection capability. These factors combine to make the WCD algorithm used on the medium resolution images the combination of choice for the PAM system.

5.7.1 Limitations

There are some limitations with the current set-up that should be discussed at this point. Firstly, as shown in the results, all the algorithms suffer from object-fade. This occurs for any object placed in the scene and for the subject if they stand still for long enough. This becomes a problem when the subject fades into the background and is not picked up. One remedy for this is to increase the number of images in the background estimation. This has the effect of delaying the object-fade and requires the subject to be stationary for a longer time period before they fade. In doing this however, a secondary problem is introduced whereby inanimate objects that are placed in the scene do not fade into the background quickly enough and are registered as changes. Since the camera is intended for use in the kitchen environment it is expected that the subject will not be still for extended periods of time, so the number of images used in the background estimation can be kept small.

A second limitation of the current system is the quality of the camera, especially in low light conditions. If two images from the camera of exactly the same scene, under the same lighting conditions are taken, there will be a difference in the two images as the camera CCD will not record exactly the same illuminance for each pixel. For the camera that is being used, under bright lighting conditions this change in pixel values is minimal. Under low lighting conditions however, the pixel change is substantial and it is this that is the main cause of the shortcomings of the HM algorithm under low lighting conditions. There is not a great deal that can be done to overcome this problem and the use of the WCD side-steps the issue altogether. Were a better camera available, the

HM algorithm could be used, but the only benefit of doing so would be to reduce the computational load of the system, which is not a priority at this time.

5.8 Summary

The work presented in this chapter has been on the selection of an appropriate image change detection algorithm for the PAM system, subject the requirements for the PAM system and to the requirements for an image change detection algorithm. The results presented show that it is possible to detect physical changes in a scene and that the performance of the WCD and HM algorithms makes both suitable for the PAM system. Both algorithms are robust to most lighting changes and by using a background estimator the effects of ghosting can be eliminated. The issue of object fade applies to both algorithms but in the context of the PAM system, where the camera will be used in the kitchen and the subject is unlikely to remain still for extended periods, this is not a big issue. The only differentiator between the two algorithms is the higher resistance of the WCD to sharp lighting changes and to light induced noise under low lighting conditions, which mean that this is the algorithm of choice for the PAM system.

The next chapter presents the technical trial. The integration of the sensors into the PAM system and the data analysis for the sensors that have not been covered so far in this thesis is discussed. The data analysis that allows for the sensor fusion and behavioural pattern detection is presented and the results of the technical trial are shown.

Chapter 6

Technical Trial

6.1 Introduction

As part of the development and testing of the PAM system, a technical trial has been carried out to test and evaluate the hardware and software in the system so that any issues can be discovered and corrected. The goals of the technical trial were as follows:

- Test the hardware components in the system.
- Test the software components in the system.
- Test the hardware and software integration.
- Identify areas that need to be improved.
- Acquire data that can be used to develop data processing algorithms.
- Evaluate the PAM system against the system requirements.

This chapter presents the work that was undertaken for the technical trial of the PAM system, the processing methods that were used and the results obtained from the data analysis. Some of this work has been presented at EMBC2010 [4] and the accompanying paper is included in Appendix A.

6.2 Experimental Protocol

The technical trial was carried out between September and December 2009 and involved the researchers on the PAM project as participants, rather than people with BD. The researchers were used because the system was still very much in a developmental stage and there were likely to be issues in the set-up and running of the system that would require technical knowledge in order to fix. Three participants were used so that the system would be tested in different environments and so that different data-sets would be available for analysis. Each trial was run according to the following criteria:

1. Preliminary interview with a semi-structured questionnaire. (A copy of which is included in Appendix C).
2. Installation of trial equipment.
3. Running of system (3 months, including any fixes and patching).
4. Removal of equipment.

The preliminary interview was conducted in order to brief the participant on what would happen throughout the trial and to gather insight as to the participants thoughts on the acceptability of the system; on their expectations about their behaviour and compliance; on their opinion about technology based monitoring; and on anything else that they think may affect the technical trial.

Equipment was installed by the participant, since they were technically versed in the system and geographically separated from other members of the team a self-installation was acceptable. The system was left to run for three months and over this time there were some fixes that needed to be installed and these were again done by the participant on their own system. After the technical trial was completed, the participants dismantled their own equipment.

Over the course of the trial, a questionnaire was used on the mobile phone to record the participant's behaviour and mood at regular intervals. However, this proved to be very obtrusive; the questioning interval was too short and did not prompt the user. The result of this was that the questionnaire program on the phone would sit unanswered until the user wanted to do something with their phone, the questionnaire now being in the way went unanswered. This intrusive behaviour led to the swift rejection of the questionnaire by the participants and no usable data has been generated with this method. Furthermore, there was supposed to be an exit interview conducted at the end of the technical trial but due to a miscommunication among the researchers this was not carried out.

6.2.1 Participant #1

Participant #1 was an occupant in multiple occupancy housing and a researcher on the PAM project. The technical trial equipment was successfully installed and run for three months. There were however communication problems in some parts of the system and some battery life issues. This resulted in data from the GPS device, the door switches and one of the PIR sensors not being of a high-enough quality to be used. Figure 6.1 shows a floor plan of Participant #1's home and the location of the sensors that were installed.

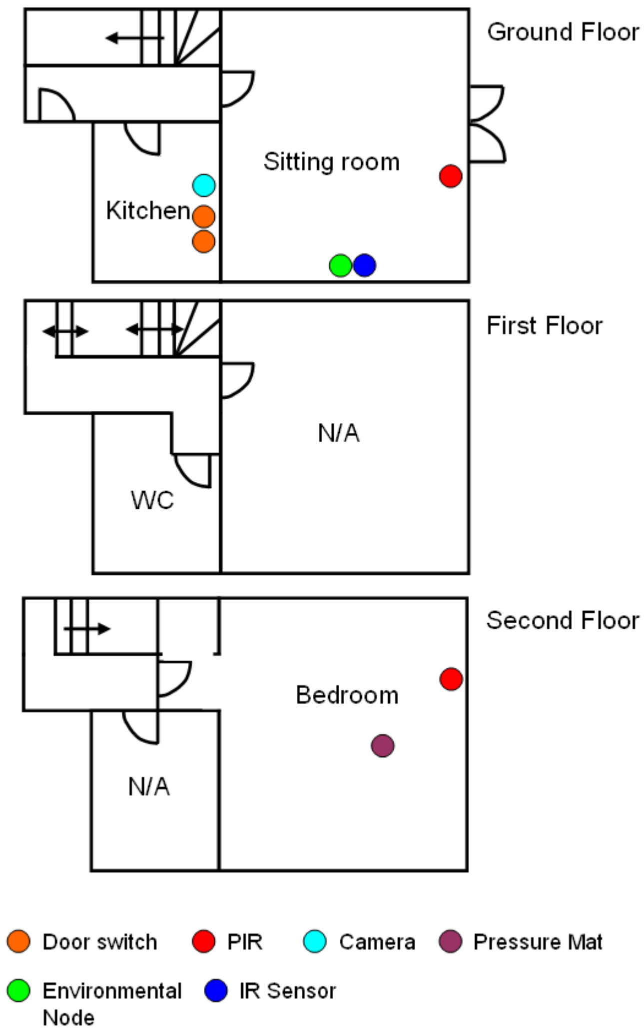


FIGURE 6.1: Floor plan of Participant #1's home showing the location of the installed sensors.

In general, Participant #1's behavioural patterns were ill defined. The participant worked from home some days and not others and had no set times that they woke or slept, aside from in the very general sense that they slept at night. Meal times were reasonably varied as far as breakfast and lunch were concerned, but evening meals tended to be eaten between 1930hrs and 2100hrs. The behavioural patterns of the other occupants were better defined. They were away from home for the vast majority of week-days and displayed much more consisted sleep and wake times. The behaviour patterns for all the occupants during the weekend was known to be different, displaying later waking time in the morning and much more time spent in the home.

6.2.2 Participant #2

Participant #2's house was under single occupancy for the most part during the week, but routinely had other occupants at the weekend. They were also a researcher on the PAM project. The technical trial equipment was successfully installed and run for around five weeks. However, significant hardware and communication issues prevented the majority of sensors from delivering usable data. In addition to this, the participant broke the wearable unit. This, combined with the hardware issues caused the participant to abandon the trial after five weeks.

The sensors that did provide usable data from Participant #2 were the environmental node, including the light sensor, sound sensor, door switches and IR receiver; the PIR sensors; and the wearable GPS sensor. Table 6.1 shows the sensors and their installed location in Participant #2's home. No floor plan is available.

TABLE 6.1: Installed sensors list for Participant #2.

Sensor	Location
PIR	Hallway 1
PIR	Hallway 2
Environmental node	Lounge
Pressure Mat	Bedroom
Camera	Kitchen

In general, Participant #2's behavioural patterns were much more defined that Participant #1's. Participant #2 woke between 7am and 8am most mornings and during the week worked away from the home on most days. Participant #2 was generally in bed by midnight most nights. A log of Participant #2's behaviour can be found in Appendix D.

6.2.3 Participant #3

The data for Participant #3 were not available to analyse due to a combination of technical error and human error on the part of the Participant.

6.3 Data Processing Methodology

The data analysis for the technical trial follows the data processing architecture described in Chapter 3. Each sensor used in the trial produces one or more data-streams, as described in section 3.5.1. The data-stream(s) from each sensor are pre-processed and aggregated using 15 minute blocks as described in further detail for each sensor in Section

6.3.1. Once the data are aggregated, they are smoothed slightly and split into 24-hour long time-series. The result of the processing to this stage is that each day is represented as a set of time-series (one for each data-stream) in which the day is split into 15 minute blocks. The data are then passed into the Continuous Profile Model (CPM), where the behaviour pattern, for each data-stream is identified. A person's collective behavioural patterns are referred to as their activity signature and this can be used to assess new days of data for differences in the person's behaviour. For the technical trial, the CPM is used on a subset of the available data in a training phase, leaving a larger subset available for testing.

For the analysis of the data from the technical trial it is not possible to show behavioural changes due to BD symptoms or prodromes so instead a surrogate behavioural change has been chosen as the target to demonstrate the effectiveness of the system. The chosen surrogate change is the one between behavioural patterns at weekends as contrasted with weekdays. To this end, the CPM was trained on weekdays only so that the weekday behaviour became 'normal' for the system. Data from weekend days should therefore, when run through the behavioural change detection part of the system be identified as different. This relies on weekend behaviour actually being different, which is not an unreasonable assumption, and is supported by the known behaviours of the participants.

6.3.1 Data-Stream Processing

The data from the data-streams are processed to turn the data from its raw format[1] into a format more suited to further processing. The processing involves converting the raw data into a score for each 15 minute block in the data-stream. Once a data-stream has been processed, it is smoothed and divided into 24-hour long (midnight to midnight) time-series. The processing for the sensors the author has investigated is processed in the manner described in the preceding two chapters and data from the other sensors is analysed 'as is'.

The 15 minute time resolution was chosen because it represents a good middle ground in the trade-off between time resolution and pattern smoothness. On the one hand, it is desirable for the 24-hour time-series to have a fine time resolution so that shorter duration behavioural changes can be accurately captured. On the other hand, it is desirable that a person's activity is recorded in such a way that differing activity levels can be determined and the 24-hour time-series produce smooth traces. If the time resolution is too small, there is not enough activity in each time bin and the resultant time-series begins to look like an impulse train and patterns of behaviour become very hard to distinguish. If the time resolution is too coarse there is not enough resolution in

[1]The word raw in this context is used to mean the data as it comes into the data processing system. This is a slight abuse of terminology because by this point, especially for the camera data, there will have been some amount of low-level pre-processing done by the sensors themselves, but this falls outside the scope of this thesis.

a single day to accurately distinguish short time duration behavioural pattern changes. The 15 minute aggregation is a reasonable compromise between these two extremes.

The smoothing of the data after the time aggregation is not strictly a necessary step, but it was discovered during the development of the data processing chain that including a smoothing step produced significantly better results from the CPM. This is due to the nature of the CPM; it performs better on smoother input signals. The nature of the data gathered for the technical trial is such that it tends to be very transient in nature and consist of a series of activity spikes, rather than a smooth curve.

A. Camera

The images from the camera were processed using the WCD described in Chapter 5. Objects in the kitchen that are important to a person's behaviour, such as the sink and cooker, were marked out as AOIs and each AOI treated as generating a separate data-stream. Activity was deemed to have taken place in an AOI when the number of changed pixels exceeded a threshold value. Figure 6.2 shows an example scene with some AOIs defined. Where activity was detected in an AOI, a notification is recorded so that the data-stream consists of a series of date/time readings corresponding to activity in an AOI. The amalgamation of all the readings in a segment is the total number of AOI activity detections recorded in that segment.

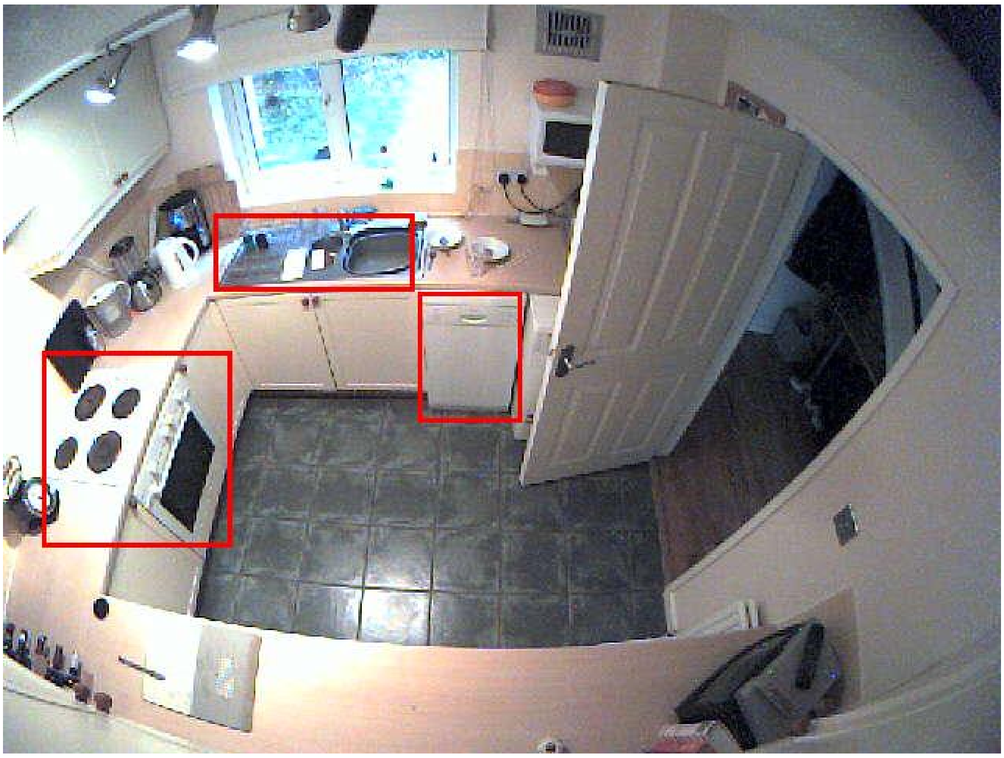


FIGURE 6.2: An example scene showing the kitchen of Participant #1. Three AOIs are defined on the cooker, the sink and the dishwasher.

B. Accelerometer

Data from the triaxial accelerometer is processed using the PCA single-dimension projection as discussed in Chapter 4. The data are split into 15-minute segments, have the features extracted for each segment and then have each segment projected from its features into a one-dimensional space. This results in each segment being reduced to a single number and it is this number that is used as the amalgamation for that segment. Some additional validation results for this approach are shown in section 6.5.1.

C. GPS

The hardware for the GPS sensor was implemented and integrated into the PAM system by the team at the University of Nottingham. Data provided by the GPS sensor contains information on the latitude and longitude of the device at the time of the reading, the number of satellites in view and the speed that the device was travelling at. It was discovered during the analysis of the GPS data that any readings where fewer than four satellites were in view was inaccurate enough to make the reading unusable. These readings are therefore ignored in the processing of the GPS data.

To incorporate the GPS data into the PAM data processing architecture the distance from 'home' is calculated for each GPS reading. Home is usually set to be the participant's home, but can be set to be any reference point. To create the 24hr long time-series, the score for each 15 minute segment is the mean distance from home over the segment.

D. Light and Sound Sensors

The hardware for the light and sound sensors was implemented and integrated into the PAM system by the team at the University of Nottingham. The light sensors, in both the wearable and environmental components, provide readings for both artificial light, being generated by a 50 Hz power source, and for natural light, defined as being light not generated from a 50 Hz power source. The separation is achieved through the use of 100 Hz notch filter, as a 50 Hz alternating current will cause light to be emitted in either polarity, giving a light flicker of 100 Hz. Each light sensor is treated as providing two data-streams, one for artificial light and one for natural light. The 15 minute amalgamation for the light readings is the mean light level over the time window.

The sound sensor provides data on the signal power and zero crossing rate that it is detecting. The sensor samples at 20 kHz and provided a processed output of 30 seconds length at 1 Hz every five minutes. The sound sensors are both treated as providing two separate data-streams and the 15 minute aggregate score for both data-streams is the mean value of the readings across the time window.

E. PIR Sensors

The PIR sensors in the PAM system are off-the-shelf X10 home automation units and were integrated into the PAM system by the author. The sensor units connect through an RF connection to an X10 base-station that is connected to the environmental sub-system smart-node. Data from the sensors is a series of non-uniformly sampled time-stamps, where each time-stamp corresponds to a detection by the PIR sensor. There is however, an enforced 10 second delay after the PIR makes a detection during which time it cannot make another. Each sensor produces its own data-stream and the 15 minute aggregate score is the total number of PIR readings in the time window.

F. Door Switches and Pressure Mat

The door switches and pressure mat were integrated into the PAM system by the team at the University of Nottingham and both sensors connect over Bluetooth to the environmental node. The data format from both these sensor types is the same as for the PIR sensors. Each sensor produces a single data-stream and the data-streams are processed in the same way as those from the PIR.

6.3.2 Activity Signature Detection

To perform the detection of activity signatures, behavioural patterns and changes in behavioural patterns, the PAM system uses CPM [53, 54], the general use of which is described here, with the mathematical explanation being provided in Section 6.4.3 in this chapter.

CPM uses an HMM to model the underlying pattern in the data, termed the latent trace, and in doing this allows us to identify the behavioural patterns from each data-stream. The PAM system uses the CPM both in a learning phase, when the behavioural patterns are being identified, and in a testing phase, when differences in behaviour are being assessed.

In the learning phase, the first T days of data are taken for each data-stream and CPM is used to find the latent trace and the total log-likelihood of generating the input time-series from the HMM model once it has been optimised. This total log-likelihood is denoted L_s^p, where s is used as a variable to indicate a particular data stream, and is divided by T to get the average log-likelihood of generating any one of the input time-series. This is calculated as

$$L\alpha_s = \frac{L_s^p}{T}, \tag{6.1}$$

and used as a point from which the testing phase log-likelihoods can be measured for that particular data-stream.

During the learning phase an estimate of the consistency of the data provided by each data-stream is also obtained. The term consistency is used to mean the consistency of the pattern of activity reported in each data-stream. A data-stream that always produces the same behaviour pattern has a high consistency, whereas a data-stream that produces random noise, for example, would have a low consistency. To establish a numeric value for consistency the correlation of the input set after it has been aligned through the CPM algorithm is used. The pairwise correlation between each pairing of data-streams in the input set is calculated and a mean average of these correlations is taken for the consistency measure. The variable ρ_s is used to denote the consistency measure for each data-stream.

The latent trace provided by the CPM is used as the behavioural pattern for each sensor and the collection of these across all the data-streams is a person's activity signature. Once the activity signature has been discovered it can be used in a testing phase to identify differences in a behaviour pattern.

6.3.3 Behavioural Pattern Change Detection

Once the HMM underlying the CPM has been optimised it can be used to assess the log-likelihood of generating an unseen time-series, which is equivalent to comparing the similarity of an unseen 24hr time-series to the latent trace. If the log-likelihood of generating the unseen data is high then it is similar to the latent trace; if the log-likelihood is low then it is dissimilar to the latent trace.

In the testing phase, data are compared on a day-by-day basis to the established behavioural patterns. For a given day, for each data-stream, the log-likelihood of generating each 24hr-long time-series is computed, denoted $L\beta_s$. The behavioural difference score D_s is calculated as

$$D_s = L\beta_s - L\alpha_s, \tag{6.2}$$

and if D_s positive it is set to 0 to effectively ignore it when the difference scores are summed. A positive D_s score indicates that the new day's time-series was more likely to be generated from the HMM than was the average from the training days. This implies that the new data are not different from the behavioural pattern and since it is difference, not similarity, that is of interest, setting D_s to 0 in these circumstances ignores the data for any subsequent processing.

The difference score is further weighted by multiplication of ρ_s, to give a weighted behavioural difference (WBD) score for each data-stream, which is designed to emphasise those sensors that have a consistent pattern and de-emphasise those that have an inconsistent pattern when the average is calculated. The average WBD score is calculated as the average of all the WBD scores for the new day's data. This can be expressed as

$$W_{\text{avg}} = \frac{1}{N} \sum_{s=1}^{N} \left(D_s \rho_s \right). \tag{6.3}$$

From this summation it can be seen that by setting positive D_s scores to 0 it is only the data that is different to the established pattern that contributes to W_{avg}. This prevents very high positive D_s values from dragging the average up and hiding any behavioural change that might otherwise have been detected.

The average WBD score is one way to obtain an indication of the difference between a person's normal behaviour patterns (defined as the first T days of the trial) and the behaviour recorded on subsequent days. Subsequently this could be used by defining a threshold for W_{avg} so that any score of W_{avg} below this threshold is indicative that the day in question had a significantly different behaviour to the normal patterns.

6.4 Data Analysis Algorithms

6.4.1 Hidden Markov Models

HMMs are discussed briefly in Section 3.9.4 C but shall be defined more formally here. The following explanation is drawn broadly from [92]. To define an HMM five elements are required

1. A set of N states, where individual states are denoted as $S = \{S_1, S_2, \ldots, S_N\}$ and the state at time t as q.
2. A set of M observable values, where individual values are denoted as $V = \{v_1, v_2, \ldots, v_M\}$ and the output at time t as v.
3. The state transition probabilities, $A = \{a_{ij}\}$ where

$$a_{ij} = P\left(q_{t+1} = S_j | q_t = S_i\right), \quad 1 \leq i, j \leq N. \tag{6.4}$$

4. The emission probabilities from state j, $B = \{b_j(k)\}$ where

$$b_j(k) = P\left(v_k | q_t = S_j\right), \quad 1 \leq j \leq N, \ \ 1 \leq k \leq M. \tag{6.5}$$

5. The initial state probabilities $\pi = \{\pi_i\}$, where

$$\pi_i = P\left(q_1 = S_i\right) \quad 1 \leq i \leq N. \tag{6.6}$$

As a result of modelling the process in this way, where the model parameters are denoted as $\Lambda = (A, B, \pi)$, three things can be achieved, as discussed in Section 3.9.4 C. Firstly to run the model an initial state is picked according to the probabilities π and the model is run emitting observations and making state transitions according to A and B. Secondly, using a sequence of observed results, $O = O_1, O_2, \ldots, O_T$, the probability of O given the model, $P(O|\Lambda)$, can be calculated and the best sequence of states, $Q = q_1, q_2, \ldots, q_T$, that produced O can be determined. Thirdly, where the model structure is known but the parameters, Λ, are not and at least one sequence of observed values is available, the optimal parameters can be calculated so that $P(O|\Lambda)$ is maximised.

A. Determining the Probability of an Observation Sequence

To calculate the probability of the observation sequence $O = O_1, O_2, \ldots, O_T$ given the model, $P(O|\Lambda)$, the forwards-backwards algorithm [13] can be used. This is a recursive approach, and can be computed either forwards or backwards. In the forwards case, let

$$\alpha_t(i) = P(O_1, O_2, \ldots, O_t, q_t = S_i|\Lambda), \tag{6.7}$$

which is the probability of observing the partial sequence $O_1, O_2, \ldots, O_t$ and ending in state S_i at time t, given the model Λ. This can be defined recursively as

1. Initialisation:
$$\alpha_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N. \tag{6.8}$$

2. Recursion:
$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1}), \quad 1 \leq t \leq T-1,\ 1 \leq j \leq N. \tag{6.9}$$

3. Termination:
$$P(O|\Lambda) = \sum_{i=1}^{N} \alpha_T(i). \tag{6.10}$$

The backwards case is defined in a similar manner, let

$$\beta_t(i) = P(O_{t+1}, O_{t+2}, \ldots, O_T|q_t = S_i, \Lambda), \tag{6.11}$$

which is the probability of the remaining sequence $O_{t+1}, O_{t+2}, \ldots, O_T$, given that the starting state at time t was state S_i. Now, $\beta_t(i)$ can be written as

1. Initialisation:
$$\beta_T(i) = 1, \quad 1 \leq i \leq N. \tag{6.12}$$

2. Recursion:
$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j), \quad t = T-1, T-2, \ldots, 1, \;\; 1 \leq i \leq N. \tag{6.13}$$

3. Termination:
$$P(O|\Lambda) = \sum_{i=1}^{N} \beta_1(i) \pi_i b_i(O_1). \tag{6.14}$$

Both of these methods provide a recursive way to obtain $P(O|\Lambda)$.

B. Identifying the Best State Sequence

To identify the best state sequence for a given observation sequence a definition of 'best' must be made. The most commonly used criterion is to find the single most likely state-sequence given the model and the observation sequence. The quantity to maximise is therefore $P(Q|O, \Lambda)$, which is equivalent to maximising $P(Q, O|\Lambda)$. The Viterbi Algorithm is a widely used dynamic programming algorithm for performing this maximization. To find the best state sequence $Q = \{q_1, q_2, \ldots, q_T\}$ given the observation sequence $O = \{O_1, O_2, \ldots, O_T\}$, let

$$\delta_t(i) = \max_{q_1, q_2, \cdots, q_{t-1}} P(q_1, q_2, \ldots, q_t = S_i, O_1, O_2, \ldots, O_t|\Lambda), \tag{6.15}$$

which is the highest probability along a state sequence until time t that accounts for the first t observations in the sequence, ending in state S_i. From this, an equation for $\delta_{t+1}(j)$ can be written as

$$\delta_{t+1}(j) = \left[\max_i \delta_t(i) a_{ij} \right] \cdot b_j(O_{t+1}). \tag{6.16}$$

To derive the best state sequence equation 6.16 has to be iterated through for each t and j and the argument that maximises equation 6.16 at each step must be kept track of, which is done in the array $\psi_t(j)$. The Viterbi algorithm now becomes:

1. Initialisation:

$$\delta_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N \tag{6.17}$$

$$\psi_1(i) = 0. \tag{6.18}$$

2. Recursion:

$$\delta_t(j) = \max_{1\leq i\leq N} [\delta_{t-1}(i)a_{ij}] \cdot b_j(O_t), \quad 2 \leq t \leq T, \ \ 1 \leq j \leq N, \tag{6.19}$$

$$\psi_t(j) = \operatorname*{argmax}_{1\leq i\leq N} [\delta_{t-1}(i)a_{ij}], \quad 2 \leq t \leq T, \ \ 1 \leq j \leq N. \tag{6.20}$$

3. Termination:

$$P^* = \max_{1\leq i\leq N} [\delta_T(i)], \tag{6.21}$$

$$q_T^* = \operatorname*{argmax}_{1\leq i\leq N} [\delta_T(i)]. \tag{6.22}$$

4. State sequence backtracking:

$$q_t^* = \psi_{t+1}(q_{t+1}^*), \quad t = T-1, T-2, \ldots, 1. \tag{6.23}$$

C. Optimising the Model

The task of optimising the model parameters (A, B, π) to maximise the probability of an observation sequence $P(O|\Lambda)$ can be performed with the Baum-Welch algorithm. This is equivalent to EM and the relationship between the two is discussed in Section 6.4.2 C. Before the Baum-Welch algorithm can be fully defined, a number of intermediate steps must be introduced. Firstly, let $\xi_t(i,j)$ be defined as the probability of being in state S_i at time t and state S_j at time $t+1$, given the model and the observation sequence

$$\xi_t(i,j) = P(q_t = S_i, q_{t+1} = S_j|O, \Lambda). \tag{6.24}$$

By using the definitions of $\alpha_t(i)$ and $\beta_t(i)$ equation 6.24 can be re-written as

$$\xi_t(i,j) = \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{P(O|\Lambda)}, \tag{6.25}$$

$$= \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}. \tag{6.26}$$

Next, let $\gamma_t(i)$ be defined as the probability of being in state S_i at time t given the observation sequence and the model,

$$\gamma_t(i) = \sum_{j=1}^{N} \xi_t(i,j). \tag{6.27}$$

Now, by summing $\gamma_t(i)$ over t, the expected number of times that the HMM is in state S_i is obtained, which is equivalent to the number of transitions out of state S_i (excluding time $t = T$). In a similar fashion, if $\xi_t(i,j)$ are summed over t the expected number of transitions from state S_i to state S_j is obtained. These two quantities are formalised as

$$\sum_{t=1}^{T-1} \gamma_t(i) = \text{ expected number of transitions from } S_i, \tag{6.28}$$

$$\sum_{t=1}^{T-1} \xi_t(i,j) = \text{ expected number of transitions from } S_i \text{ to } S_j. \tag{6.29}$$

By using the quantities defined above a set of update parameters for the parameters in the HMM can be formulated. For the initial state probabilities π_i the expected number of times the HMM is in state S_i at time $t = 1$ is used as,

$$\bar{\pi}_i = \gamma_1(i). \tag{6.30}$$

For the transition probabilities, a_{ij}, the fraction of the expected number of transitions from states S_i to S_j, over the expected number of transitions from state S_i is used as,

$$\bar{a}_{ij} = \frac{\sum\limits_{t=1}^{T-1} \xi_t(i,j)}{\sum\limits_{t=1}^{T-1} \gamma_t(i)}. \tag{6.31}$$

For the emission probabilities, the expected number of times the HMM is in state S_j and outputs v_k, divided by the total number of times the HMM is in state S_j is used as,

$$\bar{b}_j(k) = \frac{\sum\limits_{\substack{t=1 \\ \text{s.t. } O_t = v_k}}^{T} \gamma_t(j)}{\sum\limits_{t=1}^{T} \gamma_t(j)}. \tag{6.32}$$

Equations 6.30 to 6.32 can now be used to optimise the model. To do this let the current model be set as $\Lambda = (A, B, \pi)$ and use it to compute values for $\bar{\Lambda} = (\bar{A}, \bar{B}, \bar{\pi})$. By iterating between these two procedures $P(O|\Lambda)$ will increase until it reaches convergence at a local optimum. At this point the values of $\bar{\Lambda} = (\bar{A}, \bar{B}, \bar{\pi})$ represent the locally optimal model parameters.

6.4.2 The Expectation Maximisation Algorithm

The EM algorithm [24] is an algorithm that is used to solve the maximum likelihood estimation problem. The following explanation is taken from [13].

If there is a probability density function $p(\mathbf{x}|\Theta)$, which is controlled by the parameter set Θ, and a set of N observed data, $\mathcal{X}$, drawn from the distribution such that $\mathcal{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$, it can be used to give us the probability density for the set of observed data

$$p(\mathcal{X}|\Theta) = \prod_{i=1}^{N} p(\mathbf{x}_i|\Theta) = \mathcal{L}(\Theta|\mathcal{X}). \tag{6.33}$$

The function $\mathcal{L}(\Theta|\mathcal{X})$ is the likelihood function and is a function of the parameters of the distribution, keeping the observed data fixed. This is equivalent to the probability density function of the observed data given the parameters. The problem therefore is to maximise $\mathcal{L}(\Theta|\mathcal{X})$ with respect to Θ. This is formalised as

$$\Theta^* = \operatorname*{argmax}_{\Theta} \mathcal{L}(\Theta|\mathcal{X}). \tag{6.34}$$

The EM algorithm is one technique for performing the maximization of $\mathcal{L}$ when the observed data has incomplete and missing values, or when the observed data are dependent on some hidden parameters, as is the case for HMMs. In this instance $\mathcal{X}$ is referred to as the incomplete data set and it is necessary to define the hidden parameters and the joint probability distribution. Let the hidden parameters be $\mathcal{Y}$ and the complete data set be $\mathcal{Z} = (\mathcal{X}, \mathcal{Y})$. The joint density function is then specified as

$$p(\mathbf{z}|\Theta) = p(\mathbf{x}, \mathbf{y}|\Theta) = p(\mathbf{y}|\mathbf{x}, \Theta)\, p(\mathbf{x}|\Theta). \tag{6.35}$$

The definition of the joint density function leads to a new likelihood function

$$\mathcal{L}(\Theta|\mathcal{Z}) = \mathcal{L}(\Theta|\mathcal{X}, \mathcal{Y}) = p(\mathcal{X}, \mathcal{Y}|\Theta), \tag{6.36}$$

called the complete-data likelihood; the original likelihood function $\mathcal{L}(\Theta|\mathcal{X})$ is now referred to as the incomplete-data likelihood function.

A. E-Step

In the E-step, the EM algorithm evaluates the expected value of the complete-data log-likelihood, log $p(\mathcal{X}, \mathcal{Y}|\Theta)$, with respect to the current parameter estimates, the observed data $\mathcal{X}$ and the unknown data $\mathcal{Y}$. To achieve this a new function is defined as

$$Q\left(\Theta, \Theta^{(i-1)}\right) = E\left[\log p\left(\mathcal{X}, \mathcal{Y}|\Theta\right)|\mathcal{X}, \Theta^{(i-1)}\right], \tag{6.37}$$

where $\Theta^{(i-1)}$ are the current parameter settings used in the evaluation of the expectation and Θ are the new parameters that are used to maximise Q in the M-step.

B. M-Step

The M-step of the algorithm maximises the expectation that was evaluated in the E-step. That is

$$\Theta^{(i)} = \underset{\Theta}{\operatorname{argmax}} \left(\Theta, \Theta^{(i-1)}\right). \tag{6.38}$$

The two steps can be iterated as often as necessary and converge to a local maximum for the log-likelihood function.

C. Expectation Maximisation in HMMs

The Baum-Welch algorithm, described above, can be seen to be equivalent to EM for HMMs. By determining Λ and then performing updates (maximisations) of A, B, and π to get an updated $\bar{\Lambda}$ the Baum-Welch algorithm can be seen to be performing EM.

6.4.3 The Continuous Profile Model

CPM [53, 54] is a method for the simultaneous alignment of a set of time series. The time series in the set are assumed to have been generated from a noisy stochastic process and as such can be modelled as a noisy, time-varying transform from a single canonical time series, this is referred to as the latent trace. The CPM takes the set of k time series to be aligned as

$$\mathbf{X} = \left\{\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^N\right\}, \tag{6.39}$$

where

$$\boldsymbol{x}^k = (x_1^k, x_2^k, \dots, x_{n^N}^k), \tag{6.40}$$

and neither the sampling rate nor the total size of the time-series need to be uniform. That is to say that each time series in the set can have a different, non-uniform sampling rate. For notational clarity it is subsequently assumed that $N^k = N$. In order to align the time series, coping with time variance across $\boldsymbol{x}$, an HMM is used to model the latent trace. Individual time series can then be modelled by traversing along the HMM and emitting an output, that corresponds to a step in the time series being modelled, at each stage. By using an HMM with a finer time resolution than the input set, the time variances in the input can be accounted for by altering the rate at which the HMM is traversed, jumping ahead more states to speed up time and jumping ahead fewer states to slow down time. The latent trace is assumed to take the form

$$\boldsymbol{z} = (z_1, z_2, ..., z_M), \tag{6.41}$$

where M accounts for the increased resolution and is set as $M = (2+\epsilon)N$, where $\epsilon = 0.2$.

For a time series to be modelled from the latent trace, a series of hidden states is needed. This sequence of hidden states is defines to be $\boldsymbol{\iota}$; each state in this sequence maps to a time state/scale state pair such that

$$\iota_i \rightarrow \{\tau_i, \phi_i\}, \tag{6.42}$$

where the time states τ belong to the integer set $[1..M]$ and the scale states, ϕ belong to an ordered set $[1..R]$. The states ι_i and observed values are linked by the emission probability distribution

$$p(x_i^k | \iota_i, \boldsymbol{z}, \zeta^k, u^k) \equiv \mathcal{N}(x_i^k | z_{\tau_i} \phi_i u^k, \zeta^k), \tag{6.43}$$

where ζ^k is the noise level of the observed data (the variance) and u^k are scale parameters for each time series that correct for any global scale difference between time series k and the latent trace.

In order for the model to be fully specified it is necessary to define the transition probabilities. Transitions between time states and between scale states are defined separately such that the joint transition probability is defined as

$$T^k_{\iota_{i-1},\iota_i} \equiv p^k(\iota_i | \iota_{i-1}) = p(\phi_i | \phi_{i-1}) p^k(\tau_i | \tau_{i-1}). \tag{6.44}$$

There are also several constraints that are defined for CPM to work effectively. Firstly, time (as transitions through the HMM) must move forward and cannot stand still. In addition, the maximum number of states that can be jumped forward is set to J. Secondly, scale states can not be jumped to arbitrarily but instead, only neighbouring scale states may be jumped to. This has the effect of making the transition matrix very sparse and speeds up the computation speed of the algorithm. These constraints are formalised as

$$p^k(\tau_i = a|\tau_{i-1} = b) = \begin{cases} \kappa_1^k, \text{ if } a - b = 1, \\ \kappa_2^k, \text{ if } a - b = 2, \\ \vdots \\ \kappa_J^k, \text{ if } a - b = J, \\ 0, \text{ otherwise,} \end{cases} \tag{6.45}$$

$$p(\phi_i = a|\phi_{i-1} = b) = \begin{cases} s_0, \text{ if } D(a,b) = 0, \\ s_1, \text{ if } D(a,b) = 1, \\ s_1, \text{ if } D(a,b) = -1, \\ 0, \text{ otherwise,} \end{cases} \tag{6.46}$$

where $D(a, b)$ is a distance measure between a and b and equal to 1 if a is one scale state larger than b, -1 if a is one scale state smaller than b and 0 if a and b are the same. These distributions are constrained so that the sum of their probabilities is equal to 1, that is $\sum_{i=1}^{J} \kappa_i^k = 1$ and $2s_1 + s_0 = 1$.

The above definitions and constraints would be sufficient for CPM to produce a latent trace, there are however two more additions that are used to regularise the model and produce better results. The first is a smoothness penalty, which is used to encourage the latent trace to be smooth. This is done by penalising the likelihood of the HMM with the smoothing penalty

$$p(\boldsymbol{z}) \propto exp\left(-\chi\bar{u}\sum_{j=1}^{M-1}(z_{j+1} - z_j)^2\right), \text{ where } \bar{u} \equiv \frac{\sum_k {u_k}^2}{K}, \tag{6.47}$$

which encourages the latent trace to vary gradually. A higher value of χ will increase the smoothing penalty. The second regulatory addition is that of Dirichlet priors to the scale and time transition probabilities so that all non-zero transition probabilities remain non-zero. These two terms are used in the regularising term in the calculation of the likelihood function and are defined as

$$p(\boldsymbol{\kappa}^{\boldsymbol{k}}) = \mathcal{D}(\boldsymbol{\kappa}^{\boldsymbol{k}}|\boldsymbol{\eta}) \propto \prod_{v=1}^{J} \left(\kappa_v^k\right)^{\eta_v - 1}, \tag{6.48}$$

$$p(\boldsymbol{s}^{\boldsymbol{k}}) = \mathcal{D}(\boldsymbol{s}|\boldsymbol{\eta'}) \propto \prod_{v=1}^{2} (s_v)^{\eta'_v - 1}. \tag{6.49}$$

A. Likelihood Function

The likelihood function is a function that provides the likelihood of K observed time series being generated from the HMM. For CPM, the likelihood function is $L^p \equiv \mathcal{L} + \mathcal{P}$ where $\mathcal{L}$ is the likelihood arising from the HMM and is computed using the Forward-Backward algorithm described in section 6.4.1, and $\mathcal{P}$ is the regularizing term.

$$\mathcal{L} \equiv \sum_{k=1}^{K} \log \sum_{\boldsymbol{\iota} \to \{\tau_i, \phi_i\}} p\left(\iota_1\right) \left(\prod_{i=1}^{N} \mathcal{N}\left(x_i^k | z_{\tau_i} \phi_i u^k, \zeta\right) \right) \left(\prod_{i=2}^{N} T^k_{\iota_{i-1}, \iota_i} \right), \tag{6.50}$$

$$\mathcal{P} \equiv -\chi \bar{u} \sum_{j=1}^{M-1} (z_{j+1} - z_j)^2 + \sum_{k=1}^{K} \log \mathcal{D}\left(\kappa_v^k | \{\eta_v^k\}\right) + \log \mathcal{D}\left(s_v | \{\eta'_v\}\right) + \sum_{k=1}^{K} \log \mathcal{N}\left(\log u_k | 0, w\right). \tag{6.51}$$

B. Optimisation

The HMM parameters for the CPM need to be optimised in order for the HMM to be effective. This is achieved with the EM algorithm. The likelihood function is maximised with respect to the parameters that need to be optimised; specifically $\{\mathcal{Z}^c\}$, $\{u^k\}$, $\{\kappa_i^k\}$, s_0, s_1 and $\{\zeta_d^k\}$. It is assumed that v, χ and all other parameters have been set appropriately. The optimisation provides the best setting for the parameters and allows the HMM to be used to compute the alignment of the input time series and to asses new time series for alignment and likelihood of generation from the HMM.

C. Time-Series Alignment

Once the HMM parameters have been optimised, the alignment of the input time-series, or new time-series, to the latent trace can be obtained. If this is performed for all time-series in the input set the result is the simultaneous alignment of the input set. The aligned version of an input time-series is an upsampled and interpolated version of the input, with the sampling points determined by the degree of temporal adjustment that is needed to align the input to the latent trace. The aligned time-series has the same number of points as the latent trace and each point takes its value from the input time-series. Some of these points will match up exactly, and can be identified with the Viterbi algorithm. The values for the remaining points, which lie in between, have to be calculated with linear interpolation on the input time-series.

The Viterbi algorithm is used during the alignment to determine the best state sequence given the input. This state sequence corresponds to the points in the aligned time-series that have to take their values from the input time-series. Figure 6.3 shows the latent trace, the input trace, the state correlation determined by the Viterbi algorithm and the known values on the aligned trace at this stage. Red lines link the state sequence determined by the Viterbi algorithm to the points from the input from which the new values get taken.

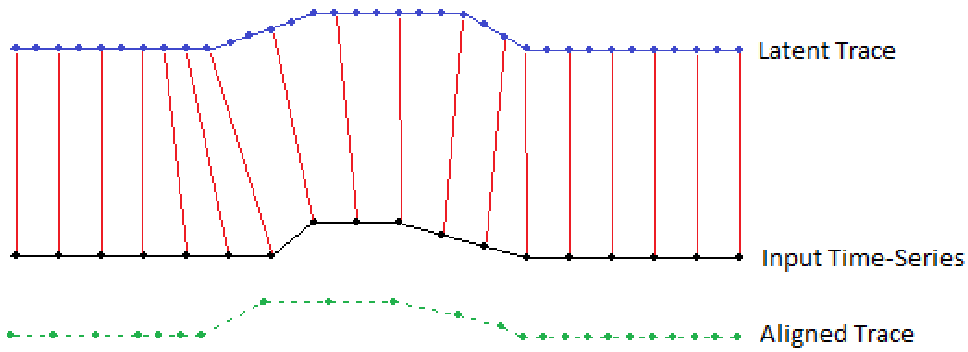


FIGURE 6.3: Diagram of the way in which the aligned trace is constructed showing the point accociations between the latent trace and input trace determined with the Viterbi algorithm (red lines) and the aligned trace once the values for the known points have been determined.

Following this there is an interpolation stage where the time points from the latent trace that have not been assigned a value are assigned an interpolated value. This value comes from a linear interpolation between points on the input time-series. For example, time points P_{L1} and P_{L4} from the latent trace are matched to points P_{I1} and P_{I2} on the input time-series. The values for time points P_{L2} and P_{L3} are derived by interpolating between P_{I1} and P_{I2}. The result of both steps is then a resampled and interpolated version of the input time-series which has been aligned to the latent trace.

6.5 Results

The validation results for the accelerometer and the data analysis of the technical trial for Participants #1 and #2 are presented in this section. As discussed in Section 6.2 the data for Participant #3 are not available for analysis and the data for Participant #2 are significantly reduced, both in terms of the number of sensors available and in the number of days of data available. To take account of the difference in days of the two data-sets, each set was analysed using a different number of days for training.

6.5.1 Accelerometry Validation

Figure 6.4 shows a scatter plot of the single-dimension PCA projection against the average signal power for the accelerometry data from Participant #1. Average signal power was calculated over 15 minute time windows with n samples as

$$\text{power} = \frac{1}{n}\sum_{i=1}^{n}\left(x_i^2 + y_i^2 + z_i^2\right). \tag{6.52}$$

In general there is a strong positive linear correlation between both values. The exception to this is in the smaller values, where the correlation becomes non-linear and the PCA projection provides a greater separation of the data than the average signal power.

Figure 6.5 shows the accelerometry data from Participant #1 projected through PCA into a single dimension and shows some sample segments of activity and the regions of the projection from which they are drawn. It can be seen from the figure that the less intense activity is lower valued than the more intense activity when the data are processed and projected with PCA.

6.5.2 Participant #1

The results of the technical trial for Participant #1 are shown here. The trial was successfully run with the majority of sensors for a full three months. The sensors that did not produce usable data are not included. In this analysis, the behavioural patterns are trained on the first 30 days of good data and the remaining data are used to test for behavioural differences. A day of good data is defined as a day where the sensor providing the data-stream is on-line and there is sufficient data to analyse.

Figure 6.6 shows the unaligned and aligned data from the camera's cooker AOI, the sitting room PIR sensor, the pressure mat and the artificial light sensor for Participant #1. The unaligned data, for all sensors, generally shows no particularly clear pattern. In the camera data, only the broadest of trends can be identified, that is to say that there is some activity from mid-morning until mid-afternoon (1000-1600hrs) and then again in the evening through into the early morning (1800-0200hrs) but this activity is vague and difficult to determine. The same observations can be made about the PIR data, but in this case the pattern is even more vague. There is no visible pattern at all in the pressure mat data. The artificial light data are the exception to the trend, in that

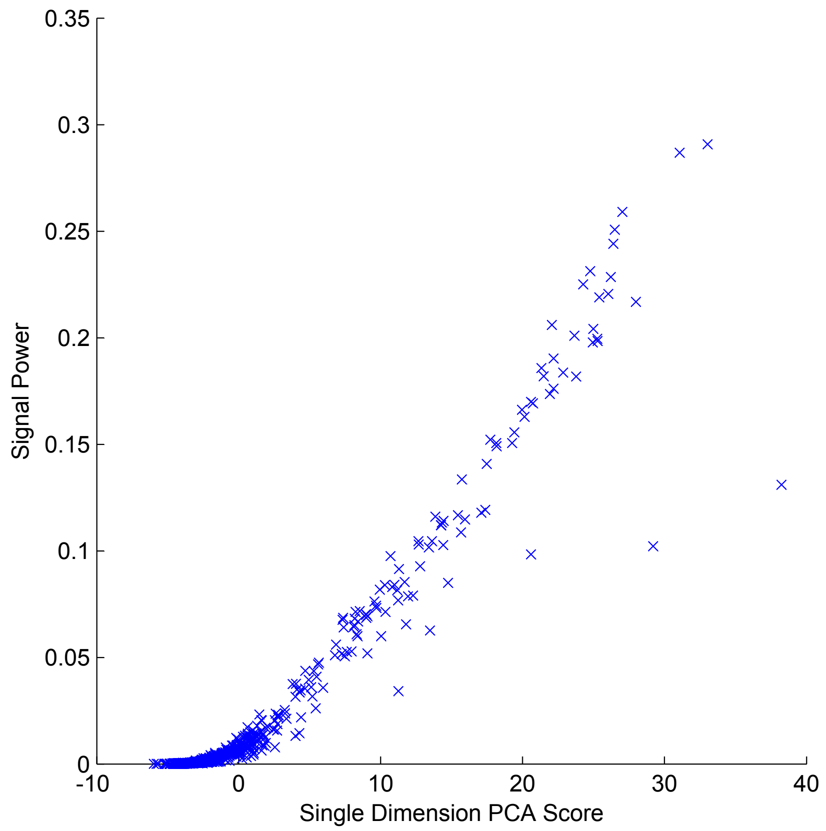


FIGURE 6.4: Scatter plot of single-dimension PCA projection against signal power for the accelerometry data taken from Participant #1. The plot shows a general correlation between the PCA projection and average signal power except in the smaller values, where the PCA projection provides a greater degree of separation.

the pattern of artificial light usage in the evening and very early morning is quite clear to see.

The aligned data makes the vague patterns in the data significantly easier to see by eye. The data for the PIR especially points to significant activity around 2000hrs and the evening and early morning presence of artificial light shows a very clear pattern around the same time.

The behavioural patterns extracted from the data are shown in Figure 6.7. The latent trace for the cooker AOI clearly identifies a peak at 2000hrs and smaller ones throughout the day and these indicate a clear pattern in activity near the cooker at these times. The traces for both the PIR and the artificial light show very definite activity from 1800hrs onwards. The latent trace from the pressure mat shows three clear activity spikes, at 0000hrs, 1000hrs and 1900hrs. The trace for the environmental artificial light

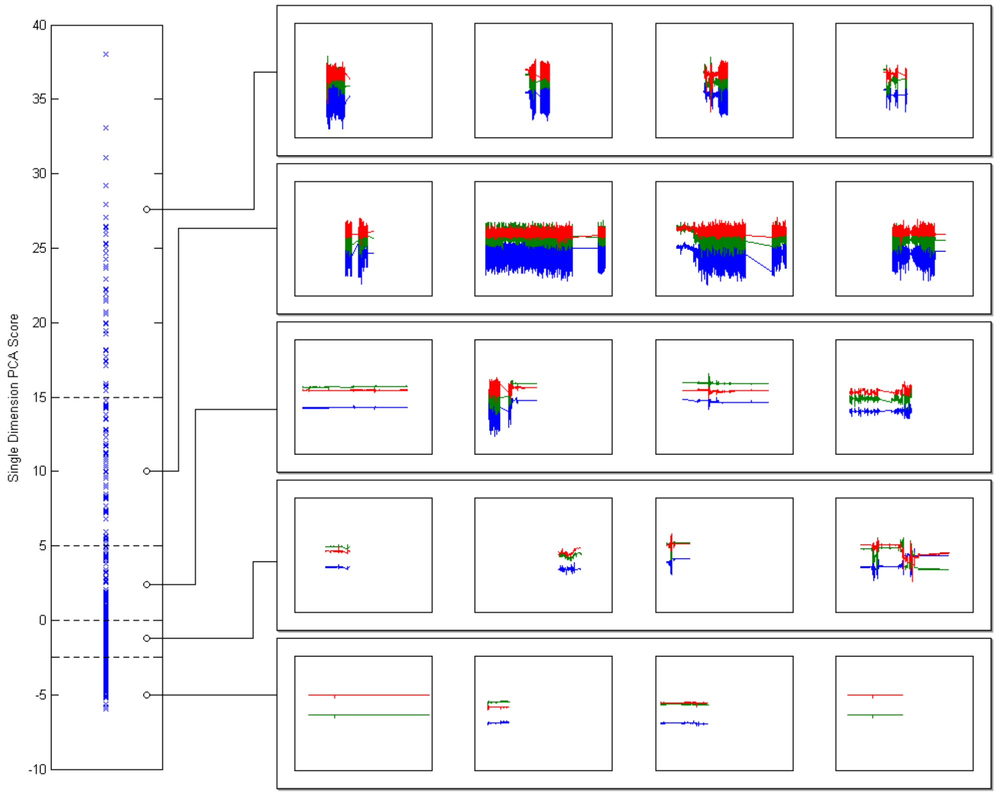
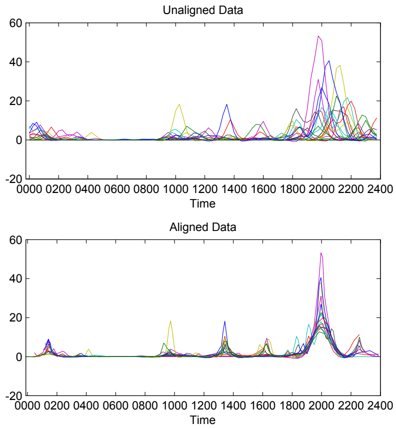


FIGURE 6.5: Single-dimension PCA projection of the accelerometry data for Participant #1 showing sample activity traces from different sections of the projection. Sections of the PCA projection have been selected by eye according to the density of data-points in the section, keeping all sections to the same density. Green, red and blue lines show the X, Y and Z axis of acceleration respectively.

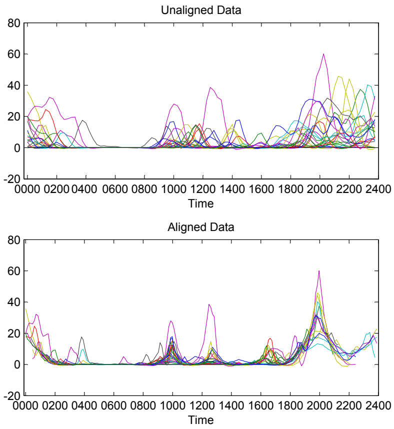
sensor shows activity from late afternoon right through the evening, peaking at around 2100hrs.

The activity signature for Participant #1 is shown in Figure 6.8. The figure shows each behavioural pattern aligned in time and has normalised the amplitude of these plots for display purposes to scale to a unit range. There is a clear correlation in the activity of the environmental sensors, which all show significant activity in the evenings. The pressure mat is the exception to this, as it shows three activity spikes, one in the very early morning, one in the mid-morning and one in the early evening. The pattern of activity in the wearable data is different to the environmental data and shows more activity during the day. The wearable artificial light sensor in particular shows activity from mid morning through to early evening.

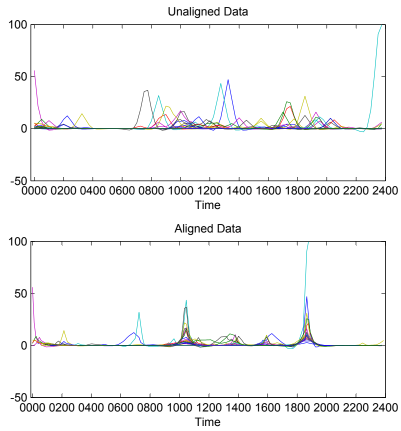
Figure 6.9(a) shows the WBD scores across all the environmental sensors used in the technical trial for participant #1. Gaps in the traces for some of the sensors are due to missing data. The traces produced are generally very disorganised, but there are some days where there is a consistent drop in several traces.



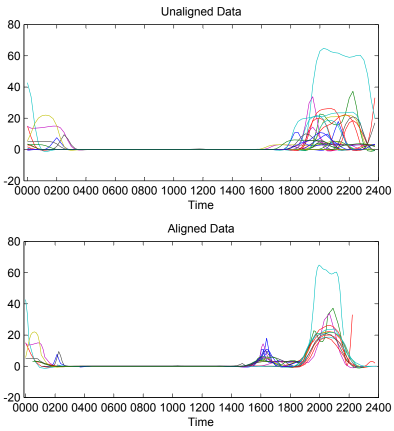
(a) Camera Cooker AOI



(b) PIR Sitting Room



(c) Pressure Mat

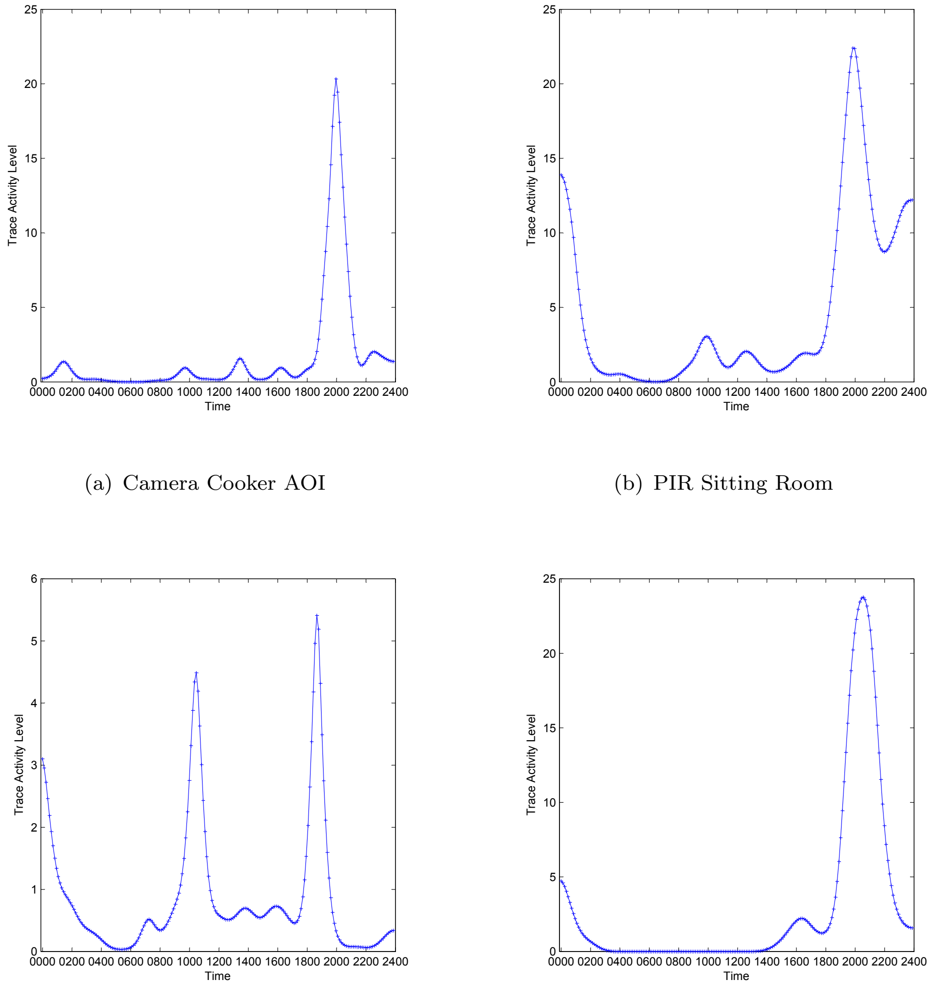


(d) Environmental Artificial Light

FIGURE 6.6: Example of the time series from four of the data-streams for Participant #1 before and after alignment with CPM, showing the cooker AOI from the camera, (a), the sitting room PIR, (b), the pressure mat, (c) and the environmental artificial light sensor (d). Different coloured lines show different day's data. After the data has been aligned the patterns in the data are significantly easier to see by eye.

The average WBD score for the test days is shown in Figure 6.9(b) and shows that the majority of weekend days have a lower WBD score than the majority of week-days. Table 6.2 summarises the contribution of each sensor to the average WBD score. The table shows that the most important sensors were the environmental natural light and both the microphone zero-crossing rate (ZCR) data-streams, whilst the least important were the wearable natural light, TV remote control and accelerometer data-streams.

A Student's T-Test was performed on the average WBD data, allowing for unequal sample size and unequal variance. The null hypothesis was that there was no difference between the week-day and weekend average WBD scores. A P-value of 0.0182 was



(a) Camera Cooker AOI

(b) PIR Sitting Room

(c) Pressure Mat

(d) Environmental Artificial Light

FIGURE 6.7: The latent traces extracted from four of the data-streams for Participant #1 using the CPM algorithm showing the cooker AOI from the camera, (a), the sitting room PIR, (b), the pressure mat, (c) and the environmental artificial light sensor (d).

obtained, which rejects the null hypothesis at the 5% significance level and shows that the difference between week-ends and week days is statistically significant.

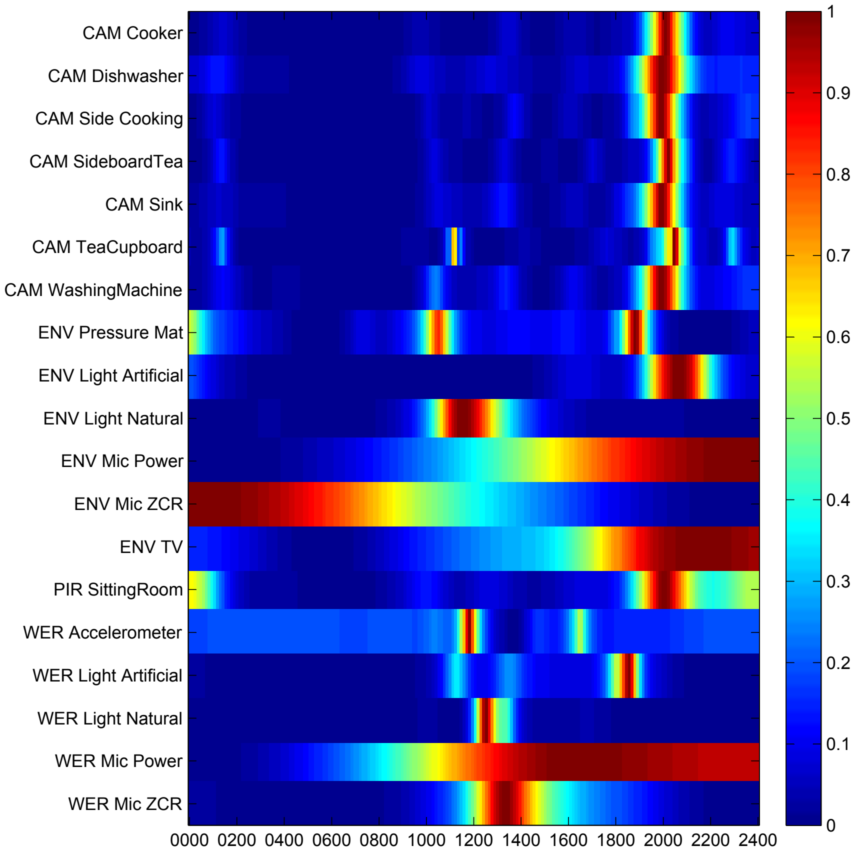
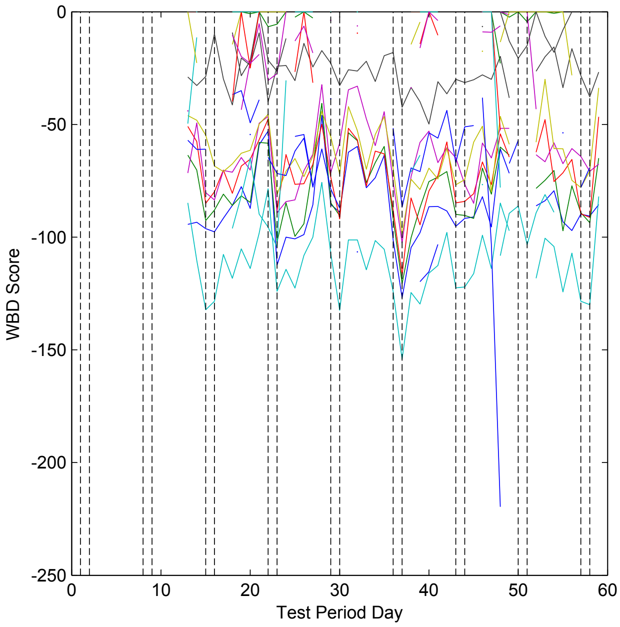


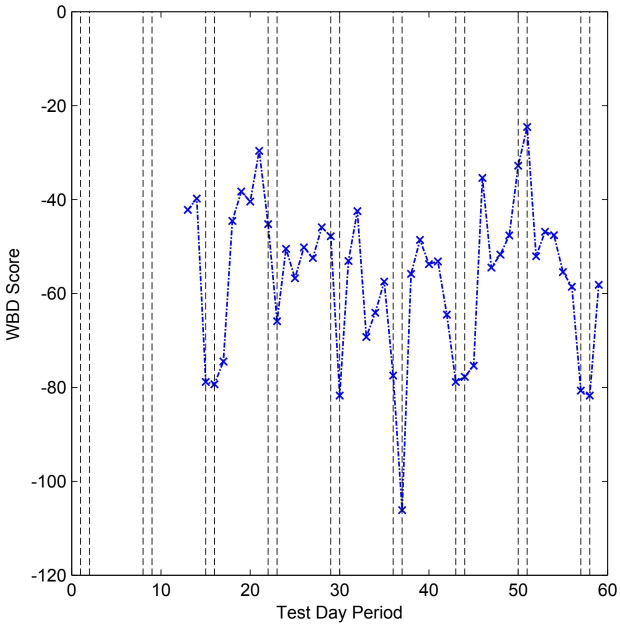
FIGURE 6.8: The activity signature for Participant #1 showing the latent traces derived from the data-streams aligned in time. Data-streams marked 'CAM' are camera AOIs, those marked 'ENV' come from the environmental node and its attached sensors and those marked 'WER' come from the wearable sensor. The two data-streams marked 'Mic ZCR' are the zero-crossing rate data from the sound sensors and the one marked 'ENV TV' is the IR receiver sensor.

Table 6.2: Weightings for each data-stream in the calculation of the average WBD score.

Data-stream	Contribution (%)
CAM Cooker	5.09
CAM Dishwasher	5.64
CAM SideCooking	4.43
CAM SideboardTea	4.86
CAM Sink	5.33
CAM TeaCupboard	5.41
CAM WashingMachine	4.98
ENV Pressure Mat	4.72
ENV Light Artificial	4.96
ENV Light Natural	11.15
ENV Mic Power	5.60
ENV Mic ZCR	7.74
ENV TV	3.17
PIR SittingRoom	3.86
WER Accelerometer	3.62
WER Light Artificial	3.89
WER Light Natural	3.15
WER Power	5.2
WER Mic ZCR	7.1



(a) All WBD scores



(b) Average WBD Score

FIGURE 6.9: The behavioural difference graphs showing all the sensors for Participant #1, (a), and the average WBD score, (b). Dashed lined indicate weekend days and gaps in the graphs indicate that no data are available for those days. The graph of the average WBD score clearly shows most weekend days have a lower average WBD score than non-weekend days.

6.5.3 Participant #2

The results of running the technical trial for Participant #2 are presented here. Due to several hardware problems with this installation, discussed in Section 6.2, the available data come from a significantly smaller subset of sensors than those for Participant #2.

Figure 6.10 shows the activity signature for Participant #2. There is no real correlated activity, but there is an overall trend for the environmental sensors to show activity in the evening. One sensor of particular interest is the PIR in the bedroom, which shows two very distinct activity spikes, one in the morning and one in the evening.

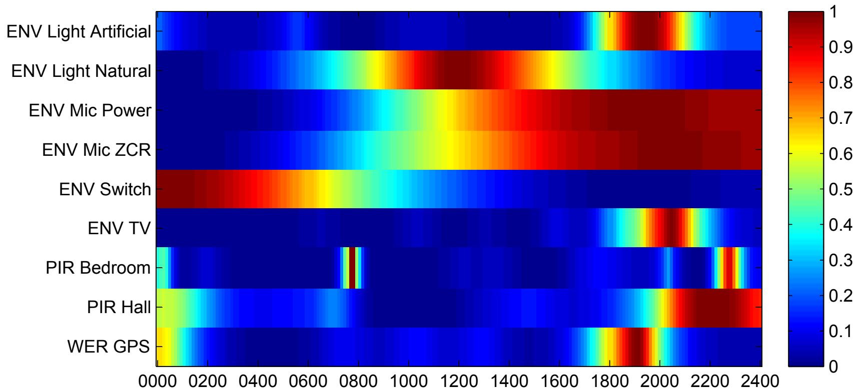


FIGURE 6.10: The activity signature for Participant #2 showing the latent traces derived from the data-streams aligned in time. Data-streams marked 'ENV' come from the environmental node and its attached sensors and those marked 'WER' come from the wearable sensor. The data-stream marked 'Mic ZCR' is the zero-crossing rate data from the sound sensor and the one marked 'ENV TV' is the IR receiver sensor.

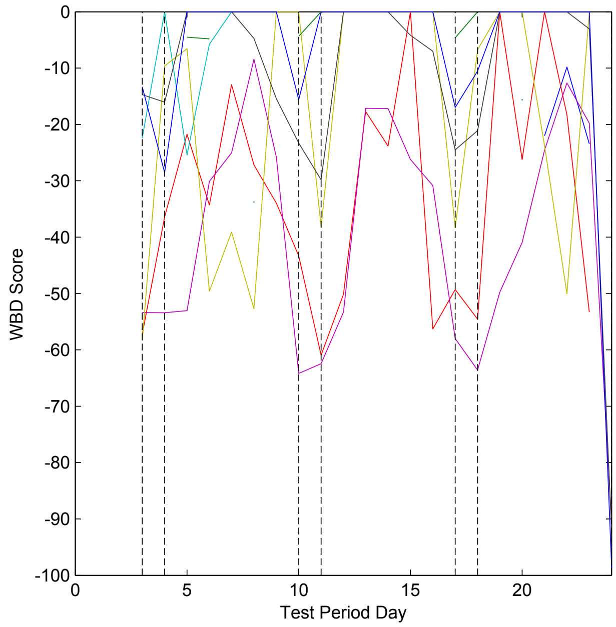
Figure 6.11(a) shows the WBD scores across all the environmental sensors used in the technical trial for Participant #2. Gaps in the traces for some of the sensors are due to communication problems between the sensor and the base-station PC, which resulted in a loss of data on some days. As with the data from Participant #1, the data here are generally unaligned and there are some days where drops across multiple sensors can be observed.

The average WBD score for Participant #2 is shown in Figure 6.11(b). It can be seen from this figure that the weekend days have been clearly identified as different and that the scores for these days are lower than the scores obtained for the week-days. Table 6.3 summarises the contribution of each sensor to the average WBD score. The table shows that the most important sensors were the environmental natural light, microphone power and GPS data-streams, whilst the least important were both the PIR and environmental door switch data-streams.

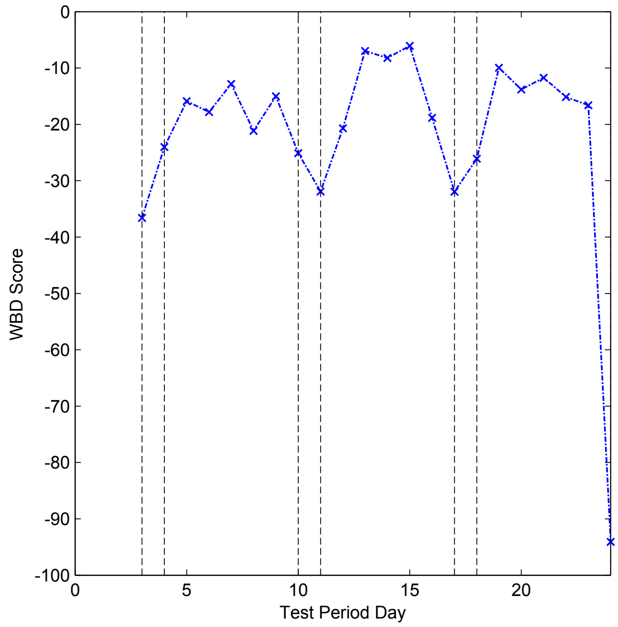
Table 6.3: Weightings for each data-stream in the calculation of the average WBD score.

Data-stream	Contribution (%)
ENV Light Artificial	9.32
ENV Light Natural	18.45
ENV Mic Power	15.25
ENV Mic ZCR	13.06
ENV Switch	7.9
ENV TV	10.08
PIR Bedroom	5.78
PIR Hall	7.54
WER GPS	12.62

A Student's T-Test was performed on these data, allowing for unequal sample size and unequal variance. The null hypothesis was that there was no difference between the week-day and weekend average WBD scores. A P-value of 0.0106 was obtained, which rejects the null hypothesis at the 5% significance level and shows that the difference between week-ends and week days is statistically significant.



(a) All Behavioural Difference Scores



(b) All Behavioural Difference Scores

FIGURE 6.11: The behavioural difference graphs for Participant #2, showing all the sensors, (a), and the average WBD score, (b). Dashed lined indicate weekend days and gaps in the graphs indicate that no data are available for those days. The graph of the average WBD score clearly shows most weekend days have a lower average WBD score than non-weekend days.

6.6 System Evaluation

The design requirements of the PAM system were discussed in Chapter 3 and serve as a benchmark for evaluating the system from a systems design point of view. The following tables list the system requirements and whether or not the current PAM system has passed or failed on these requirements. The evaluation of the requirements is split in the same way as the requirements are; into an evaluation of the data processing requirements that relate directly to the work presented by the author in this thesis and of the requirements that relate to the PAM system as a whole, including the work of the entire PAM team.

6.6.1 Data Processing Evaluation

It can be seen from an examination of Table 6.4 that the data processing requirements of the system have largely been fulfilled. The critical requirements of being able to detect activity signatures and behavioural patterns and of being able to detect changes in these patterns have both been met, as evidenced by the work presented in this chapter.

Table 6.4: PAM system requirements evaluation: data processing.

Requirement	Result
The data processing algorithms must be able to handle a modular system and allow for different configurations of sensors.	Passed. The data processing architecture shown in Chapter 3 allows for different sensors to used and this worked in the technical trial.
The low-level data processing algorithms must be implementable on the low-level system components.	Untested. The algorithms developed by the author for the processing of the accelerometer data should be implementable on the wearable node. The camera processing algorithm may be implementable on the camera as it does have some processing capability. This has not been attempted however.
The high-level algorithms must be able to cope with missing sections of data.	Passed. The high-level data processing algorithms can cope with missing data.
There should be ideally zero, or at worst absolute minimal, user involvement with the data processing.	Passed. The use of algorithms that do not require supervised training allows for there to be zero user involvement with the data processing.
The data processing should enable activity signatures and behavioural patterns to be detected.	Passed. Activity signatures and behavioural pattern detection has been demonstrated.
The data processing should enable changes in behavioural patterns to be detected.	Passed. Detection of behavioural differences at weekends has been shown.
Both the high-level and intermediate levels of data processing should be easily communicable to the user.	Passed. The average WBD graph is easy to understand, as are the graphs drilling back down into the data. Additionally, the use of single dimension PCA projection gives a clear and easy to understand analysis for the accelerometer and the change-mask and AOI detection does the same for the camera.

6.6.2 Overall Evaluation

It can be seen from an examination of the system requirements against their evaluation that in general the PAM system has succeeded for the requirements that centred around constructing a comprehensive monitoring and data analysis system. The areas in which it has failed however are those that centre around system infrastructure and user experience, for which it fared particularly poorly.

Table 6.5: PAM system requirements evaluation: monitoring for BD.

Requirement	Result
The system must capture a broad variety of behaviour.	Passed. The system has been able to capture different behaviour from different people.
The system must be able to capture data at the user's home and when the user is out of the home.	Passed. Both the environmental subsystem and wearable subsystem have been demonstrated to provide usable data.
Data must be captured with a sufficiently fine time granularity that patterns of activity are visible over the course of a day.	Passed. All of the sensors provide a high-enough sample rate that patterns can be determined in the 24-hour long time-series.
The system must be modular and able to target different aspects of behaviour.	Passed. There are a variety of sensors that can be used or not used in the system and the data processing can cope with this.

Table 6.6: PAM system requirements evaluation: usability.

Requirement	Result
The system must be as unobtrusive as possible.	Failed. Several sensors (the camera and wearable unit especially) are not particularly unobtrusive.
The battery life of any device must be sufficiently long that the user does not have to replace batteries or charge the device excessively often.	Partial pass. Many of the devices have a sufficient battery life, but several do not, the cupboard door sensors in particular.
The user-interface must be simple and easy to use.	Partial pass. The interface is usable, but could be improved.
The user must be minimally involved with maintaining the system.	Failed. Users were heavily involved in system maintenance.
The system must be modular and configurable to the needs of the individual user.	Passed. There are several different sensors that can be used and the selection of these can be tailored to an individual's needs.

Table 6.7: PAM system requirements evaluation: hardware and communication.

Requirement	Result
The system must provide appropriate sensors to capture a person's behaviour patterns.	Passed. Behaviour patterns have been captured.
The system must provide an appropriate communications infrastructure.	Partial pass. The infrastructure worked, but suffered some communication problems, particularly with the Bluetooth links having a very short range.
The system must be able to cope with any ambulatory components dropping in and out of communication range.	Majority pass. In general this was handled well, with the exception of one particular type of GPS module.
The system must provide sufficient high-level computing power that data can be processed to extract behavioural patterns.	Untested. Behavioural patterns were not extracted in an on-line manner but this should be achievable.
The system should allow for data to be stored long-term for further analysis and academic interest.	Passed. Data was stored to an backup server.

6.7 Discussion

The results presented here show that by taking a large volume of data about a participant and applying suitable processing algorithms a single metric can be obtained that gives an indication of the change in the participant's behaviour over time. This metric has been used on two different data-sets, collected from two participants using different sensor subsets, to successfully detect the difference in patterns between weekday and weekend behaviour. The results of this detection are very positive, with the majority of the weekend days generating a lower WBD score than weekdays. This is reinforced by the statistical tests, which show statistical significance for this assertion for both participants.

The weekend detection is good, but not perfect, but it is not expected to be; indeed, it is well within the realms of possibility that on some days a person's behaviour at the weekend is very similar to their behaviour during the week. It is also possible for errors to creep into the WBD score if enough sensors drop out at the same time. The weekend around test day 50 for Participant #1 is a very good example of this. Figure 6.9(b) shows that this weekend has a comparatively high average WBD score, but Figure 6.9(a) shows that most of the sensors have dropped out for this weekend, which significantly affects the resultant average WBD score.

In the current system the effect of sensor drop-out on the average WBD score is not addressed, but the addition of a confidence measure (CM) would be one way of dealing with the problem. The CM would be a numeric value for how accurate the WBD score is, given the number of on-line sensors. It would be proportional, linear or otherwise, to the fraction of the total sensors in the system that are on-line for a given day. A high CM would imply that most or all of the sensors were on-line and therefore that the WBD score can be relied upon. A low CM would imply that some sensors had off-lined and that the WBD score for that day may be less reliable.

The way in which the data are split in the time series generation phase also requires some discussion at this point. The data spilt according to calendar date, that is, from midnight to midnight. In the results for both participants this creates a discontinuity in the level of the behavioural traces either side of midnight after processing with the CPM algorithm. This is evident in Figures 6.7 and 6.6, where some traces exhibit differing values at either end, where one would expect them to be the same, and for the trace to 'wrap-around'. A different choice of separation time for the time-series generation, 0400 for example, would help remove this problem and may produce better results from the CPM stage of processing.

As has already been identified in the evaluation, it is communication drop-outs that are the primary limitation of the technical trial. The system as it was deployed was in a very prototypical state and the communications infrastructure that worked well in controlled lab situations did not cope as well in real-world settings. The main issue was with the Bluetooth communication, the range of which is severely attenuated by interior walls and floors. As such, some of the devices had to be placed on the borderline between being in and out of communication and a general lack of robustness to this issue in the system resulted in lost data. There were also significant battery life issues, particularly with the door switches, three of which ran out of battery very early in the trial and stopped sending data. Since the system provided no battery power notifications, the batteries remained unchanged resulting in a loss of data from these sensors.

6.8 Summary

The technical trial of the PAM system has provided a large volume of data on which to develop the processing tool chain for the PAM system. The method presented here first transforms the data from its raw form into sets of 24hr long time-series and then uses CPM to perform behavioural pattern detection. Once behavioural patterns are detected, new data can be assessed and the average WBD score can be computed. The results presented show that, using the average WBD score, it is possible firstly to identify, and secondly to track the changes in a person's behavioural patterns between weekdays and weekends. Furthermore, it has been shown that the data analysis technique is robust and able to produce good results for different system configurations and in the presence of lost data and system infrastructure issues.

The overall evaluation of the PAM system has identified that the system was broadly successful in terms of data capture and processing but that there is room for improvement in the user experience and system infrastructure. The evaluation of the PAM system from a data analysis standpoint, as pertains to the work in this thesis, has shown that the data analysis meets the requirements of the system in terms of being able to correctly process the data for behavioural patterns and behavioural change. The one requirement that remains untested, that of having algorithms be able to run on low level hardware, is not expected to be a significant issue moving forwards.

Chapter 7

Conclusions and Further Work

7.1 Introduction

The work depicted in this thesis has covered the use of triaxial accelerometers and wireless cameras as sensors for the PAM system; the use of these sensors and others in the system; and the data analysis from the technical trial. The use of triaxial accelerometers for the PAM system was investigated in Chapter 4. A controlled data gathering study was carried out to obtain a body of accelerometry data that was then processed with a variety of methods to determine the best way to incorporate these sensors into the PAM system. It was found that accelerometry data can be classified into one of five categories with around 79% accuracy using PCA projection and k-NN classification. However in order to incorporate accelerometers into the PAM system, the data processing needs to be independent of the user, which led to the use of PCA projection into a single dimension to obtain an activity score from a given section of accelerometry data.

The use of cameras as single sensors for complex environments was presented in Chapter 5. Two basic algorithms were implemented from first principals, the PIDX and IH algorithms, and the WCD and HF algorithms were implemented from the literature. These four detection algorithms were evaluated on a number of different image sequences in order to test their invariance to lighting induced change and ability to detect the desired changes in the scene. The most appropriate algorithm for the PAM system was deemed to be the WCD algorithm due to its superior performance, in particular its ability under low lighting conditions.

The technical trial of the PAM system was presented in Chapter 6. The sensors used in the technical trial were discussed and the methodology of the technical trial was presented. The data from the technical trial were processed in accordance with the data processing framework presented in Chapter 3. The data coming from each data-stream were converted into 24-hr long time-series, with a time division of 15 minutes. The CPM algorithm was used in a training period to identify the behavioural patterns underlying each data-stream. In the testing period, the CPM was used to assess the similarity of new days of data against the established pattern for each data-stream. These differences were weighted by a consistency factor to give the WBD score. These scores were fused across all the data-streams to give the average WBD score, which was successfully used to detect the difference between week-days and weekends.

This chapter presents general discussion points relating to the PAM system, the conclusions that have been drawn for the project and for this thesis and some areas of the PAM system that could be investigated further.

7.2 Discussion

The PAM system that has been developed to date makes use of a number of sensors to capture a person's behavioural patterns and detect changes in these patterns. The results of the technical trial of the PAM system have shown that it is possible to detect behavioural patterns pertaining to the difference between week-days and weekend days in normal controls. The ability of the system to detect behavioural patterns and changes in behavioural patterns in people with BD has yet to be tested. However, the evidence from the technical trial is strong enough that it is reasonable to assume that this is possible with the current sensor set. The evidence from the literature on BD relating mood state to behaviour provides strong evidence that detecting changing behaviour patterns will be able to provide some insight into a user's mental health trajectory, although this will have to be fully tested in a clinical setting and there are several ways in which the existing system could be used.

In any form of PAM system the average WBD score provides a very easy entry point to the system. In a very basic system the average WBD could be used as a primary indicator for which a threshold level is defined. If the average WBD drops below this threshold, an alert can be provided to the user and/or any other interested party. If necessary, the individual data-stream data still exist within the system and the WBD scores for each of these can be examined to determine the ones which are contributing most to the low average. In turn the behavioural patterns for those data-streams could be examined in conjunction with the time-series from the alert day to identify the precise behavioural changes that caused the alert. This system could be easily expanded by using the CM described in Chapter 6. By incorporating this measure into the threshold level it would be possible to reduce the number of false alerts caused by sensor drop-out.

In a more advanced system, much of the above could be automated and a layer of computational reasoning could be applied that can make inferences about the user's mental trajectory based on the observed behavioural change. This is the overall goal for the PAM system in terms of BD care.

In any PAM system however, it would be necessary to be mindful of the ways in which peoples' behavioural patterns manifest. Throughout this data analysis of the technical trial the assumption was made that all week-days are the same and that all weekend days are the same, but different to week-days. This is most probably an oversimplification and in reality, a more complex set of training and testing conditions would be used. Initially, when the system is very recently installed, the only bulk of data available would be week-days and weekend days and behavioural patterns would have to be identified for both of these. It would also be advantageous to build behavioural patterns for each individual day, (Mondays, Tuesdays, etc.), in the event that a user had certain things happening on a specific day each week that did not happen on other days. It would also be advantageous to the behavioural change detection analysis if known events were 'pre-notified' with the system so that they could be incorporated into the analysis.

The addition of an electronic diary to the system would allow for the incorporation of known events. The diary would exist to record days for which the user knows ahead of time that their behaviour is likely to result in a change being detected, birthdays and holidays are good examples of this. Provided that the diary entry is made in advance the change detection, or alerting components of the system could be relaxed for that day.

The potential expansions to the system are not limited to the inclusion of a diary or to more test cases but extends to the straight-forward addition of more sensors and to the creation of compound data-streams. As it stands, the behavioural pattern detection component of the system only needs a set of 24hr-long time-series in order to find the behavioural pattern and be able to detect changes in a single sensor. Because of this, any new sensor can be added on the fly and be incorporated into the system and any new sensor type can be added provided that the processing to turn it's data-stream into a time-series is in place. This extensibility extends not just to the addition of new sensors, but to the combination of sensor data-streams at any stage in the data processing before the CPM. It would be entirely possible to construct complex data-streams where several data-streams are combined in a non-linear way to provide new information that cannot be gained from a single sensor. A good example of this would be the combination of pressure mat, PIR sensor and light sensor described in Section 3.6.

Finally, the PAM system as it stands has been envisaged for BD, but could easily be extended to other illnesses where behavioural pattern monitoring could be useful such as chronic fatigue syndrome (CFS) and dementia. CFS is a particularly good example as one of the coping techniques used is the maintenance of an activity log. This is often done on paper and is an ideal candidate for a technology based solution. However for any extensions to specific illness, the computational logic layer would have to be specifically reworked to take into account the particularities of the illness.

7.3 Limitations

There are two main limitations in the work that has been carried out. The first of these is that the PAM system as it stands makes no distinction between different people in the home environment. In houses with multiple occupants, as used in the technical trial, this results in the behavioural patterns of the target occupant becoming obfuscated by the behaviour of the other occupants. The use of the wearable device goes a long way towards alleviating this problem, as do any devices placed in the bedroom of the target occupant in some situations. This is likely to be the cause of the improved performance in the data from Participant #2, who was not in multiple occupancy some of the time and this would have resulted in a more defined activity signature being obtained. However, even under multiple occupancy conditions a behaviour pattern still exists; it is partially that of the household, rather than the individual, and the change detection processing is still capable of detecting change and this is a limitation of the wider PAM system, rather than of the data processing in particular.

The other main limitation of the PAM system, especially in relation to the technical trial is that the communication infrastructure failed under certain conditions. Particularly when sensor units were placed on the outer extremes of Bluetooth range. In this case, the sensors would often loose contact with the environmental node and fail to re-establish contact until the environmental node was restarted. There was an unrealistic requirement on the user to observe the node and perform this restart. This has resulted in several days of lost data from the environmental node through out the technical trial and is an issue that will need to be addressed in any further development of the PAM system. This is likely to be achieved by switching from Bluetooth to a different, more reliable technology.

7.4 Further Work

There are several areas of the PAM system that could be developed in the future. Three of the most interesting areas are solving the multiple occupancy problem, adding the computational logic layer and fully exploring the way in which the end users interacts with and uses the system and how it can be used in a clinical setting.

The multiple occupancy problem is that of determining who, in a house in multiple occupation, is the target user and of monitoring their behaviour patterns to the exclusion of the patterns of the other occupants. Any wearable sub-system is already highly specific but there is much room for improvement in the environmental monitoring sub-system. There are several techniques that could be investigated as solutions or partial solutions, but the principal aim is to localise the target user in their home and treat data that is not being produced by the target user correctly. One potential solution would be the use of small embedded cameras integrated with a PIR device. Data processing on the camera images could be used to perform face recognition to recognise the target user and count the number of people in a room. With these two pieces of information the target user's location in the home can be known at all times, as can the number of people generating data from any sensor that is in the same room as the camera-PIR device. This would allow for data that is not being generated by the target to be treated properly. A second solution, similar in idea, would be to have the user wear a radio frequency identification (RFID) tag at all times whilst in their home. A series of RFID readers around the home would therefore be able to localise the user. This could be extended to visitors and cohabitants so that the locations of all parties in the home could be known.

The computational logic layer in the PAM system would provide a means of deriving information about the user's mental health trajectory. In relation to BD, this would be a three way differentiation between stable, becoming manic and becoming depressive. There are a variety of approaches that could be taken to the implementation of such a layer. One approach could be to use a machine learning mechanism to learn which behaviour patterns in the user relate to a shift in the user's mood. This would require a training period from the user in which they provide daily information about their own assessment of their mood that can act as training data for the machine learning system. A second approach would be to construct a rules base, comprising known facts about BD and knowledge about the user's specific relapse pathways. These rules would be tied in at the bottom level to the data that was being collected and when a behavioural change is detected, the rule system could be used to make logical inferences from the data up through the system to the mood state trajectory decision.

The end user interaction with the system, and the usability in particular, has been identified as a weakness in the current PAM system. In addition to this, moving forward with the project would require the presentation of processed data to both users and clinicians. Working with these groups of users could inform the development of several areas of the system, such as the user interface, the presentation of data and the modalities of use for both clinicians and users. Of these, it is the modalities of use that are of most interest as there are different ways in which the system could be used; long-term monitoring with alert generation, monitoring after hospitalisation and short-term data gathering for example.

7.5 Conclusion

The work that has been presented in this thesis has covered work that the author has undertaken on accelerometry, image change detection and data analysis for the technical trial of the wider PAM system. The development of the PAM system and the running of the technical trial have also been presented. It has been shown in Chapter 6 that the system as specified in Chapter 3 can be used in real settings to identify activity signatures in normal controls and that these signatures can be used to identify changes in a person's behaviour patterns at the weekend as contrasted with those observed during the week. Further to being able to capture and make use of behavioural information, the PAM system manages to do this in an automated and ambient way, although the unobtrusiveness of the system is debatable. Some of the sensors, the camera in particular, are fairly large and this may not be acceptable to some users. The number of separate components in the wearable sub-system (phone, wearable node and GPS) also serve to intrude upon the user as carrying all three devices can be difficult at times.

In answering the research questions posed of the PAM project as a whole no concrete decision can be made at this stage. The results from the technical trial show that behavioural patterns and behavioural changes can be detected in normal controls. This is ample evidence to suggest that behavioural patterns could be identified for people with mental illness. Furthermore, the links between mood and behaviour in BD suggest that there is information in a person's behavioural patterns that could provide information about the trajectory of their health status, at least in people with BD. Finally, the question of whether this could be used to assist their healthcare would need to be the subject of a dedicated study, however it is not unreasonable to postulate that, in people with BD, if the system can accurately predict a person's mental health trajectory then this could be used to assist their healthcare.

In answering the research questions posed by the author at the beginning of this thesis, on the detection of behavioural patterns and changes in normal controls, the answers to both of them is broadly 'yes'; it is possible to obtain activity signatures from normal controls in an automatic and ambient way and it is also possible to use these signatures to detect behavioural changes in normal controls.

7.6 Contributions

Through the work that has been performed for this thesis and for the PAM project, the author has made several contributions. These are:

- The development of the activity signature detection. The idea of extracting activity signatures in itself is a novel idea as is the application of the CPM algorithm to achieve this.
- The fusion of several time-series in an intelligent and automatically weighted way to achieve a data analysis approach that can be applied to a system independently of the sensors that are used.
- The development and integration of the camera and PIR sensors to the system.

- The data analysis for the accelerometer.
- The incorporation of the data from the PAM sensors into the high-level data analysis.

The above contributions come together in a novel way to provide a means of determining activity signatures and behavioural changes from data gathered from a set of ambient sensors. They also form part of the wider PAM project, which combines contributions from the University of Nottingham and the University of Stirling to form the PAM system that can capture data in both environmental and ambulatory settings, and provides for these data to be analysed to extract activity signatures and detect behavioural change.

References

[1] K. Alsabti, S. Ranka, and V. Singh. An efficient k-means clustering algorithm. In *In Proceedings of IPPS/SPDP Workshop on High Performance Data Mining*, 1998.

[2] J. Amor and C. James. Personalised ambient monitoring: wireless cameras for activity detection. In *Proceedings of the 5th UKRI PG Conference in Biomedical Engineering and Medical Physics*, 2009.

[3] J. Amor and C. James. Personalized ambient monitoring: Accelerometry for activity level classification. In *4th European Conference of the International Federation for Medical and Biological Engineering*, 2009.

[4] J. Amor and C. James. Behavioral pattern detection from personalized ambient monitoring. In *Proceedings of the 32nd Annual International Conference of the IEEE EMBS*, 2010.

[5] A. P. Association. *DSM-IV-TR - Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Publishing, Inc., Washington, DC, 4th - text revision edition, 2000.

[6] B. Balasko, J. Abonyi, and B. Feil. Fuzzy clustering and data analysis toolbox. Downloadable Matlab toolbox available at: http://www.fmt.vein.hu/softcomp/fclusttoolbox/.

[7] M. R. Basco and A. J. Rush. *Cognitive-Behavioral Therapy for Bipolar Disorder*. Guilford Press, New York, 2nd edition, 2005.

[8] M. Bauer, T. Glenn, P. Grof, N. Rasgon, W. Marsh, K. Sagduyu, R. Schmid, M. Adli, and P. C. Whybrow. Comparison of sleep/wake parameters for self-monitoring bipolar disorder. *Journal of Affective Disorders*, 116(3):170–175, 2009.

[9] M. Bauer, P. Grof, N. Rasgon, T. Glenn, M. Alda, S. Priebe, R. Ricken, and P. C. Whybrow. Mood charting and technology: New approach to monitoring patients with mood disorders. *Current Psychiatry Reviews*, 2(4):423–429, 2006.

[10] M. Bauer, T. Wilson, K. Nauhaus, J. Sasse, A. Pfennig, U. Lewitzka, P. Grof, T. Glenn, N. Ragson, T. Bschor, and P. C. Whybrow. Self-reporting software for

bipolar disorder: Validation of chronorecord by patients with mania. *Psychiatry Research*, 159(3):359–366, 2008.

[11] M. Baur, P. Grof, L. Gyulai, N. Rasgon, T. Glenn, and P. C. Whybrow. Using technology to improve longitudinal studies: self-reporting with chronorecord in bipolar disorder. *Bipolar Disorders*, 6(1):66–74, 2004.

[12] R. Begg, D. T. Lai, and M. Palaniswami. *Computational Intelligence in Biomedical Engineering*. CRC Press, 2008.

[13] J. A. Bilmes. A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models. Technical report, International Computer Science Institute, 1998.

[14] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[15] J. Blum and E.H.Magill. Dynamically programmable m-psychiatry system for self-management of bipolar disorder. In *Proceedings of the 5th UKRI PG Conference in Biomedical Engineering and Medical Physics*, 2009.

[16] J. Blum and E. Magill. The design and evaluation of personalised ambient mental health monitors. In *Proceedings of the 7th Annual IEEE Consumer Communications and Networking Conference*, 2010.

[17] K. Cameron, K. Hughes, and K. Doughty. Reducing fall incidence in community elders by tele-care using predictive systems. In *Proceedings of the 19th Annual International Conference of the IEEE/EMBS*, volume 3, page 1036–1039, Oct/Nov 1997.

[18] B. Celler, N. Lovell, J. Basilakis, F. Magrabi, and M. Mathie. Home telecare for chronic disease management. In *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, volume 4, page 3586–3589, October 2001.

[19] V. Cherkassky and F. Mulier. *Learning From Data*. Wiley-Interscience, 1998.

[20] M. Clarke, T. Bratan, S. Kulkarni, and R. Jones. The impact of remote patient monitoring in managing silent myocardial infarction in a residential home setting. *The Anatolian Journal of Cardiology*, 7(Supplement 1):186–188, July 2007.

[21] T. Cover and P. Hart. Nearest neighbour pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, 1967.

[22] M. Coyle. Ambulatory cardiopulmonary data capture. In *2nd Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine and Biology*, May 2002.

[23] J. Demongeot, G. Verone, F. Duchêne, G. Benchetrit, T. Hervé, N. Noury, and V. Rialle. Multi-sensors acquisition, data fusion, knowledge mining and alarm triggering in health smart homes for elderly people. *Comptes rendus biologies*, 325(6):673–682, June 2002.

[24] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

[25] P. deToledo, S. Jimeanez, F. delPozo, J. Roca, A. Alonso, and C. Hernandez. Telemedicine experience for chronic care in copd. *IEEE Transactions on Information Technology in Biomedicine*, 10(3):567–573, July 2006.

[26] C. Ding and H. Peng. Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology*, 3(2):185–205, 2005.

[27] K. Doughty, G. Williams, P. J. King, and R. Woods. Diana - a telecare system for supporting dementia sufferers in the community. In *Proceedings of the 20th annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1998.

[28] E. Durucan and T. Ebrahimi. Change detection and background extraction by linear algebra. *Proceedings of the IEEE*, 89(10):1369–1381, 2002.

[29] J. Finkelstein and R. Friedman. Home asthma telemonitoring (hat) system. In *Proceedings of the IEEE 26th Annual Northeast Bioengineering Conference*, page 103–104, April 2000.

[30] S. C. for Mental Health. Economic and social costs of mental illness in england. Available online at:
http://www.scmh.org.uk/publications/economic+social_costs.aspx?ID=332, 2003.

[31] E. Frank, H. A. Swartz, and D. J. Kupfer. Interpersonal and social rhythm therapy: Managing the chaos of bipolar disorder. *Biological Psychiatry*, 48(6):593–604, September 2000.

[32] J. D. Frost. Triaxial accelerometry: A method for quantifying tremor and ataxia. *IEEE Transactions on Biomedical Engineering*, 25(1):17–27, January 1978.

[33] S. Gabrielli, G. Marcu, O. Mayora, and J. E. Bardram. Co-designing personal healthcare solutions for the treatment of bipolar disorder. In *Proceedings of Therapeutic Strategies A Challenge for User Involvement in Design Workshop in conjunction with NORDICHI 2010*, 2010.

[34] I. Gath and A. Geva. Unsupervised optimal fuzzy clustering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(7):773–780, Jul 1989.

[35] O. J. Gibson, W. R. Cobern, P. M. Hayton, and L. Tarassenko. A gprs mobile phone telemedicine system for self-management of type 1 diabetes. In *Proceedings of 2nd IEEE EMBSS UK and Republic of Ireland Postgraduate Conference in Biomedical Engineering and Medical Physics*, page 5–6, 2003.

[36] M. E. GmbH. The solution for efficient measurement data logging. Manufacturers website available at:
http://www.msr.ch/en/product/msr145.html.

[37] A. H. Goodwin and K. J. Sher. Effects of induced mood on diagnostic interviewing: Evidence for a mood and memory effect. *Psychological Assessment*, 5(2):197–202, 1993.

[38] D. E. Gustafson and W. C. Kessel. Fuzzy clustering with a fuzzy covariance matrix. In *1978 IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes*, volume 17, page 761–766, January 1978.

[39] J. Hender, G. Pillar, S. D. Pittman, D. Zou, L. Grote, and D. White. A novel adaptive wrist actigraphy algoritm for sleep-wake assessment in sleep apnea patients. *Sleep*, 27(8):1560–1566, August 2004.

[40] HMG/DH. No health without mental health: a cross-government mental health outcomes strategy for people of all ages. Technical report, Department of Health, 2011.

[41] O. M. Ibarra. Monitoring, treatment and prediction of bipolar disorder episodes. Presentation slides from Ubihealth - Ubicomp Workshop, Copenhagen, Denmark, 26th September 2010. Available online at: http://www.monarca-project.eu/index.php?option=com_content&view=article&id=7&Itemid=31, 2010.

[42] E. C. Ifeachor and B. W. Jervis. *Digital Signal Processing: A Practical Approach*. Prentice Hall, second edition, 2002.

[43] C. James, J. Crowe, E.Magill, S. Brailsford, J. Amor, P. Prociow, J. Blum, and S. Mohiuddin. Personalised ambient monitoring (pam) of the mentally ill. In *4th European Conference of the International Federation for Medical and Biological Engineering*, 2009.

[44] S. L. Johnson and R. L. Leahy, editors. *Psychological Treatment of Bipolar Disorder*. Guilford Press, New York, 2004.

[45] I. Jolliffe. *Principal Component Analysis*. Springer Verlag, 2002.

[46] C. Kappler-Setz, J. Schumm, M. Kusserow, B. Arnrich, and G. Tröster. Towards long term monitoring of electrodermal activity in daily life. In *12th ACM International Conference on Ubiquitous Computing UbiHealth*, 2010.

[47] D. Kelly, S. McLoone, and R. Farrell. Minimal hardware bluetooth tracking for long-term at-home elder supervision. In *Proceedings of the 32nd Annual International Conference of the IEEE EMBS*, 2010.

[48] T. Kohonen. The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480, 1990.

[49] I. Korhonen, P. Paavilainen, and A. Särela. Application of ubiquitous computing technologies for support of independent living of the elderly in real life settings. In *UbiHealth 2003: The 2nd International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications*, 2003.

[50] K. V. Laerhoven and A. K. Aronsen. Memorising what you did last week: Towards detailed actigraphy with a wearable sensor. In *7th International Conference on Distributed Computing Systems*, 2007.

[51] D. H. Lam, S. H. Jones, P. Hayward, and J. A. Bright. *Cognitive Therapy for Bipolar Disorder: A Therapist's Guide to Concepts, Methods and Practice.* Wiley, Chichester, 2000.

[52] S. R. Levine and M. Gorman. "telestroke": The application of telemedicine for stroke. *Stroke*, 30(2):464–469, 1999.

[53] J. Listgarten. *Analysis of sibling time series data: alignment and difference detection.* PhD thesis, Department of Computer Science, University of Toronto, 2006.

[54] J. Listgarten, R. M. Neal, S. T. Roweis, and A. Emili. Multiple alignment of continuous time series. In *Advances in Neural Information Processing Systems*, page 817–824, 2005.

[55] P. R. Liu, M. Q.-H. Meng, P. X. Liu, F. F. L. Tong, and X. J. Chen. A telemedicine system for remote health and activity monitoring for the elderly. *Telemedicine and e-Health*, 12(6):622–631, December 2006.

[56] A. G. Logan, W. J. McIsaac, A. Tisler, M. J. Irvine, A. Saunders, A. Dunai, C. A. Rizo, D. S. Feig, M. Hamill, M. Trudel, and J. A. Cafazzo. Mobile phone-based remote patient monitoring system for management of hypertension in diabetic patients. *American Journal of Hypertension*, 20(9):942–948, September 2007.

[57] D. Lowe. Novel 'topographic' nonlinear feature extraction using radial basis functions for concentration coding in the 'artificial nose'. In *Third International Conference on Artificial Neural Networks*, page 95–99, May 1993.

[58] D. Lowe and M. E. Tipping. Feed-forward neural networks and topographic mappings for exploratory data analysis. *Neural Computing & Applications*, 4(2):83–95, 1996.

[59] D. Lowe and M. E. Tipping. Neuroscale: Novel topographic feature extraction using rbf networks. In *Advances in Neural Information Processing Systems 9*, page 543–549, 1997.

[60] G. Lyons, K. Culhane, D. Hilton, P. Grace, and D. Lyons. A description of an accelerometer-based mobility monitoring technique. *Medical Engineering & Physics*, 27(6):497–504, July 2005.

[61] I. Maglogiannis. Design and implementation of a calibrated store and forward imaging system for teledermatology. *Journal of Medical Systems*, 28(5):455–467, 2004.

[62] Y. Malliaris. Myimonitor.com v. 1: a user-friendly mobile electronic diary for bipolar patients. *Annals of General Psychiatry*, 9(Supplement 1):S12, 2010.

[63] Y. Malliaris, N. Ferrier, and J. Scott. Monitoring bipolarity in the 21st century: The mrc london/newcastle emonitoring trial. Poster presented at the Edinburgh Conference of the International Society of Bipolar Disorders at: http://www.bipolarlab.com/wp-content/isbdemonitoring020806.pdf, 2006.

[64] M. Mathie, J. Basilakis, and B. Celler. A system for monitoring posture and physical activity using accelerometers. In *Proceedings of the 23rd Annual International Conference of the IEEE*, volume 4, page 3654–3657. Engineering in Medicine and Biology Society, October 2001.

[65] M. Mathie, N. Lovell, A. Coster, and B. Celler. Determining activity using a triaxial accelerometer. In *Second Joint EMBS/BMES Conference*, page 2481–2482, October 2002.

[66] P. McCrone, S. Dhanasiri, A. Patel, M. Knapp, and S. Lawton-Smith. *Paying the Price: The Cost of Mental Health Care in England to 2026.* the King's Fund, 2008.

[67] M. McGovern and S. Whitcher. *Altschul's Psychiatric and Mental Health Nursing.* Bailliere Tindall, London, 7th edition, 1994.

[68] Mind. Understanding mental illness. UK mental health charity website at: http://www.mind.org.uk/help/diagnoses_and_conditions/mental_illness.

[69] F. Miskelly. Electronic tracking of patients with dementia and wandering using mobile phone technology. *Age and Ageing*, 34(5):497–499, 2005.

[70] T. M. Mitchell. *Machine Learning.* McGraw-Hill, 1997.

[71] V. M. Montori, P. K. Helgemoe, G. H. Guyatt, D. S. Dean, T. W. Leung, S. A. Smith, and Y. C. Kudva. Telecare for patients with type 1 diabetes and inadequate glycemic control. *Diabetes Care*, 27(5), May 2004.

[72] I. Nabney. Netlab. Downloadable Matlab toolbox available at: http://www.ncrg.aston.ac.uk/netlab/down.php.

[73] K. Nakajima and K. Sasaki. A simple system for telemonitoring the daily life of an aged person living alone. In *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, page 3724–3726, September 2005.

[74] M. Nambu, K. Nakajima, A. Kawarada, and T. Tamura. The automatic health monitoring system for home health care. In *Proccedings of the IEEE EMBS International Conference on Information Technology Applications in Biomedicine*, page 79–82, November 2000.

[75] M. Nambu, K. Nakajima, and T. Tamura. An algorithm for the automatic detection of health conditions. *IEEE Engineering in Medicine and Biology Magazine*, 24(4):38–42, July/August 2005.

[76] M. Nixon and A. Aguado. *Feature Extraction and Image Processing.* Newnes, 2002.

[77] A. C. Norris. *Essentials of Telemedicine and Telecare.* Wiley, Chichester, 2002.

[78] N. Noury, T. Herve, V. Rialle, G. Virone, E. Mercier, G. Morey, A. Moro, and T. Porcheron. Monitoring behavior in home using a smart fall sensor and position sensors. In *Proceedings of the 1st Annual International Conference On Microtechnologies in Medicine and Biology*, page 607–610, Octiber 2000.

[79] N. Noury, G. Virone, and T. Creuzet. The health integrated smart home information system (his2): rules based system for the localization of a human. In *Proceedings of the 2nd Annual International IEEE-EMB Special Topic Conference on Microtechnologies in Medicine & Biology*, page 318–321, May 2002.

[80] N. N. A. on Mental Illness. What is mental illness: Mental illness facts. Web-page available at:
http://www.nami.org/template.cfm?section=about_mental_illness.

[81] W. H. Organisation. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines.* World Health Organization, 1992.

[82] W. H. Organisation. *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research.* World Health Organization, 1993.

[83] D. W. Patterson. *Artificial Neural Networks: Theory and Applications.* Prentice Hall, 1996.

[84] H. Peng. Minimum redundancy maximum relevance feature selection. Downloadable Matlab toolbox available at: http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=14916&objectType=FILE.

[85] H. Peng, F. Long, and C. Ding. Feature selection based on mutual information criteria of max-dependancy, max-relevance and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1226–1238, 2005.

[86] R. H. Perlis, M. J. Ostacher, J. K. Patel, L. B. Marangell, H. Zhang, S. R. Wisniewski, T. A. Ketter, D. J. Miklowitz, M. W. Otto, L. Gyulai, N. A. Reilly-Harrington, A. A. Nierenberg, G. S. Sachs, and M. E. Thase. Predictors of recurrence in bipolar disorder: Primary outcomes from the systematic treatment enhancement program for bipolar disorder (step-bd). *The American Journal of Psychiatry*, 163(2):217–224, February 2006.

[87] M. Piccardi. Background subtraction techniques: a review. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 2004.

[88] C. Price, B. Archer, B. Dunk, J. Ellam, K. Gough, A. Morris, S. Morris, K. Saul, and A. Sheeham. Just checking: lessons three years on. *Journal of Dementia Care*, 17(3):12–13, 2009.

[89] J. G. Proakis and D. G. Manolakis. *Digital Signal Processing: Principals, Algorithms, and Applications*. Prentice Hall, fourth edition, 2007.

[90] P. Prociow and J. Crowe. Sensors enhancing self-monitoring for people with bipolar disorder. In *Proceedings of the 5th UKRI PG Conference in Biomedical Engineering and Medical Physics*, page 37–38, 2009.

[91] P. Prociow, J. A. Crowe, S. C. Brailsford, C. J. James, M. E., J. Amor, J. Blum, and S. Mohiuddin. Personalised ambient monitoring (pam) for people with bipolar disorder. In *Proceedings of the International eHealth, Telemedicine and Health ICT Forum For Education, Networking and Business (MedETel)*, page 531–534, 2009.

[92] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–285, 1989.

[93] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam. Image change detection algorithms: A systematic survey. *IEEE Transactions on Image Processing*, 14(3):294–307, 2005.

[94] S. J. Russell and J. L. Browne. Staying well with bipolar disorder. *Australian and New Zealand Journal of Psychiatry*, 39(3):187–193, March 2005.

[95] J. W. Sammon. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers*, C-18(5):401–409, 1969.

[96] S. Shiffman and M. R. Hufford. Ecological momentary assessment (subject experience diaries in clinical research, part 2). *Applied Clinical Trials*, 10(3):42–48, 2001.

[97] S. Shiffman, M. R. Hufford, and J. Paty. The patient experience movement (subject experience diaries in clinical research, part 1). *Applied Clinical Trials*, 10(2):46–56, 2001.

[98] L. Smith. A tutorial on principal component analysis. Unpublished tutorial at: http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf, 2002.

[99] A. Society. About dementia. UK Alzheimer's Society charity website at: http://alzheimers.org.uk/site/index.php, 2011.

[100] E. J. W. V. Someren. Actigraphic monitoring of movement and rest-activity rhythms in aging, alzheimer's disease, and parkinson's disease. *IEEE Transactions on Rehabilitation Engineering*, 5(4):394–398, December 1997.

[101] C. Stauffer and W. Grimson. Adaptive background mixture models for real-time tracking. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1999.

[102] A. A. Stone, S. Shiffman, J. E. Schwartz, J. E. Broderick, and M. R. Hufford. Patient compliance with paper and electronic diaries. *Controlled Clinical Trials*, 24(2):182–199, April 2003.

[103] A. Szot, F. L. Jacobson, S. Munn, D. Jazayeri, E. Nardell, D. Harrison, R. Drosten, L. Ohno-Machado, L. M. Smeaton, and H. S. F. Fraser. Diagnostic accuracy of chest x-rays acquired using a digital camera for low-cost teleradiology. *International Journal of Medical Informatics*, 73(1):65–73, 2004.

[104] D. Tacconi, O. Mayora, P. Lukowicz, B. Arnrich, C. Setz, G. Tröster, and C. Haring. Activity and emotion recognition to support early diagnosis of psychiatric diseases. In *Proceedings of 2nd International Conference on Pervasive Computing Technologies for Healthcare (Pervasive Health)*, 2008.

[105] M. E. Tipping and D. Lowe. Shadow targets: A novel algorithm for topographic projections by radial basis functions. *NeuroComputing*, 19:211–222, 1997.

[106] D. Toth, T. Aach, and V. Metzler. Illuminationinvariant change detection. In *Image Analysis and Interpretation, 2000. Proceedings. 4th IEEE Southwest Symposium*, page 3–7, 2000.

[107] G. Virone, N. Noury, and J. Demongeot. A system for automatic measurement of circadian activity deviations in telemedicine. *IEEE Transactions on Biomedical Engineering*, 49(12):1463–1469, December 2002.

[108] D. Wallace, S. M. Jones, C. Milroy, and M. A. Pickford. Telemedicine for acute plastic surgical trauma and burns. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 61(1):31–36, 2008.

[109] C. Xijun, M. Q.-H. Meng, and R. Hongliang. Design of sensor node platform for wireless biomedical sensor networks. In *Engineering in Medicine and Biology 27th Annual Conference*, September 2005.

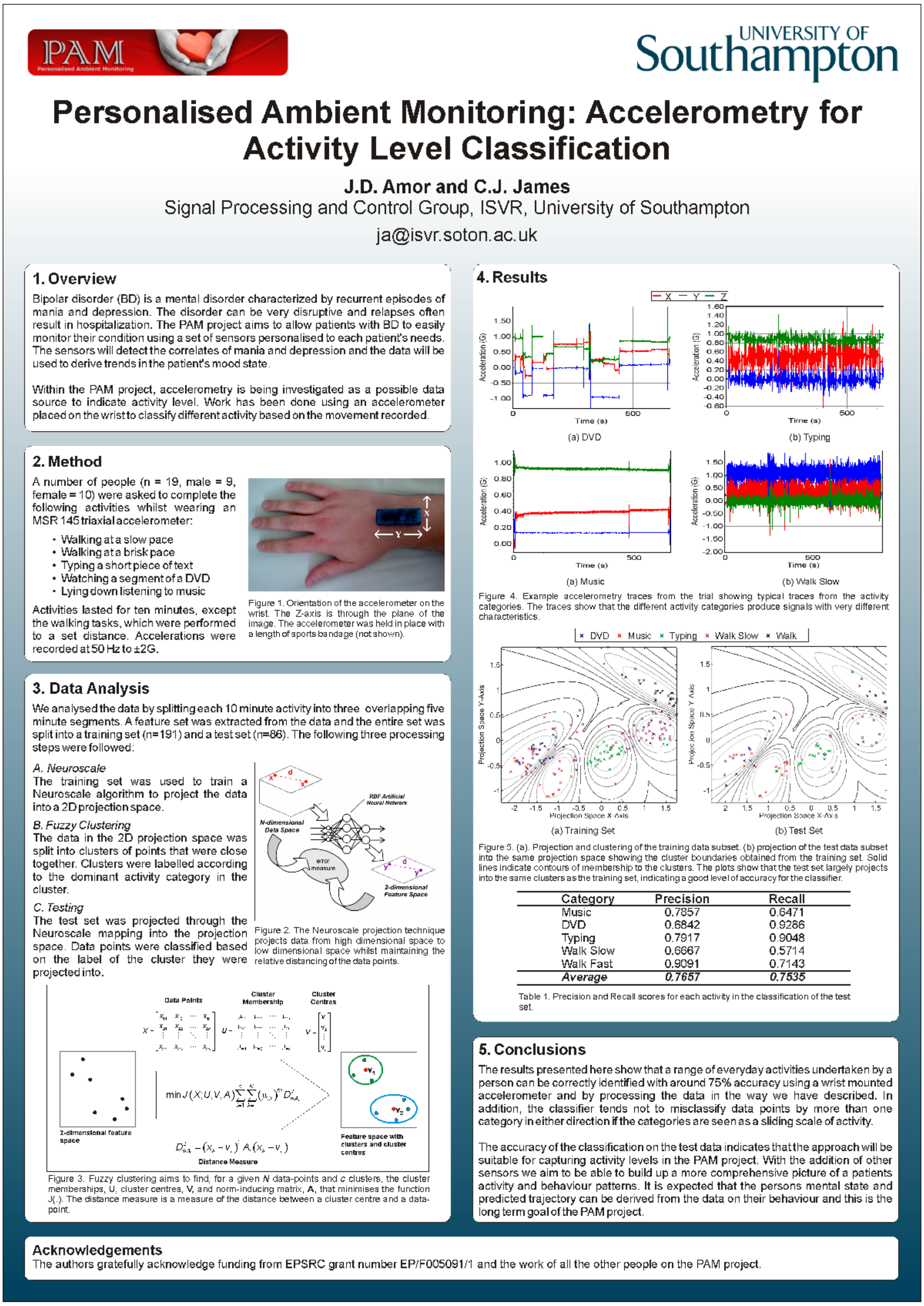
[110] J. Zheng, Y. Wang, N. Nihan, and E. Hallenbeck. Extracting roadway background image: A mode based approach. *Transportation Research Record: Journal of the Transportation Research Board*, (1944):82–88, 2006.

Appendix A

Publications

A.1 EMBEC 2008 Conference

A.1.1 Poster



PAM
Personalised Ambient Monitoring

UNIVERSITY OF Southampton

Personalised Ambient Monitoring: Accelerometry for Activity Level Classification

J.D. Amor and C.J. James
Signal Processing and Control Group, ISVR, University of Southampton
ja@isvr.soton.ac.uk

1. Overview

Bipolar disorder (BD) is a mental disorder characterized by recurrent episodes of mania and depression. The disorder can be very disruptive and relapses often result in hospitalization. The PAM project aims to allow patients with BD to easily monitor their condition using a set of sensors personalised to each patient's needs. The sensors will detect the correlates of mania and depression and the data will be used to derive trends in the patient's mood state.

Within the PAM project, accelerometry is being investigated as a possible data source to indicate activity level. Work has been done using an accelerometer placed on the wrist to classify different activity based on the movement recorded.

2. Method

A number of people (n = 19, male = 9, female = 10) were asked to complete the following activities whilst wearing an MSR 145 triaxial accelerometer:

- Walking at a slow pace
- Walking at a brisk pace
- Typing a short piece of text
- Watching a segment of a DVD
- Lying down listening to music

Activities lasted for ten minutes, except the walking tasks, which were performed to a set distance. Accelerations were recorded at 50 Hz to ±2G.

Figure 1. Orientation of the accelerometer on the wrist. The Z-axis is through the plane of the image. The accelerometer was held in place with a length of sports bandage (not shown).

3. Data Analysis

We analysed the data by splitting each 10 minute activity into three overlapping five minute segments. A feature set was extracted from the data and the entire set was split into a training set (n=191) and a test set (n=86). The following three processing steps were followed:

A. Neuroscale
The training set was used to train a Neuroscale algorithm to project the data into a 2D projection space.

B. Fuzzy Clustering
The data in the 2D projection space was split into clusters of points that were close together. Clusters were labelled according to the dominant activity category in the cluster.

C. Testing
The test set was projected through the Neuroscale mapping into the projection space. Data points were classified based on the label of the cluster they were projected into.

Figure 2. The Neuroscale projection technique projects data from high dimensional space to low dimensional space whilst maintaining the relative distancing of the data points.

Figure 3. Fuzzy clustering aims to find, for a given *N* data-points and *c* clusters, the cluster memberships, U, cluster centres, **V**, and norm-inducing matrix, **A**, that minimises the function J(.). The distance measure is a measure of the distance between a cluster centre and a data-point.

4. Results

Figure 4. Example accelerometry traces from the trial showing typical traces from the activity categories. The traces show that the different activity categories produce signals with very different characteristics.

Figure 5. (a). Projection and clustering of the training data subset. (b) projection of the test data subset into the same projection space showing the cluster boundaries obtained from the training set. Solid lines indicate contours of membership to the clusters. The plots show that the test set largely projects into the same clusters as the training set, indicating a good level of accuracy for the classifier.

Category	Precision	Recall
Music	0.7857	0.6471
DVD	0.6842	0.9286
Typing	0.7917	0.9048
Walk Slow	0.6667	0.5714
Walk Fast	0.9091	0.7143
Average	***0.7657***	***0.7535***

Table 1. Precision and Recall scores for each activity in the classification of the test set.

5. Conclusions

The results presented here show that a range of everyday activities undertaken by a person can be correctly identified with around 75% accuracy using a wrist mounted accelerometer and by processing the data in the way we have described. In addition, the classifier tends not to misclassify data points by more than one category in either direction if the categories are seen as a sliding scale of activity.

The accuracy of the classification on the test data indicates that the approach will be suitable for capturing activity levels in the PAM project. With the addition of other sensors we aim to be able to build up a more comprehensive picture of a patients activity and behaviour patterns. It is expected that the persons mental state and predicted trajectory can be derived from the data on their behaviour and this is the long term goal of the PAM project.

Acknowledgements
The authors gratefully acknowledge funding from EPSRC grant number EP/F005091/1 and the work of all the other people on the PAM project.

FIGURE A.1: The poster presented by the author at the 4th European Medical and Biological Engineering Conference, November 2008, Antwerp, Belgium, (EMBEC'08).

A.1.2 Paper

The paper included here accompanies the poster above, which was presented at the 4th European Medical and Biological Engineering Conference, November 2008, Antwerp, Belgium, (EMBEC'08) [3].

Personalized Ambient Monitoring: Accelerometry for Activity Level Classification

J.D. Amor and C.J. James

Signal Processing and Control Group, ISVR, University of Southampton, Southampton, UK

Abstract— **Bipolar disorder (BD) is a mental disorder characterized by recurrent episodes of mania and depression. The disorder can be very disruptive and relapses often result in hospitalization. With adequate training, sufferers are able to control their symptoms and reduce the disruption to their daily lives. As an aid to this self-control process the Personalized Ambient Monitoring (PAM) project is being developed.**

The PAM project aims to allow patients with BD to monitor their condition and obtain indications of their mental state. This will be achieved through the use of multiple discreet sensors, personalized for each patient's needs. The sensors will detect the correlates of mania and depression, which will be used to derive trends in the mental health state of the patient.

The major symptoms of BD center on the patient's activity level and circadian rhythm. Manic episodes are typified by increased energy and activity, often with a decreased need for sleep. Depressive episodes however often present with diminished activity. It is our aim that by measuring the patient's activity levels and circadian rhythm we can provide information that the patient can use to help control their symptoms.

Here we present some preliminary work aimed at distinguishing different activities and activity levels in normal controls, based on a small, body-mounted triaxial accelerometer. A number of participants were asked to complete some basic activities whilst wearing the accelerometer. The data was preprocessed to extract a number of salient features, which were used to train a Neuroscale algorithm. Neuroscale produces a generative mapping that visualizes high-dimensional data in a lower-dimensional space, which, with the addition of a clustering algorithm, can be used to classify unknown data points. It is expected that this approach, combined with data from other sensor types will form the backbone of the PAM approach applied to BD.

Keywords— **Activity monitoring, bipolar disorder, accelerometry, Neuroscale, data classification.**

I. INTRODUCTION

Bipolar disorder (BD) is a mental disorder that affects approximately 0.4–1.6% of the population [1]. The disorder is characterized by recurring episodes of mania and depression, which can last between 2 weeks and 5 months for mania and slightly longer for depression [2]. Symptoms of mania include elevated mood, increased activity, pressure of thoughts and speech, and decreased need for sleep. Symptoms of depression include decreased activity, sleep disorders and depressed mood [2].

There are many treatments, pharmacological and therapeutic, available for BD. The gold standard for pharmacological treatment is lithium, which can be taken both during affective episodes and as a prophylactic [3]. However, around 20–40% of the patients who take lithium do not respond to treatment [4], and in those that do there may be severe side effects. Other medications, such as antipsychotics and anti-depressants are a common alternative to lithium [3].

Therapeutic treatments for BD use a number of approaches, such as cognitive therapy [4,5], social rhythm therapy [6] and family therapy [3]. A common feature throughout many therapeutic approaches is the identification and management of the early warning signs of an oncoming episode. Research has shown that patients who are able to identify and act on the early warning signs are better able to manage their mood and maintain a more stable health condition [7].

The PAM project is being developed as a self-help tool to help identify the early warning signs and predict the patient's mood state to enable the patients themselves to better control their condition.

II. METHODOLOGY

A. PAM Approach

The PAM approach to BD centers around the use of a number of personalized, discreet, ambient sensors to capture information on the patient's behavior patterns and also on some of the specific correlates of mania and depression. It is expected that the sensor set will include both wearable and static sensors. Examples of these include accelerometers to capture information about patient activity, passive infra-red sensors to detect room occupation and magnetic contact switches to detect door opening and closing. The use of both wearable and static sensors will enable data to be gathered in both domestic and ambulatory settings.

The sensor set will be customizable to the needs of each patient, since what is acceptable for one patient will not be acceptable for another. There is also the requirement, particularly for any wearable sensors, that the sensor and associated hardware is as small as possible with low power consumption – i.e. long battery life. Further details on the PAM project and sensor set up can be found in [8].

B. System Architecture

In order to process the large volumes of data that will be produced by the sensors, a three level architecture has been proposed, as shown in Figure 1. The node level is intended to perform feature extraction and transform the raw data from a sensor into a form that can be more easily transmitted and manipulated further up the processing chain.

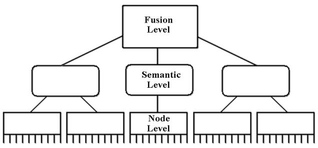


Fig. 1 PAM data processing architecture showing the Node, Semantic and Fusion levels of data processing.

The semantic level extracts basic meaning from the features passed to it. In the case of accelerometry data, the features might include measures of amplitude and frequency and the semantic meaning would be the classification of activity, such as ‘walking’ or ‘sitting’. The format of the semantic information will include several additional pieces of information, such as a time-stamp and a tag to identify the type of sensor used. Both the node and semantic levels will have to be specific to particular devices and can be seen to combine to produce useful information from a particular sensor’s raw data.

The fusion level is intended to bring together all the semantic components and provide the predictive capabilities of the system. It is expected that this will be achieved by developing specific pattern matching algorithms to identify recurring patterns in the data that can be used to help model the patient’s behavior patterns. Once a model has been generated, it will be possible to observe deviations in behavior and generate alerts appropriately.

III. Experimental work

A. Method

A number of people (n = 19, male = 9, female = 10) were asked to wear a small triaxial accelerometer (an MSR 145 [9]) on their dominant wrist and carry out a number of activities: walking (at both a slow and a brisk pace); typing a short piece of text; watching a short segment of a DVD; and lying down listening to music. The activities were timed to last for ten minutes, with the exception of the walking tasks, which were performed to a set distance.

The device was secured to the participant’s wrist using a length of elasticized sports bandage that had been modified to securely hold the accelerometer on the outside of the participant’s wrist. Accelerations were recorded to ±2G, with a sampling rate of 50 Hz, which was sufficient to capture the wrist movement whilst allowing for several runs to be stored to the device's memory.

B. Data Analysis

Each ten minute activity was split into three five minute segments, with the middle segment having a 50% overlap. Feature extraction was performed on each of the three axes (x, y and z) and on the total acceleration defined as $a=\sqrt{x^2+y^2+z^2}$. The data set was split randomly into a training set (n=191) and test set (n=86).

The training set was used to train a supervised Neuroscale algorithm [10] to produce a generative mapping from the multi-dimensional feature space into a two-dimensional projection space. Given the n-dimensional feature space containing N data points $\mathbf{x}_i$, and prior knowledge of the class association of the points in $\mathbf{x}_i$, the two dimensional space of points $\mathbf{y}_i$ is created such as to minimize the stress measure

$$E=\sum_{i<j}^{N}\left((1-\alpha)d_{ij}^{*}+\alpha s_{ij}-d_{ij}\right)^{2} , \qquad (1)$$

where d_{ij}^{*} are the inter-point Euclidian distances in the feature space, d_{ij} are the corresponding distances in the projection space, s_{ij} is a subjective dissimilarity matrix, corresponding to the prior knowledge of the classes of the data points, and α can be used to control the extent of the influence of the dissimilarity matrix.

The 2D projection of the data set from the Neuroscale algorithm was passed through the Gustafson-Kessel fuzzy clustering algorithm [11] to identify the data clusters in an unsupervised way. Fuzzy clustering is based on the minimization of the objective function

$$J(\mathbf{X};\mathbf{U},\mathbf{V},\mathbf{A})=\sum_{i=1}^{c}\sum_{k=1}^{N}(\mu_{ik})^{m}D_{ikA_i}^{2} , \qquad (2)$$

where $\mathbf{X}$ is a matrix of N data points, $\mathbf{U}$ is a matrix, from which μ is drawn, of memberships for the points in $\mathbf{X}$ to the clusters, $\mathbf{V}$ is a vector of c cluster centers that have to be found and $\mathbf{A}$ is set of norm-inducing matrices used in the calculation of

$$D_{ikA_i}^{2}=(\mathbf{x_k}-\mathbf{v_i})^{\mathrm{T}}A_i(\mathbf{x_k}-\mathbf{v_i}) \qquad (3)$$

This aims to minimize the total variance of $\mathbf{x}$ from $\mathbf{v}$ by adjusting the membership values in $\mathbf{U}$, and the cluster centers in **V**. Using a separate norm-inducing matrix **A**, which must be constrained to have a fixed determinant, for each cluster allows for the cluster shape to change whilst maintaining a constant volume.

Clusters were labeled based on the dominant class of all the data points in the cluster with a membership of greater than a half. The clustering algorithm creates a set of cluster centers and provides the functionality to assign membership to each cluster for any previously unseen data points.

The combination of Neuroscale projection with the identification of the data clusters enabled the test data set to be classified. Each data point was projected through the Neuroscale mapping and classified based on the label of the cluster to which the data point had the highest membership.

The features used for classification were selected using the maximum relevance, minimum redundancy (mRMR) feature selection method described in [12]. As a further optimization step, the number of hidden centers for the Neuroscale algorithm was optimized along with the number of features to find the optimal settings for the data analysis.

IV. Results

Figure 2 shows examples of the traces obtained from the accelerometer. The four traces are visibly different in their characteristics; for example, the 'Walk Slow' trace clearly has a higher change in amplitude than 'DVD'. There are some similarities however between the 'DVD' and 'Music' traces, which both contain significant periods with little or no movement.

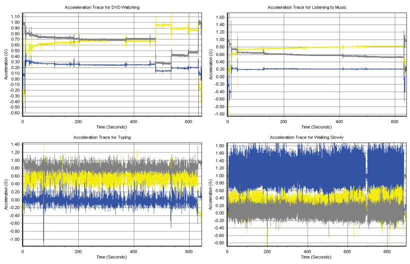


Fig 2. Acceleration traces obtained in the study. Top left = 'DVD', top right = 'Music', bottom left = 'Typing' and bottom right = 'Walk Slow'. Three axis of acceleration are shown: X (yellow), Y (blue) and Z (grey).

Figure 3 shows the result of the fuzzy clustering algorithm applied to the training set. It can be seen from the plot that three of the five classes are reasonably well defined as clusters but that the classes 'DVD' and 'Music' are quite blurred; 'DVD' contains a large number of points which should be in 'Music'. This can be seen in the measures of precision and recall in Table 1, defined as

$$\text{precision} = \frac{tp}{tp + fp}, \quad \text{recall} = \frac{tp}{tp + fn} \tag{4}$$

where tp is the number of true positives, fp the number of false positives and fn the number of false negatives, for each class. It can be seen that 'Music' has a very low recall, whilst 'DVD' has a correspondingly low precision.

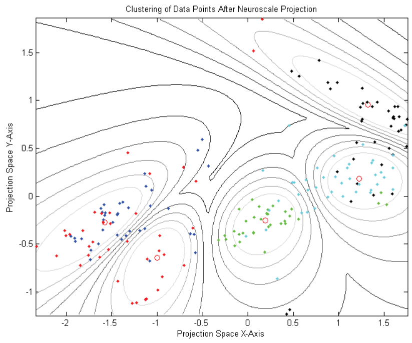


Fig 3. Clustering of the training data set after projection through Neuroscale mapping. Colored data points indicate different activity classes. (Red = Music, Blue = DVD, Green = Typing, Cyan = Walk Slow, Black = Walk Fast). Solid lines indicate contours of membership to different clusters.

Table 1. Measures of precision and recall for each class in the training set after projection through the Neuroscale mapping and subsequent clustering.

Class	Precision	Recall
Music	0.6667	0.3784
DVD	0.5714	0.8000
Typing	0.7317	0.9091
Walk Slow	0.6500	0.6842
Walk Fast	0.9000	0.7105
Average	0.7040	0.6964

Table 2 shows the classifications obtained by processing the test data set. The precision and recall for the activity classes is shown in Table 3. The overall precision and recall are reasonably high at 76.6% and 75.3% respectively. However, it should be noted that there is considerable variation in the precision and recall scores across the different classes. The 'Walk Slow' class in particular has a poor recall as a high percentage of the 'Walk Slow' data points are classed as 'Typing'.

Table 2. Absolute classification of data points in test set

Class	Total Number	Number classified as: Music	DVD	Typing	Walk Slow	Walk Fast
Music	17	11	6	0	0	0
DVD	14	1	13	0	0	0
Typing	21	2	0	19	0	0
Walk Slow	14	0	0	5	8	1
Walk Fast	14	0	0	0	4	10

Table 3. Precision and recall for the test set

Class	Precision	Recall
Music	0.7857	0.6471
DVD	0.6842	0.9286
Typing	0.7917	0.9048
Walk Slow	0.6667	0.5714
Walk Fast	0.9091	0.7143
Average	0.7657	0.7532

V. Discussion and conclusions

The results presented here show that a range of everyday activities undertaken by a person can be correctly identified using a wrist mounted accelerometer with around 75% accuracy. In addition, the major inaccuracy in the classification comes from the misclassification of some of the 'Walk Slow' data points as 'Typing' and some of the 'Music' data points as 'DVD'.

The misclassification of 'Music' as 'DVD' is understandable, since the two classes are very similar in the physical activity required. Both consist of long periods of inactivity interspersed with brief periods of movement. This leads to the accelerometry traces looking very similar, as shown in Figure 2, and to the subsequent misclassification, which is consistent with the training data shown in Figure 3.

The misclassification of 'Walk Slow' as 'Typing' is harder to explain as the two activity classes are quite distinct in the physical activity required. The poor classification could be partially due to overlaps in the frequency components of some participants but this would have to be investigated further before any precise conclusions can be drawn.

The results obtained in the experimental study presented here are very encouraging from the point of view of using this data for the node and semantic level processing in the PAM project. The classification accuracy is sufficiently high and, more importantly, tends not to misclassify data points by more than one category in either direction if the categories are seen on a sliding scale of activity. This will enable the use of accelerometry to give a general idea of the level of activity of a BD sufferer and with the addition of other sensors will form a solid backbone for the capture of activity signatures.

Further work on the PAM project, particularly from the data processing perspective, will focus on the node and semantic level processing for different sensors and their integration to the project as a whole. Once a sufficient number of sensors have been integrated, algorithms for the prediction of the patient's mood state can be developed.

Acknowledgment

The authors gratefully acknowledge funding from EPSRC grant number EP/F005091/1.

References

[1] American Psychiatric Association. *DSM-IV-TR - Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Publishing, Inc., Washington, DC, 4th - text revision edition, 2000

[2] World Health Organisation. *The ICD-10 Classification of Mental and Behavioural Disorders*. World Health Organization, 1992.

[3] Sheri L. Johnson and Robert L. Leahy, editors. *Psychological Treatment of Bipolar Disorder*. Guilford Press, New York, 2004.

[4] Dominic H. Lam, Steven H. Jones, Peter Hayward, and Jenifer A. Bright. *Cognitive Therapy for Bipolar Disorder: A Therapist's Guide to Concepts, Methods and Practice*. Wiley, Chichester, 2000.

[5] Monica Ramirez Basco and A. John Rush. *Cognitive-Behavioral Therapy for Bipolar Disorder*. Guilford Press, New York, 2nd edition, 2005.

[6] Ellen Frank, Holly A. Swartz, and David J. Kupfer. *Interpersonal and social rhythm therapy: Managing the chaos of bipolar disorder*. Biol. Psych., 48(6):593–604, September 2000.

[7] Sarah J. Russell and Jan L. Browne. *Staying well with bipolar disorder*. Aust N Z J Psychiatry, 39(3):187–193, March 2005.

[8] C.J. James, J. Crowe, E. Magill, S.C. Brailsford, J. Amor, P. Prociow, J. Blum and S. Mohiuddin. *Personalised Ambient Monitoring (PAM) of the mentally ill*. EMBEC 2008 Congress.

[9] http://www.msr.ch/en/index.php

[10] David Lowe and Michael E. Tipping. *Feed-Forward Neural Networks and Topographic Mappings for Exploratory Data Analysis*. Neural Computing & Applications, 4(2):83–95, 1996.

[11] Balazs Balasko, Janos Abonyi and Balazs Feil. *Fuzzy Clustering and Data Analysis Toolbox*. Department of Process Engineering at the University of Veszprem, Hungary. Web resource at: http://www.fmt.vein.hu/softcomp/fclusttoolbox/

[12] Hanchuan Peng, Fuhui Long and Chris Ding. *Feature Selection Based on Mutual Information Criteria of Max-Dependency, Max-Relevance and Min-Redundancy*. TPAMI, 27(8):1226–1238, 2005.

Corresponding Author

James Amor
Signal Processing and Control Group, ISVR,
University of Southampton
University Road
Southampton
UK
jda103@soton.ac.uk

A.2 PGBiomed 2009

The paper included here was presented at the The 5th IEEE EMBS UK & Republic of Ireland postgraduate conference on biomedical engineering and medical physics, July 2009, Oxford, UK, (PGBIOMED 09) [2].

PERSONALISED AMBIENT MONITORING: WIRELESS CAMERAS FOR ACTIVITY DETECTION

J.D. Amor and C.J. James[1]

[1]Signal Processing and Control Group, ISVR, University of Southampton, Southampton, UK
ja@isvr.soton.ac.uk,

Abstract – A wireless network camera is being used in the PAM project to monitor activity in the home environment. Images are processed to extract 'blobs' of activity in the scene, which are then checked for intersection with areas of interest. We describe the hardware set-up and the processing algorithms used and present some preliminary results.

I. INTRODUCTION

Bipolar disorder (BD) is a severe mental disorder characterised by recurring episodes of mania and depression. It is estimated that BD affects between 0.4–1.6% of the population [1]. Many BD sufferers monitor their condition, keeping records of their activity, to identify the onset of an affective episode so that they can take steps to lessen and counteract the severity of the episode [2].

The Personalised Ambient Monitoring (PAM) project aims to develop a system to help patients self-monitor their condition and to predict their mental health trajectory. Using a network of personalised, discreet sensors the patient's activity and behavioural patterns will be detected. These will be used to derive a 'normal' pattern of behaviour, which will be monitored for changes that indicate the onset of an affective episode.

Both the patient and their home environment will be monitored. There is difficulty however, in effectively monitoring some rooms in a patient's home, such as a kitchen, where there are many areas of interest (AOIs) that need to be monitored. The traditional approach to monitoring each AOI would involve a number of different sensors each tailored to a specific task and each requiring separate wired or wireless communication links; the system becomes very complicated very quickly.

An alternative solution is to use a wireless camera, with a wide-angle lens, to take images of the room and to process these images to extract activity in the AOIs and it is this approach that is being taken with the PAM project. This paper discusses the hardware set-up and processing algorithms that have been developed so far. Some preliminary results are also presented.

II. METHOD

A . Hardware Set-up

The camera being used for the PAM project is an Edimax IC-1520DPg with a 179° horizontal and 129° vertical field of view [3]. The camera is mounted as high in the room as possible to get a birds-eye view as this minimises the problems involved in objects passing in front of the camera and causing false activity to be registered.

The camera is connected through a secured *ad-hoc* wireless link to a base-station PC and uploads a snapshot of the room every 10 seconds.

B. Image Processing

The images that are provided by the camera are processed in sequence to extract areas of movement from the background in the scene. The idea is to render an image into 'blobs', which are a black and white representation of the scene with the white blobs showing areas of activity. Before the images are processed, AOIs are defined so that activity can be correctly identified. The image processing follows the following steps:

1. Background estimation with a temporal median of several images.
2. Background subtraction / change detection using either the Wronskian change detector or an in-house developed algorithm.
3. Activity detection within pre-defined areas of interest.

The image processing algorithms are used to detect things in an image that are different to the background image that has been derived from the preceding sequence.

The Wronskian change detection model (WM) is based on the idea of linear independence [4]. Each pixel in the image is represented as a vector comprised of itself and its neighbouring pixels. If two corresponding vectors from two separate images are linearly dependent then there has been no physical change in the scene.

The in-house (IH) algorithm is based on a simple pixel difference. Two images are differenced and where a difference is recorded a physical change in the scene is assumed to

have taken place. The resultant image is median filtered to remove speckle noise and then Gaussian filtered to blur the image and produce strong blob responses where activity has occurred.
Activity detection is performed by examining each AOI. If a sufficient number of pixels in an AOI are white (i.e. indicate a change in the scene) then an activity is recorded.

III. RESULTS

Figure 1 shows a living room in which the camera has been mounted in one corner of the room. AOIs have been marked in the image over the sofa, the patio door and a chair.
Figure 2(a) shows the result of applying the WM algorithm to the image sequence and figure 2(b) shows the result of applying the IH algorithm. The author has been clearly been identified in both images and would cause an event to be registered in the AOI set on the sofa. The IH algorithm is significantly less noisy, in terms of speckle noise, than the WM due to the filtering applied after the image differencing.

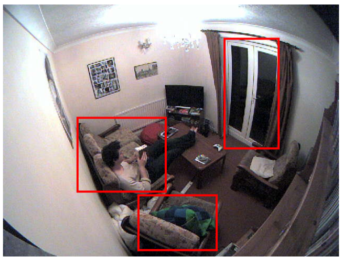
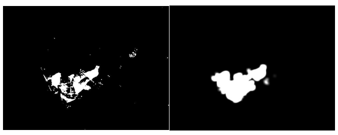


Figure 1. A camera image before processing with some AOIs marked in red.



(a) (b)

Figure 2. The same image after processing with the WM algorithm (a) and the IH algorithm (b). The presence of the author has been clearly detected by both algorithms.

Figure 3 shows a section of the image sequence where the author sits down in an empty room. The IH processed blob images are also shown. The sequence shows that the author is successfully extracted from the background in a sequence of images.

Figure 3. Five frames of the original image sequence, showing the empty room followed by the authors presence, and the blobbed images achieved with the IH algorithm.

IV. DISCUSSION

The results shown above show that changes can be detected in an image and that this detection can be used to identify activity in areas of interest in the image. Two image processing algorithms have been tested, the WM and the IH. Both algorithms successfully extract activity in AOIs but differ in their responses to lighting changes. A lighting change in the scene, either global (e.g. ambient light change) or local (e.g. reflections) can cause false changes to be registered. The IH algorithm is particularly susceptible to this; since it is based on simple pixel differences any change in light will be registered as changes in the scene. The filtering applied to the image goes some way towards eliminating this noise but does not achieve prefect results. The WM is significantly better in terms of lighting noise reduction as each pixel is considered in relation to its neighbours. This leads to less lighting induced noise but can cause the resultant blobs to be less coherent, as seen in figure 2 (a).

V. CONCLUSION

The camera and blob detection image analysis approach is still a work in progress. Promising results have been achieved so far and the approach seems to be viable as a way of monitoring a complex environment. Further improvements to the image processing algorithms are required before the approach can be fully integrated into the PAM system.

REFERENCES

[1] American Psychiatric Association. *DSM-IV-TR - Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Publishing, Inc., Washington, DC, 4th - text revision edition, 2000.
[2] Sarah J. Russell and Jan L. Browne. *Staying well with bipolar disorder*. Aust N Z J Psychiatry, 39(3):187–193, March 2005.
[3] Edimax IC-1520DPg Digital Pan / Tilt Wireless Network Camera, last accessed on 13/03/2009 at: http://www.edimax.co.uk/en/produce_detail.php?pd_id=236&pl1_id=1&pl2_id=7
[4] Durucan, E., and Ebrahimi T., *Change Detection and Background Extraction by Linear Algebra*, in Proceedings of the IEEE, 89(10):1368–1381, October 2001.

A.3 EMBC 2010

The paper included here was presented at the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, September 2010, Buenos Aires, Argentina, (EMBC 2010) [4].

Behavioral Pattern Detection from Personalized Ambient Monitoring

James D. Amor, *Student Member, IEEE,* and Christopher J. James, *Senior Member, IEEE*

***Abstract*—Bipolar disorder (BD) is a serious psychiatric condition that affects a large number of people. Many people with BD self-monitor their condition in order to try and keep the disturbances from affective episodes to a minimum. The Personalized Ambient Monitoring (PAM) project has developed a system that performs behavioral monitoring in an unobtrusive manner and can detect changes in a person's behavior. The system uses a variety of discreet sensors to gather data on the parson's behavior and this data is processed to extract behavioral patterns and detect changes in those patterns.**

In this paper present one method of data processing that takes 24hr long data-streams from the sensors, pre-processes them and uses the Continuous Profile Model to align and extract the underlying patterns from the data-streams. We present some preliminary results from a technical trial.

I. Introduction

Bipolar disorder (BD) is a mental disorder, characterized by recurring episodes of mania and depression. It is estimated to effect between 0.4-1.6% of the population [1].

There are several different treatments for BD, from both a pharmacological as well as a therapeutic angle. The first choice with pharmacological treatment is lithium, which acts as a mood stabilizer and reduces the severity of the affective episodes. However, lithium is not effective in around 20-40% of patients, and those who do respond to lithium often report unpleasant side-effects [2]. Other medications, such as anti-psychotics and anti-depressants are a common alternative to lithium [3].

Therapeutic approaches to BD take a number of routes, including social rhythm therapy [4], cognitive therapy [2] and family therapy [3]. One of the common threads through-out these treatment regimes is the identification and management of early warning signs, so that the patient can identify and deal with the earliest symptoms of an episode and in doing so lessen the effects of the episode. It has been shown that patients who are able to identify and act upon these early warning signs are better able to manage their mood and maintain a more stable health condition [5]. It should also be noted that a large number of people with BD self-monitor their condition, without following any particular therapeutic treatment.

The Personalized Ambient Monitoring (PAM) project is being developed to provide a self-monitoring tool for people with BD that will monitor their behavior patterns and provide alerts when these move outside of the normal patterns for that person, which could indicate the early symptoms of an affective episode. The system uses a number of discreet sensors, both in the home and in a wearable device, which gather data on the patient's behavior. This data is then analyzed to derive a *normal* activity signature for that patient. By comparing new incoming data to the activity signature, changes in the patient's behavior can be identified.

In this paper, we present the home based part of the system, describe the sensors used and the data processing for these sensors, and present some of the preliminary results obtained as part of a technical trial of the system.

Manuscript received May 13, 2010. This work was supported by EPSRC grant number EP/F005091/1.

J.D. Amor is with the Signal Processing and Control Group, ISVR, University of Southampton, Southampton, Hampshire, UK (e-mail: ja@ isvr.soton.ac.uk).

C.J. James is with the Institute of Digital Healthcare, WMG, University of Warwick, Coventry, Warwickshire, UK.

II. Method

The technical trial for the PAM system is being used as a test-bed for the hardware in the system and as a data-gathering exercise so that the data analysis algorithms can be written and tested. The trial was carried out in the homes of three of the researchers on the project. Each researcher was given a wearable unit (the results of this have yet to be analyzed and published) and each home was set-up with a set of sensors designed to monitor behavior in the home environment. The sensors were linked to a base-station PC that provided data logging and pre-processing capabilities as well as forwarding data to the data-storage server. This paper discusses the trial carried out, and the data set obtained, for participant #3.

A. Sensor Set

The sensor set for the PAM system is designed to capture as many facets of day-to-day behavior as possible. To this end, the sensors used are designed to detect the presence of

TABLE 1
SENSOR LOCATION AND RATIONALE

Sensor	Location	Rationale
PIR	Sitting room	Identify activity in rooms.
Light	Sitting room	Capture the presence of artificial light and the presence or absence of natural light during the day.
Microphone	Sitting room	Capture noise levels.
Camera	Kitchen	Monitor activity in various areas of interest in the kitchen.
TV IR sensor	Sitting room	Capture TV remote control usage.

the subject and their level of activity, such as being "in the sitting room", rather than capture very specific behaviors. The sensors have been placed in those locations that the PAM team felt were central to the daily routine of the participant, i.e. the kitchen, sitting room and bedroom. The list of sensors used, their location and the rationale for using them is shown in Table 1.

The Light, Microphone and TVIR sensors are all integrated into an environmental node, described in [6], which is placed in the sitting room. The light sensor provides two data-streams, one for the natural light and one for the artificial light in the room. The microphone also provides two data-streams, one for signal power and one for zero-crossing rate. The camera, an Edimax IC-1520DPg, targets several areas of interest (AOIs) in the kitchen and the data provided by each one is treated as a separate data-stream. The AOIs targeted by the camera are the cooker, the sink, the sideboard next to the cooker, the sideboard next to the tea cupboard, the tea cupboard, the washing machine and the dishwasher. The camera is further described in [7]. The PIR sensors used are X10 based, and intended for home automation systems, but are suitable for use in the PAM system.

B. Data Pre-Processing

In order for the sensor data to be used to identify behavioral patterns and changes in those patterns they first need to be pre-processed. We have 13 data-streams where each data-stream is a 90-day long time series. We break each data-stream down to a set of 24hr-long time-series, where each is split into five minute segments and for each segment and a sensor-specific score is given. For the PIR sensors, this is the number of sensor activations in each segment. The microphone data-streams are both averaged over each segment, as are the two data-streams from the light sensor. For the camera data, the images are processed as described in [7] to extract activity in each AOI in the kitchen. For each AOI, a separate time series is generated and the score is the number of times activity is registered in the AOI in a given five minute segment.

C. Behavioral Pattern Detection

The detection of behavioral patterns in the data is achieved using the Continuous Profile Model (CPM) described in [8]. CPM is a model for the alignment of multiple noisy, time-shifted time series and the generation of a latent trace which represents the underlying pattern in the data. The input set is a set of time-series which are assumed to come from a stochastic, noisy process in which time may be warped in a non-linear way for each time series. For the PAM system, we have 13 different input sets, one for each data-stream. CPM aligns the input set and allows for the variable time shifting. The latent trace that is generated is a representation of the underlying process that generated the input set and serves as a generalized estimate of the shape of the time-series in the input set.

In order to derive the latent trace, it is modeled with a Hidden Markov Model (HMM) which is optimized so that the likelihood of generating the input set from the HMM is maximized. The parameter ζ is used in the HMM to model the noise in the input set and acts as a measure of the similarity of the time-series in the input set. Detailed implementation details can be found in [8].

Once the HMM has been optimized it can be used to assess the likelihood of generating an unseen time-series, which is equivalent to comparing the similarity of an unseen 24hr time-series to the latent trace. If the likelihood of generating the unseen data is high then it is similar to the latent trace; if the likelihood is low then it is dissimilar to the latent trace.

We use CPM in the PAM system first of all to learn the behavioral patterns for each data-stream and second to test for changes in those patterns.

In the learning phase, the first thirty days of data are taken for each data-stream, where s indicates the data-stream, and CPM is used to find the latent trace, the noise, ζ_s, and the average likelihood of generating a 24hr time-series from the input once the HMM has been optimized, which is denoted Lp_s. This likelihood is generated from the total likelihood of generating the input-set divided by the number of time-series in the input set. The average likelihood of generating a time-series from the input set serves as a point from which the testing phase likelihoods can be measured for that particular data-stream.

In the testing phase, data is compared on a day-by-day basis to the established behavioral patterns. For a given day, for each data-stream, the likelihood of generating each 24hr-long time-series is computed denoted Lt_s. The behavioral difference score D_s is calculated as $D_s = Lp_s - Lt_s$ and if found to be positive, is then set to 0. This operation is designed to test the new day's data against the established pattern. If the likelihood of generating the new day's data is higher than the average likelihood for the test set it is 'more typical' of the behavioral pattern and deemed to not be different. Conversely if the likelihood of generating the new day's data is less than the average likelihood for the test set, the new day is deemed to be different and the behavioral difference score reflects this logic. The difference score is further weighted with a multiplication by $1/\zeta_s$ to give the weighted behavioral difference (WBD). This is designed to emphasize those sensors that have a consistent pattern and de-emphasize those that have an inconsistent pattern when the average is calculated. The average WBD score is calculated as the average of all the weighted behavioral scores for the new day's data. This can be expressed as

$$W_{Avg} = \frac{1}{N}\sum_{s=1}^{N}\frac{1}{\zeta_s}\left(Lp_s - Lt_s\right).$$

The average WBD score is one way to obtain an indication of the difference between the participant's normal behavior patterns (defined as the first 30 days of the trial) and the behavior recorded on subsequent days. We can make use of this by defining a threshold for W_{Avg} so that any score of W_{Avg} below this threshold indicates that the day in question had a significantly different behavior to the normal patterns.

III. Results

Figure 1 shows the result of applying the CPM model to the data from the cooker AOI from the camera. It can be seen in the figure that the unaligned data shows little in the way of a discernible pattern; there is clearly some activity from mid-morning until mid afternoon (1000-1600hrs) and then again in the evening through into the early morning (1800-

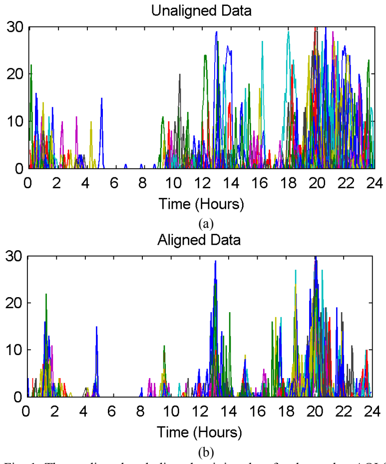


Fig. 1. The unaligned and aligned training data for the cooker AOI from the camera sensor, showing the data before alignment with CPM (a), and after alignment (b).

0200hrs). The aligned data makes these patterns significantly easier to see and the same patterns are picked up in the latent trace, shown in Figure 2. The latent trace clearly identifies a peak at 2000hrs and a lesser one at 1300hrs and these indicate a clear pattern of activity near the cooker at those times.

Figure 3 shows the WBD scores across all the environmental sensors used in the technical trial for participant #3. Gaps in the traces for some of the sensors are due to communication problems between the sensor and the base-station PC, which resulted in a loss of data on some days. The traces produced are generally very messy, but a clear dip across several sensors can be identified on days 24 and 53. This indicates that the person's behavior was very significantly different on those days.

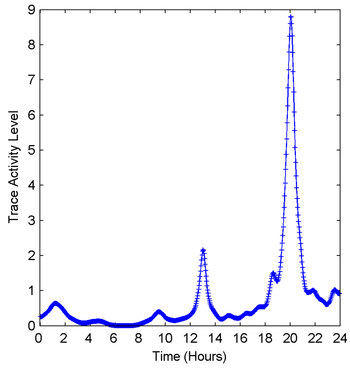


Fig. 2. The latent trace for the cooker AOI, showing the behavioral pattern derived with CPM.

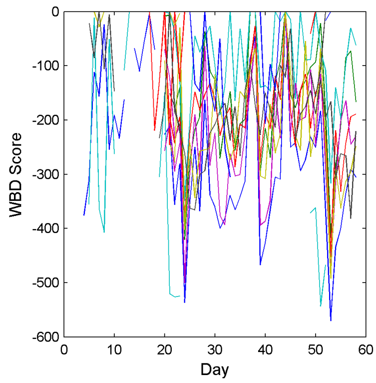


Fig 3. The behavioral difference scores across all of the environmental sensors. Gaps in the graph indicate that no data is available for those days.

Figure 4 shows the average WBD score across all of the environmental sensors with an arbitrary threshold at -300. The figure shows how the participant's behavior varied through the course of the technical trial and we see that days 3, 24 and 54 fall below the threshold, indicating a significant change in behavior on these days. These correspond to the drop in behavioral difference score on the same days shown in Figure 3 and indicate that the participant's behavior was significantly different on those days.

This is borne out by an examination of the time-series for the sensors that show a drop in difference score; we see a

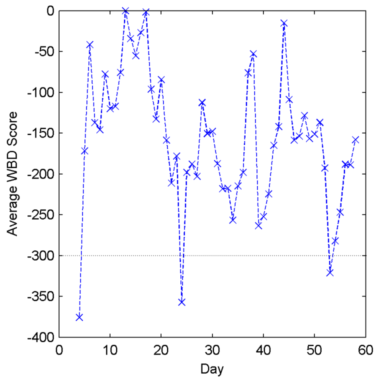


Fig. 4. Average weighted behavioral difference score across the set of environmental sensors, with a threshold level at -300, showing how the behavior patterns of participant #3 differed from the normal rhythms across the test days in the technical trial.

distinctly different behavior than the established behavior pattern on the days in question. The cooker data-stream from the camera for example shows a very large activity spike between 1200hrs and 1400hrs, which is out of character for that data-stream.

IV. Discussion

The results presented here show that by taking a large volume of data about the participant and applying suitable processing algorithms a single metric can be obtained that gives an indication of the change in the participants behavior over time. When this metric falls below the threshold, on day 24 for example, it is possible to drill down through the layers of data processing to identify the sensors that show these behavioral differences and to identify the exact aspects of the participants behavior that have differed.

Following this method, it would be possible for a clinician, or other concerned party, to track a patient's behavior with the average WBD score and only have to access the detailed behavioral information when necessary.

Despite showing some good results, there are some aspects of the system that need to be addressed. Firstly, because the system is very prototypical, there are sections in the data-streams where data is missing. In the data analysis above we attempt to resolve this issue by extending the training period for each data-stream until 30 days of good data have been collected. This has the knock on effect of delaying the start of the testing phase and is the reason that some of the sensor readings in Figure 3 do not start at day 1. In the testing phase, data-streams that have missing data are ignored and the average WBD score ignores these data-streams when the average is calculated. By dealing with the missing data in this way the disruptive effects of the missing data can be minimized.

A second consideration is that in the data analysis shown here we have made the assumption that all days will follow the same behavioral pattern. In reality we expect that this is not the case and that better results could be obtained by looking at week-day vs. weekend patterns, and patterns for each day of the week.

A final consideration is that participant #3 was only one resident in a house with 2 other residents. The data shown here is therefore representative of the entire house and not solely of participant #3. We expect that once data from the wearable sensor is included in the data analysis it can be used to cross-reference with the environmental data to provide a much more personalized system. Nevertheless, the data processing methodologies shown here have picked up behavioral changes and remain a suitable approach.

V. Conclusions

The technical trial of the PAM system has provided us with a large volume of data on which to develop our processing architecture. The method presented in this paper shows that it is possible to first identify and secondly track the changes in a person's behavioral patterns using the average WBD score. We have shown that this system is capable of identifying behavioral changes and that the sensor data on the days identified as changed does show a deviance from the normal pattern of behavior from those sensors.

The next stage for the project is to add the data from the wearable device into the data analysis. This will provide us with data that is unambiguously related to the participant and we expect this addition to enhance the ability of the system to correctly identify behavioral changes.

References

[1] American Psychiatric Association. *DSM-IV-TR - Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Publishing, Inc., Washington, DC, 4th - text revision edition, 2000.

[2] D. H. Lam, S. H. Jones, P. Hayward, and J. A. Bright. *Cognitive Therapy for Bipolar Disorder: A Therapist's Guide to Concepts, Methods and Practice*. Wiley, Chichester, 2000.

[3] S. L. Johnson and R. L. Leahy, editors. *Psychological Treatment of Bipolar Disorder*. Guilford Press, New York, 2004.

[4] E. Frank, H. A. Swartz, and D. J. Kupfer. Interpersonal and social rhythm therapy: Managing the chaos of bipolar disorder. *Biol. Psych.*, 48(6):593–604, September 2000.

[5] S. J. Russell and J. L. Browne. Staying well with bipolar disorder, *Aust N Z J Psychiatry*, 39(3):187–193, March 2005.

[6] P. Prociow and J. Crowe, Sensors Enhancing Self-Monitoring for People with Bipolar Disorder, in *Proceedings of the 5th UKRI PG Conference in Biomedical Engineering and Medical Physics*, 2009.

[7] J. D. Amor and C. J. James, Personalised Ambient Monitoring: Wireless Cameras for Activity Detection, in *Proceedings of the 5th UKRI PG Conference in Biomedical Engineering and Medical Physics*, 2009.

[8] J. Listgarten, Analysis of sibling time series data: alignment and difference detection, Ph.D. Thesis, Department of Computer Science, University of Toronto, 2006.

Appendix B

Supplemental Camera Processing Images

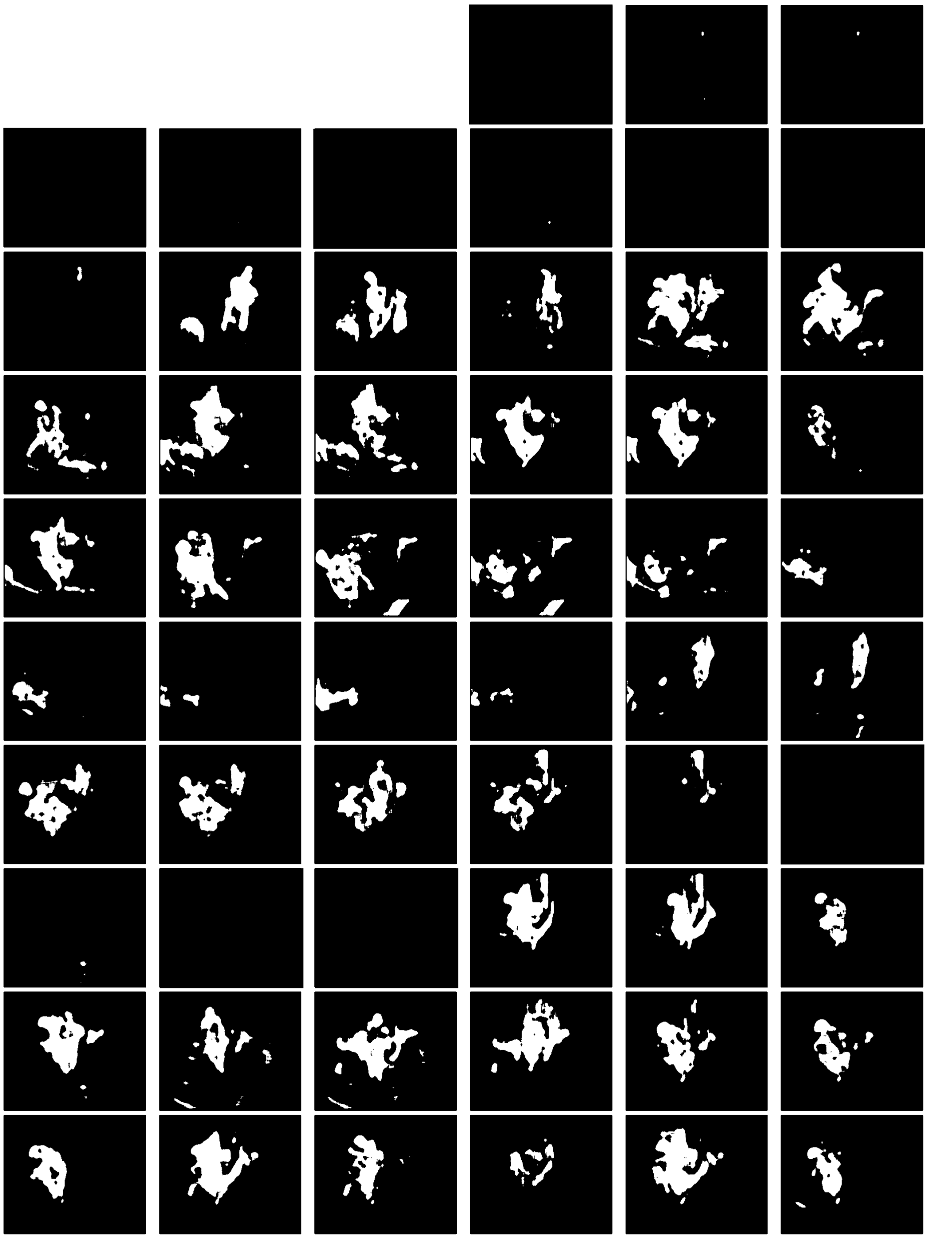


FIGURE B.1: The bedroom image sequence processed with the IH algorithm. There is a lot of lighting induced noise and object fade present throughout this image sequence.

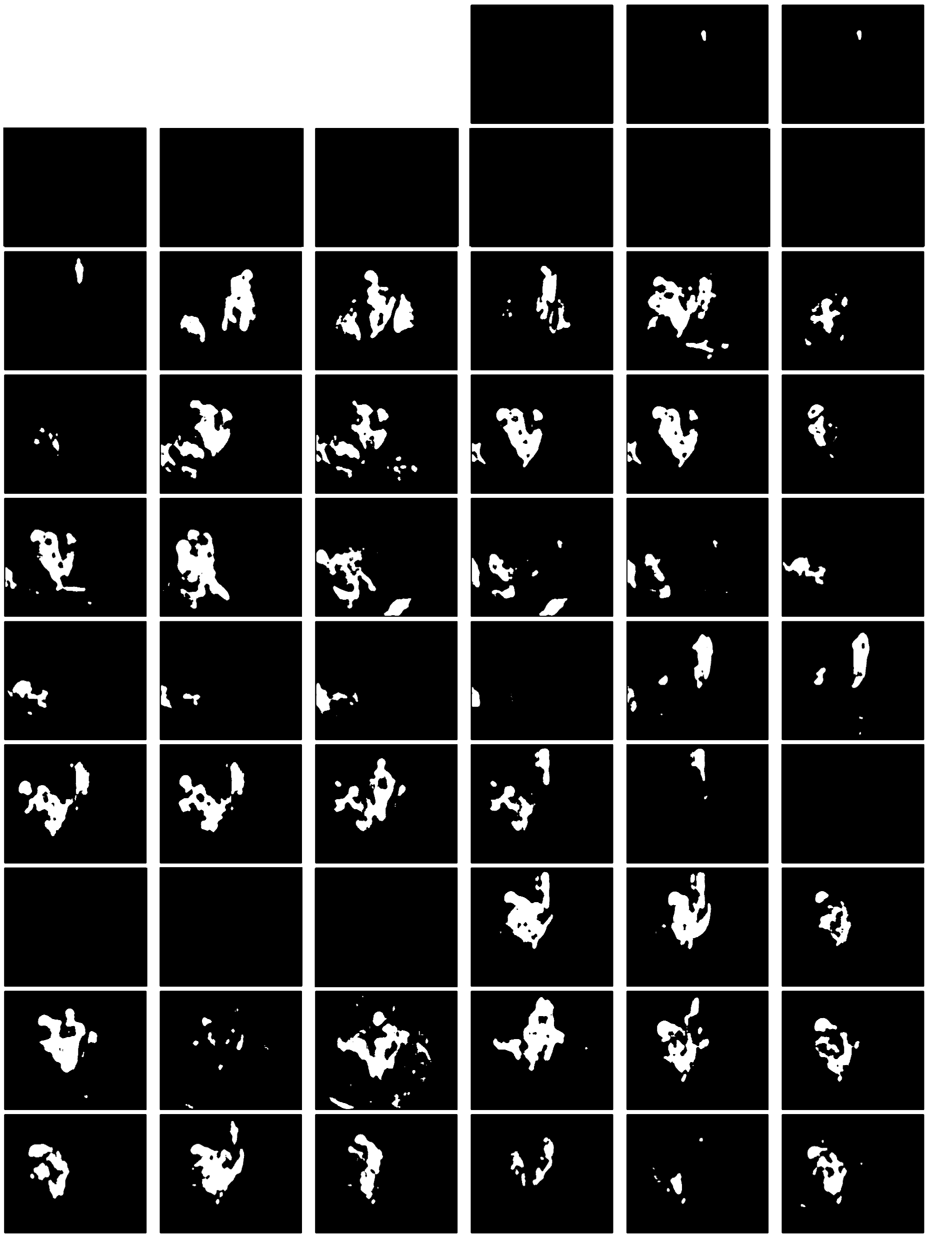


FIGURE B.2: The bedroom image sequence processed with the WCD algorithm. Object fade is present throughout this image sequence and there is minimal lighting induced noise.

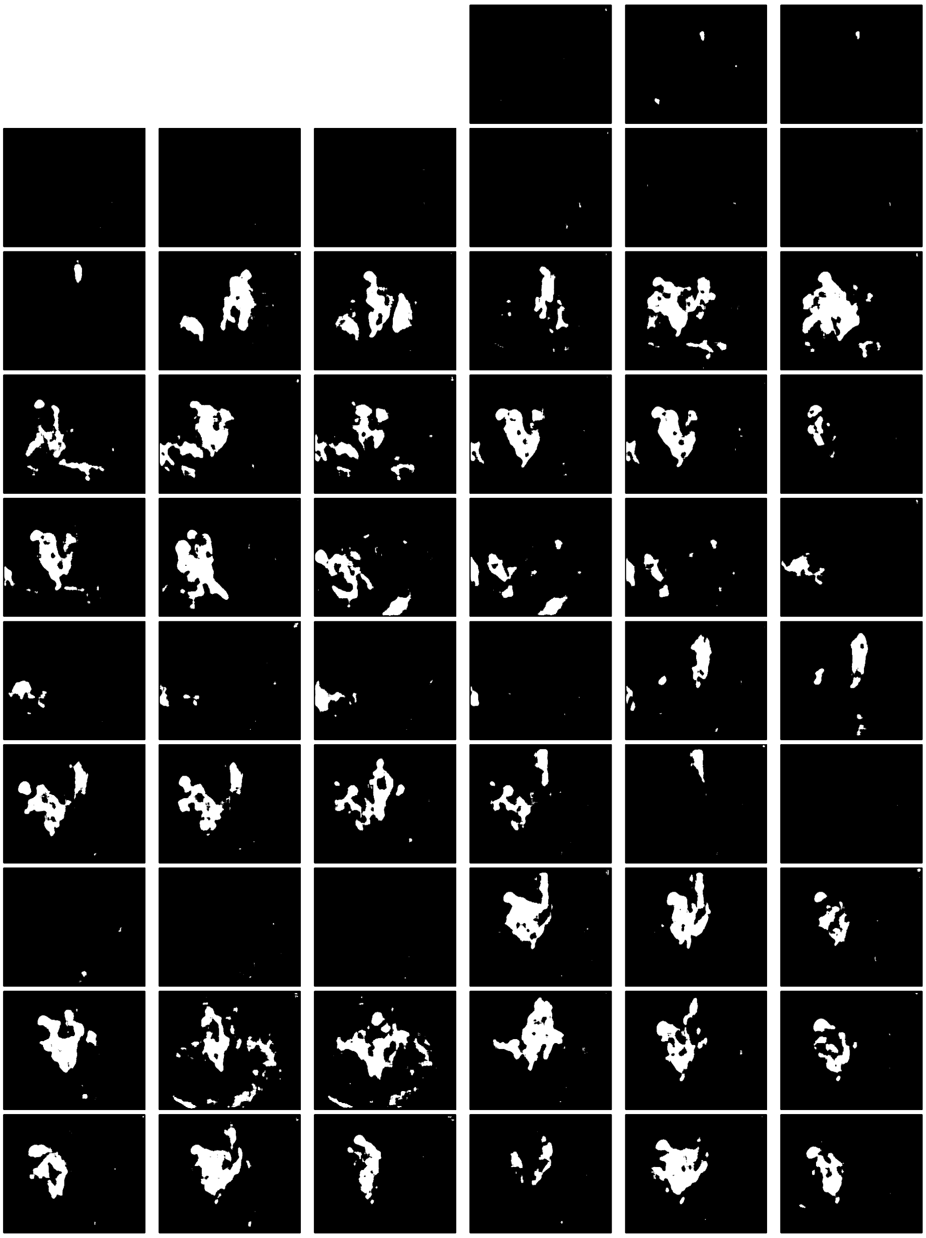


Figure B.3: The bedroom image sequence processed with the HM algorithm. Object fade is present throughout this image sequence and there is minimal lighting induced noise.

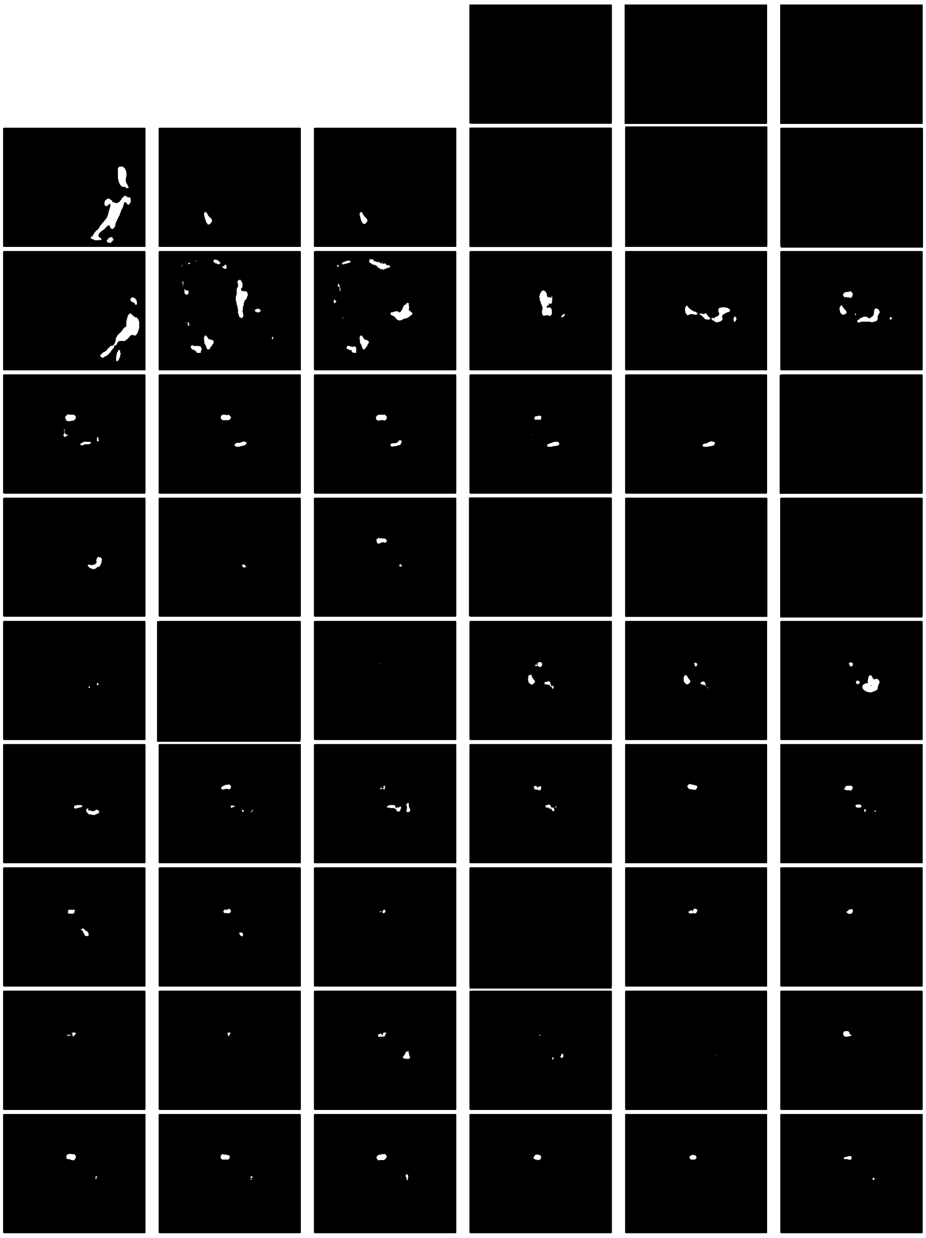


FIGURE B.4: The sitting room image sequence processed with the IH algorithm showing good change detection results.

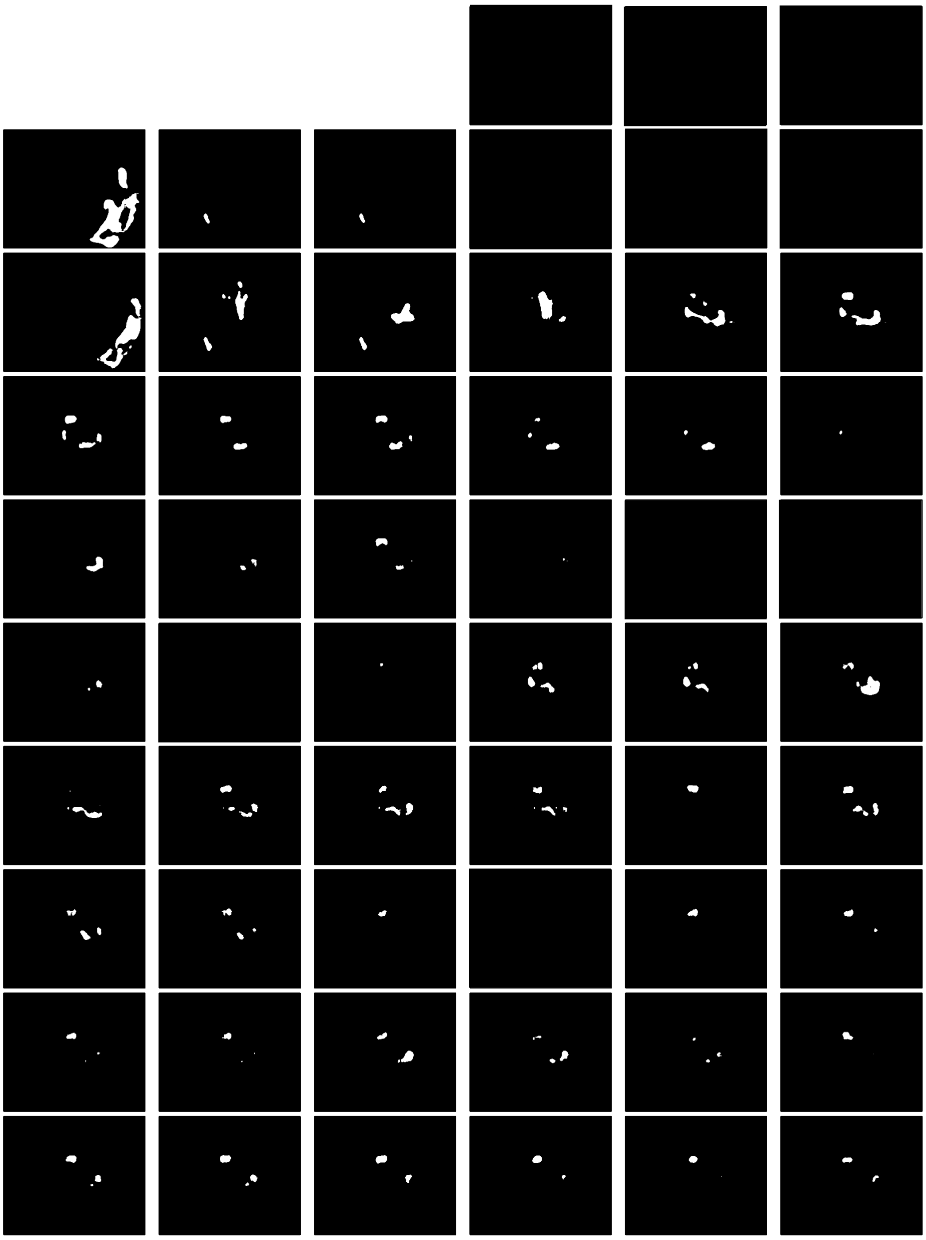


FIGURE B.5: The sitting room image sequence processed with the WCD algorithm showing good change detection results.

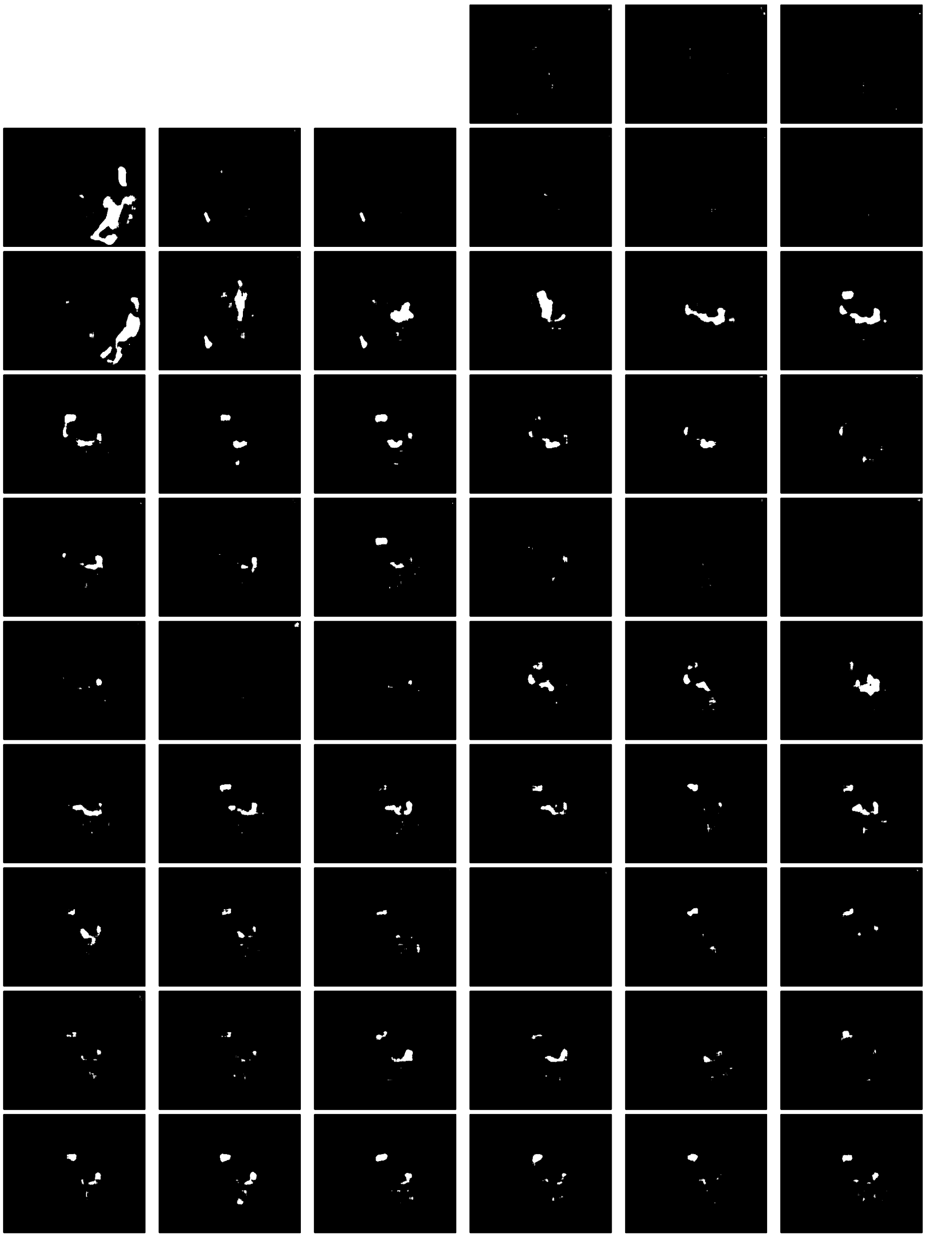


FIGURE B.6: The sitting room image sequence processed with the HM algorithm showing good change detection results.

Appendix C

Technical Trial Interview Form

Document Name			
Document Title	PAM Technical Trial Entrance Interview		
Author(s)	Jesse Blum, Syed Mohiuddin		
Date	Aug, 2009	**Filename**	techTrialEntranceInterview_1_1.doc
URL			
Access	☑ PAM Internal		☐ General dissemination

Document History			
Version	**Date**	**Author**	**Comments**
1	March, 2009		Initial draft
1.1	Aug,2009		Updated with interview consent form & start of interview explanatory note

Introduction

The Personalised Ambient Monitoring project (PAM) is a three year project attempting to establish new technology to help people that suffer from Bipolar Disorder using their mobile phones and discrete sensors. The PAM project team is developing technology to allow sensors to automatically detect behaviour, define patient activity signatures, and combine this with mood diaries. A technical trial has been designed to test device and network considerations, as well as system acceptability and user compliance. The technical trial will involve collecting data at four different sites. The specific objectives of the trial are to find out:

- Is the system acceptable to the users?
- Are there user compliance issues?
- How reliable is the technology (black spots, communication faults)?
- Can behaviour patterns be captured unobtrusively?
- Can the data be used to inform a model?
- Do the system components integrate properly?

As part of the trial evaluation criteria (described in the PAM Technical Trial Evaluation Criteria document), the trial participants shall twice be interviewed using semi-structured interview format, once at before the start of the trial (the entrance interview), and once upon the completion of the trial (the exit interview). The semi-structured interview format is discussed in [1] and [2].

This document constitutes the interview guide for the PAM Technical Trial entrance interview. A second document is available for the interview guide for the PAM Technical Trial exit interview.

Entrance Interview Purpose

The purpose of the entrance interview is to discover prior to the trial the interviewees':

- Values, beliefs and emotions regarding system acceptability
- Thoughts about their trial behaviour and probable compliance
- Opinions regarding technological monitoring and other cultural factors that may impact the project

Understanding the participants in this way may help with the analysis of their behaviours during the trial and help to improve system acceptability and operation during the full trial. Therefore it is important for the interview to capture detailed points of view, reference framework of meanings and insights in to the PAM project, its technology and monitoring in general.

Interviewer Conduct

A qualitative semi-structured format has been chosen to accomplish the goals of the entrance interview. Semi-structured interviewing provides the interviewer with flexibility regarding the order that questions are asked, the wording used to ask the questions, and opportunity to introduce new questions whilst exploring what the interviewees are saying in more detail. A guide containing a list of open ended topics to be covered in the interview is provided in the next section.

Successful interviewer conduct will be engaging and interactive, directive at times (but without imposing interviewer assumptions on the interviewees), but mostly will be based on attentive listening, sensitive and responsive to what the interviewee is saying and doing. Avoid embarrassing questions, providing advice, and presenting your own views.

In preparation for the interview:

- Set up a mutually convenient time and place for conducting the interview prior to the technical trial start date
  - Choose a quiet place and limit distractions (turn off phones; hide from students, children, etc.)
  - Make sure that both the interviewer and the interviewee allocate enough time to the interview (at least one hour)
- Prepare to record the interview
  - Seek permission from the interviewee to use an audio recorder
    - Inform the interviewee that the recording and transcription will only be made available to the PAM group
    - Also seek permission to publish anonymised interview question responses
    - Inform the interviewee that they may have access to the recording and transcription of their interview if they so request
  - If permission to record is granted then acquire the recording device and enough storage for at least an hour's conversation
  - If permission to record is not granted then you will need to record the interview as notes

At the start of the interview:

- Explain the purposes of the interview
- Ask the participant to complete the interview consent form. Only proceed with the interview after the consent form has been signed.

Upon interview completion:

- Reflect on the interview and take note of any additional insights you may have regarding the process
  - How did the interview go?
  - Was the setting appropriate?
  - How did you do as an interviewer?
  - Are there additional insights that you have regarding the PAM project that weren't recorded?
- Transcribe your recordings and/or notes
- Store your notes, recordings and transcriptions securely so as not to compromise interviewee privacy
- Only share your notes, recordings and transcriptions of the interview with the PAM project team and no one else except the interviewee if they so request

Topic Guide

The following is a list of topics to help guide the interview. The topics are meant to guide the interview; please phrase and order the questions as appropriate.

- Introduction -- Ask the interviewee to describe who they are and their relation to the PAM project
- Past experiences with monitoring technology and being monitored
- Feelings related to being observed by the system
- Data access and privacy concerns
- Sensor installation issues
- Wearing/carrying sensor issues
- Mobile phone issues
- Tell other people about this participation?
    - What reactions are expected?
- How may constant monitoring change behaviour?
- PAM team service expectations
- Objections to particular devices
- Reactions and concerns regarding placement, form-factor and usage of different devices
    - How much time do you consider it appropriate to spend (charging batteries, answering questionnaires, etc)
- Assumptions regarding how life might change using the system or particular devices
- Thoughts and feelings about external monitoring
- Anything else that might affect participation

Interview Consent Form

PAM Technical Trial Entrance Interview - Personalised Ambient Monitoring Project

The purpose and nature of the PAM Technical Trial entrance interview has been explained to me. I agree to be interviewed for the purposes of the PAM project.

I understand that:

1. Taking part in this interview is entirely voluntary.
2. It is my right to decline to answer any question that I am asked during this interview.
3. I am free to end the interview at any time.
4. My name and identity will remain confidential in any publications or discussions about the interview.

I agree to:

1. The recording of the interview.
2. Recordings and transcripts of the interview being stored electronically for the duration of the PAM project and that these can only be accessed by the members of the PAM project team and me.
3. I agree that my interview responses can be used in publications and discussions related to the PAM project and the research conducted for the purposes of the PAM project research assistants attaining their education qualifications.

I HAVE READ THIS CONSENT FORM. I HAVE HAD A CHANCE TO ASK QUESTIONS CONCERNING ANY AREAS THAT I DID NOT UNDERSTAND.

____
(Signature of Interviewee)

____
(Printed name of Interviewee)

____
(Date)

Bibliography

[1] Britten, N. *Qualitative research: qualitative interviews in medical research*, British Medical Journal, vol. 311, num. 6999, pp. 251-253, 1995, http://www.bmj.com/cgi/content/full/311/6999/251, last accessed 06 Mar, 2009.

[2] Bryman, A. *Social research methods: Chapter 15 Interviewing in qualitative research*, Oxford University Press New York, 2004, http://fds.oup.com/www.oup.co.uk/pdf/0-19-874204-5chap15.pdf, last accessed 06 Mar, 2009.

Appendix D

Participant #2: Behavioural Log

D.1 Introduction

The following behavioural log is reproduced as received from Participant #2 and describes their behaviour over the course of the technical trial.

D.2 A month in the life of Participant #2.

My behaviour over the course of the technical trial, to the best of my recollection was as follows. I'm a single dad with two kids that stay with me from Friday afternoon to Monday morning. The remaining days I was on my own except when my fiance visited me, which was from the 16th to the 20th. Also, I think a friend crashed at mine on either 8th, 9th or 10th. On non-weekend days I tended to spend a half hour or so in the kitchen in the mornings and probably a similar amount of time in the evenings. I usually got between 6.5 to 8 hours of sleep per night.

Generally my schedule could be broken down into 4 sections: Mondays, Fridays, Tues-Thurs, and weekends.

Monday schedule: Woke around 7am to get the kids up and ready for school. After dropping them off I'd head up to uni where I tended to work/study until I'd return home in the evenings (ranging from 18:30 - 23:00ish). I tended to be in bed before midnight.

Tues-Thurs schedule: I would tend to wake up a little later than Mondays (closer to 8am) and commute to uni; usually to be in for around 9am but a few days I'd work form home in the mornings until about 10am. On Tuesdays (except for 29th) I left uni at 14:00 to take my son to an appointment and would return home by around 16:30 and would work from home for a couple of hours. Weds and Thurs, I tended to stay at uni until some time between 18:30–21:00 (except when my fiance was here).

Fridays schedule: I tended to be awake and at uni for 9am. I left uni around 2pm to collect my kids from school and the three of us would be at home by 4pm. We'd head off again to do the weekly shop around 5pm and be home by 6pm. We'd have our evening meal around 6:30pm. My daughter went to bed around 19:30 and my son around 21:00. I tended to stay up until around midnight.

Weekends schedule: We'd be up and out of the house on Saturdays by 8:30am to take my daughter to a class. My son and I would return briefly in the morning and then set off again by 11:30. The three of us would return by 12:30 and have lunch. Afterwards we'd sometimes stay in and sometimes go out some more. We tended to eat our evening meal between 5:30pm and 6:15pm. My daughter went to bed around 19:30 and my son around 21:00. I tended to stay up until around midnight. On Sundays the children tended to be awake by 7 or 7:30, but I'd lie in until up to 9ish. We'd have breakfast a little while later and be in and out of the house until our evening meal and routine, which tended to match Saturdays.