A Machine Learning Framework for Automatic Human Activity Classification from Wearable Sensors

Edmond J. Mitchell, B.Eng.

School of Electronic Engineering

Dublin City University

Doctor of Philosophy(Ph.D)

Supervisor

Prof. Noel. E. O'Connor

December 2014

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of PhD is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed:

Edmond Mitchell

ID No.: 59108002

Date: 10/12/2014

Contents

List of Figures

Abstract

A Machine Learning Framework for Automatic Human Activity Classification from Wearable Sensors

Edmond Mitchell

Wearable sensors are becoming increasingly common and they permit the capture of physiological data during exercise, recuperation and everyday activities. This work investigated and advanced the current state-of-the-art in machine learning technology for the automatic classification of captured physiological data from wearable sensors. The overall goal of the work presented here is to research and investigate every aspect of the technology and methods involved in this field and to create a framework of technology that can be utilised on low-cost platforms across a wide range of activities. Both rudimentary and advanced techniques were compared, including those that allowed for both real-time processing on an android platform and highly accurate post-processing on a desktop computer. State-of-the-art feature extraction methods such as Fourier and Wavelet analysis were also researched to ascertain how well they could extract discriminative physiological information. Various classifiers were investigated in terms of their ability to work with different feature extraction methods. Consequently, complex classification fusion models were created to increase the overall accuracy of the activity recognition process. Genetic algorithms were also employed to optimise classifier parameter selection in the multidimensional search space. Large annotated sporting activity datasets were created for a range of sports that allowed different classification models to be compared. This allowed for a machine learning framework to be constructed that could potentially create accurate models when applied to any unknown dataset. This framework was also successfully applied to medical and everyday-activity datasets confirming that the approach could be deployed in different application settings.

Acknowledgements

I would like to take this time to thank Prof. Noel E. O'Connor for providing me with opportunity to conduct research in a field that I have always been interested in. I would also like to thank him and Dr. David Monaghan for their constant guidance and advice over the past couple of years. None of this work would of occurred without the funding provided by Clarity and Science Foundation Ireland so I am immensely grateful towards those bodies. I would also like to thank all my colleagues who I have collaborated with and thank everyone else whom I have interacted with in DCU who made me feel welcome here. Most importantly I wish to thank my parents for their constant support and encouragement I have received throughout my whole life.

Publications

List of Publications Journal

- E Mitchell, D Monaghan and N.E. O'Connor. Classification of Sporting Activities Using Smartphone Accelerometers. Sensors. 2013; 13(4):5317-5337.
- A Ahmadi, E Mitchell, C Richter, F Destelle, M Gowing , N.E. O'Connor and K Moran. Towards Automatic Activity Classification and Movement Assessment During a Sports Training Session. IEEE Internet of Things Journal.

Peer reviewed international conference/workshop

- A Ahmadi, E Mitchell, C Richter, F Destelle, M Gowing , N.E. O'Connor and K Moran. Automatic activity classification and movement assessment during a sports training session using wearable inertial sensors. In Body Sensor Networks, 16-19 June 2014, Zurich, Switzerland.
- K Sweeney, E Mitchell, J Gaughran, T Kane, S Coyle and N.E. O'Connor and D Diamond. (2013) Identification of sleep apnea events using discrete wavelet transform of respiration, ECG and accelerometer signals. In: Body Sensor Networks 2013, 5-10 May, Boston, MA.
- E Mitchell, S Coyle, NE O'Connor, D Diamond, T Ward. Breathing feedback system with wearable textile sensors. Body Sensor Networks (BSN), 2010 International Conference on. ISBN: 978-1-4244-5817-2

- E Mitchell, S Coyle, A Doherty, A Smeaton, N.E. O'Connor and D Diamond (2010) Smart tablecloths - ambient feedback of domestic electricity consumption. In: iHCI 2010 - 4th Irish Human Computer Interaction Conference, 2-3 September 2010, Dublin, Ireland. ISBN 1872327885

- S Coyle, E Mitchell, T Ward, G May, N.E. O'Connor and D Diamond (2010) Textile sensors for personalized feedback. In: IAPMA2010 - ECIR2010 workshop on information access for personal media archives, 28 March 2010, Milton Keynes, UK.

- S Coyle, E Mitchell, T Ward, G May, N.E. O'Connor and D Diamond (2009) Wearable sensors and feedback system to improve breathing technique. In: UNCSR 1st Annual Symposium, 22 October, Dublin, Ireland.

1 Introduction

Up until the beginning of the millennium computers had been predominantly associated with the traditional desktop personal computer with the mouse and keyboard being the main methods of input. Today computers are becoming more pervasive and are embedded within our smartphones, personal music players, auto-mobiles, clothing, watches and even glasses. In fact it is difficult to imagine an object that will not contain a computer at some stage in the future. The vision of ubiquitous computing is that, eventually, computers will disappear and become part of our environment, fading into the background of our everyday lives [5]. Therefore the ultimate goal in ubiquitous computing is that these unseen computers will autonomously gather contextual information from users and their bodies to enhance their everyday life. With computers today being embedded in smaller and smaller devices and being made available in more and more aspects of daily living, the goal is to find novel methods for humans to interact with this new era of computing technology

One avenue of research in this regard would be to curtail the amount of explicit interaction that users are forced to endure in order to communicate with computers. Implicit interaction implies that instead of the user having to directly communicate with the computer, the computer itself can determine the users requirements. In order for a computer to calculate the users requirements, it first must be able to sense the raw contextual information it can use to infer judgement within a particular situation. Sensors allow computers to detect and record changes in the users environment or physiology. These sensors are able to capture the current state of the user, their environment and their context, i.e their name, their location relative to the sensor, their current activity and the current state of their environment.

Current research in computers that are contextually aware focuses on the use of sensors either on the user or in their environment to capture data that can be used to infer the users context. This contextual information can be used by a computer to automatically detect the current requirements of the user, thus eliminating the requirement of the user to interact with the device. This, not only, increases productivity and efficiency but also allows users to react to different scenarios naturally without having to input information directly into a device.

As technology advances and more efficient manufacturing techniques are discovered the cost to invent and construct sensors is lowered. In turn, the quantity of data available from different sensor modalities increases. With this increase the potential to fuse data from different modalities allows researchers to infer new contextual information that could not be discovered from a single modality.

In this work a new class of context aware computing is explored. Automatic activity recognition aims to recognise the actions and events of a user utilising either physiological data captured from the body or data captured from the users environment. This research area has reached prominence in machine learning communities due to its ability to provide unobtrusive information to the user in areas such as medicine, human computer interaction and sport. The main research contribution outlined in this thesis focuses on using physiological data captured from sensors worn on the body to recognise a wide range of human activities.

Firstly, the capability of a system to recognise a single activity is introduced, which would allow the system to evaluate a users performance. This type of technology has many applications in health and in sport where patients need to perform rehabilitation exercises and athletes want to hone their techniques and skills. Secondly, a single sensor system is employed to automatically recognise a range of complex activities. This technology again has various applications in health and sport. Thirdly, multiple sensors are combined in order to provide higher levels of accuracy in specific applications that involve more complex activities, making them very difficult to distinguish using a single sensor. Then sensors of different modalities are fused together so that any weakness in contextually recognizing human actions or activity in one sensor can be remedied from data from another sensor of a different modality. Finally a search-space parameter optimisation algorithm is introduced that allows the whole classification process to be sped up as well as increasing the overall classification accuracy.

Wearable sensors are used along with advanced feature extraction techniques and machine learning methods to capture, train and test classification models in order to automatically recognize the user's activity. One of the main advantages of using wearable sensors, as oppose to sensors built into the infrastructure of a stadium is the ability to observe the world from a 1st person perspective, continuously, and lacking the requirement of any outside infrastructure. In this thesis activity recognition with wearable sensors will be shown to have the capacity to create new applications and to enhance current ones used in the areas of sport and health.

The main goal of this work is to create a machine learning framework for automatic human activity classification. This framework will allow the creation of accurate classification models for any annotated dataset. These models will be created from a variety of advanced feature selection parameters as well as a multitude of classifiers. This framework will also contain state of the art optimisation procedures to ensure efficient parameter selection.

1.1 Research Objectives

There are several research objectives associated with creating a machine learning framework for automatic human activity classification. The initial objective is to outline the classification process from the sensing of physiological information to the output of a recognised activity. To simplify this, the first step is to examine whether it is possible to identify a single desired activity. The capability to evaluate a desired activity accurately permits the creation of a wide range of rehabilitation applications for people and also the creation of training applications for athletes. The chapter also presents the research contributes of this work and cites examples where the framework was applied successfully.

The next research objective is to investigate novel approaches to automatically identify various different activities with a single wearable sensor. This would allow users to have a permanent record of activities accomplished at a specific time or time intervals. To accomplish this objective, novel data mining and machine learning techniques will have to be inspected to ascertain their ability to perform this task. The category of sensor and the location where it will be placed on the user is paramount to the success of this goal. This is a challenging task as a single sensor can only capture physiological from a signal modality. Every classification model requires features from the raw data to be inputted into it in order to yield accurate results. Therefore extracting the most discriminative features from the raw sensor data is an extremely important objective for this work.

Another objective is to investigate whether attaching more than one sensor of the same type to different locations on a user can yield greater classification accuracy results than a single sensor approach. With multiple sensors, the issue of fusion arises, of which there are two different approaches when fusing sensor data together for classification purposes, therefore both early fusion and late fusion need to be investigated.

One more research objective is to ascertain the plausibility of combining one or more sensors of a different modality and investigating whether any weakness in recognizing a human action or activity by one sensor can be remedied from data from another sensor.

A machine learning framework for automatic human activity classification will have many parameters that need to be selected before a model can be generated. Researchers must generally limit the parameters investigated otherwise the process will take an unreasonable amount of time. This limitation in parameters can lead to the situation where potential models that provide higher accuracy are not investigated and therefore this is highly undesirable. Parameter selection optimisation procedures will examined to see whether they can negate this outcome.

1.2 Research Contributions

A wearable sensor network is a network of intelligent physiological sensors can be integrated into a wearable computing network. Wearable sensor networks have an important role to play in future healthcare delivery and management by sensing the body and interpreting physiological data. To this end an interactive system was developed that helped patients perform respiratory exercises, and maintained the interest of children during these exercise sessions. This wearable system classifies breathing technique into separate grades and provides visual feedback to the user through a graphical user interface. The exercise sessions can be repeated using the same reference signal, which means that medical staff need not be present for the exercise, thereby improving efficiency in the hospital.

Utilizing the accelerometer in the commonly available smartphone to recognise human activities has added various novel contributions to this thesis. The accelerometers in smartphones are less accurate than more expensive professional units, which adds to the difficulties in processing this data. Smartphones have been used to identify common day to day activities in the literature however to the best of our knowledge it is the first time they have been used accurately in a sporting context. This technology is easily available to athletes and allows them to track their sporting performance over the course of a training season. Using a smartphone placed on athletes back, this thesis introduces a set of algorithms to detect key movement and sporting activities including walking, standing still, jogging, sprinting, tackling, hitting a ball and soloing with a ball. These algorithms were shown to create accurate models for two different field sports.

An analysis into the use of cheap and easy to use respiration, ECG and accelerometer sensors for the classification of sleep apnea events was concluded. Standard polysomnography tests can be prohibitively expensive therefore the ability to pre-test for sleep apnea using a cheap and accurate system could allow the tests to be preliminarily carried out outside of the clinic environment. Therefore a cheap reusable wearable system consisting solely of a t-shirt was created that allowed different apnea events to be recognised. Additionally a comparison of early and late fusion approaches for multimodal sensor fusion was carried out. It was concluded that the early fusion approach outperformed the late fusion approach. There are very few instances of this test in literature so it is beneficial to have conducted it.

Finally having a large set of possible parameters for any problem presents a significant time investment problem. The more permutations that must be investigated increases the length of time to conduct any experiment. To overcome this problem, a final contribution is presented, in which a novel classification parameter selection procedure is presented which allows users to search through a list of permutations for the most accurate solution with a significant increase in speed. This genetic algorithm approach was shown to decrease the time to locate the best classification parameters with an 87.5% decrease in time required on average. Additionally the novel approach presented shows how the GA can permit a much larger parameter search space to be investigated which was shown to identify model parameters with higher accuracy than the model found within a smaller space. This in turn allows more accurate models to be created than a standard brute force approach.

1.3 Thesis Outline

Chapter 2 explores the technical background of preprocessing the raw sensor data before looking at the literature for advanced feature extraction. Then physiological sensors and their relevant applications are described in detail for classification purposes followed by a look at the different data fusion techniques which are investigated on their applicability for multimodal sensor fusion. After this chapter there is an exploration at the state of the art machine learning techniques, which are used for to create classification models in this work. Finally state of the art parameter selection optimisation techniques are introduced, which were to hasten the classification model creation process. Relevant literature is presented as a basis for the framework design choices presented in this thesis.

Chapter 3 describes the challenges which were overcome in order to perform human activity recognition using a single sensor. This chapter outlines the feature extraction techniques required to initially evaluate a users activity performance before creating algorithms to identify various different sporting activities. A number of experiments are undertaken in order to ascertain the best approach to creating a classification model. A black box experiment is compared to a thorough investigation of all parameters. These two approaches are then compared to a final approach where each activity has its own specialised classifier. All methods proposed are compared to a literature benchmark for evaluation purposes. Different feature extraction techniques such as DWT, FFT and some simple time domain techniques were implemented for comparison purposes.

Chapter 4 explores the challenges encountered when creating a multimodal human action recognition system. Advantages from using two accelerometers versus a single accelerometer to identify different training activities performed by a subject are presented. As the results prove, sensor fusion can significantly improve the accuracy rate for classification models. Early fusion and late fusion are the two techniques used in this chapter to fuse data from different sensors. Experiments are conducted that use both early and late fusion to fuse the data from ECG,respiration and accelerometer sensors. Results prove that even though early fusion requires less computational time, it is similarly accurate at detecting human activities as a late fusion approach. After evaluating those two approaches, results obtained when using different permutations of three sensors of different modality are presented in this chapter. Results indicate that adding a sensor which captures physiological data that is already being accurately measured by a different modality can decrease classification accuracy.

Chapter 5 takes the results of all the model parameter permutations discovered in section 3.3 and uses them to test different genetic algorithms. This chapter explains why genetic algorithms can help optimise the parameter selection process. It also goes into detail on the role of each parameter that makes up the genetic algorithm and process behind it. Three experiments are conducted to investigate the genetic algorithms suitability to optimise the process of parameter selection. The first experiment explores the use of different population sizes and compares each GA to the brute force approach. Experiments were conducted 100 times each to give a fair representation of each populations ability. The second experiment increases the number of parameter permutations by a factor of ten but the number of possible solutions investigated was limited at the same amount as in section 3.3. This showed that the genetic algorithm could locate a new superior solution in the same amount of time it took the brute force algorithm to search through a search space one tenth of the size. The final experiment investigated whether this new superior models could of been extrapolated from original best performing models in section 3.3.

Chapter 6 succinctly presents the overall conclusions and briefly explores the research contributions and discusses directions for future work.

2 Technical Background

2.1 Introduction

This chapter presents an overview of the background literature required to understand the techniques employed in the core work of this thesis. Initially signal processing techniques that enhance the ability to discriminate between different signal types are presented. Then discrete wavelet transforms and fast Fourier transforms are introduced as they are required for signal feature extraction for automatic human activity classification. The current state of the art in inertial sensing is investigated which allows the fusing of sensor data from separate and different sensors. Automatic activity classification using signal processing and sensor fusion is greatly strengthened by utilizing the best machine learning methods available. Finally genetic algorithms are presented as a means to search large parameter spaces.

2.2 Digital Signal Filtering

2.2.1 Filtering

A filter is a component that is devised to change the spectral content of an inputted signal in a required manner. Filtering is commonly used to improve the quality of a signal and extract relevant information from a signal. Common filtering objectives include improving signal quality and de-interlacing previously combined signal components. In this work filtering is employed to improve the classification process by removing unwanted noise from sensor data.

Filtering is a process that allows a desired range of frequency components to pass in a signal called the passband while attenuating all other frequency components called the stopband. A digital filter is a mathematical algorithm implemented in software that manipulates an inputted digital signal to create a digital output signal that achieves a wanted criteria.

Filters can be classified in terms of their magnitude response. These are lowpass, highpass, bandpass, and stopband filters. A low-pass filter passes signal components under a desired frequency while attenuating signal components over this cutoff frequency. A high-pass filter passes signal components over a desired frequency while attenuating signal components lower than this cutoff frequency. A bandpass filter combines a low-pass and high-pass filter to only allow frequencies components within a desired range to pass. A stopband filter again combines a low-pass and high-pass filter to only attenuate frequencies components within a desired range. Filters can also be categorized in terms of their impulse response. The two possibilities are a finite impulse response (FIR) filter and a infinite impulse response (IIR) filter. The FIR filters impulse response is of finite duration because it settles to zero in finite time. For a linear and time-invariant FIR filter of order N, each value of the output sequence is a weighted sum of the most recent input values. This can be seen in equation 1 where $x[n]$ is the input signal, $y[n]$ is the output signal, N is the filter order and b_i is the value of the impulse response at the i'th instant

$$y[n] = \sum_{i=0}^{N} b_i \cdot x[n-i] \tag{1}$$

The advantages of a FIR filter over a IIR filter include the fact that it has no feedback. This causes no rounding errors to be compounded by repeated iterations. This also makes FIR filters inherently stable. $y[n]$ will always be a finite number since the largest value it can attain is

$\sum_{b_i} I$ where I is the largest input.

The IIR filters impulse response is of infinite duration because it does not become exactly zero past a certain point, but continues indefinitely. It can be seen in equation 2 where P is the feedforward filter order, b_i are the, Q is the feedback filter order feedforward filter coefficients, a_i are the feedback filter coefficients, where $x[n]$ is the input signal and $y[n]$ is the output signal.

$$y[n] = \frac{1}{a_0}(\sum_{i=0}^{P} b_i x[n-i] - \sum_{j=1}^{Q} a_j y[n-j]) \tag{2}$$

The main advantage digital IIR filters have over FIR filters is that they require less memory and cpu iterations to achieve a required filter response characteristic for any magnitude. IIR filters require a lower order Q than a FIR filter to achieve the same results. Also, certain desired responses are not practical to construct with FIR filters.

2.2.2 Butterworth Filter

A Butterworth filter is a popular IIR filter used in DSP. It is a maximally flat magnitude filter and was first invented in 1930 by the engineer and physicist Stephen Butterworth[6]. Its frequency response is maximally flat in the passband and slopes towards zero in the stopband. Butterworth filters have a monotonically changing magnitude function with ω whereas other popular filters such as Chebyshev or Elliptic filters that have non-monotonic ripple in the passband or the stopband. The Butterworth filter was employed in this work to reduce background noise such as when using captured ECG data

2.3 Signal Feature Extraction

Once data is captured, features must be extracted that will allow the identification of activities. Feature extraction is the process of simplifying the amount of information required to describe data accurately. Features with higher correlation between similar patterns (intra-class variation) and poorer correlation between dissimilar patterns (inter-class variation) are desirable. Analysis with a large number of variables can require a large amount of computer memory and computation power. More importantly a large input into a classification algorithm can cause overfitting in the training sample which produces models which respond poorly to new samples. Choosing the most discriminative features is key otherwise the model will not be able to distinguish between samples and the classification accuracy will be unsatisfactory. There are many well known general dimensionality reduction techniques used in signal feature extraction such as principal component analysis, Fourier analysis, semidefinite embedding and wavelet analysis.

2.3.1 Fourier transform

The Fourier transform is the cornerstone of discrete signal processing due to its ability to deal with linear time-invariant operators because the output does not depend on the particular time the input is applied. It is also able to deal with uniformly regular signals. While not as complex as the DWT is has been employed to extract features for activity classification problems with success. The discrete Fourier transform (DFT) transforms a definite array of samples of a function into a array of coefficients of a fixed length of complex sinusoids, arranged by their frequencies. Therefore it is able to convert information in the time domain to the frequency domain and vice versa. The fast Fourier transform (FFT) is an algorithm that calculates the discrete Fourier transform (DFT) and its inverse. The FFT is used instead of the DFT because it reduces the number of complex multiplications from N^2 to $N \log_2(N)$. For these reasons the FFT was investigated in this thesis.

Use of mean and energy of FFT components has been shown to result in accurate recognition of certain postures and activities [7][8][9]. Frequency-domain entropy is calculated as the normalized information entropy of the discrete FFT component magnitudes of the signal. This feature supports discrimination of activities with similar energy values as more uniform movement patterns may show a single dominant frequency component and very low magnitude for all other frequencies. On the other hand complex movements may show various frequencies of a similar magnitude[10].

2.3.2 The Wavelet Transform

The wavelet transform has been used with much success in extracting discriminative features from data to aid in classification [11][12][13]. The wavelet transform is a commonly used function [14] in signal processing applications such as decomposing, compression, feature extraction, encoding, and signal reconstruction. The Fourier transform is the cornerstone of discrete signal processing due to its ability to deal with linear time-invariant operators or uniformly regular signals but for signals that have transient properties, the Fourier transform is not ideal as it requires a large number of coefficients to represent a localized event. Wavelet bases, like Fourier bases, reveal the signal regularity through the amplitude of coefficients, and their structure leads to a computationally efficient algorithm. However, wavelets require few coefficients to represent local transient structures due to being well localized. The technical computing software MATLAB [15] has toolboxes which allow for the extraction of DWT coefficients from a data signal [1].

A wavelet is an oscillating function about zero which includes both the analysis and window functions. By shifting the wavelet over the signal and correlating the two, time information can be calculated. To decompose a signal into a set of basis functions called wavelets requires a pair of waveforms that represent the high frequencies and low frequencies. The wavelet function corresponds to the high frequency details of the signal while the scaling function corresponds to the low frequency parts of the signal [1].

2.3.3 The Discrete Wavelet Transform

Using the Discrete Wavelet Transform (DWT), any signal can be decomposed into a group of discrete wavelet coefficients. Almost all DWTs use filter banks for the analysis and reconstruction of a signal which may contain either finite impulse response or infinite impulse response filters. The filter banks contain high and low frequency filters to derive the frequency content of the signal in the sub-bands. Therefore the DWT decomposes a discrete signal into two sets of coefficients; approximation and detail. After the filtering, half of the samples can be eliminated according to the Nyquist's rule, since the signal now has a highest frequency of $\frac{\pi}{2}$ radians instead of π. The signal can therefore be subsampled by 2 by discarding every other sample [1].

Using the same method the resulting approximation coefficients are then split into new approximation and detail coefficients. This procedure is iteratively executed to create a group of approximation coefficient vectors A_i and detail coefficient vectors $D_1, D_2, \ldots, D_i$ at the ith level,

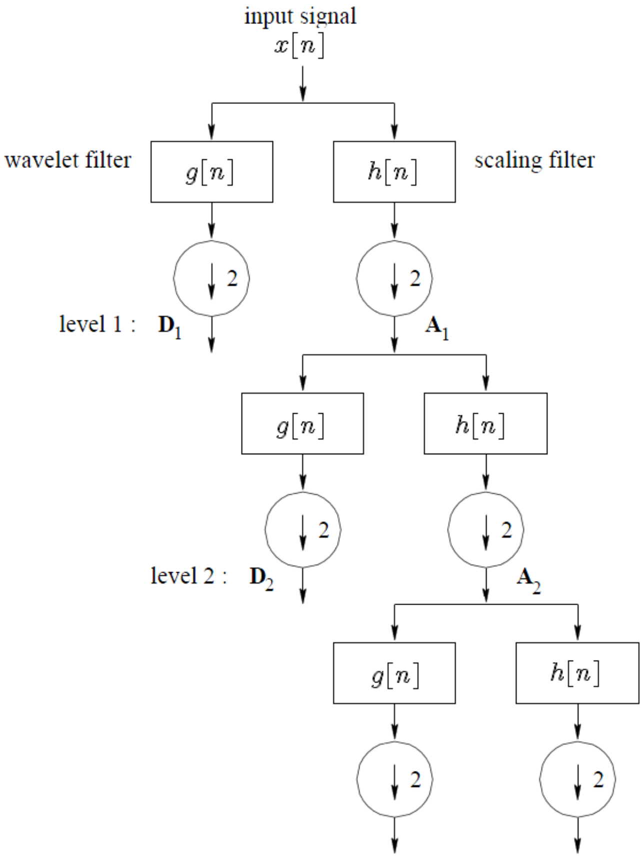

Figure 1: DWT decomposition of signal $x[n]$

as outlined in Figure 1. There are $\frac{N}{2^i}$ elements in the approximation vector A_i and $\frac{N}{2^j}$ elements in the detail vector D_j (where $j = 1, \ldots, i$) when the original signal has N elements[1].

The choice of the so called " mother wavelet" is crucial as it generates all the wavelet functions that determine the properties of the resulting wavelet transform which in turn relates to the transform's performance in any application. Currently there is no standardized way to select the mother wavelet and the choice depends on the application. The most important and commonly considered parameters when choosing a wavelet are its number of vanishing moments, its regularity, compactness and symmetry. A wavelet has p vanishing moments only if the wavelet scaling function can generate polynomials up to degree $p - 1$. More vanishing moments means that the scaling function can represent more complex functions which allows for a sparser set of wavelet coefficients. The regularity gives an approximate measure of the number of continuous derivatives that the wavelet function possesses. The regularity therefore gives a measure of the smoothness of the wavelet function with higher regularity implying a smoother wavelet[1].

The compactness of the wavelet (size) is also important. For example, Daubechies second order is fast to compute but the narrowness in time implies a very large width in frequency. Alternatively, wavelets with large compact support such as the Coiflet order 22 are smoother, have finer frequency resolution and are usually more efficient at denoising. Thus, a balance between analysis accuracy and computational time is required. The symmetry properly indicates whether the filters have a linear phase, which is an important characteristic to provide perfect reconstruction. Symmetric wavelets show no preferred direction in time, while asymmetric wavelets give unequal weighting to different directions[1].

In signal processing the most commonly used wavelets are Haar, Daubechies, Coiflet, Symlet, bi-orthogonal and reverse bi-orthogonal. Coiflets and Symlets evolved from the Daubechies wavelet. Daubechies, Coiflet and Symlet are orthogonal and compactly supported wavelets. Daubechies wavelet is asymmetric, is compactly supported and has minimum-phase associated scaling filters. Coiflet is near symmetric, is complactly supported and has the highest number of vanishing moments. Symlet has the least asymmetry, is compactly supported and has linear-phase associated scaling filters. These wavelets suffer poor regularity[1].

Bi-orthogonal and reverse bi-orthogonal wavelets are bi-orthogonal and are compactly supported wavelet pairs. Bi-orthogonal wavelets are also symmetrical, resulting in linear-phase filters, which are needed for perfect signal reconstruction. Bi-orthogonal (BO) wavelets utilize two different wavelets; one for decomposition and the second for reconstruction. Exact reconstructions are possible with FIR filters. Reverse bi-orthogonal wavelets swap the BO wavlets synthesis and analysis parts. Due to being bi-orthogonal processing time is increased[1].

2.3.4 Energy of the DWT

In this work features of sensor signals are extracted using the DWT and fed into various classification algorithms in order to correctly identify a person's current activity. Various DWT decomposition levels can be explored however each increment in level increases the overall computational time required. The total energy E_T at level i of the DWT decomposition is given by [16]:

$$E_T = A_i A_i^T + \sum_{j=1}^{i} D_j D_j^T \tag{3}$$

where A_i is the approximation coefficient at level i and D_i is the detailed coefficient at level i. One feature that can give discriminating results is the energy ratio in each type of coefficient [16]. EDR_A represents the energy ratio of the approximation coefficients while EDR_{Dj} represents the energy ratio of the detail coefficients.

$$EDR_A = \frac{A_i A_i^T}{E_T} \tag{4}$$

$$EDR_{D_j} = \frac{D_i D_j^T}{E_T} \qquad j = 1, \ldots, i \tag{5}$$

With the EDRs calculated a foundation has been created for detailed information features to be extracted. In [16] Ayrulu-Erdem and Barshan found that the normalized variances of the DWT decomposition coefficients and the EDRs provided the most informative features for a different albeit similar problem to that investigated in this thesis. Specifically they investigate extracting the informative features of gyroscopic signals using the DWT decomposition and provide them as input to multi-layer feed-forward artificial neural networks for leg motion classification. They investigated the performance of their approach using different informational features such as normalized means, minimums and maximums of the EDRs and obtained superior performance. As such it is proposed to adopt the approach in this work. The variances of the coefficients are calculated over each DWT coefficient vector at the ith level

$$A_i, D_1, D_2, D3, ..., D_i \tag{6}$$

Therefore, at the ith level there are $i+1$ variance values calculated for each axis segment, totalling $3(i+1)$ features for an accelerometer signal. The amount of EDR features is equal to the amount of DWT coefficients. Adding these features to the variances gives a total of $6(1+i)$ features at level i. Figure 2 provides an overview of how this approach fits into a complete classification pipeline [1].

2.3.5 Applications

Industrial

Tools that can accurately predict the lifespan of equipment in industry allow the optimization of resources and reduce the number of delays.

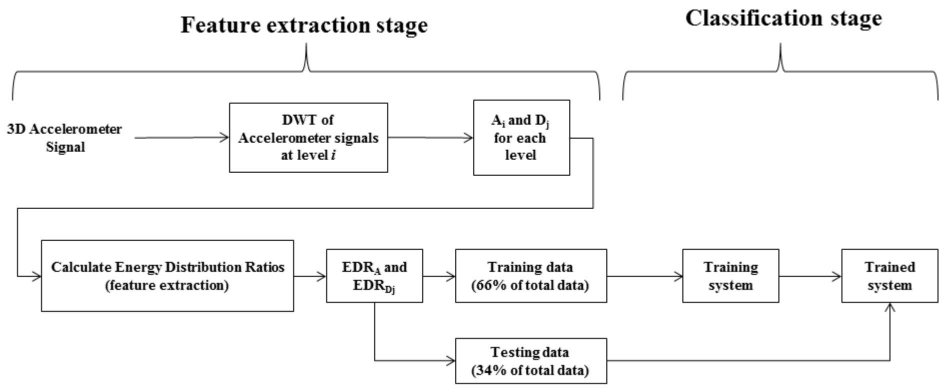

Figure 2: System overview of the DWT decomposition and classification process [1]

Saravanan et al. in [17] utilizes extracted features from the DWT to diagnose faults in a gear box. The vibration signals of a spur bevel gear box in different conditions were recorded. The DWT was used to extract features from all possible types of transients in the vibration signals and these features were passed into an artificial neural network for classification. Vibration signals obtained during the microdrilling process were used in [18] for Drill Wear Monitoring. The DWT with statistical estimations of the signal energy distribution was again employed to extract features describing energy spikes quantitatively. Non-destructive automatic identification of defects in equipment or produced goods improves safety and lowers costs. The authors in [19] present a method to detect a variety of rail-road wheel-bearing faults using audible acoustic signals at a variety of different train speeds. FFT and DWT features were both implemented with each achieving accurate results.

Healthcare

Signal feature extraction can aid healthcare related applications such as the automatic detection and classification of cardiac abnormalities [20]. This can facilitate the diagnosis of cardiac disorders and diseases more easily and quickly than current methods.

In [20] the authors propose a method to accurately classify ECG arrhythmias through a combination of wavelets and artificial neural networks. The capability of the DWT to decompose a signal at various resolutions allows accurate extraction of features from non-stationary signals like ECG. Their method is efficient at differentiating the natural sinus rhythm and 12 different arrhythmias. Heart rate variability (HRV) is a widely employed quantitative marker of the autonomic nervous system and can be used as a predictor of a person's risk to cardiovascular diseases. In [21] Acharya et al. uses the FFT to extract the power spectral densities of the HRV to determine to which of nine cardiac classes a person belongs.

Detailed contextual information regarding a user's activities can be used to encourage people to lead a less sedentary lifestyle and therefore become more active and healthy. In [10] the authors use the FFT to extract features from five small biaxial accelerometers worn simultaneously on different parts of the body. A dataset was created with twenty subjects who self annotated twenty different daily activities. The authors then investigated the performance of different classification algorithms achieving high results. Nyan et al. in [22] created a system that used DWT features to detect falls. When a fall was detected a SMS was automatically sent indicate someone has fallen and to summon professional medical assistance.

2.4 Sensing

Motion detection is the process of detecting a change in position of an object relative to its surroundings or the change in the surroundings relative to an object. Calculating a person's change in momentum involves deciphering their movement in a three dimensional space and is a challenging undertaking. The amount of academic literature which deals with wearable inertial sensors in the area of automatic activity classification has began to grow in recent years largely due to the relatively recent drop in cost of inertial sensors.

While chapter 3 describes the state of the art in using inertial sensors for automatic human activity classification, the following section more generally introduces inertial sensing and its applications.

Accelerometers have been used for human activity recognition in a large amount of existing work [23][24][25]. Research has shown that accelerometers can be used to identify human activity for high energy actions such as walking, jogging, jumping, etc [26]. In sports, accelerometers have been used to monitor elite athletes in competition or training environments. In swimming applications, accelerometers have allowed the comparison of stroke characteristics for a variety of training strokes and therefore have helped perfect swimming technique [27]. When used in competitive rowing and coupled with other monitoring techniques such as impeller velocity, they allow for the study of intra and inter stroke phases as a means to assess performance and this has been used by competition rowers to improve performance at national and international competitions [23].

2.4.1 Accelerometers

The authors in [2] describe a single axis accelerometer as a mass, suspended by a spring in a housing as seen in Figure 3. The mass is permitted to shift in one direction which is the sensitive direction of the accelerometer. The deracination of the mass is the contrast between the acceleration a and gravity g along the selected axis given by the unit vector n. $s_{A,n}$, a electrical signal is directly connected to these physical properties as seen in equation 7.

$$s_{A,n} = k_{A,n}(a - g) \cdot n + o_{A,n} \tag{7}$$

An, k defines the scaling factor while An, o defines the offset. Therefore a tri-axial (3D) accelerometer can be constructed by combining three single axis accelerometers.

Parvis et al. in [28] created an algorithm that allows the change in orientation to be described by the changes by the change in the three axes. The output vector ${}^{S}y_A$ can therefore be related to the starting acceleration and gravity by equation 8.

$$^{S}y_A =^{S} a -^{S} g \tag{8}$$

The vector is using a coordinate system and therefore there is a S on the left side of a vector to indicate this. A 3D accelerometer can be used to the calculate angle for activities in which the acceleration is lower than the gravity vector. The angle is calculated by measuring the angle of the sensor axes to the gravity vector. One negativity is that the rotation around the vertical cannot be quantified since if the device is spun around the gravity vector ${}^{S}y_A$ stays uninterrupted[2]. Figure 4 illus-

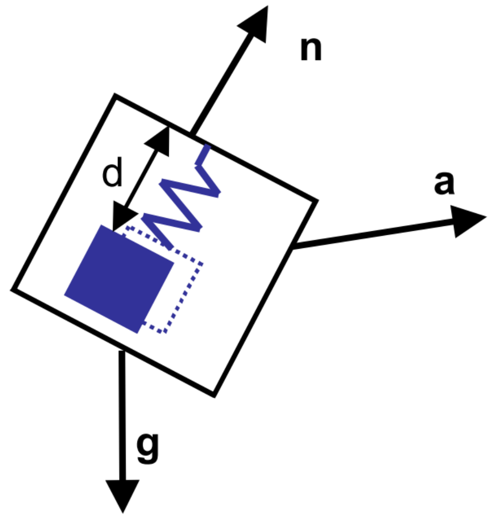

Figure 3: A single axis accelerometer, containing a mass suspended by a spring. The distance d of the mass with respect to the sensor housing is calculated and is a function of acceleration and the direction of gravity with respect to the direction of distance measurement. The unit vector n represents the sensitive axis of the sensor. [2]

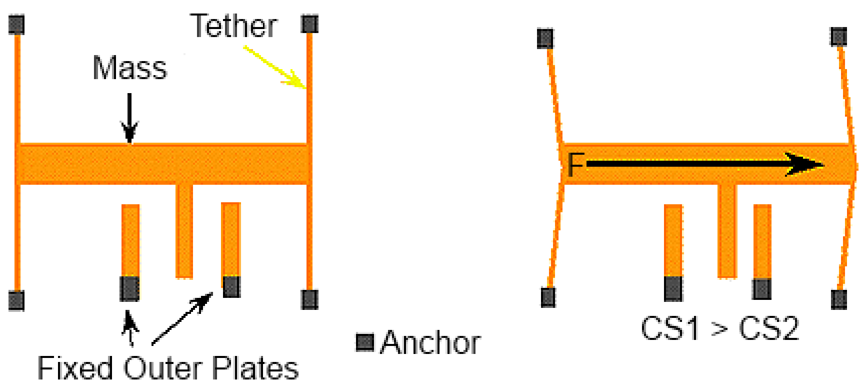

Figure 4: A single axis accelerometer showing how energy generated by a force charges an electrical circuit which can be measured [3]

trates a single axis capacitive accelerometer. Capacitive accelerometers measure the difference in electrical capacitance, caused by acceleration, to alter the output of an powered circuit. The measuring node comprises of two parallel plate capacitors (CS1 and CS2) performing in a differential mode. These capacitors work in a bridge circuit, together with two defined capacitors, and change the peak voltage created by an oscillator when the device experiences acceleration. This value is captured by a detection circuit which is then amplified and outputted [3]. While undergoing a consistent acceleration, the capacitance is constant, which results in a signal corresponding to uniform acceleration. By layering three capacitive accelerometers at alternate angles in an XYZ fashion, a 3D accelerometer can be assembled. The investigation of accelerometer signals in this work are gathered on a 3D accelerometer device with an output detailed by Equation (8), after being calibrated on the logic of Equation (7).

2.4.2 Electrocardiography

Electrocardiography (ECG) is a transthoracic interpretation of the electrical activity of the heart over a fixed period of time [29]. This activity is determined by electrodes affixed to the skin. This noninvasive sensor is used to calculate and record the regularity of heartbeats. This information can be used to infer the size and position of the heart chambers, whether the heart has suffered any damage and the effects of medication or devices used to regulate the heart such as a artificial cardiac pacemaker.

When the heart muscle depolarizes during each heartbeat, a small electrical development on the skin transpires. This development can be monitored and amplified by a ECG device. initially each cell in

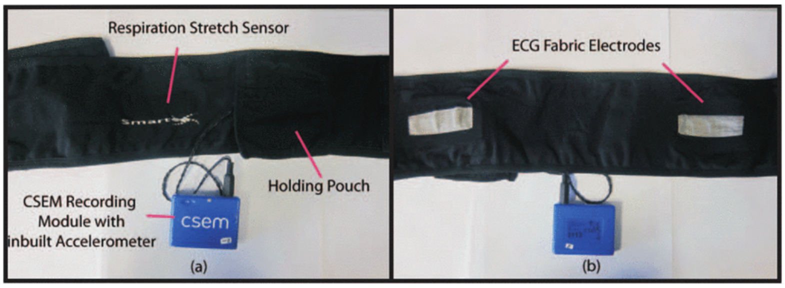

Figure 5: Smartex Wearable Wellness System. (a) Respiration sensor positioned at the front centre of the band. Accelerometer located in the CSEM recording module which is housed in the indicated pouch. (b) Fabric ECG electrodes located on the inside of the chest strap.

the heart has a negative charge across it cell membrane. The arrival of positive ions, Na+ and Ca++ boosts the negative charge to zero which is called depolarization. This kicks off the processes in the cell that make it contract. A healthy heart will have a regular progression of depolarisation waves that are formed by the sinoatrial node. These waves advance through the atrium before making their way through the ventricles. The ECG detects this as small rises and falls in the voltage between the two electrodes. An example of an ECG that was used in this work can be seen in Figure 5.

The human heart rate can alter widely according to the body's current physical activity. This includes the current oxygen absorption rate and need to excrete carbon dioxide. Physical exercise, sleep, anxiety, stress, illness and ingesting are some of the many activities that can instigate change in the heart rate [30].

2.4.3 Applications

This section presents applications for automatic human activity classification systems where unobtrusive monitoring is desired. First applications for healthcare and assisted living are described. Following that applications are presented. Additionally for completeness two other application groups are presented; industrial application uses and applications for the entertainment and gaming industry.

Healthcare and Assisted Living

Automatic human activity classification systems and context-aware computing is often motivated by the desire to create new healthcare applications and technologies for the ageing population. People all over the world are living longer due to recent improvements in public health, nutrition and medicine. Ironically this is creating new problems in the healthcare system as the percentage of elderly people in society is increasing due to longer life expectancy. It is foreseen that emerging technology can solve these problems for example by allowing the elderly to become more independent and therefore require less direct assistance by medical professionals.

Systems have been designed that aim to prevent conditions and accidents prevalent in the elderly population before they occur. Sensors which have gathered physiological data over a long period of time to identify changes or unusual patterns in a user's daily activities which can indicate early symptoms of diseases such as alzheimer's disease and thus help prevent them. Automatic recognition of small changes in daily activities is a popular area of research. Accelerometers can log vast amounts of data over long periods of time and thus have the capability to give user's summaries of daily activities [31] or applications which accumulate data from physiological sensors [32] [33] allow doctors and care givers to give better service as they can better determine the current health state of a patient

Detailed contextual information of a users activities can be used to encourage people to lead a less sedentary lifestyle and therefore become more active and healthy. Accelerometers in mobile phones were used by Maitland et al in [34] to increase the awareness of daily activity levels. By monitoring daily activity levels they can provide regular detailed summaries and use this information to encourage the user when appropriate. In [35] Consolvo et all describes a similar method where inertial sensors are employed for automatic activity recognition so that when a positive activity such as walking a certain distance durning the day is achieved by the user, a feedback system displays virtual rewards on a mobile phone screen. Andrew et all in [36] utilizes localisation data as well as inertial data on a mobile phone to suggest contextual physical activities. One such example is suggesting to walk to the next bus stop when there is ample time before the next bus will arrive.

Human activity classification systems can be used to diagnose diseases and disorders. This can lower healthcare related costs and speed up the time taken to diagnose a patient. In [37] the authors create a real-time monitoring system for cardiovascular disease using a wearable system. They fuse the portability of Holter monitors with a start-of-the-art Smartphone that can provide an helpful diagnosis solutions. Capturing data via a Smartphone instead of a large, bulky contemporary ECG machine allows much more data to be recorded without drastically interfering with daily activities.

Recent advancements in wireless communication technology have helped to improve non invasive wearable sensors, which can be used in the home or at a health institution. Wearable sensors that allow motion to be captured can be used for healthcare treatment and diagnosis [38][39]. Accelerometers and gyroscopes are the most frequently used inertial sensor used to distinguish between motor movements in the healthcare system [40]. They can be attached to any location on the body as well as specialised equipment to collect inertial data created by patients or even recuperating athletes. Glaros et al [41] incorporates a portable virtual reality interface to help optimize treatment and training procedures during rehabilitation for athletes. It also provides them with instant feedback on mistakes made any time during a training session.

Sporting & Leisure Applications

Sporting applications is another area where inertial sensors have found significant use. In [42] Ermes et al constructed a wearable system that could identify basic everyday activities as well as sporting activities such as playing soccer, riding a bicycle and performing exercise routines such as rowing. A Neural Network classifier was employed to distinguish between different activities. In [26] the authors calculated the amount of energy expended when a selection of daily activities such as walking, running, cycling and driving are performed. Along with these daily activities some sporting activities such as soccer, volleyball, badminton, boxing and table tennis, were analysed .

Motion detection in the context of martial arts activities can be identified by fixing a 3D accelerometer to the torso of a subject to capture the unique body acceleration performed. The research conducted in [43] uses accelerometers and gyroscopes attached to the body to identify different actions in Wing Tsun to increase immersion in video games of martial arts. Additionally this same technology could create similar systems for martial arts instruction. In [44], the authors present an on-body wireless sensor system for measuring snowboarding specific activities in real-time (see Figure 6).

Figure 6: A on-body wireless sensor system for measuring activities during snowboarding in real-time

Accelerometers are used to calculate force impacted on the snowboard along with an intelligent wireless network that captures and analyses the posture and motion of the snowboarder.

Ghasemzadeh et al. use signal progressing algorithms to calculate the angular rotations of a subject's wrist during a golf swing in [45]. As in [43] where Heinz is able to find the quality of the martial arts movements performed Ghasemzadeh et al. is able to quantify the users expertise and skill level of the person making the golf swing. The system can then recommend appropriate feedback for the user. Arvind et al. use a double pendulum system to model the golf swing and use accelerometers placed along the body and golf club in [46]. This setup allows them to determine how closely the movements of the user follow a predetermined motion and give an appropriate score.

Commercial systems which employ inertial sensors for sporting applications include Nike+, which monitors an athlete's sporting activities. A small transmitter device is either embedded or placed on the shoe, which can log all running and jogging exercises. This data can be aggregated over time to allow the user to observe the change in their performance over time. This sensor can be integrated with additional devices such a smartphone that allows them to engage in challenges with other runners and walkers. The F50 adiZero mi-Coach was first released in 2011 and it features mi-Coach match analysis technology which provides feedback on user performance. Mi-Coach is a three part system including a stride sensor, a heart monitor and a receiver. It tracks a subject's max speed, number of sprints, distance travelled and number of sprints. This data can be aggregated over time to allow the user to track his/her fitness training performance.

Industrial Applications

In an industrial setting, automatic human activity classification systems can potentially help workers in their responsibilities, reduce accidents, improve productivity and increase overall safety in the workplace. Xybernaut has been creating wearable monitoring systems since 1990 that support workers in industry by allowing them to conveniently access relevant information and to collect suitable data. In [47] Stanford explains that the shipping, airline and telecommunications industries were the first organisations to incorporate intricate wearable technology successfully into their businesses.

The authors in [48] show an example of on body sensors being used in emergency response units, hospitals, aircraft maintenance and motor manufacturing assembly lines. Data collected by the wearable sensors allow activity classification software to provide hands free interactions to data which speeds up the training of new workers and creates a summary of worker activities. The authors in [49] combine information from body worn microphones and accelerometers to recognise activities that are

Figure 7: XBee accelerometer sensor box for integrating dance motion with interactive visualizations (with a quarter shown for size comparison)[4].

characterized by a hand motion and an accompanying sound. They describe a method for the continuous recognition of activities such as sawing, hammering, drilling and grinding.

Entertainment and Games

Increasingly, entertainment systems like home gaming consoles and smartphones are incorporating player activity classification technology to allow the creation of a wide range of customised applications for entertainment gaming. Accelerometers have been employed to distinguish between different activities in various entertainment contexts. One such area is the performing arts where sensors have been secured to dancers to enhance audience interaction. In [4] the authors investigate lightweight methods for integrating dance motion with interactive visualizations and enhancing audience interaction. The sensors allow dancers to add an extra dimension to their performance with interactive multimedia content that correlates with their movements Their proposed lightweight system can be seen in Figure 7.

The authors in [50] introduce a system for augmented reality-based evaluations of Salsa dancer performances. Their system enables an enhanced dance visualisation experience, through the augmentation of the original media with the results of their automatic analyses by fusing data from wireless inertial measurement units and audio sensors.

Visualising dancer's motions is a popular research area as seen in [51][52][53]. Inertial sensors are used in correlation with advanced machine learning technology to recognise different body movements. The data captured by the sensors is logged and then once the performance is finished this logged data is classified which in turn allows the dancers movements to be visualised.

The gaming industry has pushed the introduction of inertial sensors into their products for some time. The hugely successful Nintendo Wii which has sold more than 100 million units has an accelerometer in its controller which allows gesture recognition which is the main method of interacting with the system. Zhang et al. in [54] introduce a system to control a computer game using a wearable motion sensor. Similarly in [43] the author's employ body worn accelerometers to detect movements which control an avatar in the game.

Other Application Areas

Inertial sensors have been used in other fields for automatic activity classification. In [55] the authors employ wearable accelerometers to recognise soldier activities. These sensors can record the list of activities performed by a soldier during a mission or training exercise which then can be used to automatically generate field or training reports. Additional this technology can automatically supply vital information to military command to aid timely strategy decisions. Sala et al explains in [56] that activity recognition with inertial sensors can be used to target mobile advertising.

2.5 Sensor Fusion

The vestibular system in the inner ear of biological creatures supplies inertial information which is required for movement, body position and orientation. This system allows humans to achieve efficient head stabilisation and perform visual tasks. The knowledge acquired by the vestibular system is required to perform eye movements such as tracking and gaze holding [57]. It has been established that human vision and the vestibular system fuse neurological signals at a very early processing stage [58]. The inertial information increases the accuracy of the vision system and the visual cues help spatial orientation. This approach of using combining complementary sensor signals to aid a system can be used in computer science. There are many advantages to sensor fusion and in this thesis, fusion of multiple sensor streams is used, in one example the fusion of accelerometers, ECG and respiration sensors are proposed to increase the overall classification accuracy of a system compared to a solitary sensor stream.

Inertial sensors have been employed for navigation systems as well as guidance of defence systems. Orientation, speed and altitude are calculated using accurate accelerometer and gyroscope sensors whose data is fused with localisation technology such as the Global Positioning System (GPS) as well as data from radar stations. Each of these sensors is capable of creating vasts amounts of data so therefore intelligent algorithms are required that can extract and combine the pertinent pieces of information. All sensors have advantages and disadvantages and no one sensor is a 100% accurate at measuring a physical quantity.

Sensor fusion is the process of fusing the sensory information from two or more noisy sensors to acquire useful information, where data captured is unique. Sensor fusion has many applications such as automatic vehicle guidance, automatic target recognition in missile guidance, combat surveillance and automated threat detection systems, such as identification friend-foe-neutral (IFFN) systems [59]. Non military applications include movement of materials in manufacturing processes, robotics [60], automatic vehicle guidance such as robotic vacuum cleaners and also within healthcare devices.

The techniques used to combine the data is drawn from areas such as digital signal processing, statistical evaluation and artificial intelligence [61] [62]. In 1985 the Joint Directors of Laboratories (JDL) created the Data Fusion Group which published a model which separated the different processes associated with data fusion into 6 levels [63]. This model is still in use today and provides researchers important guidelines for data fusion. Other widely used approaches to fuse sensor data are Bayesian Fusion [64] and Kalman Filtering [65]. These methods can be employed to fuse data from various indirect and noisy sensors.

Sensor fusion allows the advantages of one sensor to overcome the disadvantages of another. One example of this is how magnetometers are utilized to reduce integration drift that occurs in gyroscopes. Iron in magnetic equipment interferes with the local magnetic fields and this affects the orientation measurement. This drift problem can be rectified by examining the errors in the gyroscope drift as it will have a different pattern than found in local magnetic field.

There are two approaches to fusion; early and late. The difference between this two approaches is simply when in the classification pipeline results are fused together. Early fusion merges the features of each modality before any machine learning is conducted whereas late fusion merges the features of each modality after some machine learning has been conducted. These two approaches are described in more detail in section 4.4.

2.6 Machine Learning

Data mining is a methodology which can infer new knowledge by extracting information from a data set and transform it into an understandable structure for further use. These structures can lead to potentially useful new information that is not always apparent without intelligent data analysis. Machine learning is a technology which concerns the creation and study of systems that can learn from data and is therefore a branch of data mining. It requires computer programs which are trained to locate patterns in data. The following sections describe the main principles of machine learning. First, Bayes theorem is introduced, which is the fundamental equation for statistical learning. This is then followed by an introduction to classification and regression before presenting the main differences between generative and discriminative models. Finally supervised and unsupervised approaches in machine learning are described.

2.6.1 Bayes Theorem

Bayes theorem is widely used to find probabilities in machine learning and is fundamental to Bayesian Networks. In mathematical terms, Bayes theorem states the relationship of the probabilities of A and B, $P(A)$ and $P(B)$, and the conditional probabilities of A given B $P(A|B)$ and B given A $P(B|A)$ [66]. In its simplest form, it is:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{9}$$

2.6.2 Regression & Classification

Regression and classification are two machine learning methods to construct prediction models from a dataset. The goal of classification is to determine a route from the feature space, F, to a label space, L. For instance, where the feature space is $F \in R^d$ and the label space is $L = \{0, 1\}$, the function $f : R^d \to L$ can be used to classify each instance in F to its most likely discrete value in L. It is this mapping function, f, which is the classifier and each classifier seeks to reduce the generalisation error. The generalization error of a model is a function that calculates how well a learning machine generalizes to unseen data. The aim of regression is to determine a route from the input space, I, to the output space, O. This f, is the mapping function. For instance, we might have, $I \in R^d$ and $O = R$, then regression will use the mapping function $f : R^d \to R$, to determine which indiscrete output a given instance belongs to. The main difference between regression and classification is that in regression the output variable takes continuous values whereas in classification the output variable takes class labels. Classification has dependent variables that are categorical and unordered. Regression has dependent variables that are continuous values or ordered whole values.

2.6.3 Generative & Discriminative Models

There are two different types of models traditionally used in machine learning; generative and discriminative. A generative model calculates the joint probability distribution $p(x, y)$ and a discriminative model (also called conditional models) calculates the conditional probability distribution $p(x|y)$, which is the probability of y given x[66].

Discriminative models do not allow the generation of samples from the joint distribution of x and y but since classification and regression do not require it discriminative models can yield superior performance [67].

Generative models grant a degree of ambiguity, uncertainty and abstraction. They perform well while modelling time-series data as they tend to be efficient in manipulating large mounts of data [68].

Prominent generative methods include Hidden Markov Models, Naive Bayes, Gaussian mixture models, Latent Dirichlet allocation and Bayesian Networks. Popular discriminative methods include Support Vector Machines, K-Nearest Neighbour, Neural Networks and Linear regression. While these methods have differences, they have the common goal of constructing the perfect decision hypothesis that reduces classification errors to zero. Both approaches use feature representation to estimate the class label [69].

The judgement criteria for choosing either a generative or discriminative supervised approach has been a constant source of discussion in the machine learning community, culminating in an array of research on the area being published in the literature.

For instance Ng et al. puts forward in [70] that discriminative models have issues with over-fitting when the amount of training data is low so therefore a generative mode is more applicable Yeom et al. concurs in [71] that generative models are most suitable when the data contains a large amount of ambiguity and there is not enough data to sufficiently train against. The authors in [70] put forward that discriminative models lack the complexity of generative models. This can be an obstacle for users since they might need manual calibrate using kernel functions, regularization and penalty methods and that the connection between parameters are not well defined so they act more like black boxes. In this work both approaches are investigated to ascertain which is more appropriate for autonomic activity classification.

2.6.4 Supervised & Unsupervised learning

There are two fundamental approaches used in Machine Learning; supervised learning and unsupervised learning.

Supervised learning is the machine learning method of creating a function from labelled training data. The training data consist of a set of annotated training examples. In supervised learning, each example consists of two pieces of data. The first is an input object which is normally a vector and the second is its associated label often called the supervisory signal

A supervised learning algorithm evaluates the training data and constructs a function, which can be used for mapping new examples. The goal is to allow the algorithm to correctly identify the appropriate class labels to unseen examples. This requires the learning algorithm to not over-fit when constructing models from the training data and to permit unseen situations to be correctly classified. This process occurs in human and animal psychology and is often termed concept learning.

Unsupervised learning is the machine learning method of trying to locate hidden structure in unlabelled training data Since the examples in-putted into the classifier are unlabelled, there is no error or reward signal to evaluate the constructed classifier. The aim of unsupervised learning systems is to locate obscure structures from unlabelled data.

In the supervised machine learning approach, there are three fundamental areas. The first is binary or binomial classifiers where the elements of a given set are split into two groups, on the basis that the examples of each class have some similar traits. The second supervised learning approach is known as multiclass or multinomial classification which aims to classify examples into two or more classes. The third method is regression which has been discussed above. Some commonly used terminology in machine learning is introduced in the next section.

2.6.5 Concepts, Attributes and Instances

In machine learning, an input is the sum of three parts; concepts (classifier), instances (examples) and attributes (features). The aim of the learning process is to produce a distinct characterization of what the data represents in a form that the classifier can use to locate analogous traits. This work employs supervised learning techniques to construct models from a training dataset that can receive unknown examples that can be classified accurately.

The data the classifier is supplied with is refereed to as an instance. Each instance is a unique example of the concept to be analysed. Each training data set comprises of a set of these instances. The traits of each instance are judged by the attribute values which are contained in every instance. These definitive properties are a measurement of some desired trait that are used to distinguish between classes.

2.6.6 Classifiers

A classifier refers to a mathematical function, that maps input data to a category. In this section, we discuss five popular families of classifiers that were employed in this thesis. The classifiers choosen are those most commonly used in the state of the art and collectively represent a broad range of different approaches.

Support Vector Machine (SVM)

Support vector machines have secure theoretical foundations, strong regularization properties and excellent empirical successes. They have been applied to tasks such as image classification [72], speech processing [73], protein classification [74] and human activity classification [75]. Support Vector machines can be defined as systems which use hypothesis space of a linear functions in a high dimensional feature space which are trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory [76]. SVM performs well on data sets that have a large amount of attributes, even data sets which contain very few cases on which to train the model. In fact there is no upper limit on the number of attributes a data set can contain and hardware poses the only constraints. The SMO algorithm is used to efficiently solve the optimization problems which occur during the training of SVMs. SVMs are often described as a "black box" classifier as the user does not need to choose many parameters. In this thesis, John C. Platt's Sequential Minimal Optimization (SMO) optimization algorithm was used for the training of the SVM classifier.

K Nearest Neighbour (K-NN)

K-NN algorithms are used for classifying data based on closest training examples in the feature space. K-NN is a class of instance-based learning techniques where the function is only approximated locally and calculations are suspended until classification. The K-NN algorithm finds a group of k objects in the training set that are nearest to the inputted object, and judges the allocation of a label on the predominance of a class in this neighbourhood. For this method there are three basic components: a set of labelled attributes, a distance measure to compute distance between objects, and the value of k, the number of nearest neighbours. To classify unknown data, the distance of this data to the known data is computed, its k-nearest neighbours are determined, and the class labels of these neighbours are then used to identify the class label of the unknown object. K-NN algorithms can handle missing values, are robust to outlying data points, and have a good history as predictors. They tend to only handle numeric variables, are sensitive to monotonic transformations of features, are not immune to insignificant inputs, and provide models that are difficult to interpret. The K-NN algorithm sets equal weighting to all inputs therefore it is sensitive to noise and redundant features. It has been used in many applications in the field of data mining, statistical pattern recognition, image processing and many others. Some successful applications include recognition of handwriting [77], text classification [78] satellite imagery analysis [79] and ECG pattern analysis [80]. The IBk classifier is a simple instance-based learner that uses the k nearest neighbour (k-NN) algorithm for training and was used in this work.

Bayesian Network

One very important probability-band classifier is the naive Bayes method which is also known as idiot's Bayes, simple Bayes, or independence Bayes. It assumes that the presence or absence of a particular feature of a class is unrelated to the presence or absence of any other feature, given the class variable. This method is significant for many reasons. It does not need any complicated iterative parameter estimation schemes therefore it is simple to construct. This means it may be applicable to large datasets. One advantage of the naive Bayes classifier is that to calculate the parameters (means and variances of the training data) required for classification, it only requires a small amount of training data. Only the variances of the variables for each class need to be determined because independent variables are assumed, and not the whole covariance matrix. This classifier has been used in a large range of applications such as medical diagnosis [81], data mining [82] and musical style recognition [83] and was employed in this thesis.

Classification tree

Classification trees create a model that predicts the value of a target variable based on several input features. In these tree structures, leaves represent class labels and branches represent conduits that allow features to lead to class labels. A logistic model tree, which is a classification tree with logistic regression functions at the leaves was employed for this work. This method has been shown to give better results than standard decision trees and simpler logistic methods[84]. A stage-wise fitting process is used that selects relevant attributes in the data. This incrementally refines the leaves constructed at higher levels in the tree. The logistical model tree has been used in applications such as ECG arrhythmia studies [85], textual entailment classification [86] and real-time human movement classification using accelerometers [24] and for this reason was choosen in this work.

Artificial Neural Network

Biological neural networks have inspired mathematical models called artificial neural networks (ANNs). A multilayer perceptron (MLP) [87] is a feedforward ANN that consists of multiple layers of nodes that each have the same destination, with each layer completely connected to the adjacent layers. Apart from the input and output nodes each node is a neuron, that is to say a processing element with a nonlinear activation function. MLP uses backpropagation for training the network which allows the network to converge on a satisfactory feature weighting and flow. MLP is an adaptation of the standard linear perceptron and can analyse data that is not linearly separable. The MLP has been used in a wide array of classification problems such as skin segmentation [88], classification of multispectral satellite images [89] and recognizing human motion with multiple acceleration sensors [90]. For these reasons the MLP was choosen as one of the classifers employed in this work.

Classification Evaluation

F-measure gives a measure of a classifiers accuracy. It uses both precision p and recall r of a test to calculate the score. Precision is calculated as the number of correct results divided by the number of total results while recall is the number of correct results divided by the number of results that should have been returned positive. These metrics are often described in terms of the metrics true positive (T_p), false positive (F_p) and false negative (F_n). The F-measure score is a harmonic mean of precision and recall, where an F-measure score reaches its best value at 1 and worst score at 0. In this work all results are presented using the F-measure algorithm unless otherwise stated.

$$F_1 = 2.\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

$$\begin{aligned} \text{Precision} &= \frac{T_p}{(T_p + F_p)} \\ \text{Recall} &= \frac{T_p}{(T_p + F_n)} \end{aligned} \tag{11}$$

2.7 Heuristic Approach to Optimisation

A heuristic approach refers to a experience-based method used for problem solving, learning, and discovery where an exhaustive search is impractical. The solution acquired by a heuristic approach is not guaranteed to be optimal but typically used to accelerate the procedure of locating a adequate solution. Humans use heuristic methods such as rule of thumb, an educated guess, intuition, discrimination or common sense to hasten decision making. In the context of a heuristic algorithm, a heuristic will be a process of performing a slight alteration, or a sequence of alterations, of a given or partial solution in order to achieve a contrasting solution or partial solution. A heuristic algorithm iteratively applies one or more heuristics in conformance with a specified design strategy.

2.7.1 Genetic Algorithm

The genetic algorithm (GA) is a search heuristic that mimics the mechanism of natural selection. It is frequently used to achieve effective solutions to search and optimisation problems [91]. In an genetic algorithm a population of candidate solutions to an optimisation problem are iteratively altered in an evolutionary manner aiming for improved solutions. Each of these solutions has a known set of parameters called chromosomes which are mutated throughout the process. Occasionally, the chromosomes may be weighted in areas where optimal solutions are expected to be found. In this thesis, a genetic algorithm is used to select the optimal parameters which create the most accurate classification models.

The initial population set is generated randomly and the evolution of this population is an iterative process. The population in each subsequent population is called a generation. All candidate solutions have their fitness evaluated upon creation. Fitness is defined by the objective function in the optimization process being solved. In this work the fitness of the candidate solution is the classification accuracy of the models created.

During each subsequent generation, a selected proportion of the previous generation is selected to breed the next generation. Candidate solutions are created through a fitness-based procedure where the fitter the candidate solution the more probable it is to be selected for a genetic operation. A genetic operator is an operation used in genetic algorithms to create candidate solutions. One example is mutation which is a genetic operator used to preserve genetic diversity from one generation to the next and behaves like its biological counterpart. Mutation transforms one or more chromosomes from one state to another based on a mutation probability.

Crossover is another genetic operator that fuses two candidate solutions called parents together to create a new one. Fitness proportionate selection also know as roulette wheel selection is a genetic operator used for identifying potentially effective candidate solutions for recombination. In fitness proportionate selection the probability of selecting an individual candidate solution as a parent is based on its fitness. The higher the fitness the more likely it will be selected. If f_i is the fitness of individual i in the population, its probability of being selected is

$$p_i = \frac{f_i}{\sum_{j=1}^{N} f_j} \tag{12}$$

where N is the number of individuals in the population. In contrast to simpler selection algorithms such as truncation selection there is a probability that some less fit candidate solutions may withstand the selection process. This can be advantageous as though a candidate solution may be unfit, it could include a chromosome that could prove effective after the recombination process. Additionally the probability that a candidate solution undergoes either crossover or mutation can be set prior to the commencement of the genetic algorithm.

One popular genetic operator due to its success[92] is elitism. This is where a number of the most fit candidate solutions are kept unaltered and guaranteed a place in the next generation. They are able to be selected as parents and allow mutations created from them in addition to their original form being brought into the next generation. All three of these genetic operators, crossover, mutation and elitism are utilized in this work.

There are three commonly used criteria employed to judge when the genetic algorithm process should stop. These are achieving a required solution, reaching a candidate solution count limit and reaching a convergence break limit. Achieving a required solution occurs when the user has selected a specified target solution for the genetic algorithm for example creating a classification model of 100%. A candidate solution count limit is where the user has set a maximum number of solutions that can be calculated and once this value is reached the genetic algorithm stops. Usually this is calculated as a percentage of the maximum number candidate solutions that can occur. Finally the convergence limit is where the best performing candidate solution has ceased developing for a set number of candidate solutions. This again is usually calculated as a percentage of the maximum number candidate solutions that can be created. The genetic algorithm uses stochastic processes but the result is non-random. It can be used for a number of different multidimensional optimization problems in which the chromosomes are known and are non-infinite.

2.7.2 Applications

Many practical classification task require learning of an appropriate classifier function that assigns a given input usually a series of attributes to one of a finite set of classes. The choice of features, attributes, feature size and classifier type all have a direct impact on the overall accuracy of a model. This presents a parameter selection problem in automated design of activity classification. The feature subset parameter selection problem refers to the task of recognizing and determining a efficient set of parameters to be used to represent patterns from a larger set which contains often redundant, possibly irrelevant parameters with different associated measurement costs and or risks.

Huang et al. in [93] uses a GA gene-based feature selection and parameters optimization for support vector machines. The SVM, a popular technique for pattern classification, had its parameters and feature subset optimized without degrading classification accuracy. Creating classification models for large datasets is a time consuming task as finding the most productive parameters is computationally expensive. Punch et al. in [94] employs a GA in conjunction with a K-nearest neighbour algorithm to optimize classification by searching for an optimal feature weighting. This basically warps the feature space to coalesce individuals within groups and to isolate groups from each another. The GA can also be used to implement efficient methods of fusing classification models for one overall prediction goal. In [95] the authors use a genetic algorithm to design a multiple-classifier system. They tested their methods on four real data sets. They found that GA design was less prone to overfitting compared to classifiers using: all features; the best feature subset found by the sequential backward selection method; and the best feature subset found by a GA (individual classifier).

2.8 Conclusion

This chapter presented the technical background necessary for the experiments designed to investigate automatic activity classification. A comprehensive state of the art review of all relevant areas provided the basis for certain key decisions taken. The Butterworth filter is chosen as the filter of choice. It was decided to investigate both Fourier and Wavelet Transforms for feature extraction. A sensor fusion approach is suggested in this work. A variety of classification models including SVMs, Decision Trees, ANNs, Bayesian Networks and k-NNs were selected for evaluation. Finally a genetic algorithm was investigated to optimise the selection of parameters that would yield the highest classification accuracy.

3 Unimodal Human Action Recognition

In this chapter novel approaches are examined to automatically recognise human actions using a single sensor worn by a human subject. Firstly, a method is proposed that allows for the classification of a subject's ability to perform a single desired action. Then a number of different techniques are investigated to ascertain each techniques ability to accurately distinguish between certain types of activities. Algorithms are proposed that can be employed on a wide range of human activity classification problems. The validity of the approaches is demonstrated by applying them to real world noisy data collected as part of this thesis.

3.1 Introduction

This chapter gives an introduction to inertial sensors for identifying specific activities before giving an overview of the current research challenges faced in this field. Following this a novel method for automatically classifying human activities with a signal accelerometer is given.

One of the fundamental goals of wearable computing is to allow for the creation of personal software that can adapt and respond to the current context of the user appropriately. "Context" in this work is defined as all types of information about a user or the objects that are surrounding him or her. Context aware computing takes into account a user's state and surroundings, and the mobile computer modifies its behaviour based on this information [96].

The capacity of a system to be able to quantify a subject's ability to perform an activity has many applications in sport and health. For the case of sporting applications the authors of [97] present a golf swing training system, which incorporates wearable motion sensors to obtain inertial information and provide feedback on the quality of movements. Similarly in [44] Spelmezan et al. present a wireless prototype system used to detect some common mistakes during snowboarding and to provide students with immediate feedback on how to correct their mistakes. In the field of healthcare Jovanov et al. designed a wireless system that is composed of intelligent motion sensors for computer assisted physical rehabilitation [98]. Glaros et al. present a wearable intelligent system for monitoring health condition and rehabilitation of running athletes [41].

The first work completed in activity recognition with wearable sensors was in 1993 when breakthroughs in hardware technology allowed sensors to constructed that were light enough for portable automated systems to be developed. These systems could be fastened to a human subject for a long duration of time[99]. These initial research prototypes were somewhat cumbersome and interfered with a subject's movement and comfort more than desired. However it was predicted that these systems would "vanish into the background" [5]. It was envisaged by Weiser in [5] that the potential of allowing a computer to sense human actions would allow the creation of genuinely personal applications.

Early work focused on the established text and keyboard based applications before exploring advanced approaches of capture and communication. Some examples of these new approaches were the placement of cameras on subjects to capture contextual information from visual data [100] [101], or placing microphones [102] to extract context from environmental audio. Other contextual data would be incorporated with existing audio or visual data such as the subjects current location, the topic of a conversation and the identity of who the subject is communicating with. This information can be used to provide the subject with pertinent real-time information about their current activity or to record this information for future analysis [103].

Calculating the physical activity of a subject using technology that works for subjects of all shapes and sizes has been a consistent aim for researchers across numerous scientific disciplines. Accelerometers have primarily been employed for this ambition for several decades as they can be attached easily and securely to various locations on the body [104][105].

The goal of these early systems was to calculate global measures such as the total energy expenditure or the oxygen requirements of the subject while he or she was completing an array of contrasting activities. Portable systems that incorporated accelerometers that could distinguish particular physical activities began to emerge at the turn of the last decade. The growth in the prevalence of these systems was facilitated by advances in hardware electronics and new machine learning approaches as well as the expected usefulness for the new paradigm of context-aware computing [106] [107]. Unfortunately these early devices contained prohibitively costly hardware and difficult to use user interfaces [108].

Current research in automatic human activity classification from wearable sensors covers many areas such as activity recognition in daily living for healthcare [109], automatic recognition of activities in unlabelled data [110], semi-automatic or unsupervised learning of activities [111] [112] or combining various sensor modalities to increase recognition accuracy [113].

There have been various techniques applied to raw sensor data in order to recognise human activities. Due to the electronic nature of "on body" sensors, it is common for the captured data to undergo a pre processing step to eliminate noise. The detection and removal of high frequency noise in acceleration data is an important step as without it feature extraction and classification become more difficult. Some successful methods to remove high frequency noise are low-pass median [24], Laplacian [114], and Gaussian filters [115].

The capacity of a activity recognition system to remove unnecessary information from raw data while preserving pertinent information is critical. This process has a direct impact to the efficiency, computational time and success of activity recognition systems. Sensors can capture immense volumes of data, which if unregulated can overwhelm the whole activity recognition system. Consequently it is vital to locate abstractions in the raw data via relevant features. The feature vector allows the detection of independent actions using dimensionality reduction [116] [117] and these vectors are then employed for classification.

One of the most common approaches used in order to extract features is the Fourier Transform, which has the ability to discriminate between useful and redundant information by representing the data in frequency clusters and thus reducing the amount of dimensions of the sensor data [118]. Discrete-Fourier Transforms are an application specific version of Fourier Transform that take discrete data as an input [119]. Since it deals with a finite amount of data, it can be implemented in computers by numerical algorithms or even dedicated hardware. These implementations usually employ efficient fast Fourier transform (FFT) algorithm.

A more rudimentary approach to extract discriminative features from sensor data is to use time-domain features. These can be signal statistics and basic waveform attributes, which are directly calculated from a raw data segment. In [120] Ward et al. attach a triaxial accelerometer and a microphone to a subject to recognise human actions. They extract the mean and variance and compute the number of peaks for each accelerometer axis as the features to represent each action to be recognised. Their aim was to recognise construction activities such as hammering, drilling and sawing and these time domain features allowed them to accomplish this to a high degree of accuracy. Calculating the variance of raw accelerometer signals has also been successfully used to recognise human actions in [107] [121] [122] [49] [116] and as such is one of the more popular time-domain features.

Sensors have been placed on various locations on the human body in order to record different physiological data. The placement of the sensor is highly dependent on the actions that are to be recognised. In [120] the author placed an accelerometer on the subject's wrist when attempting to recognise human construction actions as each tool required a unique wrist motion. In [123] the authors successfully employ an accelerometer placed on the upper back to recognise dynamic and static human activities. Additionally this placement generally does not interfere with movement and is unlikely to get damaged during sporting contact as it is illegal in the vast majority of contact sports to exert a force on an opponent's back that would damage the sensor.

3.1.1 Discussion

This section deals with the specific case where only one sensor is attached to a subject. Using a single sensor reduces interference with a subject's normal body movements, reduces the chance of injury as the less sensors attached to a subject the less chance a sensor could cause damage and finally lowers the cost of the system. Embedding sensors into normal clothes is also advantageous as it requires subjects just to simply get dressed before important physiological data can be captured and studied.

The sensor and its communication platform can be coated in a non conductive polymer that makes the clothes machine washable. This makes the system reusable thus reducing cost further.

3.2 Evaluating a Subjects Performance

3.2.1 Introduction

Wearable sensors have an important role to play in future healthcare delivery and management by sensing the body and interpreting physiological data[124]. In the experiment outlined in this section the aim is to develop an interactive system to help patients perform respiratory exercises, and particularly to maintain the interest of children during exercise sessions. This comprises a wearable system which classifies breathing technique into separate grades and provides visual feedback to the user through a graphical user interface. An additional advantage of this system is that exercise sessions can be repeated using the same reference signal, which means that medical staff need not be present for the exercise, thereby improving efficiency in the hospital. Typically these types of exercises require some degree of expert supervision and monitoring. In this section a classification framework is created that classifies a subject's ability to perform a desired action.

3.2.2 Motivation

The average human takes approximately 7×10^9 breaths in their lifetime and each one is critical to maintaining homeostasis. Breathing is essential for our survival and yet it is generally something people perform without conscious thought. Breathing behaviour can have a profound impact on a person's health[125]. Furthermore, despite the fact that a person's breathing technique can affect overall health, we receive very little information this in everyday life. For example, by breathing in a slow, steady and deep manner, a person's heartbeat slows and relaxes, blood pressure normalises, stress hormones drop and muscles loosen. These techniques are availed of by athletes to improve performance and reduce stress before and during competitions [126].

There are many diseases that affect the respiratory system in humans. One such disease is asthma, an inflammation of the lungs which causes narrowing of the airways. Asthma affects an estimated 300 million people worldwide. Another is Cystic Fibrosis(CF), a life-threatening disease, which has a high occurrence ratio in Ireland, with one in nineteen people being carriers of the CF gene [127]. Whilst this hereditary disorder affects the entire body, breathing complications are the most serious symptom and frequent lung infections typically occur. One of the symptoms of CF is the accumulation of large amounts of phlegm the lungs. A widely accepted technique to provide relief in such cases is the loosening and removal of phlegm through various breathing practices.

The Active Cycle of breathing is one such popular technique that uses breathing exercises to remove phlegm from the lungs. Clearing secretions from peripheral airways is the most important defence mechanism of the respiratory system. The technique involves 4-5 deep breaths coupled with holding periods in-between to allow air to be transported behind obstructed areas in the lungs. These exercises need to be performed regularly six to seven times a day [128]. Training the lung muscles can provide significant benefits for patients with respiratory diseases. Regular short shallow breaths can lead to frailties in chest muscle, reduce oxygen circulation, induce shortness of breath and cause poor lung capacity. Proper breathing exercises can help to resolve the severity of these symptoms, increase muscle strength, and improve posture and mental ability [129]. Therefore a system was created to encourage patients to carry out their prescribed breathing exercises. The system was designed for use in both a clinical setting and also for home use where the patient can be remotely monitored.

3.2.3 Feature Extraction System

One key element of the system was that it had to be low-cost and compatible with existing computer systems found in hospitals or homes. Sensors embedded in clothing must be comfortable to wear. Therefore, this system consists of a textile sensor which tracks breathing patterns via an embedded arduino microcontroller that is comfortable to wear. The microcontroller was used to sample and transfer the user's breathing signal to a computer for signal processing and analysis. The Digi Xbee wireless radio frequency module was used to sample data from the sensor. This allowed data to be transferred to a PC while granting unobtrusive monitoring of the lung muscle. The user interacts with a Graphical User Interface (GUI), which gives real-time feedback of the breathing technique and facilitates continuous assessment of the patient's performance. This system can be seen in Figure 8.

The textile-based sensor was developed using a piezoresisistive material specifically created to detect body movements [130]. As the material is stretched, its resistance increases. This effect can be used to detect joint movement and also to measure breathing rates. Breathing rate may be measured by monitoring the expansion and contraction of the ribcage. Placement of the sensors at the chest and abdomen allows patterns of breathing to be monitored, e.g. shallow breathing versus deep abdominal breathing. The software allows the patient and doctor/therapist to

Figure 8: Wireless Breathing Monitoring T-Shirt

perform a breathing exercise for a specified length of time which the patient must attempt to emulate. This respiratory signal is recorded as a reference breathing signal. Once the reference signal has been recorded it is then used to instruct the user for future exercises. Instruction is given to the user by means of an avatar, whose mouth expands and contracts in time with the breathing sequence. The user's breathing signal is compared in real-time to the reference signal and immediate feedback is given to the user.

3.2.4 Filtering

Piezoresisistive sensors can capture an abundance of data and generally the data rate is limited by the microprocessor. However in order to procure a clean signal which gives an accurate representation of the activity being performed, some signal pre-processing steps are required. In this work filtering is employed to remove the noise emanating from the wireless sensor. Filtering high frequency noise from the breathing

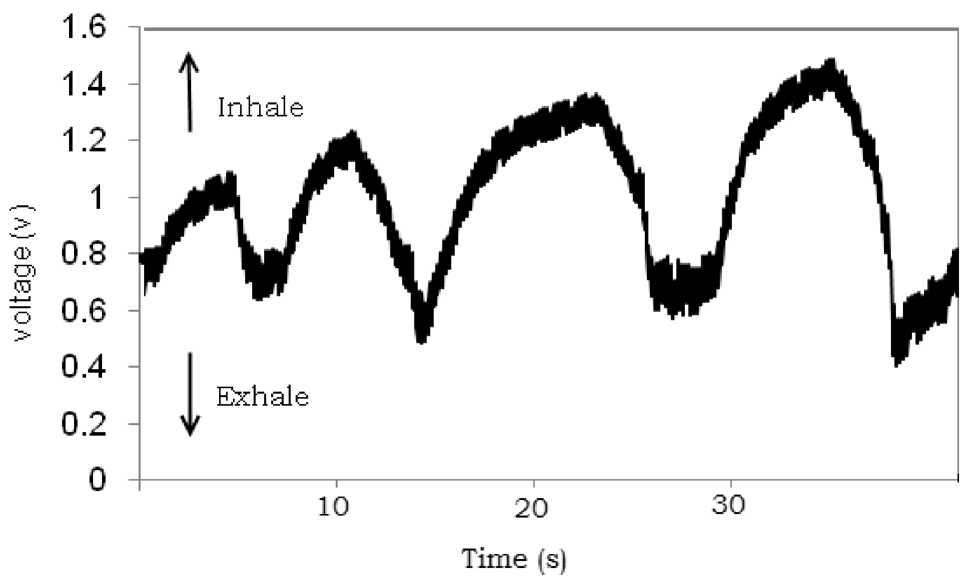

Figure 9: Unfiltered Breathing Signal

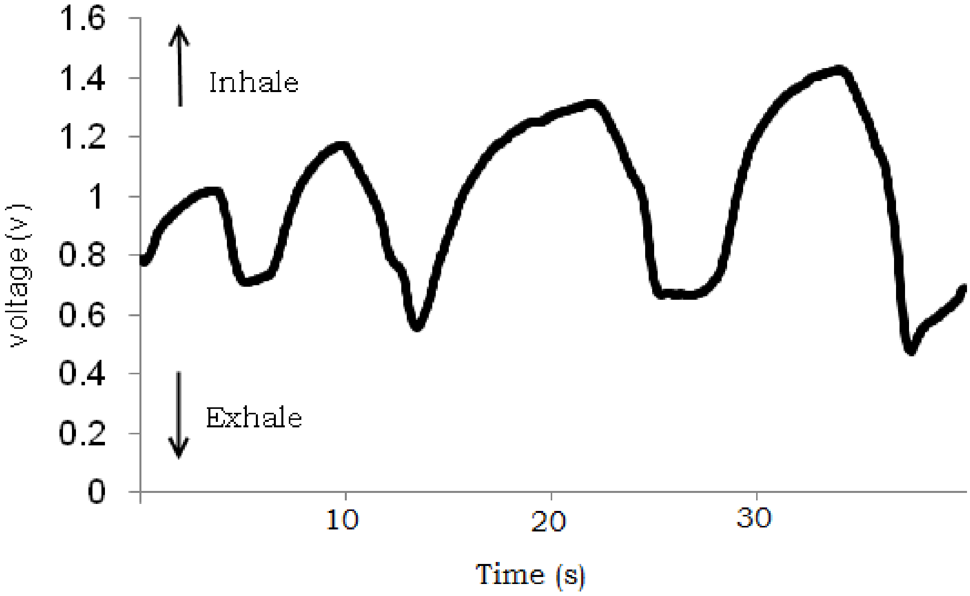

Figure 10: Filtered Breathing Signal

Figure 11: Reference Signal Recording with Real-time graphing.

sensor signal is a vital step as the subject most emulate the reference signal and a noisy signal would be difficult to follow. In this work a digital filter is employed to achieve this[1].

The signal is filtered with a 3rd order digital low-pass Butterworth filter with a cut off at 1Hz. This filter was found to adequately reduce the noise emanating from the breathing sensor. The raw data signal as shown in Figure 9 has been filtered and is displayed in Figure 10. Filtering the data ensures a smooth transition of the avatar's feedback allowing the user to emulate it more easily. Figure 11 shows the real-time graphing on the patients respiration signal in the lower right hand corner. Figure 12 shows an avatar with two different coloured mouths. The cyan mouth represents the patients current respiration signal while the outer darker mouth represents the respiration that the subject wishes to emulate. Each mouth expand and contract based on the respiration signals inputted.

[1]http://www-users.cs.york.ac.uk/ fisher/mkfilter/

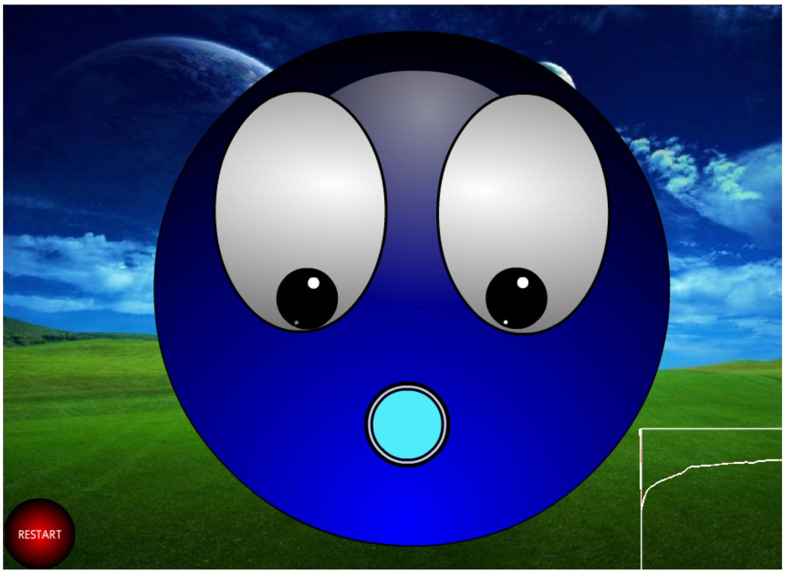

Figure 12: Graphical User interface - User attempting to emulate reference signal

3.2.5 Grade Classification

When the patient has completed the exercise, they are presented with a number of stars depending on their performance. In order to calculate the number of stars a user receives, a measurement of how similar the signals were would need to be determined. One such way to ascertain this was to calculate the correlation coefficient. Equation 13 was used to correlate the reference signal and the emulated signal.

$$R = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{N}}{\sqrt{(\sum x_i^2 - \frac{(\sum x_i)^2}{N})(\sum y_i^2 - \frac{(\sum y_i)^2}{N})}} \tag{13}$$

Where R is the normalised cross correlation coefficient, x is the reference signal, y is the emulated signal and N is the number of samples. R values are most sensitive to similarities and discrepancies in shape. It is also very sensitive to timing and can be used to find correlations between signals with inherent delay. Figure 13 shows three different breathing signals recorded by

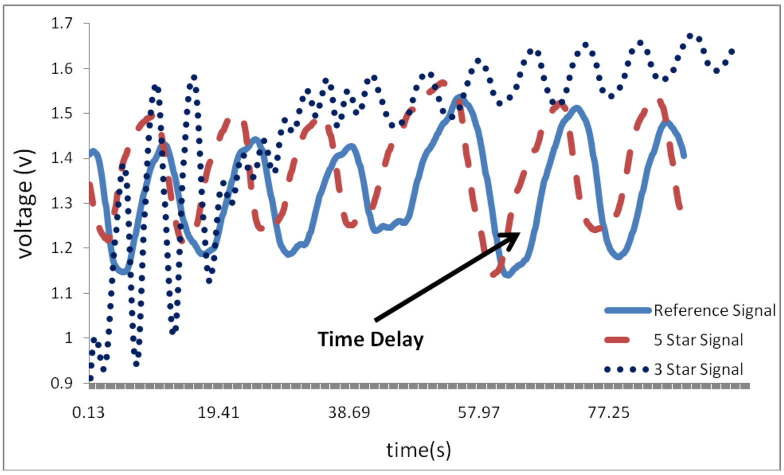

Figure 13: Comparisons of two signal classifications

the system. The solid (light blue) line is the reference signal which consisted of seven slow deep breaths over the course of 90 seconds. Two emulated signals were compared to it. The dashed (red) line represents a trial where the user emulated the avatars breathing very accurately. This performance was given a five star grading. The dotted (dark blue) line represents a trial where the user performed short shallow breaths that did not synchronise well with the avatar's breathing. The amplitude change of this signal is smaller than the five star signal as the sensor is not stretched as much during shallow breathing. This performance was given a three star grading. The grading system was designed to be supportive so that it would be encouraging and provide positive feedback to the user for carrying out the exercise in the first instance.

From Figure 13 it can be seen that there is a slight delay between the reference signal and the user's performance of the task. Therefore the program first needed to calculate this time delay before grading the user. The correlation function was calculated for the reference and emulated signals. The system implements a sliding window to identify the time lag between the signals by finding the index of the maximum correlation coefficient. For the reference signal and five star trial, the maximum correlation coefficient was

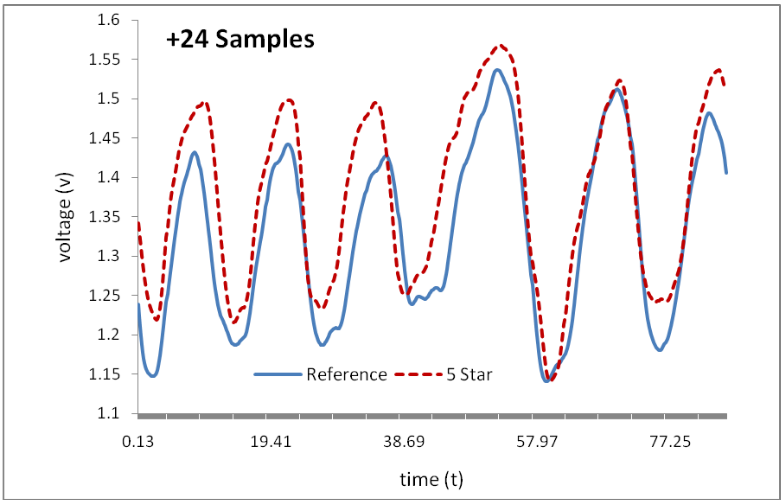

Figure 14: Plot of Reference and 5 Star Signal once time delay has been removed using correlation coefficient

found to be 0.9242 at a time lag of 24 samples (3.08 sec). There were two sources to this delay. Firstly there is a delay between the user seeing and performing the breathing action. Secondly the sensor data captured had to be inputted into the flash application via simulated keypresses which has a low maximum data throughput. Figure 14 show the maximum correlation value achieved during the sliding window.

Various tests were conducted in order to ascertain the grading of user signals based on the normalised cross correlation coefficient R. Three grades were created based on star-grading, with 5 stars corresponding to full marks and 3 stars corresponding to the poorest grade. From Figure 13 and Figure 14 it can be seen that the 5 star signal closely resembles the reference signal. With this data and similar data from other tests it was empirically decided that users who received over 0.85 for R deserved 5 stars. The other grades were calculated in the same way. From Figure 13 it can be seen that there is little resemblance between the 3 star signal and the reference signal. Even with signal time shift the maximum cross correlation coefficient was low. Any user that receives a coefficient of less than 0.7 receives 3 stars while if they

Grade Classifications	R value
5 star	$R \geq 0.85$
4 star	$R \geq 0.70$ & $R < 0.85$
3 star	$R < 0.70$

Table 1: Breathing Monitor Grade Classifications

receive a score between 0.7 and 0.85 a 4 star grade is awarded. Three stars were assigned to the poorest grade in order to encourage the user to keep continuing with the breathing training. The primary target population for this feedback system is children, and it is extremely important to keep them well motivated and focused on improving their technique. Table 1 shows the R values required for grade classification. The values were chosen after extensive testing on a single subject.

3.2.6 Conclusion

In this section a classification framework is introduced that classifies a subject's ability to perform a desired action. Creating a breathing feedback system motivates patients to train their lungs sufficiently which improves life expectancy and quality of life. The advantage of this system is that it is low-cost and the sensor garment is flexible and comfortable to wear. This low-cost sensor could be mass-produced by screen-printing processes commonly used in the textile industry. The entire system is very easy to deploy as it only consists of the wireless sensor and a small software package. Creating a framework that is able to accurately ascertain the competency of a subjects performance has a wide range of applications in health and sport. In this work one such application is explored but this approach could be easily applied to various problems across the healthcare and elite performances spheres. The state machine diagram for this framework can be seen in Figure 15 and it shows how this framework could be applied to other performance grading applications.

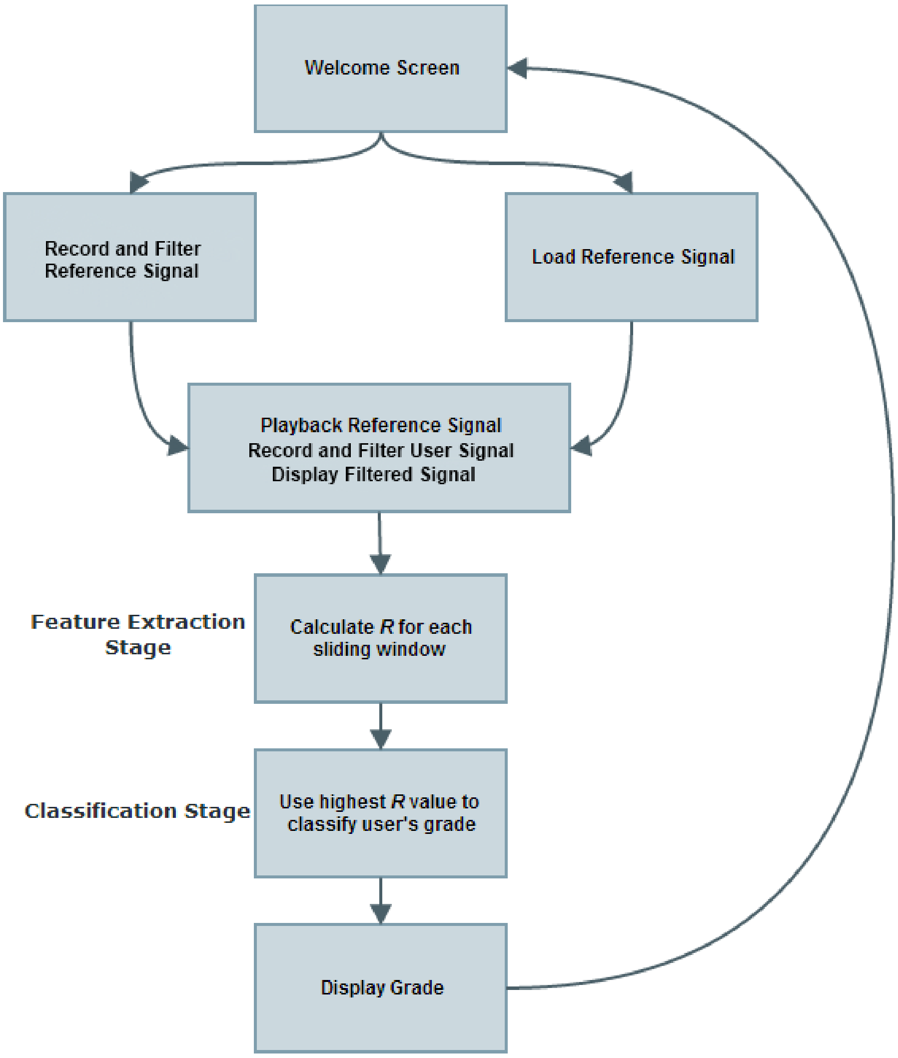

Figure 15: State Machine Diagram of Breathing Feedback System

3.3 Recognising Specific Activities

3.3.1 Data Capturing System

Smartphone usage has grown dramatically since their introduction over a decade ago. Over 50% of adults in the United States and over 40% of adults in Europe own a smartphone [131]. By 2016 it is expected that there will be one billion smartphone owners worldwide [132]. A smartphone is a mobile phone with a purpose built mobile operating system with advanced computing ability and inter connectivity compared to a standard mobile phone. Smartphones have more advanced Application Programming Interfaces (APIs) for running third party applications. They also contain technology which standard phones lack, such as portable media players, digital cameras, GPS navigation systems and modern web browsers. One key feature provided by smartphones relevant to this work, is access to embedded sensors, such as gyroscopes, magnetometers and accelerometers.

Most approaches in human activity recognition have relied on multiple expensive sensors. With the increase in smartphone ownership there has been more research conducted utilizing the sensors embedded within smartphones. Human activity recognition using smartphones have been employed to support patient monitoring [133], to identify the user's current mobility [134] and for monitoring daily activities [135]. However in this work we show how smartphones can be used to recognize human activity in sport.

In this work, the embedded accelerometer within a smartphone is used Whilst there is a large amount of literature for activity recognition in general, it is quite limited for classifying sporting activities. Most of this literature uses custom, albeit commercially available, sensors requiring athletes to purchase these sensors such as the miCoach and the Nike+. However over 40% of adults in Europe alone has a smartphone in their pocket. Performing classification using the smartphone potentially makes the technology available to everyone at all levels without additional hardware bar a cheap vest.

Whether or not player monitoring technology is allowed in competition varies from sport to sport, however, both low-cost solutions, e.g. miCoach or Nike+, and high-end offerings, e.g GPSports, are used widely at all levels in training sessions and competition (when allowed). The GPSports device contains sensors that allow the recording of speed, distance, heart rate and acceleration. The device contains an advanced accelerometer that records at a rate of 100Hz. This technology is used by some of the world's top sporting teams including Barcelona F.C, Real Madrid F.C and Liverpool F.C however its retail price restricts its widespread adoption for non-elite athletes. However, the level of automatic data analysis provided for understanding player activity is quite limited. The technology proposed here can be considered to be a low-cost solution that provides finer grained information about player's activity based on an automatic classification framework.

Athletes can take advantage of this technology to judge their overall match and training participation, physiotherapists could be notified of potential injuries and coaches could factor this information into their team selection. In sports where the wearing of sensors is forbidden during competitive matches, this technology can still be used in training environments to assess an athlete's performance. In the work presented here the sample rate was set to a low value that current smartphones can easily accommodate (16-22 Hz) when logging raw accelerometer signals.

The placement of a sensor on a subject is a decision that requires careful consideration. Incorrect placement of the sensor can result in inadequate data being captured or can result in the subject becoming uncomfortable which may affect their ability to perform activities correctly. In this work sensor placement is given considerable consideration in order to capture suitable data while not impeding or endangering the user. Also this location is the least impeding position i.e the location that interferes with the athlete the

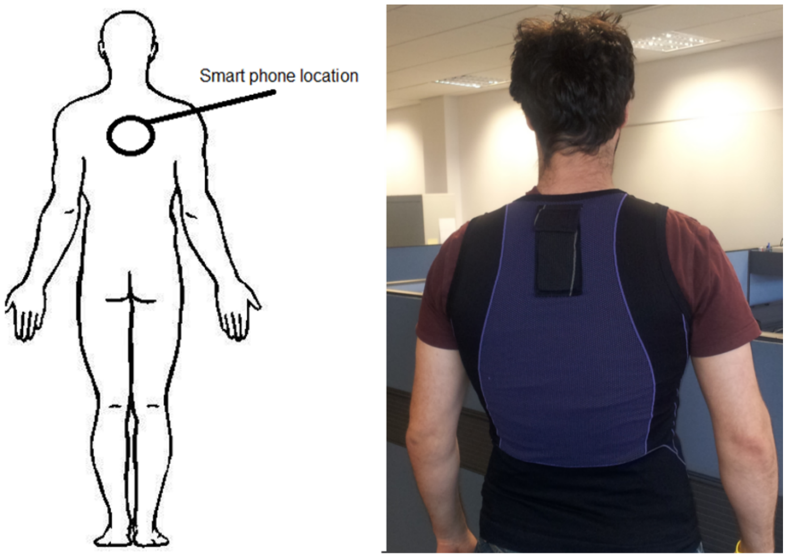

Figure 16: Location of Smartphone.

least. Therefore a single a single sensor is attached to a subjects upper back crevice. Current research suggests that placing a sensor on the upper back has produced excellent results as well has reducing the interference to the subject to an appropriate level and also reducing the chance of injury to the subject from the device [136].

3.3.2 Targeted Activities and Experimental Methodology

Accelerometer data was captured from two different field sports, five-a-side soccer and field hockey. Hockey players regularly change their back position when performing field hockey activities. For this reason field hockey represents a moderately difficult scenario whereas five-a-side soccer was chosen as it was envisaged that it would present significant difficulty in attempting to recognize activities and would provide a solid stress test for the technology. This was due to the smartphone being placed upon the upper area of a user's back as shown in Figure 16. Players wearing the vest reported that wearing it did not affect their performance due to the placement of the smartphone and the light weight nature of the vest.

In all forms of soccer the primary appendage used are the feet therefore deciphering actions executed by the feet from the upper back presents a difficult challenge. Consequently, achieving accurate results with five-a-side soccer is an ideal challenge for classification approach. Seven different sporting activities common to both five-a-side soccer and field hockey are targeted for classification. In this context, an activity is defined as a quantifiable action preformed by the user that is deemed significant. With this definition we identify the following activities:

A1 Subject is stationary

A2 Subject is walking

A3 Subject is jogging

A4 Subject is sprinting

A5 Subject is hitting the ball

A6 Subject is attempting a standing tackle

A7 Subject is dribbling the ball

It was concluded that these activities were comprehensive yet generic as they cover both inertial (A1-A4) and game (A5-A7) activities. Examples of these signals can be viewed in Appendix A. Table 2 displays the specification of the smartphones employed to collect our dataset. The intrusive size of the HTC Desire can be seen in Figure 17. From experimental observation we have found that the constant recording of the sensors on a smartphone of this nature for a period of 1 hour uses approximately 20% of the battery. When the dataset was being constructed these smartphones were at the more expensive range of the smartphone market. Cheaper smartphones with less advanced hardware would not be able to capture accelerometer data at a

Figure 17: HTC Desire Smartphone with a €2 coin for scale

Smartphones used		
	Google Nexus One	HTC Desire
Sampling Rate	16Hz	25Hz
Accelerometer	Tri-axial	Tri-axial
Resolution	8-bit	8-bit

Table 2: Smartphone specifications

high rate so we chose the standard rate for sampling so that in principle any smartphone could be used.

Five-a-side soccer data was recorded during five matches with each lasting one hour. From these five matches, the accelerometer data from 15 players was recorded. For field hockey, six matches were recorded with a total of 17 different players. Each match was video recorded with a Sony DCR-SR50 which allowed player activities to be accurately annotated by synchronising the video data with the accelerometer data. When annotating inertial activities, each annotator was required to determine the speed of the player from video data alone. Ideally the speed of the player would be available to the annotater as mislabling could occur otherwise. When logging an activity, nine seconds of data was collected, with the activity being placed in the centre of this window. This allowed us to experiment with different window sizes for feature extraction. Nine seconds was chosen as it was large enough for these sporting activities to be completed and small enough that it did not drastically increase computational time.

There were 30 instances of activities A5-A7 in the recorded matches and all of them were added to the dataset. As mentioned in section 2.4.3 there is an abundance of technology that can detect different inertial movements however there is very little work in the literature for identifying more complex movements such as dribbling with a ball. Adding every inertial movement annotation to the dataset would create a class imbalance. In [137] Chawla et al. states that “Learning algorithms that do not consider class-imbalance tend to be overwhelmed by the major class and ignore the minor one”. Therefore the number of instances of A1-A4 added to the dataset was limited to 30 to prevent this.

Datasets were created that contained 30 examples of each activity from both five-a-side soccer and field hockey data. These 30 examples were chosen randomly from activities logged from the matches recorded. These datasets contained activities from various players and allowed comparisons based on varying classification model parameters. In activity classification problems one important aspect is the changes in performance when different people perform the same activity. This inter-subject variability can have a distinct effect on classification accuracy. Each individual performs an activity differently due to their weight, height, sex and strength. In this study we captured data from a variety of players in order to get a realistic classification result. One limitation of including samples from every player in the training data is that the classification results do not give an indication on how the model respond to a new subjects data. However this was not judged to be an issue as each person's performance of an activity will at vary each time they execute it.

Subjects whose data was captured while playing soccer were amateur enthusiasts whereas subjects whose data was captured while playing field hockey were elite athletes. By capturing data in a naturalistic environment we reduce the possibility that a player's activities have been altered by the experiment. Ethical approval was granted for the data capture.

Normalisation

Across all humans, elite athletes and amateurs, there is a wide range of acceleration magnitudeS and therefore it is necessary to normalise the accelerometer signals to account for variances in actions performed by different subjects. A simple but powerful normalisation approach is utilized, which normalises the sensor signal so that the maximum is equal to 1 and the minimum is equal to 0. This approach is applied to accelerometer signals and aids the classification process as it helps to find similarities in different human subjects. Each action performed is then manually annotated and these annotations are required for extracting features and training a classifier. Equation 14 shows the process used

$$e_i^N = \frac{e_i - E_{min}}{E_{max} - E_{min}} \tag{14}$$

where E_{min} is the minimum accelerometer value, E_{max} is the maximum accelerometer value and e_i is the value to be normalised.

3.3.3 Approach and Results

Initially other popular classification approaches from the literature are applied on the collected datasets for benchmarking reasons. Three separate experiments are then conducted to discover the most accurate approach to creating the best classification framework. Finally a publicly available accelerometer based human activity recognition dataset is employed to compare the methods proposed here against other popular and successful classification approaches

Benchmarking

In [138], Kwapisz et al extracts forty-three time domain features from a smartphone accelerometer and employs them, along with a Multilayer Perceptron artificial neural network classifier (ANN) for activity recognition. These activities are walking, jogging, walking upstairs, walking downstairs, sitting and standing. They achieved an overall recognition accuracy of over 90%. Due to this high result we employed Kwapisz et al methods on our dataset. Their approach achieved an average accuracy rate of 73% for soccer and 79% for field hockey. It took 2 ms to compute the time domain features for a nine second data window. Table 3 explains in detail the time domain features extracted and fed into the ANN.

As mentioned earlier, the Fast Fast Fourier Transform (FFT) is a popular method for extracting informational features from a data signal. In [13] Preece et al. investigates various FFT feature extraction techniques using accelerometer data captured from subjects performing various daily activities such as walking, jogging, walking upstairs, walking downstairs, running, hopping on left leg, hopping on right leg and jumping. Preece et al. concluded that extracting the DC component and the magnitude of the first five components of FFT analysis produce the most accurate models. Therefore this technique was also applied to the soccer and field hockey datasets. It achieved an average accuracy rate of 78.1% for soccer and 78.5% for field hockey. It

Feature[No]	**Description**
Average [3]	Average acceleration (for each axis)
Standard Deviation [3]	Standard deviation (for each axis)
Average Absolute Difference [3]	Average absolute difference between the value of each of the 200 readings within the ED and the mean value over those 200 values (for each axis)
Average Resultant Acceleration [1]	Average of the square roots of the sum of the values of each axis squared $\sqrt{x_i^2 + y_i^2 + z_i^2}$ over the ED
Time Between Peaks [3]	Time in milliseconds between peaks in the sinusoidal waves associated with most activities (for each axis)
Binned Distribution [30]	The range of values is determined for each axis (maximum – minimum), divide this range into 10 equal sized bins, and then record what fraction of the 200 values fell within each of the bins.

Table 3: Benchmark Time Domain Features

Activity	A1	A2	A3	A4	A5	A6	A7
A1	28	2	0	0	0	0	0
A2	0	30	0	0	0	0	0
A3	0	1	29	0	0	0	0
A4	0	0	0	29	0	1	0
A5	0	0	0	0	22	3	5
A6	0	1	0	0	5	13	11
A7	0	0	0	0	9	8	13

Table 4: Confusion matrix for FFT benchmark for Soccer Smartphone data

took 25 ms for the FFT features to be extracted using MATLAB from a ten second data window. All computation duration tests in this work were completed on a Intel Core 2 Quad CPU Q9650 processor with 4 gigabytes of RAM.

Experiment 1

A black box experiment is using a system which is viewed solely based

Activity	A1	A2	A3	A4	A5	A6	A7
A1	29	1	0	0	0	0	0
A2	0	30	0	0	0	0	0
A3	0	1	29	0	0	0	0
A4	0	0	0	29	0	1	0
A5	0	0	0	0	23	3	4
A6	0	1	1	0	4	14	10
A7	0	0	1	0	8	9	12

Table 5: Confusion matrix for FFT benchmark for Hockey Smartphone data

on terms of input and output. In this scenario parameters are selected based on their popularity or their ease of use. SVMs are one of the most popular classifiers used in human activity problems as it is relatively simple to understand and fast. Therefore for this baseline experiment an SVM was used and the selected parameters can be seen in Table 6. Daubechies 4 wavelet "db4" is a popular motherwavelet choice in signal analysis problems due to its regularity and fast computational time. A level two DWT was chosen to keep computational time fast while still extracting discriminative features. A window length of five seconds was chosen as every activity had concluded by then. It took 10 ms for the DWT features to be extracted from the 5 second window. It took 4 ms for this approach to classify the extracted DWT features with the SVM.

For field hockey this experiment achieved a 65.9% F-Measure score while for soccer it achieved a 62.7% F-Measure score. Field hockey and soccer models both suffered from high mean absolute error, 21.13% and 21.41% respectively. Table 7 and 8 give the confusion matrix for this experiment. Both models identify inertial activities (A1 - A4) adequately, but perform poorly when trying to identify game activities (A5 - A7). This approach is the fastest to create and train, however this is outweighed by its relatively poor performance compared to other approaches.

Classifier	Mother-wavelet	DWT level	window size
SVM-SMO	db4	2	5 seconds

Table 6: Parameter specifications for black-box approach

Activity	A1	A2	A3	A4	A5	A6	A7
A1	30	0	0	0	0	0	0
A2	3	27	0	0	0	0	0
A3	0	0	30	0	0	0	0
A4	0	0	0	30	0	0	0
A5	0	15	3	0	7	4	1
A6	4	4	2	0	7	8	5
A7	3	4	4	0	7	4	8

Table 7: Confusion matrix for Soccer Smartphone data for Experiment 1

Experiment 2

In this experiment a diverse range of classifiers is investigated and the input parameters are varied to understand to what extent they influence the classification procedure. The classifier, DWT decomposition level, window length and motherwavelet are inspected to see how adjusting them affects the accuracy of the models the that framework produces.

In Figure 18 the average accuracy for each classifier investigated can be seen. Interestingly all classifiers perform similarly except for the SVM-SMO classifier during soccer. Further investigation showed that this classifier could

Activity	A1	A2	A3	A4	A5	A6	A7
A1	30	0	0	0	0	0	0
A2	3	27	0	0	0	0	0
A3	0	0	30	0	0	0	0
A4	0	0	0	30	0	0	0
A5	0	6	5	0	11	1	7
A6	0	1	0	0	16	1	12
A7	0	0	1	0	9	3	17

Table 8: Confusion matrix for Hockey Smartphone data using Experiment 1

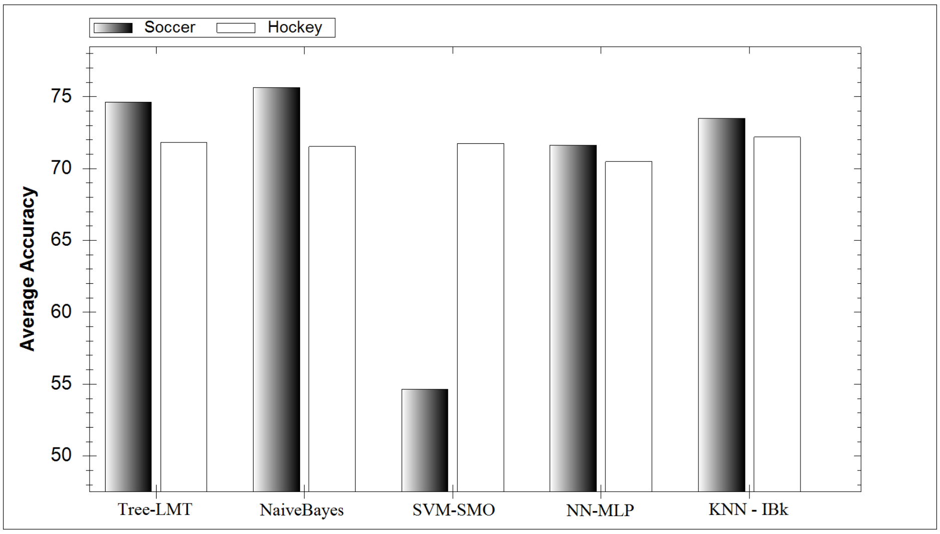

Figure 18: Average classifier family accuracy for experiment 2

not reliably distinguish between lower extremity game activities. SVM classifiers themselves have many parameters and therefore require tweaking to reach their full potential. During hockey activity classification the SVM-SMO performed well as the game activities were much more distinct. Interestingly the average soccer model outperforms its hockey counterpart, which was an unintuitive result. A reason for this could be that the range of parameters investigated favoured soccer classification. However the highest accuracy hockey models created performed better than the highest accuracy soccer models, which was intuitively expected.

In Figure 19 the overall average accuracy for each DWT level over all classifiers can be observed. It is interesting to note that there is an increase in average accuracy with every level increase during soccer while during hockey classification each level performs similarly well. As mentioned earlier five a side soccer was envisaged to be much more difficult to classify due to the position of the smartphone. Therefore it was concluded that retrieving more features with additional decomposition levels allows classifiers to rectify difficult to interpret data. DWT decomposition levels ranging from one through

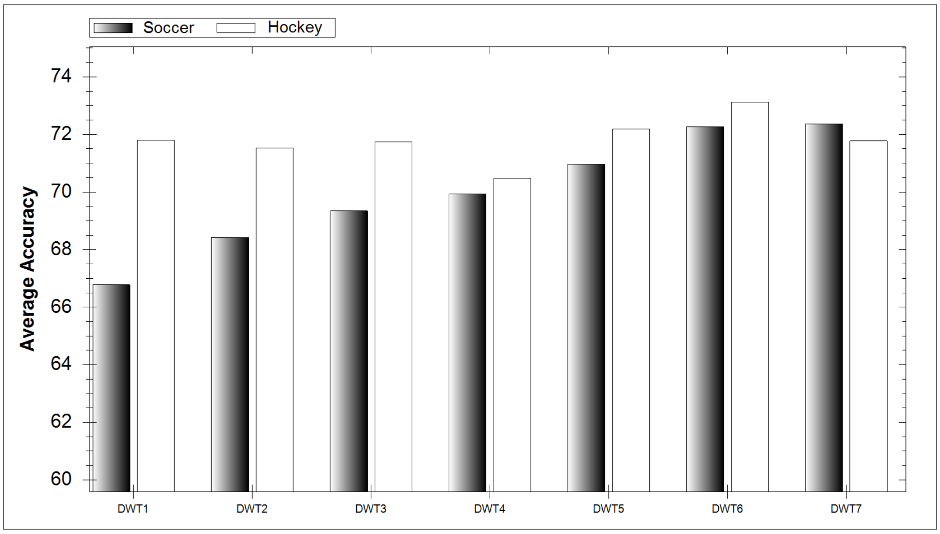

Figure 19: Effect of DWT Levels on classification accuracy

seven were investigated as further levels would increase computational time significantly.

In Figure 20 the accuracy average for each window length can be seen over all classifiers. For soccer the accuracy of the model decreases with an increase in window length. This makes sense as soccer activities have a shorter duration than their hockey counterparts. With shorter activities the longer the window the more activities can occur. If two or more activities occur in a window then classification difficulty is dramatically increased. When selecting a time window it is vital that it is long enough to contain the whole activity being performed and short enough that it does not include additional events.

In Figure 21 the average accuracy for each mother wavelet family can be seen. Each family performs well and no one family out performs the rest. This result reinforces the conclusion in the literature that it is almost impossible to prejudge what motherwavelet will perform well in an application. Similarly, results from individual wavelets show no discernible difference between their performance. However the mother wavelet itself is important due to its

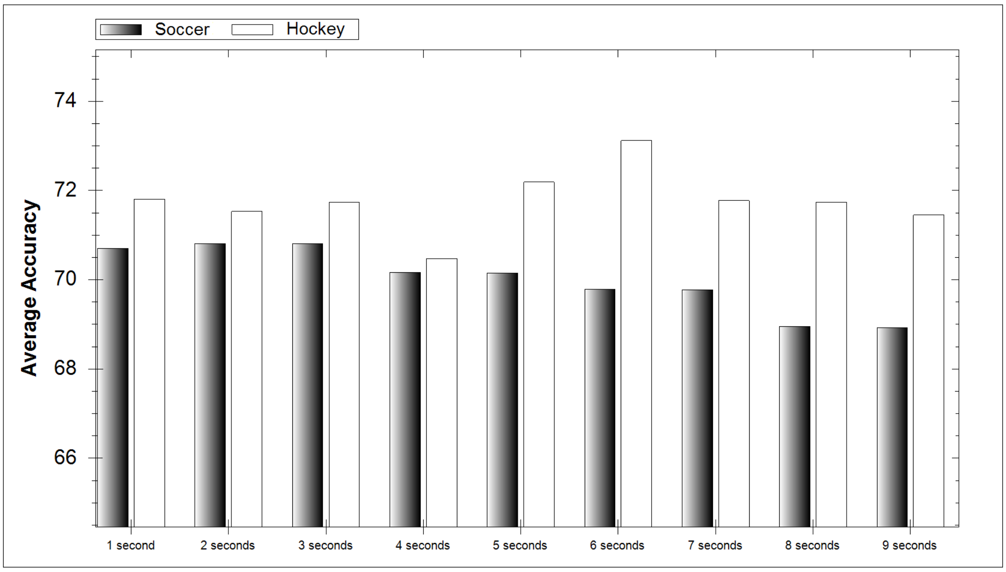

Figure 20: Effect of window length on average accuracy

Device	Sport	Classifier	DWT lvl	Mother W.	Length(sec)	F-Measure
Smartphone	Soccer	NaiveBayes	6	rbio1.1	3	0.799
Smartphone	Hockey	MLP	6	bior1.1	7	0.823

Table 9: Highest classification accuracies attained for Experiment 2

integral part in the DWT process.

As with Experiment 1 classifiers are very competent at identifying inertial movement activities (A1 - A4) however game activities (A5 - A7) pose more of a challenge. The inertial activities are very distinct as the energy during these activities is unique. They range from zero energy outputted when the player is stationary to maximum energy when the player is sprinting. The confusion encountered between the game activities is due to the similar motions being performed. In soccer these motions involve lower leg movement while in hockey these game activities involve the upper arm movement. Table 9 provides the parameters for the highest classification accuracy attained for each respective sport. Tables 10 and 11 display their confusion matrix data. It took 5 ms for this approach to classify the extracted DWT features.

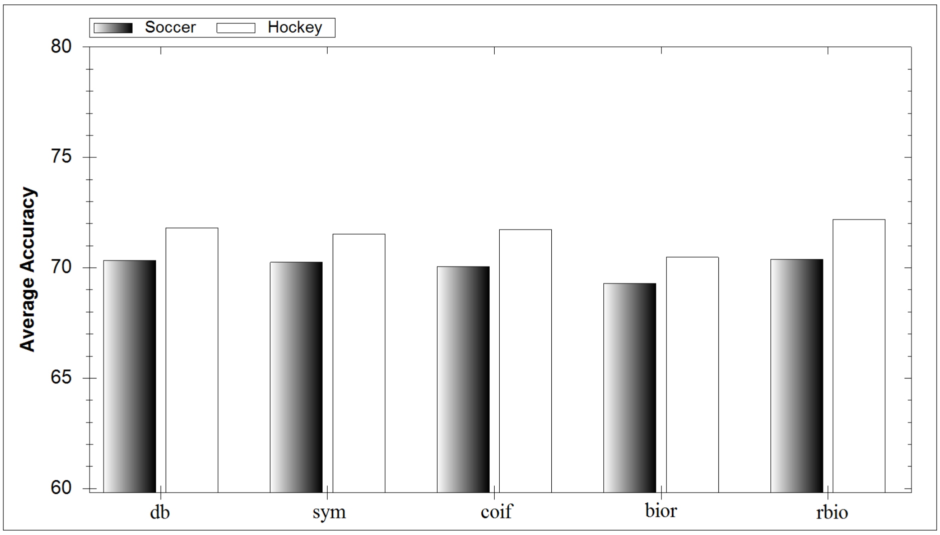

Figure 21: Effect of choice of wavelet

Activity	A1	A2	A3	A4	A5	A6	A7
A1	28	0	0	0	0	0	0
A2	0	30	0	0	0	0	0
A3	0	0	30	0	0	0	0
A4	0	0	0	30	0	0	0
A5	0	1	0	0	24	4	1
A6	0	2	0	0	9	12	7
A7	0	1	0	0	12	2	15

Table 10: Confusion matrix for Football Smartphone data for Experiment 2

Activity	A1	A2	A3	A4	A5	A6	A7
A1	30	0	0	0	0	0	0
A2	1	29	0	0	0	0	0
A3	0	0	30	0	0	0	0
A4	0	0	0	30	0	0	0
A5	0	0	0	0	19	7	4
A6	0	0	0	0	7	15	8
A7	0	0	0	0	4	6	20

Table 11: Confusion matrix for Hockey Smartphone data for Experiment 2

Experiment 3

In this experiment separate classification model are created for each activity and the results investigated. This allows the creation of a fusion of classifiers whereby the classifier result with highest confidence dictates the result. This late fusion method is described in more detail in section 2.5. The best performing classifiers from experiment two were used, NaiveBayes for Soccer and MLP for Hockey. The average F-measure score for soccer data yielded a result of 86.3%, while for hockey data it yielded 88.8%. Both results had low mean absolute error, 4.42% for football and 4.55% for hockey. Figure 22 compares the accuracy of all three experiments and includes the absolute mean error. Figure 23 compares the ability of each experiment to identify specific activities. It took 27 ms for this approach to classify the extracted DWT features. The increased computational time compared to experiment 2 is due to testing the extracted DWT features of a signal with each activity model.

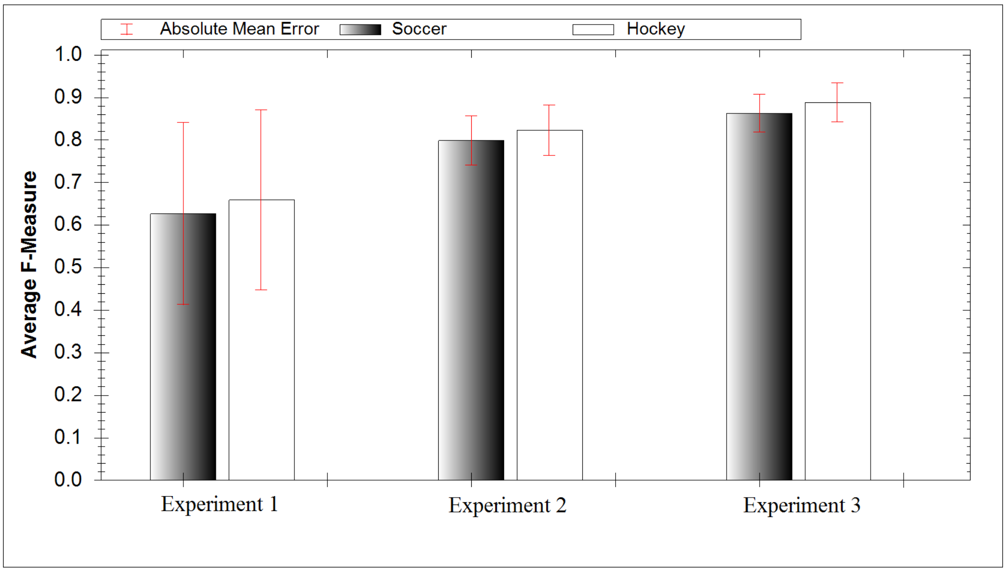

Figure 22: Average model accuracy for each experiment

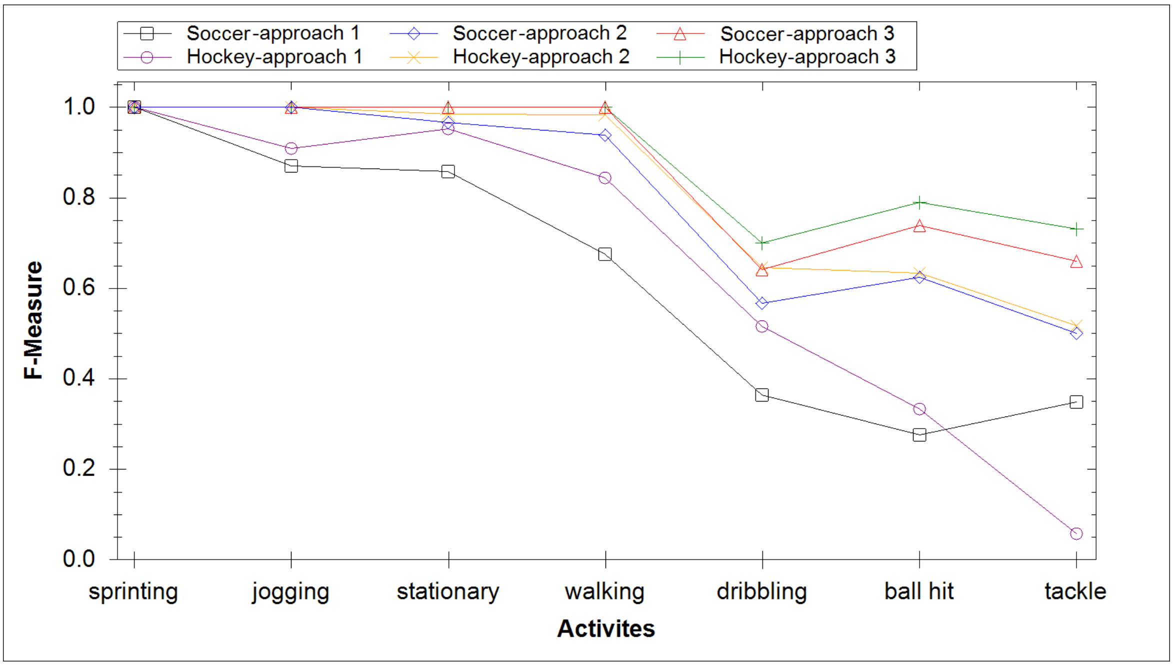

Figure 23: Single activity accuracy results for each approach

Experiment 4

In order to facilitate research of accelerometer-based human activity recognition, Xue et al[139] built a naturalistic 3D acceleration-based activity dataset, SCUT-NAA. It provides researchers in the field of acceleration-based activity recognition with a naturalistic activity dataset with training and testing samples. It also allows for comparing and evaluating performance of different algorithms. It was for this reason this dataset was chosen to apply the automatic human activity classification approach described above in Experiment 3. SCUT-NAA dataset is the first publicly available 3D acceleration-based activity dataset. It contains 1278 samples of 44 individuals (34 males and 10 females) which were collected in naturalistic settings with only one tri-axial accelerometer located in the pants pocket. Each sampling person is asked to perform ten activities. To collect activity data, they developed a sampling device comprising an accelerometer ADXL330, microprocessor ADuC7026,

Notation	Activities Captured
a	climbing downstairs
b	climbing upstairs
c	jump 45 seconds
d	relax
e	run 100m
f	step walking 45s
g	walking 50m
h	walking backward
i	walking quickly 50m
j	bicycling

Table 12: List of Classified Activities in SCUT-NAA

Bluetooth transceiver module, FLASH data storage module and keyboard module. The ADXL330 is a tri-axial accelerometer capable of sensing acceleration between minus and plus 3g with tolerance within 10%. The output signal of the accelerometer was sampled at 100 Hz[139].

In [139] Xue et al performs four different feature extraction techniques on the SCUT-NAA dataset. They are the Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), Time Domain (TD) features and Autoregressive (AR) processing. The FFT approach achieved an average accuracy of 84%, the DCT 83%, the TD 47% and the AR 48%. The real-time approach presented here on the smartphone achieved an accuracy of 78% which is comparable to the offline FFT and DCT methods. There is little comparable research of using advanced algorithms on smartphone devices in real-time as it is only recently that smartphones processing power has developed sufficiently to employ them. Table 13 shows an in depth look at the accuracy levels that the fusion of classifiers method obtained for each activity performed. This approach was more successful than any of the four methods proposed by the authors in [139].

Notation	Precision	Recall	F-Measure	Class
a	0.889	0.821	0.855	climbing downstairs
b	0.648	0.686	0.667	climbing upstairs
c	0.931	0.942	0.936	jump 45 seconds
d	0.988	0.965	0.976	relax
e	0.885	0.895	0.89	run 100m
f	0.728	0.872	0.794	step walking 45s
g	0.635	0.628	0.632	walking 50m
h	0.641	0.477	0.547	walking backward
i	0.878	0.802	0.84	walking quickly 50m
j	0.944	0.85	0.895	bicycling
Weighted Avg	0.884	0.842	0.863	

Table 13: Classification Results

3.3.4 Conclusion

In this section a framework is presented that allows for the automatic identification of sporting activities from a single smartphone worn on the upper body. Discriminative informational features were extracted from smartphone accelerometer signals using the Discrete Wavelet Transform (DWT) decomposition. These features were very informative as they were able to reduce accelerometer signals to a much less complex input. For example with the first model in Table 9 we were able to reduce the accelerometer segment from 75 samples into 42 descriptive features. Training and classifying activities would take much longer and be prone to over-fitting without reducing the signal to a small set of features. One disadvantage of using the DWT when extracting features from signals is that is there is no widely accepted method of picking the most suitable motherwavelet for a particular application. We investigated five prominent motherwavelet classes Daubechies, Coiflets, Symlets, Biorthogonal and reverse Biorthogonal. Daubechies provided the motherwavlet for half of the best accuracy models however the overall results show that performance differences between motherwavelets are not very significant.

No one classifier family has to date been shown to have a direct advantage in activity classification problems so classifiers from different families were examined. Investigating different window lengths, DWT levels, classifiers and motherwavelets allowed us to create classification models that achieved F-measures of 86.3% and 88.8% for five a side soccer and field hockey respectively. This DWT extraction process is much more accurate compared to state of the art time domain and FFT feature extraction methods mentioned earlier. Figure 18 illustrates the average performance for each of the classifiers investigated. None of the classifiers parameters were explored as this would of increased the overall feature investigation time exponentially.

Overall the worst performing classifier was the support vector machine (SVMs). Hsu et al. in [140] explains that the main disadvantages of SVMs is that they require several steps before an acceptable classification result can be achieved. The first step he mentions is to conduct simple scaling on the extracted features. This eliminates greater numeric ranges from dominating smaller numeric ranges. The second step requires testing different kernel parameters. The SVMs kernel is responsible for analysing patterns and creating relations in the dataset. In general, the RBF kernel is a reasonable first choice [140] and was the kernel chosen in this work. However there are other popular kernels such as the polynomial and sigmoid kernels. Both polynomial and sigmoid kernels have had success classifying activities using smartphones. Anguita et al. in [141] achieved an overall accuracy of 89.3% using a SVM with a sigmoid kernel detecting six everyday activities such as walking upstairs and downstairs. Similarly Fleury et al. in [142] achieved an overall accuracy of 75.9% recognising complex everyday activities such as dressing oneself and using the toilet with a polynomial kernel. Both these kernels have independent variables that can heavily influence the performance of the SVM. Hsu recommends searching through these variables in order to create the best possible classifier.

In experiment two the highest soccer accuracy attained using a SVM was

76.33% and it was found using a bior3.3 mother wavelet, a DWT decomposition level of 6 and a 3 second window. By employing the best performing parameters mentioned on the mother wavelet, DWT decomposition level and window size, the overall accuracy was improved from 76.33% to 82.7% which makes it the new best performing classifier. The difference between the default SVM and the new improved SVM can be seen in Table 14.

	Kernel	**c**	**Scaled features**	**F-Measure**
Default SVM	RBF	1	no	76.33%
Explored SVM	polynomial	10	yes	82.7%

Table 14: Improved SVM Results

Decision tree classifiers are immune to scaling problems therefore large ranges of numeric data do not dominate smaller ranges. This can explain why on average decision trees performed well whereas the SVM did not. Bayesian classifiers such as NaiveBayes perform well when its assumption that all inputted features are independent is true [143]. Since all extracted features are based of the X,Y and Z axis of an accelerometer this assumption holds true and can explain why the NaiveBayes classifier performs very well overall in this work. The neural network multilayer perceptron (MLP) achieved success due to its ability to automatically adjust its own parameters using back-propagation. Therefore since the MLP automatically optimizes itself its high overall average accuracy is explained. The KNN algorithm is relatively simple compared to the other classification algorithms mentioned and therefore has very little adjustable parameters. Due to its capacity to handle outlying data points it is more robust when working with smaller datasets which prevents overfitting. This could explain why the KNN-IBk classifier also had a high overall average accuracy.

In this work the effect of changing several of the DWT input parameters are investigated, including motherwavelets, window lengths and DWT decomposition levels. During the course of this work we created a unique sports activity analysis dataset, comprised of five a side soccer and field hockey activities. All experimental results presented in this section are based on this dataset. The average maximum F-measure accuracy of 87% was achieved using a fusion of classifiers which was 6% better than a single classifier model and 23% better than a standard SVM approach. However this relatively modest 6% increase comes with significant increase in computation.

Most approaches in human activity recognition rely on multiple expensive sensors. With the increase in smartphone ownership there has been more research conducted utilizing the sensors embedded within smartphones. Human activity recognition using smartphones have been employed to support patient monitoring [133], to identify the users current mobility [134] and for monitoring daily activities [135]. In this work we have shown that smartphones can be used to recognize human activity in sport.

Performing classification using data gathered by a smartphone potentially makes the technology available to everyone at all levels without additional hardware bar a cheap vest. Currently all processing is performed offline after data gathering. If real time processing is required then the preferred solution would be a continuous connection to a server rather than performing the analysis on the smartphone itself.

The approach proposed here for human sporting activity classification can be applied to other human motion activity problems. Experiment 4 shows that movements that humans perform daily are able to be accurately recognised using a fusion of classifiers. Now that the framework has been set up the key problem when creating classification models is acquiring sufficient training data. Additionally this method is not confined to an offline setup especially with smart phones which posses the ability to communicate over the web. The smartphone also has many other embedded sensors that could be used to capture physiological information. Future work will focus on investigating this and also comparing other feature extraction methods to the

DWT. The investigation of other feature dimensionality reduction techniques such as principal component analysis is another area that warrents future research. Furthermore there are other sensors which have grown in popularity such as the miCoach by Adidas. It would be interesting to investigate their performance compared to smartphones. Planned future work will look to examine and to compare accuracy of the current major available commercial devices.

4 MultiModal Human Action Recognition

4.1 Introduction

This chapter introduces the methods used to fuse data from a collection of sensors. This sensor fusion approach will allow contextual information from diverse sensors to be combined to allow the automatic classification of human activity. In the next section the literature concerned with the fusion of sensors of the same type is explored first before literature concerned with fusion sensors that detect different physiological data is explored. Thereafter an experiment is conducted to investigate whether adding sensors to the human body can aid in automatic activity classification. Methodologies and results are presented on whether early or late fusion is more beneficial for the specific application of classifying sleep apnea events.

4.2 Related Work

In [144], Zhu et al. uses two inertial sensors in their classification system. The first sensor was placed on the subject's waist and the other sensor on the foot. Each inertial measurement unit compromised of an accelerometer and magnetometer. Initially the data from the two inertial sensors are fused for coarse-grained classification. This allows the classification of the of activity performed into one of the three groups; zero displacement activity (standing or sitting), transitional activity (sitting to standing, standing to sitting), or strong displacement activity (walking upstairs, walking downstairs). Following this step a fine-grained classification is performed on each of the three groups. This allows classification between activities within the same groups e.g sitting to standing versus standing to sitting. The first classification step uses a neural network to identify the group whereas the second step uses a Hidden Markov Model (HMM) to further distinguish the activity. They achieved an overall classification accuracy of 89.75%. Aminian et al. use a different method to fuse inertial sensor data [8]. They attached one inertial sensor to a subject's chest and a second sensor was attached to the rear leg. They then fused the acceleration samples from the two sensors. By fusing the data they were able to recognise activities such as sitting, lying, standing and dynamic activities such as walking. Before fusing the data they were unable to accurately distinguish each activity.

In [145], Lawrence et al. use a MicroLEAP2(ULEAP2) system that would aid in automatically recognizing rehabilitation activities performed by people suffering from chronic obstructive pulmonary disease. This second generation of the energy-aware wearable sensor system incorporates various sensors for physiological monitoring[146]. It contains an accelerometer and a gyroscope to capture motion data. It also includes a piezoelectric belt that generates a small voltage in response to mechanical stretching in order to measure respiration. Finally it houses a heart rate sensor that locally computes the user's heart rate. The N-point(FFT) was used as all activities to be identified exhibited a certain periodicity. The FFT extracted the F_{peak} and F_{energy} from the sensor signals every second. Each sensor modality was 16-bit and acquired at 128 Hz. HMMs were used to classify the features extracted by the FFT. They were able to accurately identify activities like resting, taking the stairs, walking and running 85% of the time.

Early fusion is a fusion scheme which merges the features of each modality before any machine learning is conducted. In [147], extracted features from each sensor are concatenated to create a fused multimedia representation of visual, textual and audio sensors. After these multimodal features are concatenated a supervised learning approach is adopted to classify the semantic concepts. The main advantage of early fusion is that the concatenated vector is a true representation of all classes and also that only one machine learning stage is required to classify the sensors.

Snoek et al. in [147] define late fusion as a scheme which initially learns the concept scores from individual unimodal features and afterwards these results are fused to learn concepts. Late fusion approaches have also been utilized with much success in the field of multimodal analysis. The authors in [148] use individual probabilistic models to classify text and video using a late fusion scheme. In their approach, the text model is based on the language modelling approach to text retrieval and the visual information is modelled as a mixture of Gaussian densities. The scores are linked after individual classification to give the final classification accuracy score. This means that with the late fusion approach it is possible to ascertain the classification accuracy of each individual sensor since eachs data is untouched before the first machine learning stage. However as mentioned in [147], late fusion is expensive with regards to learning effort and computational time, as a classification stage is required for each unique modality. In addition to this, after the learned concepts are merged, some forms of late fusion have an additional learning step before final fusion prediction is obtained.

4.3 Experiments with a single type of sensor

In the first experiment conducted in this chapter the goal is to investigate the potential of combining raw data from the same type of sensor but attaching them to different locations on the body. Two wireless/wearable inertial motion units (WIMUs) were placed on various subjects of different athletic ability. The ability of using two separate sensors to distinguish different training movements and activities was compared to just using a single sensor. A dataset was created for a real world application which required two WIMUs to gather the required data. The methods and results for both approaches is presented in the next section.

4.3.1 Target & Application

Sport and physical activity have important cardiovascular, musculoskeletal and mental health benefits [149] and are enjoyed by large numbers. However, associated lower body musculoskeletal injuries are very common [150], [151], [152]. Almost all injuries are caused by relative excessive loading on the tissues i.e. high loading relative to tissue strength. One factor that significantly influences this loading is movement technique. Athletes can be biomechanically screened to determine an athlete's predisposition for injury [153] by recording and quantifying both their movement technique (i.e. joint angle and angular velocity) and some measure[2] of loading on their lower body during a series of actions common to their sport and known to be related to injury (e.g. running [151], jumping and landing [154], agility cuts [155]).

Generally, the athlete completes 3 - 5 maximum effort trials of each action [154] and their results are compared to normative values, if available [156]. These tests are almost exclusively completed in a laboratory since biomechanics based motion analysis systems tend to be camera based (6+ cameras typically) which must remain spatially fixed during the testing session and tend to be negatively affected by changing lighting conditions. This screening process creates several assessment and comparison challenges, which significantly reduce its ecological validity and usefulness.

A solution to the above assessment challenges would be to use sensors that could be worn throughout a training session or competitive event, detecting an athlete's joint angular motion and impact accelerations. With the use of these sensors more opportunities are presented to make the process of tracking the change in a subject's athletic ability for training and health purposes.

With the recent development of more accurate and relatively cheap WIMUs, combined with improved algorithms to more accurately determine sensor ori-

[2]Direct loading on individual tissues cannot be measured in a non-invasive fashion but this is possible for aggregate loading on a region of tissues or structures.

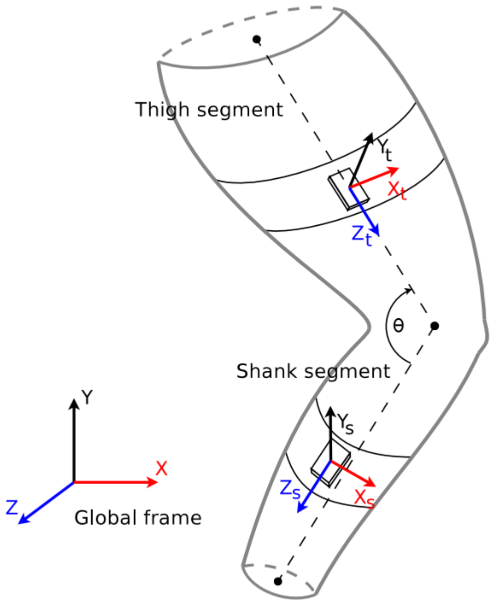

Figure 24: Placement of two inertial sensor units on the thigh and shank as well as their local coordinate system in a global coordinate system is illustrated.

entation [157], [158] , it has become feasible to deploy wearable body sensor networks in training sessions.

Automatic activity classification is used to identify different training activities as this would allow training sessions to be more quickly evaluated by sporting and health professionals. It would allow them to quickly segment an athlete's training session by activity and thus allow the desired data to be more easily located. This approach also facilitates the creation of a database containing the evolution of an athlete's movements within and across training sessions.

4.3.2 Data Collection

To evaluate the proposed framework, recordings of ten subject whose actions were captured using four wearable inertial sensors. The injured subject was experiencing lower back pain that did not WIMUs were placed on the left/right shank and left/right thigh of a subject as shown in Figure 24. The location of the sensor on each body segment was chosen to avoid large muscles; as soft tissue deformations due to muscle contractions and foot-ground impacts may negatively affect the accuracy of joint orientation estimates.

The sensors were affixed to the subject with double sided tape and velcro straps with some elasticity in the fabric, so as not to restrict the subject's movement and performance in any way. Next, the subject was asked to perform a series of actions as they normally do during outdoor training sessions. Each subject performed a predefined exercise routine on a large outdoor grass soccer pitch. The exercise routine consisted of the following motions: agility cuts, walking, sprinting, jogging, box jumps and football free kicks. Each motion lasted approximately 60 seconds for a total of approximately 9 - 10 minutes for the entire session.

The data from each sensor was recorded to an internal SD card on board the device. As each sensor recorded data independently, a physical event was required to synchronize all devices together. This was achieved by instructing each subject to perform five vertical jumps, ensuring large acceleration spikes would occur simultaneously on each device, that would be clearly visible in the accelerometer stream. In a post processing step, peak alignment was automatically performed and all data streams were cropped to two seconds before the first vertical jump landing. Video footage of each data capture session was also recorded and annotated, to be used as ground truth for the automatic segmentation and recognition of movements categories (i.e. jogging, agility cuts, sprinting etc.). Samples of the signals can be seen in Appendix B. Ethical approval was granted for the data capture.

4.3.3 Methodology and Results

In order to develop an approach to activity classification the exercise routine performed by each athlete was segmented and annotated for all activities and used to create a training set. The acceleration data from the two WIMUs

Activity	a	b	c	d	e	f
a = Agility Cut	166	0	6	4	3	1
b = Walking	0	399	0	0	0	0
c = Jumping on box	4	2	17	3	1	2
d = Jogging	0	0	0	205	0	0
e = Sprinting	1	0	0	0	27	0
f = Ball Kicking	2	4	5	5	3	68

Table 15: Confusion Matrix for the classifier using a single sensor on the shank

was isolated and features extracted for classification purposes. An early fusion approach was adopted to fuse the two feature vectors from each accelerometer. More detail about different fusion schemes is described in the next section. A window length of three seconds was chosen as this was sufficient time for each of the selected training activities to be completed. The DWT was used with much success in extracting discriminative features from accelerometer data in section 3.3.3 and thus was used to extract features for classification. The Daubechies 4 wavelet is a popular mother wavelet choice in signal analysis problems due to its regularity and fast computational time, and was chosen in this work.

The F-measure scores when using data solely from the leg shank sensor is presented in Table 18 and the confusion matrix associated is presented in Table 15. Similarly the scores from the thigh sensor is presented in Table 16 and its confusion matrix is in Table 19. Table 20 shows the F-measure scores when data from both the leg shank sensor and leg thigh sensor are both used to classify the activity being performed. All values in this experiment were computed using a ten-fold cross validation. Since the classifier was trained with classes which had different instance populations the F-measure scores are shown. The F-measure score gives a better indication of a model's ability to correctly identify an activity than standard classification accuracy alone.

Comparing Table 15, Table 16 and Table 17, it is possible to see the activities that require two sensors for accurate classification. Agility cut and

Activity	a	b	c	d	e	f
a = Agility Cut	176	0	2	0	0	2
b = Walking	0	399	0	0	0	0
c = Jumping on box	3	2	21	0	0	3
d = Jogging	0	0	0	205	0	0
e = Sprinting	0	0	0	0	28	0
f = Ball Kicking	4	3	5	4	1	70

Table 16: Confusion Matrix for the classifier using a single sensor on the thigh

Activity	a	b	c	d	e	f
a = Agility Cut	180	0	0	0	0	0
b = Walking	0	399	0	0	0	0
c = Jumping on box	0	0	27	2	0	0
d = Jogging	0	0	0	205	0	0
e = Sprinting	0	0	0	0	28	0
f = Ball Kicking	3	5	2	3	1	73

Table 17: Confusion Matrix for the classifier using two sensors

jumping on a box are much more complex activities than walking or jogging and more information is required to distinguish those activities from the rest. The sensor on the thigh would capture the more pronounced movement required before a jump is undertaken. The thigh sensor also aids in classifying agility cuts as the single sensor approach confuses this activity with walking, jumping on a box, jogging and sprinting whereas the two sensor approach has a 100% success rate recognising this activity. Both approaches perform similarly when attempting to recognise a subject kicking a ball. One reason why this confusion could occur is the varied kicking style between subjects. No specification was made hence right foot, left foot, inside of the foot strike, laces strike, passing, shot, cross, chip etc. are all viable methods that lie in the “ball kicking” label. A larger number of subjects in the dataset and a more specified activity would help account for the variation in kicking styles.

Tables 18 and Table 20 show the precision, recall and F-measure scores for both approaches. The two sensor approach has a consistent classification accuracy rate across all activities unlike the single sensor approaches. Figure

Activity	Precision	Recall	F-Measure
Agility Cut	0.96	0.922	0.941
Walking	0.985	1	0.993
Jumping on box	0.607	0.586	0.596
Jogging	0.945	1	0.972
Sprinting	0.794	0.964	0.871
Ball Kicking	0.958	0.782	0.861

Table 18: Precision, Recall and F_1 score obtained post classification using a single sensor on the shank.

Activity	Precision	Recall	F-Measure
Agility Cut	0.962	0.978	0.97
Walking	0.988	1	0.994
Jumping on box	0.75	0.724	0.737
Jogging	0.981	1	0.99
Sprinting	0.966	1	0.982
Ball Kicking	0.933	0.805	0.864

Table 19: Precision, Recall and F_1 score obtained post classification using a single sensor on the thigh.

25 illustrates the difference in accuracy between using a single sensor and using two strategically placed sensors.

4.3.4 Conclusion

In this section a novel body worn inertial sensor framework capable of automatically segmenting and classifying various actions in outdoor unconstrained environments is described. Sensors have been used extensively in body monitoring applications. With sensors becoming cheaper and more available, there

Activity	Precision	Recall	F-Measure
Agility Cut	0.984	1	0.992
Walking	0.988	1	0.994
Jumping on box	0.931	0.931	0.931
Jogging	0.976	1	0.988
Sprinting	0.966	1	0.982
Ball Kicking	1	0.839	0.913

Table 20: Precision, Recall and F_1 score obtained post classification using two sensors.

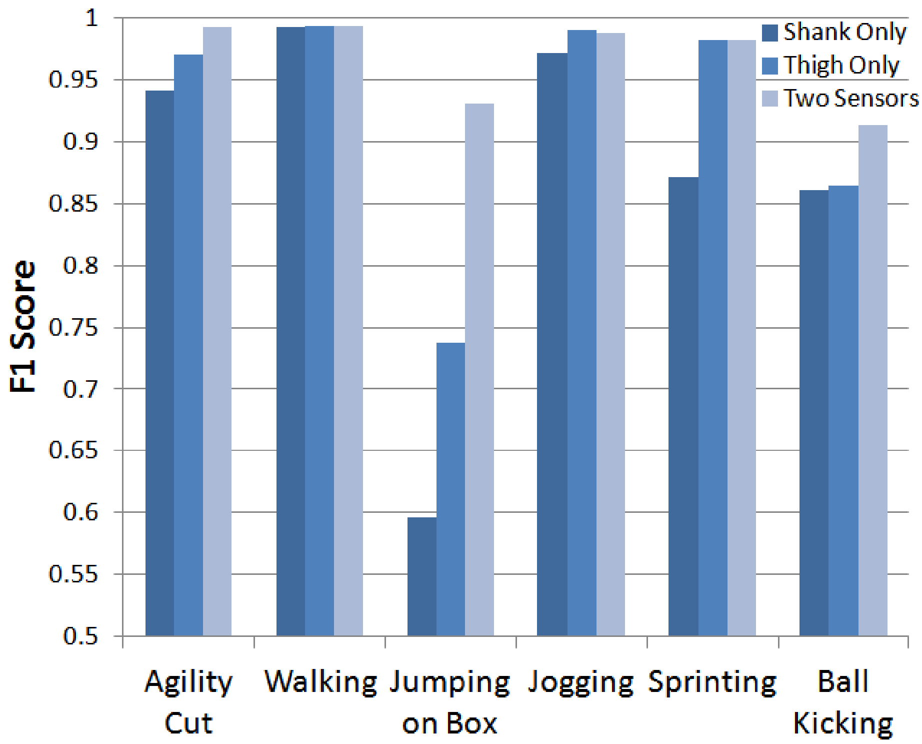

Figure 25: F_1 score comparison between one sensor and two sensors

are a variety of applications that can benefit from using two or more sensors that can capture physiological information from strategic locations. In this experiment the use of one inertial sensor versus two strategically placed inertial sensors is evaluated. The use of a second sensor has allowed in this experiment for complex activities to be recognised.

4.4 Experiment with multiple types of sensors

In this experiment different types of sensors are fused together to aid in the classification of a human activities. When fusing data from two different sensors there is two different methods; early fusion and late fusion. In this section both methods are investigated to see which performs better when using sensors of different types. By fusing these different sensor modalities it is hoped that any weakness in contextually recognizing a users actions or activity in one sensor can be remedied from data from another sensor of a different modality.

4.4.1 Early Fusion

In the early fusion scheme used throughout in this work, after the features are extracted from each sensor modality, each set of features is concatenated into a single vector. In this work feature vectors from multiple accelerometers, respiration sensors and ECG signals are concatenated to get a current context awareness of the subject. After the multimodal features are concatenated an instance based learning scheme is employed to classify the physiological data captured as shown in Figure 26.

4.4.2 Late Fusion

In this thesis research is conducted to determine whether late fusion is a suitable approach for automatically classifying human activities. In this approach, which is illustrated in Figure 27, the data from each sensor is only fused with each other at the very end of the classification process. As Figure 27 illustrates that each individual sensor is treated as a unique modality, therefore concepts are learned from three individual modalities. The prediction made from each modality is then inspected and the modality with the highest confidence is assumed to have predicted correctly or they could be weighted statistically.

Early fusion and late fusion may be combined to form a hybrid fusion approach. This occurs where a certain number of sensor modalities are fused together with a early fusion scheme to create a new fused modality. This fused modality acts like a unique modality and outputs a prediction with an associated confidence value. The remaining modalities are treated as unique modalities and each output their prediction with an associated confidence value. As in late fusion the modality with the highest confidence value is then assumed to have predicted correctly.

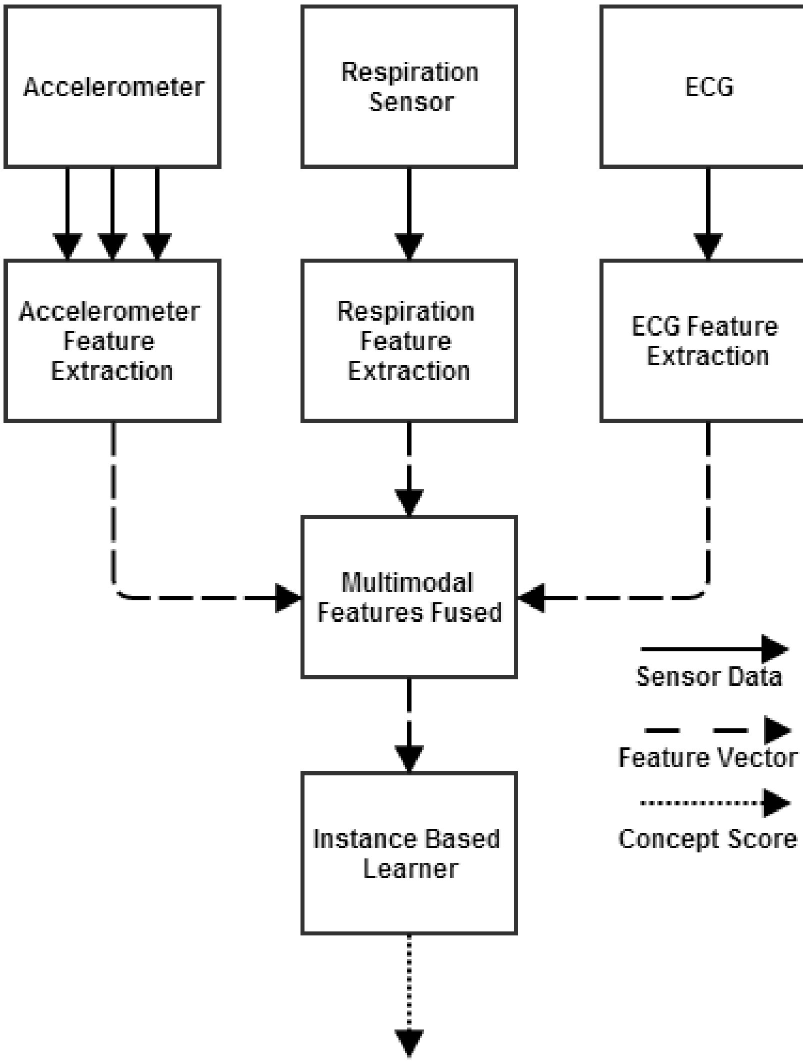

Figure 26: Early Fusion scheme. Features are fused before a concept is learned

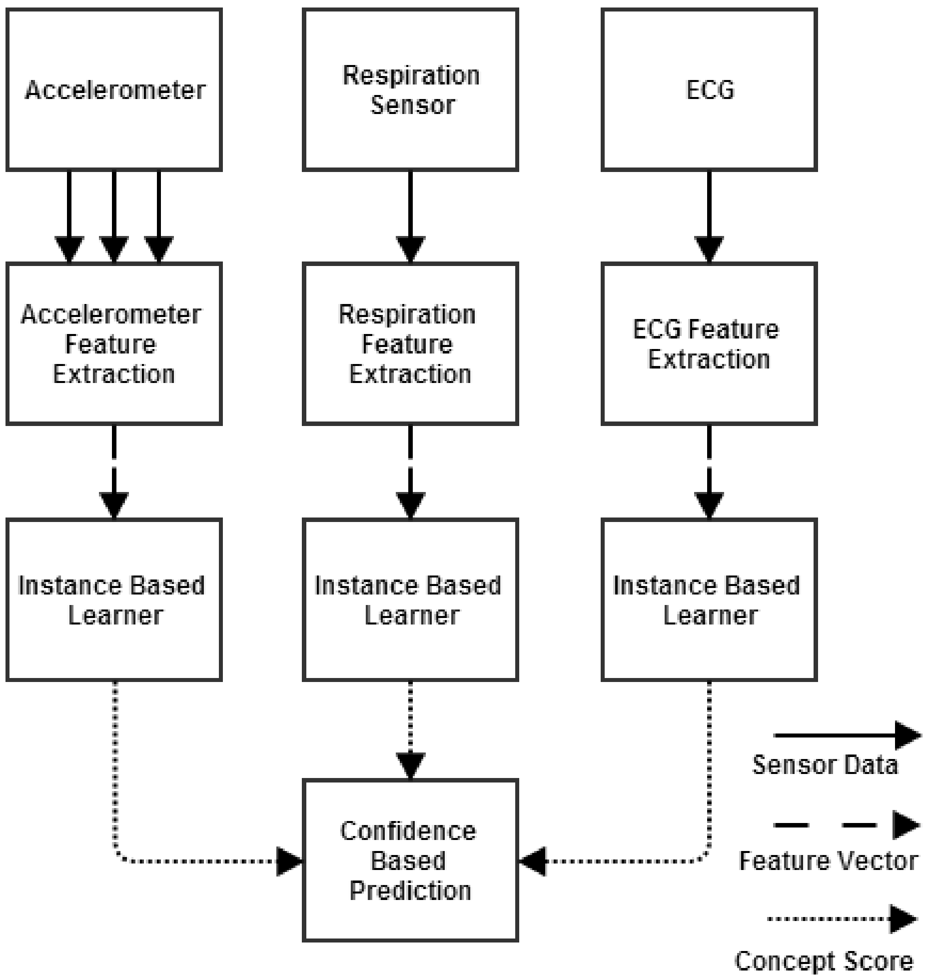

Figure 27: Late Fusion scheme. Features from three individual sensors are used to learn four individual concepts. Confidence scores determine the outputted class

4.4.3 Target Application & Motivation

Healthy sleep patterns have long been proven to be essential for maintaining both mental and physical health [159]. Sleep apnea is a common disorder which seriously degrades sleep quality and is characterised by recurrent pauses in breathing (apnea) or by instances of abnormally low breathing during sleep (hypopnoea). Apnea events can be classified into two main groups: obstructive apnea (OA) is the cessation of airflow due to the collapse of the upper airway while central apnea (CA) is due to the lack of neural input from the central nervous system [160].

Patients suffering from sleep apnea have been shown to be more prone to a number of different health complications. The associated reduction in sleep quality has been proven to increase the likelihood of accidents both at home and at work [161][162] whilst patients suffering from sleep apnea have also been shown to be more susceptible to more major health risks, including cardiovascular related deaths [163][164]. In conjunction with this increased health risk is the high cost to national healthcare systems. The U.S. National Commission on Sleep Disorders Research [164] estimated that the annual cost to the American taxpayer, for disorders related to sleep apnea, is in excess of 42 million dollars. This sum shows the clear requirement for systems which are both capable of accurately detecting sleep apnea events whilst remaining low cost.

To date there have been a number of independent systems employed to aid in sleep apnea classification [165][166], however the widely accepted gold standard diagnostic method is known as a polysomnograph (PSG). During a PSG the patient attends a specialised sleep clinic and is monitored over the course of a single night using multiple different monitoring systems. These systems commonly include a measurement of the heart (electrocardiogram (ECG)), the skeletal muscles (electromyogram (EMG)), eye movement (electrooculogram (EOG)), respiratory airflow, respiratory effort and oxygenation saturation of the blood (PPG) [166]. The position of the patient in the bed and the snoring level is also often recorded to aid in the diagnosis.

This requirement for multiple recording modalities results in a high cost (private patient test can cost up to €1,000) as well as a large quantity of data which must be examined post-recording by a trained technician for each patient. This commonly leads to long waiting lists for patients requiring testing. Current research is continuing to examine the use of less complex systems to accurately classify sleep apnea events. Examples of this research includes the use of the ECG to classify between obstructive and central apnea events [167][168] and the use of accelerometers placed on the suprasternal notch to screen for sleep apnea events [169].

This section explores the possibility of accurately identifying sleep apnea events by combining data from three different types of sensors. Classification results are obtained using a combination of electrocardiogram (ECG), respiration and acceleration sensors. By fusing these different sensor modalities it is hoped that any weakness in contextually recognizing a human actions or activity in one sensor can be remedied from data from another sensor of a different modality.

4.4.4 Data Collection

Data was collected using the "Smartex WWS" [170] from 5 adult patients (3 female, mean age 52 years, standard deviation 5.89 years) during routine PSG recording. All analysed data was recorded as part of routine sleep apnea diagnosis, therefore additional ethical approvement was not required. Patient data was analysed post recording and an exclusion criterion was implemented based on a positive indication of the presence of sleep apnea events. One female patient was discovered to not suffer from sleep apnea and was thus excluded from the study.

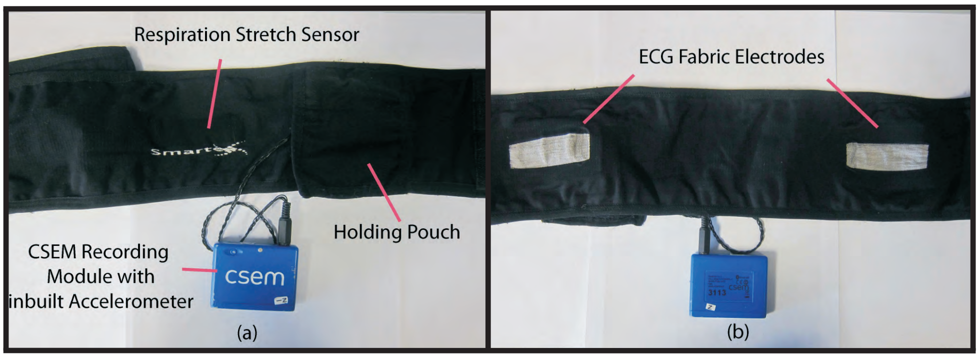

Figure 28: Smartex Wearable Wellness System. (a) Respiration sensor positioned at the front centre of the band. Accelerometer located in the CSEM recording module which is housed in the indicated pouch. (b) Fabric ECG electrodes located on the inside of the chest strap.

Patient data was recorded overnight in St. Joseph's Clinic, Raheny, Dublin, Ireland. During recording the gold standard PSG monitoring was performed while concurrent measurements were made using the implemented "Smartex" system. The gold standard PSG provided accurate information as to the time points and duration of all apnea events observed overnight, allowing for an detailed measure of the efficacy of post classification results using the "Smartex" system.

As stated previously, PSG recordings are regarded as the gold standard method for determining the presence of sleep apnea events. During the routine recording, a number of patent physiological signs were monitored including respiration flow, thorax effort, oxygen saturation, heart rate and breaths per minute. The body position and snoring output of the patient was also monitored. These signals provided a database from which a trained clinician could analyse the data, post recording, and manually tag the epochs relating to apnea events. The detected apnea events were tagged as either obstructive apnea, central apnea, hypopnoea or mixed apnea depending on their nature. Any epochs in which one or more of the recorded signals were observed to be noisy were labeled as an artifact epoch. All remaining epochs were then marked as clean. Data was recorded for a period of between 6-8 hours per subject, during which period the patient slept in a bed situated in an isolated room.

As described earlier, three separate signal modalities were measured concurrently with the PSG recording. Figure 28 illustrates the "Smartex Wearable Wellness System (WWS)" chest strap [170] used to house the three recording sensors. This WWS is a wearable system based on textile knitted sensors [171]. The electrocardiography signal (ECG) is used to monitor the electrical activity associated with the pumping of the heart. The ECG signal was recorded using two moistened fabric sensors located at either side of the ribcage (Figure 28 (b)). The use of these fabric electrodes eliminates the requirement for adhesive electrodes which can be cumbersome to apply and have been shown to occasionally cause skin irritation [172], while also allowing for unlimited use. The ECG signal was recorded at a sampling rate of 250 Hz.

The acceleration signal was recorded using a tri-axial accelerometer located in the recording module shown in Figure 28. This recording module was securely stored in a pouch located on the front of the chest strap. This accelerometer was capable of determining patient body position as well as being a proxy for the respiration signal due to the movement of the chest. The sampling rate of the accelerometer was set at 25 Hz.

The respiration signal was also monitored using the chest strap. The respiration signal was recorded using a piezoresistive knitted textile stretch sensor located on the front of the chest strap as can be seen from Figure 28 (a). As the subject both inhales and exhales, the force on the stretch sensor alters, presenting a recordable change in resistance. This resistance change can then be related to a change in lung volume. The respiration sensor was also sampled at the lower frequency rate of 25 Hz. All data recorded was stored on an on-board SD card for post processing using MATLAB.

The data from the Smartex chest strap was triggered against the PSG data post recording using the available information regarding the patient's body position. Using the accelerometer data, so that any change in patient body position could be determined, the time points relating to positional change could be aligned with the positional data available from the PSG analysis.

The "Smartex WWS" chest strap was secured to the patient's chest below the pectoral muscles and above the base of the ribcage using the available velcro. This allowed for a similar position to be attainable for both the male and female patients. It should be noted that the only pre-test requirement was the wetting of the ECG electrodes. This allowed for a very quick and easy application of the recording sensors.

4.4.5 Methodology and Results

This section describes the post processing performed on the data. First the initial filtering, tagging and windowing performed on the data is described. Then it is explained how the DWT is used to extract features from the various sensors before an explanation of how the regression tree classifier is used to classify the individual signal epochs as either clean or as containing an apnea event. Finally classification accuracy results are presented,

A number of post processing steps were completed prior to the feature selection stage of classification. Initially each signal was filtered to remove any unwanted frequencies. The ECG signal was bandpass filtered between 0.05 Hz and 20 Hz. The DC components were filtered to remove any DC offset from the signal while the ECG frequencies below 20 Hz have been shown to contain the majority of the desired ECG components. Frequencies above this frequency are required if detection of arrhythmias is desired [173]. The respiration signal bandpass cut-off frequencies were 0.05 Hz and 0.8 Hz as advised by Hejjel et al. in [173]. This upper limit was chosen as the maximum frequency of human breathing is unlikely to exceed this value [174]. Finally, the accelerometer data was low-pass filtered with a cut-off frequency of 0.8 Hz to again be capable of representing the respiration signal. All filtering was completed using 2^{nd} order butterworth filters. The DC offset was not removed from the accelerometer signals to allow their use in the determination of all positional changes. Following the signal filtering, each signal was normalised to ensure no biasing during classification.

Using the event information, available post analysis of the PSG data, the epochs of ECG, respiration and acceleration data relating to apnea events were tagged for each patient. This tagged data was truncated into individual windows, each 20 seconds in length. As an apnea must have a duration longer than 10 seconds to be classified as an apnea event [175], a window length of 20 seconds was chosen to allow for adequate representation. Each window was individually tagged as either clean or as containing an apnea event to allow for classification.

Following the windowing of the data, a total number of 1082 windows containing an apnea event were available. As the number of clean windows (9087) was much higher than the number of apnea contaminated windows, a random selection of 864 clean windows was chosen to ensure the data was not biased. These 864 clean windows were chosen from sleep segments where patients did not move in their sleep thus reducing the noise imposed on the accelerometer.

The DWT was again employed to extract discriminative features from raw sensor data due to its success in section 3.3.3. More detail on the DWT process employed here can be found in section 2.3.3 and any modifications made to the process are explained next. The energy of the signals at each decomposition level were chosen as the features for classification. To ensure independent features, the signals chosen to generate the features were the detail signals at each level and the final approximation signal. The energy of each signal $s(n)$ was calculated as:

$$E = \sum_n |s(n)|^2. \tag{15}$$

The Daubechies 5 mother wavelet was implemented [176] and the signals were decomposed to the 5^{th} level. Therefore for each window of data, only 6 features were calculated for each signal modality. Additional tests were run whilst applying additional signal features, but results were not observed to improve significantly in order to warrant their inclusion.

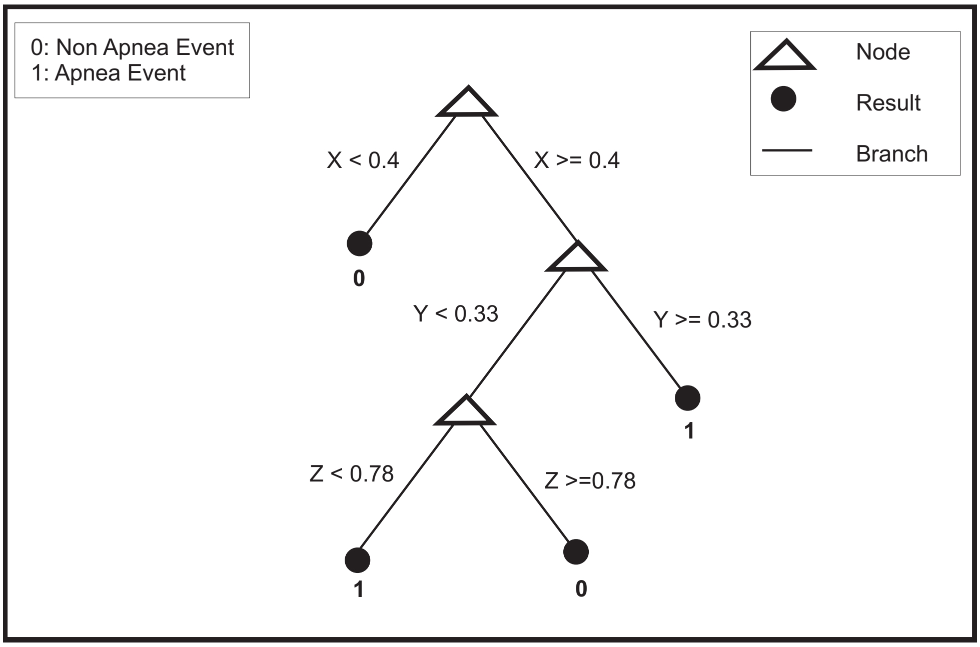

Figure 29: Simple example of a decision tree with three input features X, Y and Z.

In this work, classification is performed using a decision tree [177]. In a decision tree the class labels (i.e. Apnea/Non-Apnea) act at the leaves and the logical conjunctions (nodes) act as the branches that lead towards class labels. An example of a simple decision tree is presented in Figure 29. When a decision tree is being trained, it analyses the inputted feature set from each individual observation and develops a weighted path to every class label. Therefore, as each observation in the training set is analysed the tree becomes incrementally refined. Once the decision tree has been trained, any new observation from the test data-set can be assigned to a particular class according to its particular feature set. Classification algorithms are then evaluated by establishing how accurately they can determine the correct class label for each observation in the test data-set.

The decision tree classification algorithm has previously been implemented successfully on all three of the signals analysed in this paper (acceleration [26], ECG [178] and respiration [179]). This particular classifier was chosen due to its robustness, success in similar work and ease of use.

The purpose of the analysis performed in this section was to determine the highest classification results obtainable when classifying between apnea and non-apnea events when using only simple available sensors. The signals from the three independent sensor modalities (ECG, respiration and acceleration) were available to generate the feature set from which the class regression tree classifier was trained. Due to the availability of the three signals, seven individual combinations of feature sets could be obtained, as can be seen from Table 21.

Early Fusion Experiment

Employing information from all three signals (i.e. row 1 in Table 21) resulted in an 18 element feature set, using a combination of any two of the signal modalities (i.e. row 2, 3 or 4) produced 12 features while only 6 features were available when using the signals independently (i.e row 5, 6, 7). Table 21 presents the F_1 score results obtained when employing the seven different feature sets. Figure 30 presents the results from Table 21 visually and in descending order.

A number of interesting conclusions can be inferred from the information presented in Figure 30. Primarily, the results obtained provide a high classification accuracy, similar to that achieved by de Chazal et al. (89 %) [180]

#	Signals Employed			F_1 Score
1	Respiration	ECG	Acceleration	0.912
2	Respiration	ECG	˜	0.831
3	Respiration	˜	Acceleration	0.890
4	˜	ECG	Acceleration	0.914
5	Respiration	˜	˜	0.750
6	˜	ECG	˜	0.830
7	˜	˜	Acceleration	0.879

Table 21: F_1 score obtained post classification using a early sensor fusion approach. Table presents the change in F_1 score when employing different combinations of the three available signals to generate features.

and Yilmaz et al. (80-90 %) [168]. De Chazal et al. used power spectral density estimates of the R-wave maxima and R-R intervals from ECG data to identify apnea events. Yilmaz et al used an R-peak detection algorithm on PSG recordings (ECG included) with a SVM to identify apnea events. This result proves that the employment of the simple sensors with a low number of features is a viable option for sleep apnea classification.

Of the three signals, the respiration signal can be seen to have the lowest individual classification accuracy whilst also adding little in terms of classification improvement when added to other signal modalities. The inclusion of the respiration signal features with the acceleration signal features sees a rise in classification accuracy of only 0.011, whilst its inclusion with the ECG signal features results in a lower accuracy improvement of just 0.001. Interestingly the highest classification results are obtained when the respiration signal is omitted during feature selection. This result may be due to the accelerometer signal being capable of more accurately representing the subject's respiration, thus the respiration signal obtained using the stretch sensor does not provide any additional useful information.

Of the three sensors, the accelerometer signal proved to be the best when employing the sensors independently. By using only a single sensor, the accuracy dropped by a mere 0.035 compared to the highest accuracy obtained.

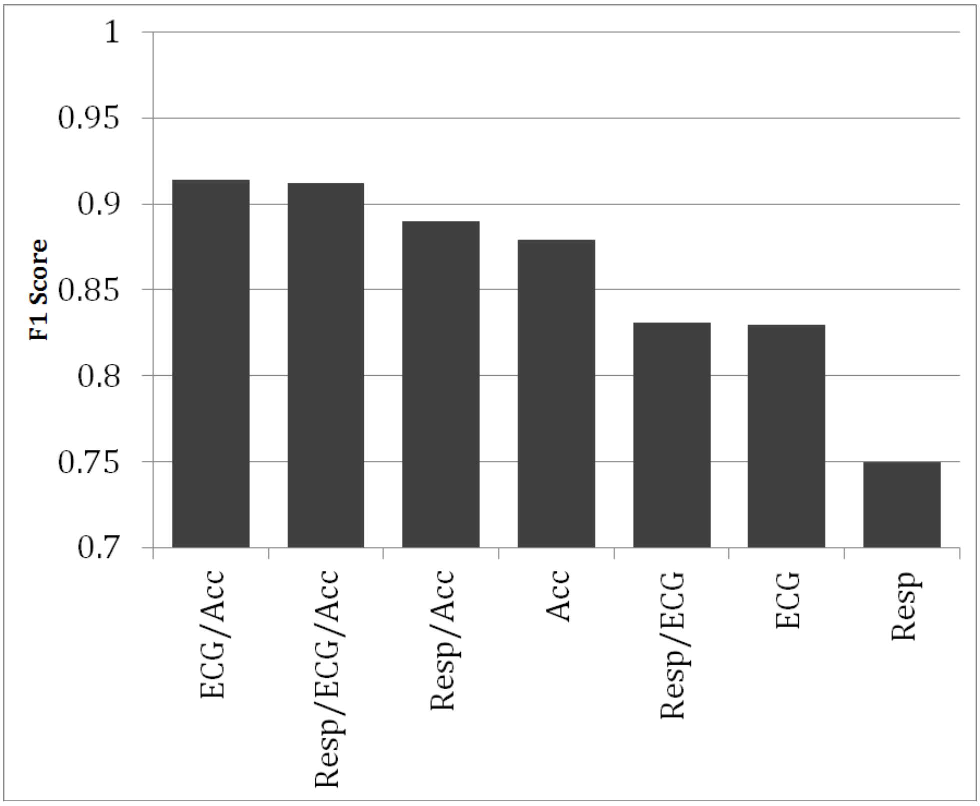

Figure 30: Impact of changing the input signals on the determined F_1 score using an early fusion approach.

This result demonstrates the functionality of solely using the accelerometer signal to attempt to classify sleep apnea events. Accelerometer sensors are very cheap to produce and can be easily attached to the subject using either a strap or an adhesive. Also, as the sensors do not require direct contact with the skin (as ECG does), it is less likely to output inaccurate or false results over a full night of testing due to the motion of the subject causing movement of the sensor with respect to the skin.

The combination of the accelerometer signal with the ECG signal provides the best results when employing the full "Smartex" system. This combination results in an accuracy of 0.914 allowing for a high confidence rate when applied over a large number of apnea events.

Late Fusion Experiment

In this experiment the same methodology was used as in the previous section with the exception that a late fusion approach was applied instead of an early fusion approach. Again seven individual combinations of feature sets could be obtained, as can be seen from Table 22.

As expected classification results when only using a single modality was the same with the late fusion approach as the early fusion approach. Classification results are lower across all combinations when compared to their early fusion equivalents. This could occur for a variety of reason's one of which could be that decision tree classifiers perform better when they have more features that they can directly compare to one another when creating rule-sets for class prediction.

Classification results from different sensor combinations are very similar between late fusion and early fusion experiments with the exception that the respiration sensor adds noise to its combination with the ECG sensor and lowers the classification accuracy compared to using the ECG sensor on its own. This is akin to how the classification performance is improved in the early fusion experiment when the respiration sensor is omitted.

#	**Signals Employed**			**F_1 Score**
1	Respiration	ECG	Acceleration	0.910
2	Respiration	ECG	~	0.857
3	Respiration	~	Acceleration	0.892
4	~	ECG	Acceleration	0.908
5	Respiration	~	~	0.786
6	~	ECG	~	0.851
7	~	~	Acceleration	0.890

Table 22: F_1 score obtained post classification using a late sensor fusion approach. Table presents the change in F_1 score when employing different combinations of the three available signals to generate features.

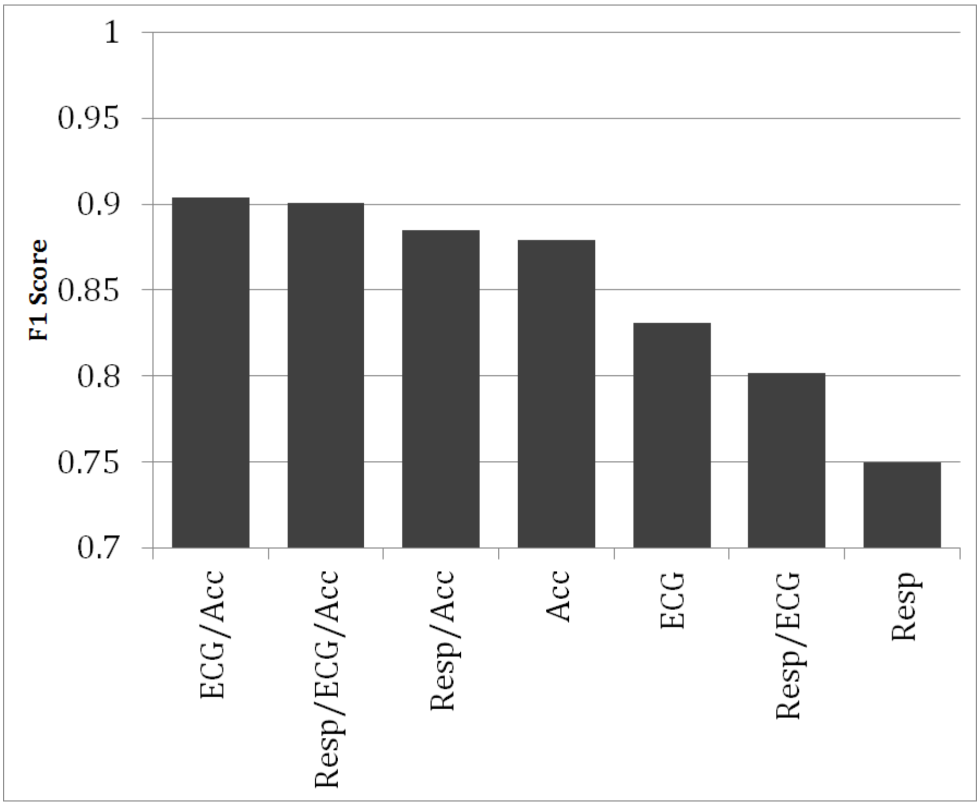

Figure 31: Impact on the determined F_1 score of changing the input signals using a late fusion approach.

4.4.6 Conclusion

In this section, an analysis into the use of cheap and easy to use respiration, ECG and accelerometer sensors for the classification of sleep apnea events has been investigated. Results have shown that the accelerometer signal provides the best results when a single sensor is used to classify the data. This result has, to the author's knowledge, never previously been highlighted. Many authors have discussed the sole use of an ECG signal to classify similar data [180][167][168] however results shown within this work instead propose the use of a simple accelerometer signal for classification purposes. However more accurate results are available if features from the ECG data are also included in the analysis. The respiration signal was determined to not improve the classification results and thus should not be included in analysis.

The ability to accurately classify the sleep apnea data using cheap and easy to use sensors can allow for pre-PSG testing to take place. As PSG testing can be prohibitively expensive, the ability to pre-test for sleep apnea using a cheap and accurate system could allow the tests to be preliminarily carried out outside of the clinic environment. This pre-test would reduce the number of patients who do not suffer from sleep apnea applying for PSG testing, thus saving them money and free up medical resources and reducing he strain on the public sector. The availability of this pre-test would also reduce the waiting time for PSG tests which currently can be up to 6 months between initial referral and testing. As these systems continue to become more accurate, the requirement for final analysis using the full PSG system may eventually become redundant.

Early and late fusion of sensor data was also investigated in this section. Early fusion requires less computational time than late fusion and achieved a similar classification accuracy.

4.5 Conclusion

Several experiments have been conducted in this chapter. Initially it was shown that two strategically placed sensors that capture the same physiological information outperform a single sensor. In the fusion section, results prove that even though early fusion requires less computational time, it is as accurate at detecting human activities as a late fusion approach. Additionally it was proven that adding a sensor of a different type can hinder a classification system instead of aiding it therefore it is important to accurately gauge whether the data that one sensor is capturing is not also being captured more accurately by a sensor already in use.

5 Parameter Selection Optimisation using a Genetic Algorithm

5.1 Introduction

Genetic algorithms are one of the most active research areas today as they are a state of the art approach to generating useful solutions to optimization problems. Their popularity is motivated by a wide range of potential applications in numerous areas. Applications where genetic algorithms have been successfully applied include chemistry [181], manufacturing [182], engineering [183], economics [184], pharmacometrics [185], bioinformatics [186], mathematics [187], phylogenetics [188], physics [189] and computational science [190]. Genetic algorithms have been used in the literature to select the optimum parameters for specific classifiers [93][94] they have not to the authors knowledge been been widely applied to the complete model creation process.

Genetic algorithms are a type of optimization methodology inspired by the mechanisms of biological evolution and behaviours of natural organisms[91]. The search for parameters mimics the mechanism of evolution by forced selection. Evolution by natural selection is the well known process where certain evolutionary traits are passed to the next generation if they prove useful to survival based on a myriad of external influences. However the genetic algorithm more closely mimics evolution by forced selection, such as breading a sub-species of dog for a particular purpose. The GAs are lead by a cost function that are imposed at the outset. Genetic algorithms are a sub-group of evolutionary algorithms that formulate solutions to optimization problems using methods inspired by natural evolution. These methods include mutation, inheritance, crossover and selection. In a genetic algorithm a proposed solution to an optimization problem is called a chromosome and a set of these chromosomes is termed a population. The population evolves over numerous generations and the algorithm converges when a particular chromosome in the current population is deemed to be the most optimum solution.

Genetic Algorithms were first described by John Holland in the 1960s and were continually developed by him and his colleagues into the 1970s. His goal was not to create an algorithm to solve a specific problem, but rather to formally study the phenomenon of adaptation as it occurs in nature and to establish a method in which the mechanisms of natural adaptation could be imported into computer systems[191]. In the late 1980s practical applications of the GA were feasible due to the considerable increase in computational power that had occurred in that time.. General Electric then started selling the world's first genetic algorithm product, a mainframe-based toolkit designed for industrial processes. Since then GA use has spread to various fields with much success.

There are many preprocessing and postprocessing steps required in order to create the most accurate activity classification models. These include filtering the signal, extracting discriminative identifying features and selecting an appropriate classifier. Each of these steps can have a large number of parameters. Choosing which parameters to investigate is complex and often arbitrary in practice. This gives researchers a limited search space in which to achieve optimum accuracy. In this chapter genetic algorithms are investigated in terms of whether they can improve this cumbersome approach.

Figure 32 gives an overview of the GA process employed to achieve this goal. The initial population is compromised of candidate solutions whose parameters are generated purely randomly. The fitness of these candidates are then evaluated by a fitness or cost function and ranked accordingly.

A certain percentage of the top performing candidate solutions are selected to remain in the population, this process is known as elitism. A certain percentage of brand new random solutions are also generated for the next

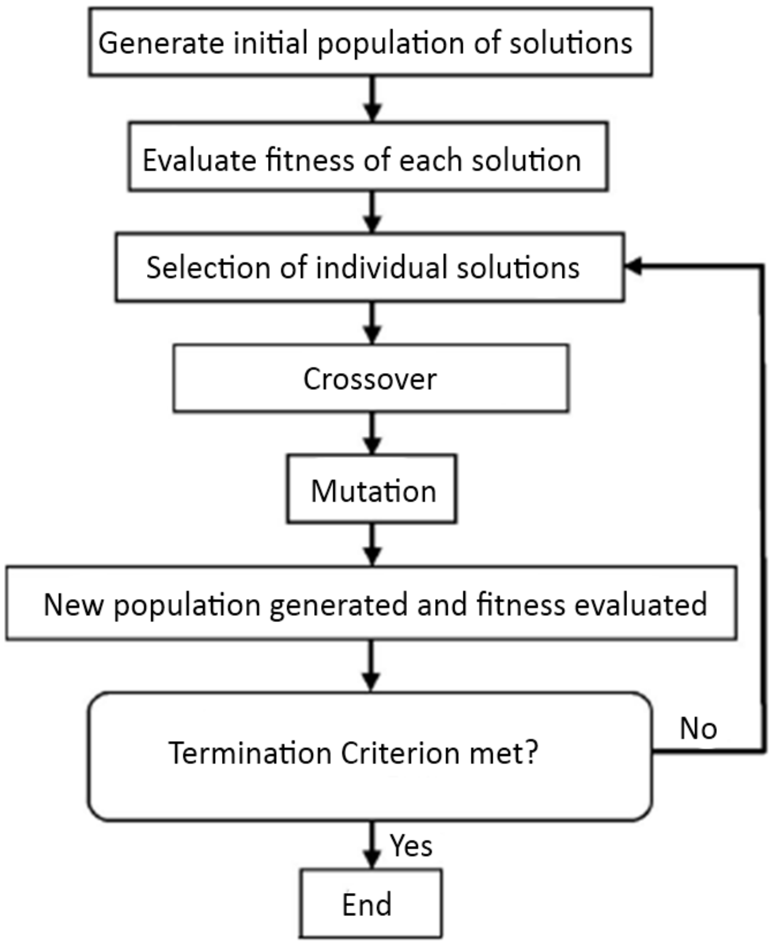

Figure 32: Genetic algorithm process

generation, to avoid a stagnant gene pool. The remaining solutions in the next generation are "children" of the previous generation. These "children" are created by selecting two parents, mixing or crossing the properties from the parent, then allowing for the random mutation of single parameters of these "children". Once this new generation of candidate solutions set has been fully created, each solution is evaluated with a fitness function. If any of the solutions meet a termination criteria determined by the cost function, such as required solution, generation count or stagnation of overall fitness then the process is stopped. If non of these criteria are met then the process reverts back to the selection stage and starts again.

5.2 Related Work

Many practical classification tasks require learning of an appropriate classifier function that assigns a given input, which is usually a series of attributes, to one of a finite set of classes. The choice of features, attributes, feature size and classifier type all have a direct impact on the overall accuracy of a model. This presents a parameter selection problem in automated design of activity classification. The feature subset parameter selection problem refers to the task of recognizing and determining an efficient set of parameters to be used to represent patterns from a larger set which contains often redundant, possibly irrelevant parameters with different associated measurement costs and or risks.

Creating classification models for large datasets is a time consuming task as finding the most productive parameters is computationally expensive. Huang et al. in [93] uses a GA gene-based feature selection and parameters optimization for support vector machines (SVM). They showed that the parameters and features subset of a SVM, a popular technique for pattern classification, could be optimized using a genetic algorithm based approach without degrading classification accuracy. The GA was compared against other parameter search methods such as the Grid algorithm and found that the GA based approach significantly improved the classification accuracy whilst also having fewer input features for the SVM which decreased classification time.

The work reported by Punch et al. in [94] employs a GA in conjunction with a K-nearest neighbour algorithm to optimize classification by searching for an optimal feature weighting. This warps the feature space to coalesce individuals within groups and to isolate groups from each another. The GA can also be used to implement efficient methods of fusing classification models for one overall prediction goal.

In [95] the authors use a genetic algorithm to design a multiple-classifier system. They tested their methods on four real data sets. They found that the GA design was less prone to overfitting compared to classifiers using: all features; the best feature subset found by the sequential backward selection method; and the best feature subset found by a GA (individual classifier). They concluded that their GA-designed system used in their experiments was more accurate than the best individual classifier in the system. designs.

The literature cited in this section focuses on using the GA to obtain the best combination of parameters for a specific classifier which will allow the most accurate classification model to be created. Similarly, in this work, the goal is to employ the GA to locate the best combination of parameters to generate the most accurate classification model however in this instance input vector size, feature extraction parameters and different types of classifiers will constitute the search space. This allows the GA access to the entire classification process instead of focusing on a specific area.

5.3 Experiments

In this section various experiments were conducted using a GA for parameter selection for classification problems. Firstly the GA is applied to a known classification dataset for which model accuracy has been previous calculated for all parameter possible permutations. This allows for a ground-truth to be established and from this it is possible to ascertain whether the GA is able to consistently optimize the choice of parameters for classification problems more quickly than brute searching all possible combinations.

The results obtained using the ground-truth dataset can then be extrapolated for classification problems where the highest accuracy model is unknown. In other words, by establishing the most optimum parameters of the GA and exploring the average time it takes the GA to converge on the highest accuracy model for a given known ground-truth dataset then these GA settings can be applied to an unknown dataset and the statistical probability for the G.A finding the most accurate solution can be calculated. In order to evaluate the GAs performance with this unique dataset different population sizes for the GA are examined to ascertain their convergence rate for the most accurate model in the dataset. The second experiments aim was to investigate whether the GA could locate a superior solution from a larger parameter search space in the same amount of time taken as the brute force approach earlier. The final experiment was to investigate whether the most accurate models for the larger search-space could be extrapolated from the best performing models from section 3.3.3. If these newly created models do not contain the most accurate model then the combination of the best parameters is stochastic.

Table 23 shows the parameters chosen when implementing the GA. The elitism percentage is the percentage of fittest candidate solutions from each population that will be carried over to the next generation. The crossover probability is the likelihood that a new candidate solution will be generated using crossover principles. The mutation probability is the likelihood that a new candidate solution will have mutation applied to it. The new gene percentage is a percentage of each generation that will consist of purely randomly generated candidate solutions. Chakraborty et al. argues in [192] that the use of elitism hinders the possibility for such algorithms becoming trapped in local minimum and demonstrates that values around 20% normally prove the most effective. Experimental results in [193] have shown that large values for crossover and mutation probability have a larger success than lower values, hence 80% was chosen for both. Dedicating 20% to brand new candidate solution is common in GA optimization problems and as such is implemented in this work.

Elitism Percentage	20%
Crossover Probability	80%
Mutation Probability	80%
New Gene Percentage	20%

Table 23: Parameters chosen for the GA

5.3.1 Benchmarking

In section 3.3.3 16,380 different models were created and investigated in order to achieve the highest possible classification accuracy. These models comprised of four different parameters which were window length, choice of motherwavelet, decomposition level and classifier. Decomposition levels between 1 and 7 were investigated. Five popular families of classifiers were employed for sports activity classification using the discrete wavelet transform decomposition of accelerometer signals.

John C. Platt's Sequential Minimal Optimization (SMO) optimization algorithm was used for the training of the support vector machine (SVM) classifier. The IBk classifier is a simple instance-based learner that uses the k-nearest neighbour(k-NN) algorithm for training. The Naive Bayes classifier applies Bayes' theorem with strong (naive) independence assumptions to train its classification models. A logistic model tree (LMT) is a decision tree with logistic regression functions at the leaves for supervised learning tasks. A multilayer perceptron (MLP) is a feedforward artificial neural network that utilizes back-propagation for training a network. Window lengths between 1 and 9 seconds were investigated with a step size of one second. Five motherwavelet groups were employed with a total of 52 separate motherwavelets investigated. These were the first 10 Daubechies, 7 Symlet, 5 coiflet, 15 biorthogonal and 15 reverse biorthogonal wavelets. Table 24 shows the highest accuracy results obtained using a single model classification model.

Sport	Classifier	DWT lvl	Mother W.	Length(s)	F1
Soccer	NaiveBayes	6	rbio1.1	3	0.799
Hockey	MLP	6	bior1.1	7	0.823

Table 24: Highest classification accuracies attained before optimisation

5.3.2 Experiment 1

Aim

In this initial experiment the goal was to investigate whether the GA was able to reduce the amount of time needed to find the parameters required to create the most accurate model. GAs with different population sizes were used to find the optimum parameters required to create the most accurate model. In practice users can select a number of criteria that would determine if the algorithm should terminate such as desired accuracy level, amount of candidate solutions constructed or if the current best result has not improved for a certain amount of candidate solutions.

Methodology

In section 3.3.3 16,380 different models were created and investigated in order to achieve the highest possible classification accuracy. Four parameters were investigated; window length, DWT decomposition level, classifier and motherwavelet choice. This brute force approach took approximately 5 days to complete and achieved an accuracy of 79.9% for soccer and 82.3% for field hockey. The GA was employed on this same dataset and the length of time it took to find the highest classification accuracy model was investigated. The maximum amount of candidate solutions to be created was set at 50% of the total possible parameter permutations. This figure was chosen after initial exploratory experiments when the majority of population experiments concluded before 50%. Therefore it was assumed that if the GA had not identified the most accurate model by 50% of the total possible permutations then the algorithm was stuck in a local minimum. In Figure 33 the final bar in each graph indicates the number of times the GA did not locate the optimum parameters for the most accurate classification model.

Results

Figure 33 shows the performance of different population sizes used in this experiment. Nine population sizes were investigated ranging from 10 to 120. Each population size experiment was conducted one hundred times in order to give a unbiased account of each population size performance. The X-axis of each graph indicates what percentage of the total possible parameter permutations had to be calculated before the most accurate model was found. The Y-axis of each graph shows the percentage of times the most accurate model was created in the X-axis percentage range. For example in the graph which demonstrates the performance of the GA whose population size was 30 it can be seen that 15% of the time the most accurate model was created within the first 5% of total possible models. When brute searching every possible parameter permutation(100%,) must be investigated. Figure 33 illustrates that the majority of population sizes located the optimum parameters on average well before 50% of the time

Figure 34 shows the average percentage of total possible solutions required to be investigated before the optimum solution is located for each population size. Lower population sizes must construct a larger amount of permutations before the optimum solution is found. Population sizes from forty onwards required a similar amount of candidate solutions to be constructed before the optimim solution was located however the GA with a population size of 50 located the optimum parameters fastest on average.

From Figure 33 it can be observed that while the smaller population sizes perform better than a standard brute force approach, they do not converge on the best combination of parameters on average as quickly as the larger

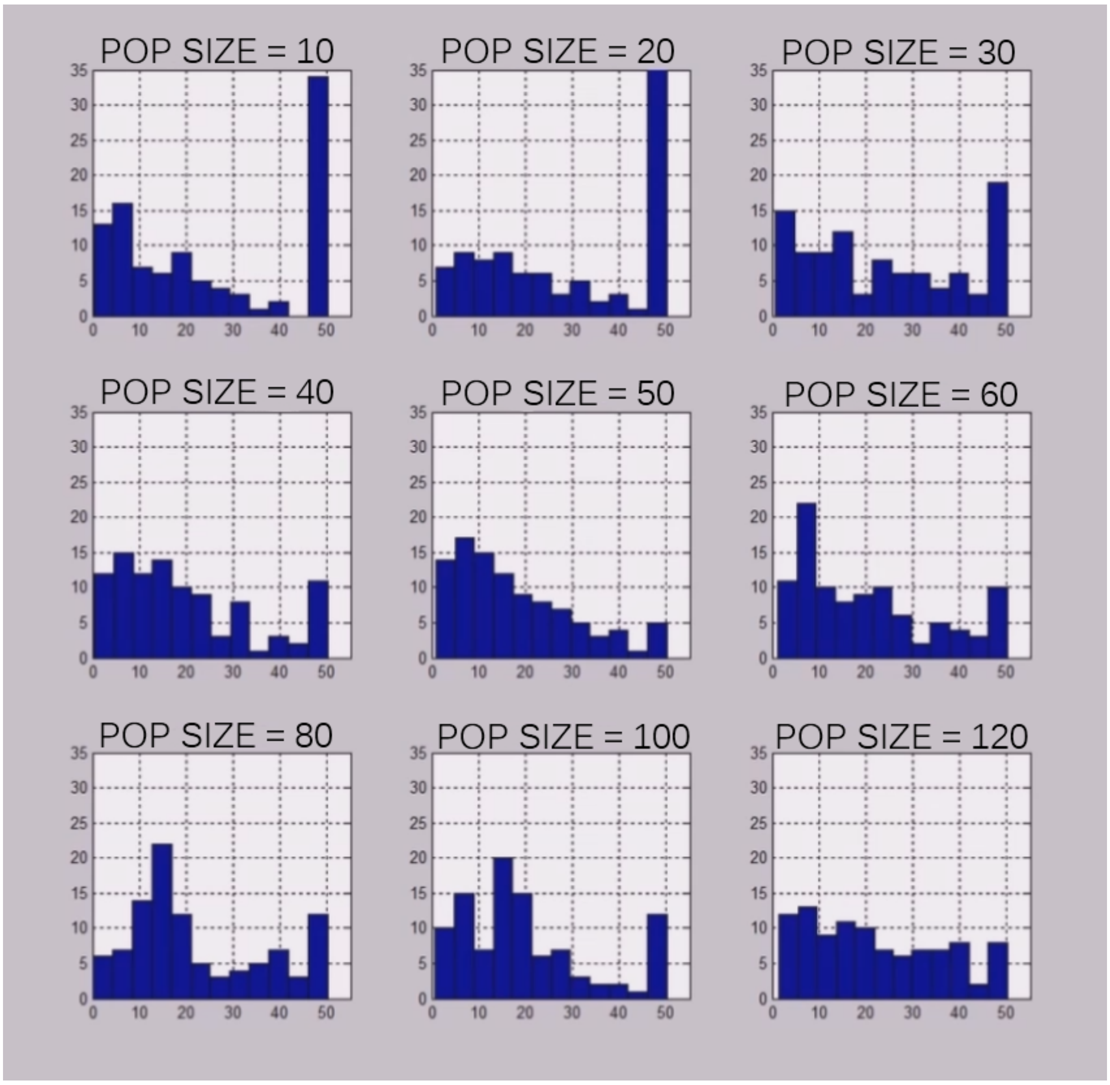

Figure 33: Comparison of GAs with different population sizes

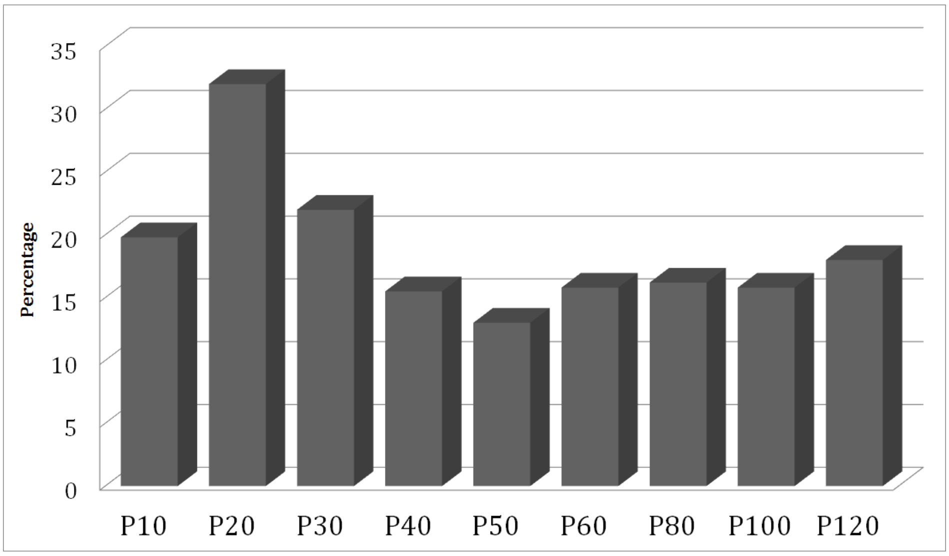

Figure 34: Average amount of solutions required before optimum solution found

population sizes. Larger population sizes performed well but did not converge on the optimum set of parameters on average as quickly as the middle population sizes.

Figure 35 shows the comparison between the two best performing population sizes, 50 and 60, from Figure 33. It can be observed that the GA with a population size of 60 only outperforms the population size of 50 when 8% of the possible permutations have been investigated. Therefore the GA with a population size of 50 was found to locate the optimum parameters most reliably. It can also be observed that the GA with a population size of 50 will find the parameters for the most accurate model 31% of the time within the first 10% of parameter permutations. On average however it will also find the parameters for the most accurate mode 75% of the time within the first 25% of the total possible parameter permutations. In addition the GA with a population of size of fifty only failed to locate the optimum parameters for the most accurate classification model 5% of the time before the 50% termination criteria was reached. This 5% figure was the lowest failure

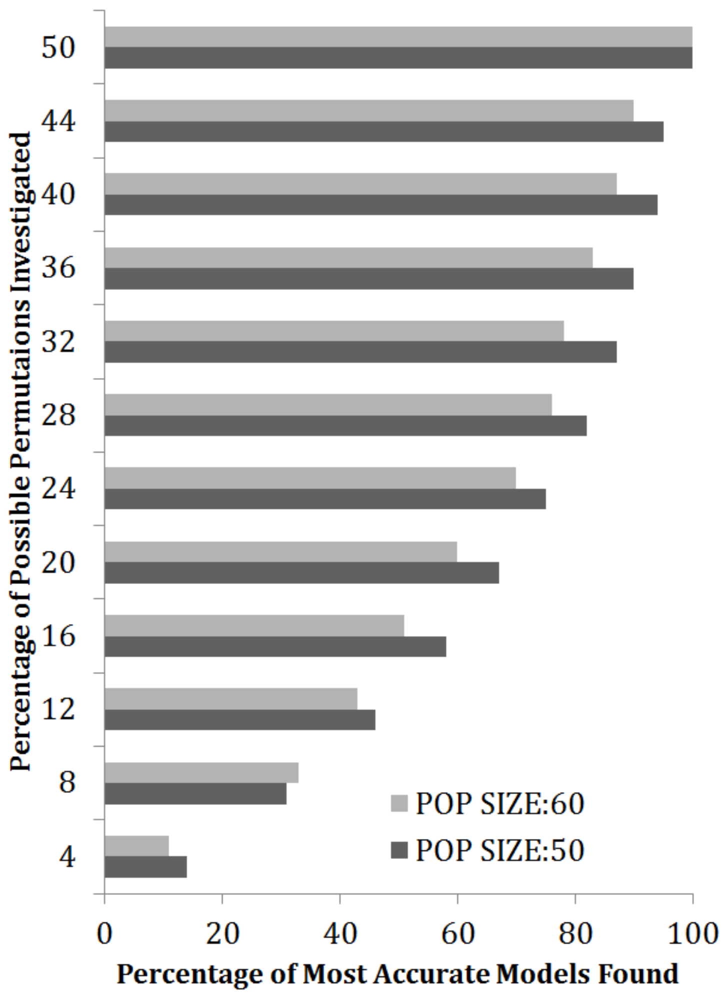

Figure 35: Performance comparison of population sizes 50 and 60

rate compared to the rest of the population sizes investigated. This result presents many opportunities for researchers of various classification areas to utilize generic algorithms to optimize the creation of optimum classification models.

5.3.3 Experiment 2

Aim

In this experiment the aim was to investigate whether the GA could locate a superior solution from a larger parameter search space in the same amount of time taken as the brute force approach earlier.

Methodology and Results

In experiment two the number of window length permutations was increased by a factor of ten by reducing the step size from one second to one tenth of a second for all models. This increased the number of permutations from 16,380 to 163,800. Implementing a brute force for this problem would be infeasible as it would take approximately one and half months to complete. This figure was estimated from the amount of time it took to generate the initial 16,380 models as that took approximately five days. The targeted model accuracy was set at 100% and it was not expected to be achieved. The candidate solution count limit was set at 10% of the maximum number of possible candidate solutions. This value was chosen as it computes the same amount of candidate solutions as was computed in section 3.3.3 while investigating a much larger search space. There was no convergence break limit set. The population size was set at 50 as results from experiment one had shown that it was the most productive figure. The GA population also reserved 20% for

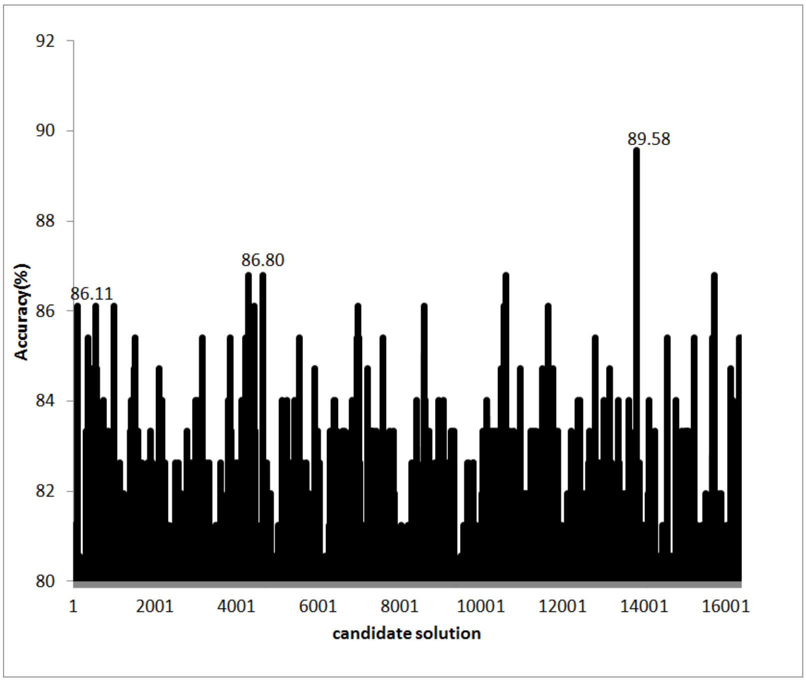

Figure 36: Experiment 2 - Football Dataset Results

Sport	Classifier	DWT lvl	Mother W.	Length(s)	F1
Soccer	Tree-LMT	3	db7	2.1	0.8958
Hockey	Tree-LMT	1	sym5	1.5	0.8888

Table 25: Highest classification accuracies attained with optimisation

elite candidate solutions, 20% for new random genes and the remaining 60% for possible crossover candidates and mutation candidates. There was a 80% chance that a candidate solution would have its genes mutated or crossed over.

Figures 36 and 37 show the amount of candidate solutions required to be calculated before a more accurate model was created. The most accurate Football candidate solution was 13810th while the most accurate Hockey candidate solution was the 9440th. An improved model accuracy of 89.58%(+9.59%) was achieved in the Football Dataset while 88.88%(+6.5%) was achieved in Hockey Dataset with the use of a genetic algorithm. Table 25 shows the parameters for the best performing model for each sport.

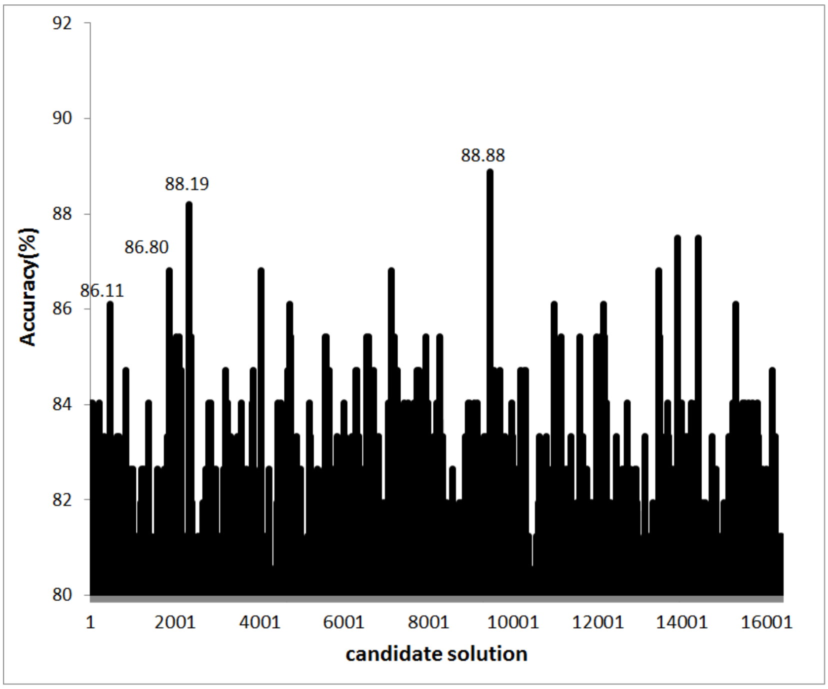

Figure 37: Experiment 2 - Hockey Dataset Results

5.3.4 Experiment 3

Aim

The aim of this experiment was to investigate whether the most accurate models for the larger search-space could be extrapolated from the best performing models from section 3.3.3. This is accomplished by varying each old model's window length with the smaller step size. If these newly created models do not contain the most accurate model then the combination of the best parameters is stochastic.

Methodology and Results

In experiment three the top twenty performing classification models for each sport created by the brute force method described in section 3.3.3 were investigated. The size of the window length permutations was increased by a factor of ten by reducing the step size from one second to one tenth of a second as with experiment two. The accuracy of every model permutation was calculated to see if there was any correlation between the previous top performing models and the new larger search space. Investigating all permutations created 1602 models for both datasets.

The maximum accuracy achieved in the Football dataset was 86.11% which is less than the 89% achieved in experiment two. The maximum accuracy achieved in the Hockey dataset was 86.80% which is less than the 88.8% achieved in experiment two. This indicates that that the optimum parameters for this larger classification search-space cannot be inferred from the best parameters from a smaller search-space.

5.4 Conclusion

Genetic algorithms are one of the most active research areas today to generate useful solutions to optimization problems yet their potential in classification parameter reduction has not been fully explored. In this section the genetic algorithm was employed to optimize the discovery of a combination of parameters required to create the most accurate classification model from a dataset. GAs have been utilized in many different application areas for optimization purposes however in the machine learning field the literature has focused on using the GA to obtain the best combination of parameters for a specific classifier which will allow the most accurate classification model to be created. However in this work, the goal is to employ the GA to locate the best combination input vector size, feature extraction parameters and different types of classifiers. This allows the GA access to the entire classification process instead of focusing on a specific area.

From the first GA experiment it can be observed that the GA is regularly able to substantially reduce the time needed to locate the optimum parameters for a classification model. This is significant as attempting to locate optimum model parameters for a large dataset requires substantial processing power.

Researchers generally speculate on what they believe are the most influential parameters in a classification problem and limit their investigation as it is not feasible to investigate very large search spaces. The second GA experiment shows how the GA can permit a much larger parameter search space to be investigated. This in turn allows more accurate models to be created. In experiment two the GA termination criteria was the amount of candidate solutions allowed to be constructed and it was set at the same amount of solutions generated by the brute force approach and the search space was expanded ten fold. The GA improved the classification accuracy for soccer by over 9.5% and improved the classification accuracy for hockey by over 6.5%.

Finally, experiment three was conducted to investigate whether the parameters located for this new highest accuracy model could of been extrapolated from the previous top candidate solutions. This was not the case therefore the GA is required to locate the new best performing parameters.

In summary three experiments were conducted to examine whether the GA would be able to optimise the parameter selection process for the classification framework purposed in this thesis. The results indicate that this is possible but also that the application of the GA can lead to the creation of even more precise classification models.

6 Conclusion

Context-aware computing is an expansive area of research. This thesis examined one particular field of context-aware computing, namely activity recognition with wearable sensors. It explains that sensor technology combined with the power of data mining and machine learning techniques, can respond to various needs of users in a context aware way. To facilitate research in this area, a machine learning framework was constructed that allows automatic human activity models to be generated for classification purposes. The automatic activity recognition framework aims to recognise the actions and events of a user utilising physiological data captured from sensors attached to the body. However the algorithms, methods and framework that is presented in this work can be applied to any classification problem.

6.1 Thesis outline

Chapter 1 introduces this work, providing a brief overview of context aware computing and its future role in society. It presents the several research objectives associated with creating a machine learning framework for automatic human activity classification. It gives a indication on the steps required to be investigated before this outcome could be realised. It explains that physiological sensors were chosen as the primary means of capturing user data and that data will allow user context to be deciphered. It briefly presents the research contributions of this work before giving an overview of each chapter in this work.

Chapter 2 explores the technical background of preprocessing the raw sensor data before looking at the literature for advanced feature extraction. Then physiological sensors and their relevant applications are described in detail for classification purposes followed by a look at the different data fusion techniques which are investigated on their applicability for multimodal sensor fusion. After, this chapter there is an exploration at the state of the art machine learning techniques, which are used for to create classification models in this work. Finally state of the art parameter selection optimisation techniques are introduced, which were to hasten the classification model creation process. Relevant literature is presented as a basis for the framework design choices presented in this thesis.

Chapter 3 describes the challenges which were overcome in order to perform human activity recognition using a single sensor. This chapter outlines the feature extraction techniques required to initially evaluate a users activity performance before creating algorithms to identify various different sporting activities. A number of experiments are undertaken in order to ascertain the best approach to creating a classification model. A black box experiment is compared to a thorough investigation of all parameters. These two approaches are then compared to a final approach where each activity has its own specialised classifier. All methods proposed are compared to a literature benchmark for evaluation purposes. Different feature extraction techniques such as DWT, FFT and some simple time domain techniques were implemented for comparison purposes. High class recognition scores are given as evidence that the methodology presented in this work has scientific merit.

Chapter 4 explores the challenges encountered when creating a multimodal human action recognition system. Advantages from using two accelerometers versus a single accelerometer to identify different training activities performed by a subject are presented. As the results prove, sensor fusion can significantly improve the accuracy rate for classification models. Early fusion and late fusion are the two techniques used in this chapter to fuse data from different sensors. Experiments are conducted that use both early and late fusion to fuse the data from ECG, respiration and accelerometer sensors. Results prove that even though early fusion requires less computational time, it is similarly accurate at detecting human activities as a late fusion approach.

After evaluating those two approaches, results obtained when using different permutations of three sensors of different modality are presented in this chapter. Results presented in this chapter indicate that adding sensors of a different modality to a activity recognition system can help increase the accuracy of said system. Results also indicate that adding a sensor which captures physiological data that is already being accurately measured by a different modality can decrease classification accuracy.

Chapter 5 takes the results of all the model parameter permutations discovered in section 3.3 and uses them to test different genetic algorithms. This chapter explains why genetic algorithms can help optimise the parameter selection process. It also goes into detail on the role of each parameter that makes up the genetic algorithm and process behind it. Three experiments are conducted to investigate the genetic algorithms suitability to optimise the process of parameter selection. The first experiment explores the use of different population sizes and compares each GA to the brute force approach. Experiments were conducted 100 times each to give a fair representation of each populations ability. The second experiment increases the number of parameter permutations by a factor of ten but the number of possible solutions investigated was limited at the same amount as in section 3.3. This showed that the genetic algorithm could locate a new superior solution in the same amount of time it took the brute force algorithm to search through a search space one tenth of the size. The final experiment investigated whether this new superior models could of been extrapolated from original best performing models in section 3.3. The main result presented in this chapter suggests that using a GA on a similar classification parameter selection problem as presented in this chapter will yield significant savings in computational time or required processing power.

6.2 Suggestions for Future Work

The experiments conducted in this work to create a classification framework suggest there are many research areas which merit additional research. This section analyses these possible research areas which could add to the literature in automatic activity classification. As previously stated one area of future work will focus on investigating utilizing all sensor in smartphones which can capture physiological information. While various forms of feature extraction methods are analysed and compared in this thesis there is still scope for future research into other feature dimensionality reduction techniques such as principal component analysis, discrete cosine transform and the Walsh-Hadamard transform.

Furthermore there are other sensors which have grown in popularity such as the miCoach by Adidas. It would be interesting to investigate their performance compared to smartphones. Planned future work will look to examine and to compare accuracy of the current major available commercial devices and investigate whether the physiological data captured by each sensor could be fused together to improve classification accuracy.

The success of the GA on optimising the parameter selection process warrants the investigation of other search heuristic algorithms. Particle Swarm Optimization (PSO) is one such relatively new heuristic search method whose mechanics are inspired by the swarming or collaborative behaviour of biological populations. Successfully implementation of parameter selection optimisation techniques when creating classification models allows much larger parameter search spaces to be investigated thus allowing for more accurate models to be created.

A Sample Signals from Section 3.3.2

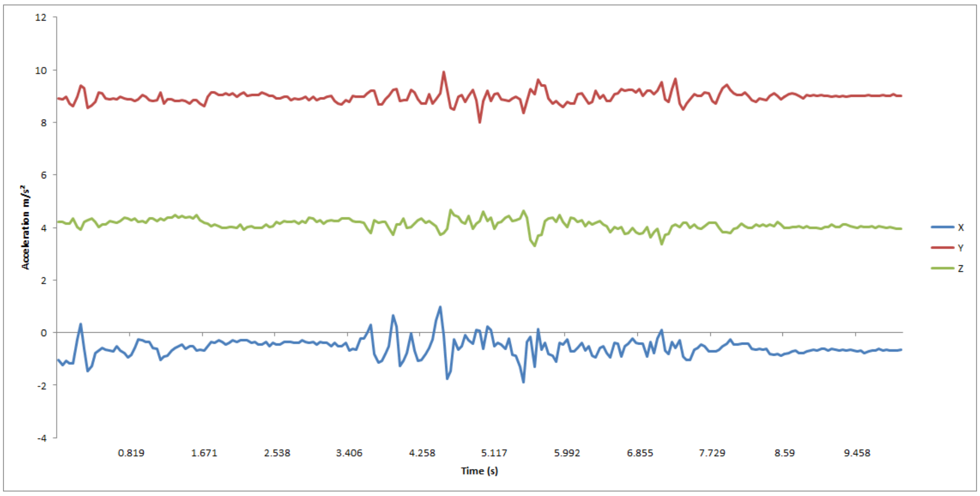

Figure 38: Player Stationary

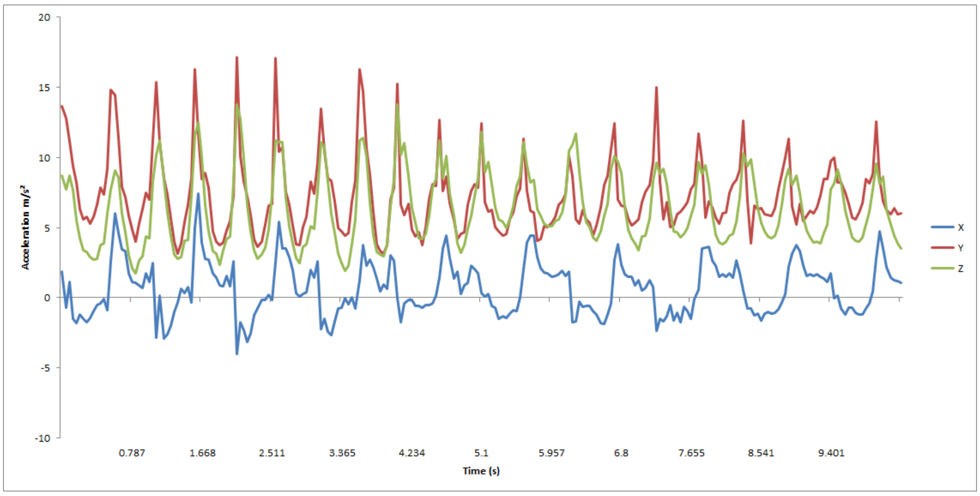

Figure 39: Player Walking

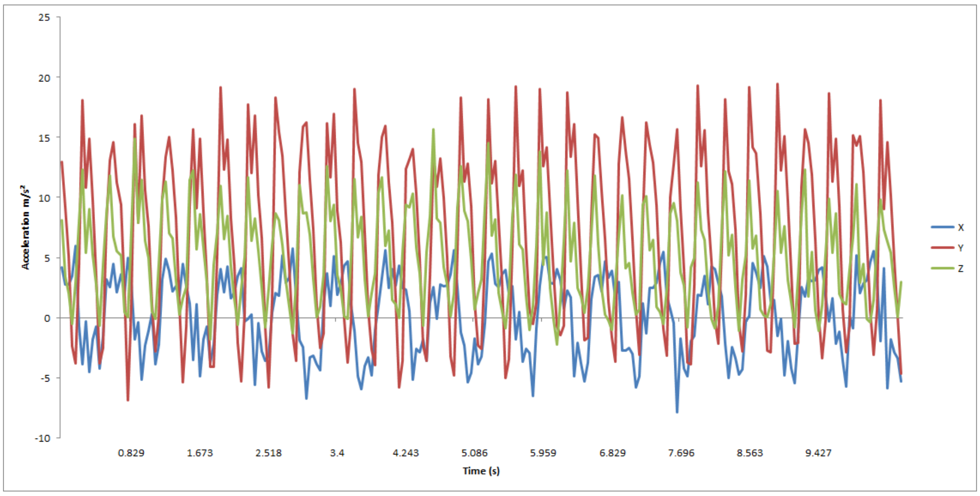

Figure 40: Player Jogging

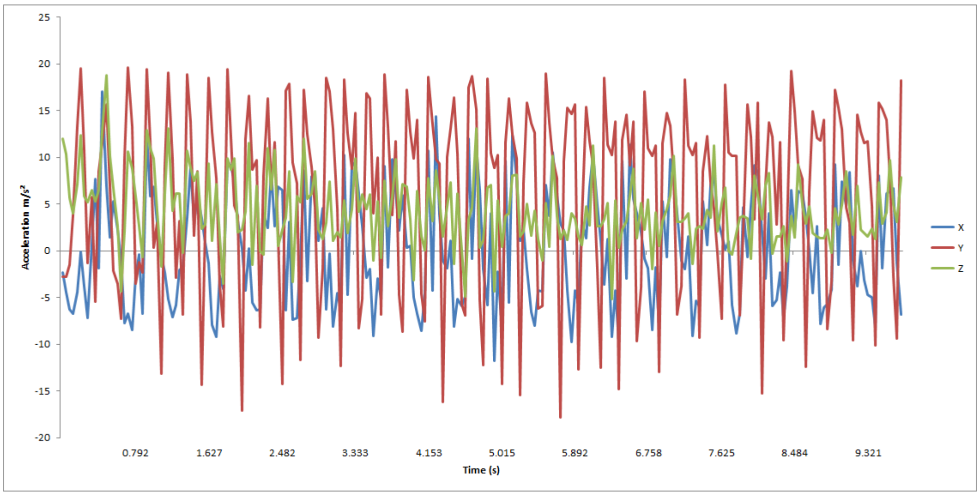

Figure 41: Player Sprinting

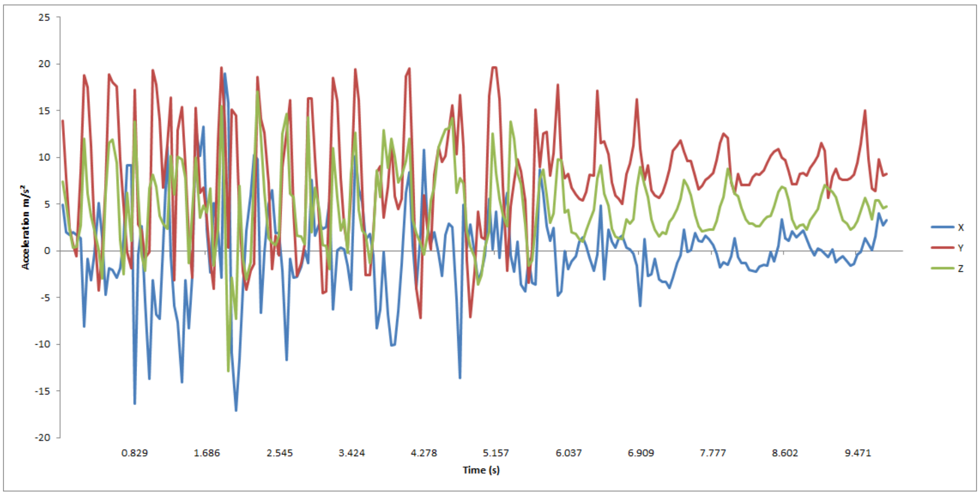

Figure 42: Player Hitting the Ball

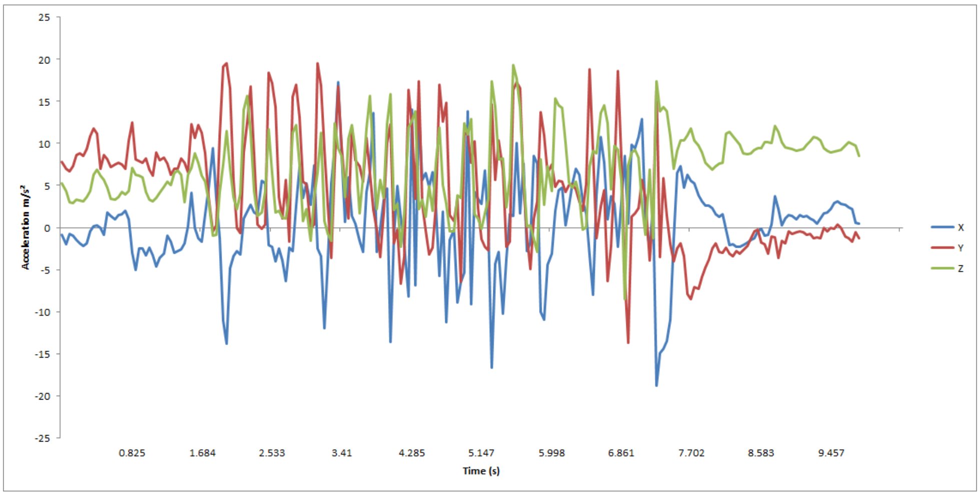

Figure 43: Player Tackling

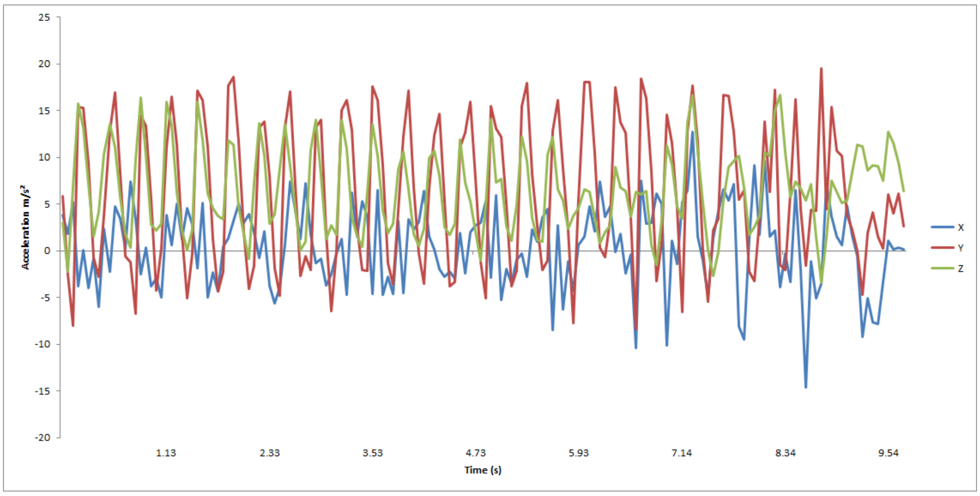

Figure 44: Player Soloing with the Ball

B Sample Signals from Section 4.3.1

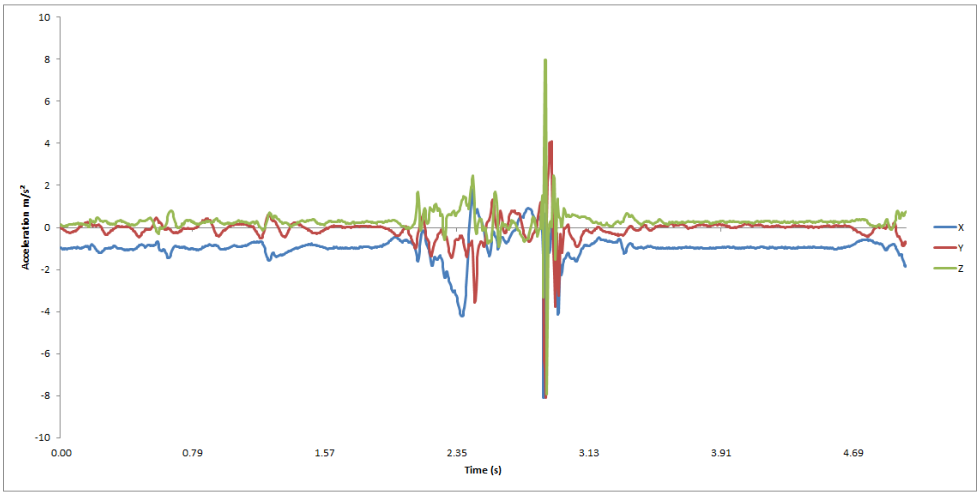

Figure 45: User Jumping on a Box

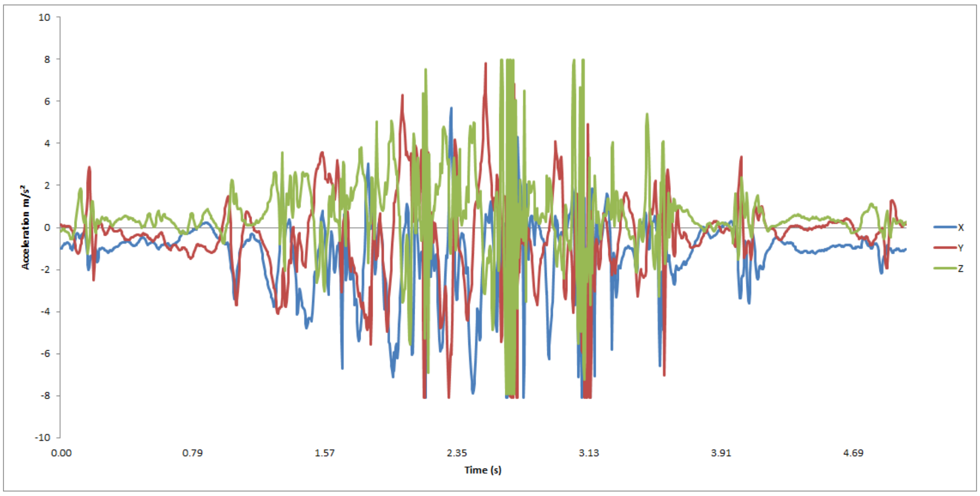

Figure 46: User Sprinting

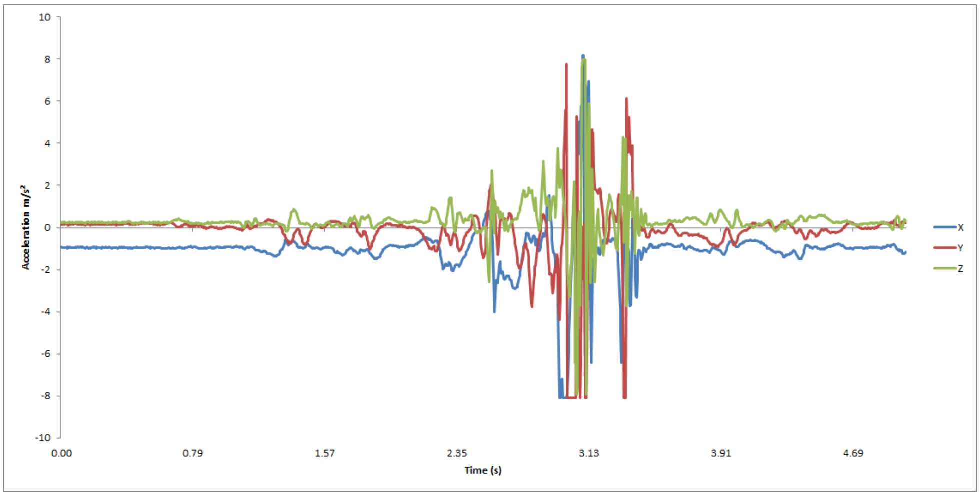

Figure 47: User Hitting the Ball

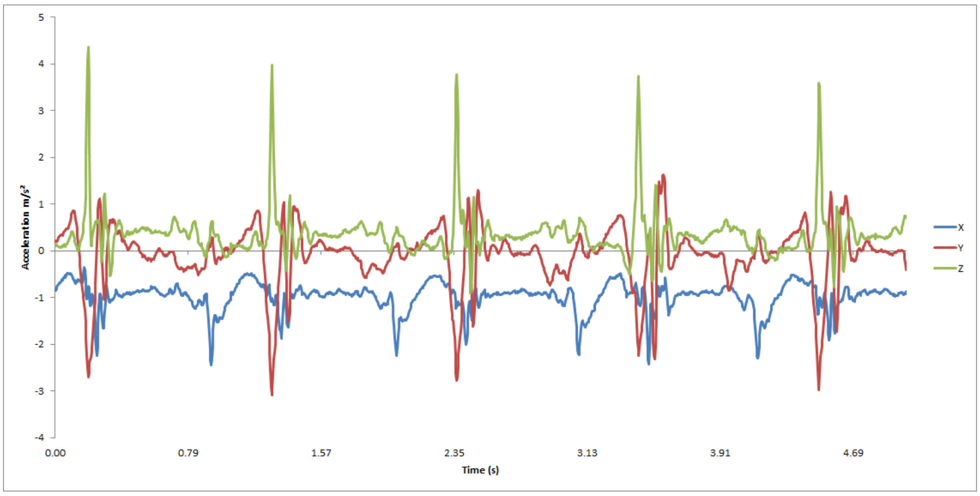

Figure 48: User Walking

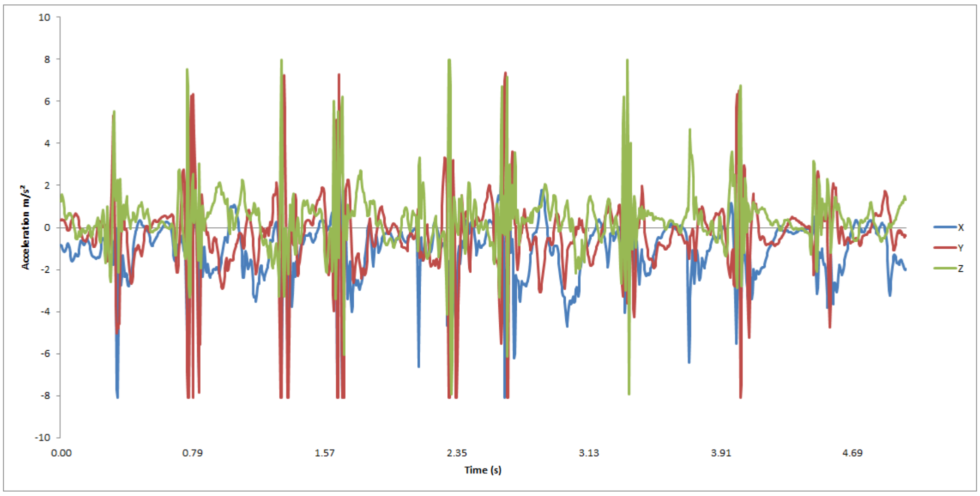

Figure 49: User performing Agility Run

References

[1] E. Mitchell, D. Monaghan, and N. E. O'Connor, "Classification of sporting activities using smartphone accelerometers," *Sensors*, vol. 13, no. 4, pp. 5317–5337, 2013.

[2] H. Johannes, "Inertial sensing of human movement," Ph.D. dissertation, Enschede, October 2002. [Online]. Available: http://doc.utwente.nl/38637/

[3] R. OReilly, A. Khenkin, and K. Harney, "Sonic nirvana: Using mems accelerometers as acoustic pickups in musical instruments," *Analog Dialogue*, vol. 43, 2009.

[4] C. Latulipe, D. Wilson, S. Huskey, M. Word, A. Carroll, E. Carroll, B. Gonzalez, V. Singh, M. Wirth, and D. Lottridge, "Exploring the design space in technology-augmented dance," in *CHI'10 Extended Abstracts on Human Factors in Computing Systems.* ACM, 2010, pp. 2995–3000.

[5] M. Weiser, "The computer for the 21st century," *Scientific american*, vol. 265, no. 3, pp. 94–104, 1991.

[6] S. Butterworth, "On the theory of filter amplifiers," *Wireless Engineer*, vol. 7, pp. 536–541, 1930.

[7] F. Foerster, M. Smeja, and J. Fahrenberg, "Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring," *Computers in Human Behavior*, vol. 15, no. 5, pp. 571–583, 1999.

[8] K. Aminian, P. Robert, E. Buchser, B. Rutschmann, D. Hayoz, and M. Depairon, "Physical activity monitoring based on accelerometry: validation and comparison with video observation," *Medical & biological engineering & computing*, vol. 37, no. 3, pp. 304–308, 1999.

[9] A. Sugimoto, Y. Hara, T. Findley, and K. Yoncmoto, "A useful method for measuring daily physical activity by a three-direction monitor." *Scandinavian Journal of Rehabilitation Medicine*, vol. 29, no. 1, pp. 37–42, 1997.

[10] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive computing.* Springer, 2004, pp. 1–17.

[11] L. Bruce, C. Koger, and J. Li, "Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 40, no. 10, pp. 2331–2338, 2002.

[12] S. Pittner and S. Kamarthi, "Feature extraction from wavelet coefficients for pattern recognition tasks," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 21, no. 1, pp. 83–88, 1999.

[13] S. Preece, J. Goulermas, L. Kenney, and D. Howard, "A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data," *Biomedical Engineering, IEEE Transactions on*, vol. 56, no. 3, pp. 871–879, 2009.

[14] A. N. Akansu, W. A. Serdijn, and I. W. Selesnick, "Emerging applications of wavelets: A review," *Physical communication*, vol. 3, no. 1, pp. 1–18, 2010.

[15] MATLAB, *version 7.10.0 (R2010a).* Natick, Massachusetts: The MathWorks Inc., 2010.

[16] B. Ayrulu-Erdem and B. Barshan, "Leg motion classification with artificial neural networks using wavelet-based features of gyroscope signals," *Sensors*, vol. 11, no. 2, pp. 1721–1743, 2011.

[17] N. Saravanan and K. Ramachandran, "Incipient gear box fault diagnosis using discrete wavelet transform (dwt) for feature extraction and classification using artificial neural network (ann)," *Expert Systems with Applications*, vol. 37, no. 6, pp. 4168–4181, 2010.

[18] Q. Sun, Y. Tang, W. Y. Lu, and Y. Ji, "Feature extraction with discrete wavelet transform for drill wear monitoring," *Journal of Vibration and Control*, vol. 11, no. 11, pp. 1375–1396, 2005.

[19] H. C. Choe, Y. Wan, and A. K. Chan, "Neural pattern identification of railroad wheel-bearing faults from audible acoustic signals: Comparison of fft, cwt, and dwt features," in *AeroSense'97*. International Society for Optics and Photonics, 1997, pp. 480–496.

[20] G. K. Prasad and J. Sahambi, "Classification of ecg arrhythmias using multi-resolution analysis and neural networks," in *TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region*, vol. 1. IEEE, 2003, pp. 227–231.

[21] U. R. Acharya, M. Sankaranarayanan, J. Nayak, C. Xiang, and T. Tamura, "Automatic identification of cardiac health using modeling techniques: A comparative study," *Information Sciences*, vol. 178, no. 23, pp. 4571–4582, 2008.

[22] M. Nyan, F. E. Tay, M. Manimaran, and K. Seah, "Garment-based detection of falls and activities of daily living using 3-axis mems accelerometer," in *Journal of Physics: Conference Series*, vol. 34, no. 1. IOP Publishing, 2006, p. 1059.

[23] D. James, N. Davey, and T. Rice, "An accelerometer based sensor platform for insitu elite athlete performance analysis," in *Sensors, 2004. Proceedings of IEEE*. IEEE, 2004, pp. 1373–1376.

[24] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 10, no. 1, pp. 156–167, 2006.

[25] A. Khan, Y. Lee, S. Lee, and T. Kim, "Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis," in *Future Information Technology (FutureTech), 2010 5th International Conference on.* IEEE, 2010, pp. 1–6.

[26] X. Long, B. Yin, and R. M. Aarts, "Single-accelerometer-based daily physical activity classification," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE.* IEEE, 2009, pp. 6107–6110.

[27] Y. Ohgi, H. Ichikawa, and C. Miyaji, "Microcomputer-based acceleration sensor device for swimming stroke monitoring." *JSME International Journal Series C*, vol. 45, no. 4, pp. 960–966, 2002.

[28] M. Parvis and F. Ferraris, "Procedure for effortless in-field calibration of three-axial rate gyro and accelerometers," *Sensors and Materials*, vol. 7, no. 5, pp. 311–330, 1995.

[29] I. Kaur *et al.*, "Electrocardiogram signal analysis-an overview." *International Journal of Computer Applications*, vol. 84, 2013.

[30] R. Kitney and O. Rompelman, *The study of heart-rate variability.* Oxford University Press, USA, 1980, no. 3.

[31] T. Choudhury, M. Philipose, D. Wyatt, and J. Lester, "Towards activity databases: Using sensors and statistical models to summarize people's lives." *IEEE Data Eng. Bull.*, vol. 29, no. 1, pp. 49–58, 2006.

[32] K. J. Liszka, M. A. Mackin, M. J. Lichter, D. W. York, D. Pillai, and D. S. Rosenbaum, "Keeping a beat on the heart," *IEEE Pervasive Computing*, vol. 3, no. 4, pp. 42–49, 2004.

[33] U. Anliker, J. A. Ward, P. Lukowicz, G. Troster, F. Dolveck, M. Baer, F. Keita, E. B. Schenker, F. Catarsi, L. Coluccini *et al.*, "Amon: a wearable multiparameter medical monitoring and alert system," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 8, no. 4, pp. 415–427, 2004.

[34] J. Maitland, S. Sherwood, L. Barkhuus, I. Anderson, M. Hall, B. Brown, M. Chalmers, and H. Muller, "Increasing the awareness of daily activity levels with pervasive computing," in *Pervasive Health Conference and Workshops, 2006.* IEEE, 2006, pp. 1–9.

[35] S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby *et al.*, "Activity sensing in the wild: a field trial of ubifit garden," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2008, pp. 1797–1806.

[36] A. Andrew, Y. Anokwa, K. Koscher, J. Lester, and G. Borriello, "Context to make you more aware," in *Distributed Computing Systems Workshops, 2007. ICDCSW'07. 27th International Conference on.* IEEE, 2007, pp. 49–49.

[37] J. J. Oresko, Z. Jin, J. Cheng, S. Huang, Y. Sun, H. Duschl, and A. C. Cheng, "A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 14, no. 3, pp. 734–740, 2010.

[38] I. Lee, G. J. Pappas, R. Cleaveland, J. Hatcliff, B. H. Krogh, P. Lee, H. Rubin, and L. Sha, "High-confidence medical device software and systems," *Computer*, vol. 39, no. 4, pp. 33–38, 2006.

[39] I. P. Pappas, T. Keller, S. Mangold, M. R. Popovic, V. Dietz, and M. Morari, "A reliable gyroscope-based gait-phase detection sensor embedded in a shoe insole," *Sensors Journal, IEEE*, vol. 4, no. 2, pp. 268–274, 2004.

[40] N. Kern, B. Schiele, and A. Schmidt, "Multi-sensor activity context detection for wearable computing," in *Ambient Intelligence*. Springer, 2003, pp. 220–232.

[41] C. Glaros, D. Fotiadis, A. Likas, and A. Stafylopatis, "A wearable intelligent system for monitoring health condition and rehabilitation of running athletes," in *Information Technology Applications in Biomedicine, 2003. 4th International IEEE EMBS Special Topic Conference on*. IEEE, 2003, pp. 276–279.

[42] M. Ermes, J. Parkka, J. Mantyjarvi, and I. Korhonen, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 12, no. 1, pp. 20–26, 2008.

[43] E. A. Heinz, K. S. Kunze, M. Gruber, D. Bannach, and P. Lukowicz, "Using wearable sensors for real-time recognition tasks in games of martial arts-an initial experiment," in *Computational Intelligence and Games, 2006 IEEE Symposium on*. IEEE, 2006, pp. 98–102.

[44] D. Spelmezan and J. Borchers, "Real-time snowboard training system," in *CHI'08 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2008, pp. 3327–3332.

[45] H. Ghasemzadeh, V. Loseu, and R. Jafari, “Wearable coach for sport training: A quantitative model to evaluate wrist-rotation in golf,” *Journal of Ambient Intelligence and Smart Environments*, vol. 1, no. 2, pp. 173–184, 2009.

[46] D. Arvind and A. Bates, “The speckled golfer,” in *Proceedings of the ICST 3rd international conference on Body area networks.* ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008, p. 28.

[47] V. Stanford, “Wearable computing goes live in industry,” *Pervasive Computing, IEEE*, vol. 1, no. 4, pp. 14–19, 2002.

[48] P. Lukowicz, A. Timm-Giel, M. Lawo, and O. Herzog, “Wearit@ work: Toward real-world industrial wearable computing,” *Pervasive Computing, IEEE*, vol. 6, no. 4, pp. 8–13, 2007.

[49] J. A. Ward, P. Lukowicz, G. Troster, and T. E. Starner, “Activity recognition of assembly tasks using body-worn microphones and accelerometers,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 10, pp. 1553–1567, 2006.

[50] M. Gowing, P. Kell, N. E. O’Connor, C. Concolato, S. Essid, J. Lefeuvre, R. Tournemenne, E. Izquierdo, V. Kitanovski, X. Lin *et al.*, “Enhanced visualisation of dance performance from automatically synchronised multimodal recordings,” in *Proceedings of the 19th ACM international conference on Multimedia.* ACM, 2011, pp. 667–670.

[51] M. Barry, J. Gutknecht, I. Kulka, P. Lukowicz, and T. Stricker, “From motion to emotion: a wearable system for the multimedia enrichment of a butoh dace performance,” *Journal of Mobile Multimedia*, vol. 1, no. 2, pp. 112–132, 2005.

[52] R. P. Aylward, "Sensemble: a wireless inertial sensor system for the interactive dance and collective motion analysis," Ph.D. dissertation, Massachusetts Institute of Technology, 2006.

[53] E. Lee, U. Enke, J. Borchers, and L. de Jong, "Towards rhythmic analysis of human motion using acceleration-onset times," in *Proceedings of the 7th international conference on New interfaces for musical expression*. ACM, 2007, pp. 136–141.

[54] H. Zhang and B. Hartmann, "Building upon everyday play," in *CHI'07 extended abstracts on Human factors in computing systems*. ACM, 2007, pp. 2019–2024.

[55] D. Minnen, T. Westeyn, D. Ashbrook, P. Presti, and T. Starner, "Recognizing soldier activities in the field," in *4th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007)*. Springer, 2007, pp. 236–241.

[56] M. Sala, K. Partridge, L. Jacobson, and J. Begole, "An exploration into activity-informed physical advertising using pest," *Pervasive Computing*, pp. 73–90, 2007.

[57] S. Martinez-Conde, S. L. Macknik, and D. H. Hubel, "The role of fixational eye movements in visual perception," *Nature Reviews Neuroscience*, vol. 5, no. 3, pp. 229–240, 2004.

[58] A. Berthoz, *The brain's sense of movement*. Harvard University Press, 2002.

[59] D. Hall, R. Linn, and J. Llinas, "A survey of data fusion systems," in *Proc. SPIE Conf. on data structure and target classification*, vol. 1470, 1991, pp. 13–36.

[60] M. Kam, X. Zhu, and P. Kalata, "Sensor fusion for mobile robot navigation," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 108–119, 1997.

[61] D. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 6–23, 1997.

[62] R. C. Luo and M. G. Kay, "Multisensor integration and fusion in intelligent systems," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 19, no. 5, pp. 901–931, 1989.

[63] F. E. White, "Data fusion lexicon," DTIC Document, Tech. Rep., 1991.

[64] R. Manduchi, "Bayesian fusion of color and texture segmentations," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 2. IEEE, 1999, pp. 956–962.

[65] R. G. Brown, *Introduction to random signal analysis and Kalman filtering.* Wiley New York, 1983, vol. 8.

[66] R. Routledge, *Bayes's theorem.* Encyclopedia Britannica, 2013.

[67] P. Singla and P. Domingos, "Discriminative training of markov logic networks," in *AAAI*, vol. 5, 2005, pp. 868–873.

[68] K. Abou-Moustafa, M. Cheriet, and C. Y. Suen, "Classification of time-series data using a generative/discriminative hybrid," in *Frontiers in Handwriting Recognition, 2004. IWFHR-9 2004. Ninth International Workshop on.* IEEE, 2004, pp. 51–56.

[69] I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques.* Morgan Kaufmann, 2005.

[70] A. Y. Ng and M. I. Jordan, "On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes," *Advances in neural information processing systems*, vol. 2, pp. 841–848, 2002.

[71] K.-W. Yeom and J.-H. Park, "An approach of information extraction from web documents for automatic ontology generation," in *Computational Intelligence and Security.* Springer, 2005, pp. 450–457.

[72] J. Malobabić, H. Le Borgne, N. Murphy, and N. O'Connor, "Detecting the presence of large buildings in natural images," 2005.

[73] Y. Li and C. Dorai, "Svm-based audio classification for instructional video analysis," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on*, vol. 5. IEEE, 2004, pp. V–897.

[74] C. Leslie, E. Eskin, and W. Noble, "The spectrum kernel: A string kernel for svm protein classification," in *Proceedings of the pacific symposium on biocomputing*, vol. 7. Hawaii, USA., 2002, pp. 566–575.

[75] Y. Kim and H. Ling, "Human activity classification based on micro-doppler signatures using a support vector machine," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 47, no. 5, pp. 1328–1337, 2009.

[76] V. Jakkula, "Tutorial on support vector machine (svm)," *School of EECS, Washington State University*, 2006.

[77] D. Willems, S. Rossignol, and L. Vuurpijl, "Features for mode detection in natural online pen input," in *Proceedings of the 12th Biennial Conference of the International Graphonomics Society*, 2005, pp. 113–117.

[78] W. Shang, H. Huang, H. Zhu, Y. Lin, Y. Qu, and H. Dong, "An adaptive fuzzy knn text classifier," *Computational Science–ICCS 2006*, pp. 216–223, 2006.

[79] C. Christodoulou, S. Michaelides, and C. Pattichis, "Multifeature texture analysis for the classification of clouds in satellite imagery," *Geo-*

science and Remote Sensing, IEEE Transactions on, vol. 41, no. 11, pp. 2662–2668, 2003.

[80] S. Karimifard, A. Ahmadian, M. Khoshnevisan, and M. Nambakhsh, "Morphological heart arrhythmia detection using hermitian basis functions and knn classifier," in *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE.* IEEE, 2006, pp. 1367–1370.

[81] W. Wei, S. Visweswaran, and G. Cooper, "The application of naive bayes model averaging to predict alzheimer's disease from genome-wide data," *Journal of the American Medical Informatics Association*, vol. 18, no. 4, pp. 370–375, 2011.

[82] N. Lavrač, "Selected techniques for data mining in medicine," *Artificial intelligence in medicine*, vol. 16, no. 1, pp. 3–23, 1999.

[83] R. Dannenberg, B. Thom, and D. Watson, "A machine learning approach to musical style recognition," 1997.

[84] N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," *Machine Learning*, vol. 95, no. 1-2, pp. 161–205, 2005.

[85] V. Mahesh, A. Kandaswamy, C. Vimal, and B. Sathish, "Ecg arrhythmia classification based on logistic model tree," *Journal of Biomedical Science and Engineering*, vol. 2, no. 6, pp. 405–411, 2009.

[86] E. Agichtein, W. Askew, and Y. Liu, "Combining lexical, syntactic, and semantic evidence for textual entailment classification," *Proceedings of TAC*, vol. 31, 2008.

[87] M. Gardner and S. Dorling, "Artificial neural networks (the multilayer perceptron)–a review of applications in the atmospheric sciences," *Atmospheric environment*, vol. 32, no. 14-15, pp. 2627–2636, 1998.

[88] S. Phung, A. Bouzerdoum Sr, and D. Chai Sr, "Skin segmentation using color pixel classification: analysis and comparison," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 1, pp. 148–154, 2005.

[89] Y. Venkatesh and S. Kumar Raja, "On the classification of multispectral satellite images using the multilayer perceptron," *Pattern Recognition*, vol. 36, no. 9, pp. 2161–2175, 2003.

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

[91] D. E. Goldberg *et al.*, "Genetic algorithms in search, optimization, and machine learning," vol. 412, 1989.

[92] D. Thierens, "Selection schemes, elitist recombination, and selection intensity," in *ICGA*, 1997, pp. 152–159.

[93] C.-L. Huang and C.-J. Wang, "A ga-based feature selection and parameters optimizationfor support vector machines," *Expert Systems with applications*, vol. 31, no. 2, pp. 231–240, 2006.

[94] W. F. Punch III, E. D. Goodman, M. Pei, L. Chia-Shun, P. D. Hovland, and R. J. Enbody, "Further research on feature selection and classification using genetic algorithms." in *ICGA*, 1993, pp. 557–564.

[95] L. I. Kuncheva and L. C. Jain, "Designing classifier fusion systems by genetic algorithms," *Evolutionary Computation, IEEE Transactions on*, vol. 4, no. 4, pp. 327–336, 2000.

[96] D. Siewiorek, A. Smailagic, and T. Starner, "Application design for wearable computing," *Synthesis Lectures on Mobile and Pervasive Computing*, vol. 3, no. 1, pp. 1–66, 2008.

[97] H. Ghasemzadeh, V. Loseu, E. Guenterberg, and R. Jafari, "Sport training using body sensor networks: A statistical approach to measure wrist rotation for golf swing," in *Proceedings of the Fourth International Conference on Body Area Networks*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2009, p. 2.

[98] E. Jovanov, A. Milenkovic, C. Otto, and P. C. De Groen, "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation," *Journal of NeuroEngineering and rehabilitation*, vol. 2, no. 1, p. 6, 2005.

[99] T. Starner, B. Rhodes, J. Weaver, and A. Pentland, "Everyday-use wearable computers," in *International Symposium on Wearable Computers*, 1999, p. 9.

[100] B. Schiele, N. Oliver, T. Jebara, and A. Pentland, "An interactive computer vision system dypers: Dynamic personal enhanced reality system," in *Computer Vision Systems*. Springer, 1999, pp. 51–65.

[101] T. Starner, J. Weaver, and A. Pentland, "A wearable computer based american sign language recognizer," in *Assistive Technology and Artificial Intelligence*. Springer, 1998, pp. 84–96.

[102] B. Clarkson and A. Pentland, "Extracting context from environmental audio," in *Wearable Computers, 1998. Digest of Papers. Second International Symposium on*. IEEE, 1998, pp. 154–155.

[103] B. J. Rhodes, "The wearable remembrance agent: A system for augmented memory," *Personal Technologies*, vol. 1, no. 4, pp. 218–224, 1997.

[104] H. J. Montoye, R. Washburn, S. Servais, A. Ertl, J. G. Webster, and F. J. Nagle, "Estimation of energy expenditure by a portable accelerometer." *Medicine and Science in Sports and Exercise*, vol. 15, no. 5, pp. 403–407, 1982.

[105] T. C. Wong, J. G. Webster, H. J. Montoye, and R. Washburn, "Portable accelerometer device for measuring human energy expenditure," *Biomedical Engineering, IEEE Transactions on*, no. 6, pp. 467–471, 1981.

[106] C. Randell and H. Muller, "Context awareness by analysing accelerometer data," in *Wearable Computers, The Fourth International Symposium on.* IEEE, 2000, pp. 175–176.

[107] K. Van Laerhoven and O. Cakmakci, "What shall we teach our pants?" in *Wearable Computers, The Fourth International Symposium on.* IEEE, 2000, pp. 77–83.

[108] A. R. Golding and N. Lesh, "Indoor navigation using a diverse set of cheap, wearable sensors," in *Wearable Computers, 1999. Digest of Papers. The Third International Symposium on.* IEEE, 1999, pp. 29–36.

[109] J. Lester, T. Choudhury, and G. Borriello, "A practical approach to recognizing physical activities," in *Pervasive Computing.* Springer, 2006, pp. 1–16.

[110] D. Minnen, T. Starner, I. Essa, and C. Isbell, "Discovering characteristic actions from on-body sensor data," in *Wearable computers, 2006 10th IEEE international symposium on.* IEEE, 2006, pp. 11–18.

[111] D. Wyatt, M. Philipose, and T. Choudhury, "Unsupervised activity recognition using automatically mined common sense," in *AAAI*, vol. 5, 2005, pp. 21–27.

[112] T. Huynh and B. Schiele, "Unsupervised discovery of structure in activity data using multiple eigenspaces," in *Location-and Context-Awareness.* Springer, 2006, pp. 151–167.

[113] T. Stiefmeier, G. Ogris, H. Junker, P. Lukowicz, and G. Troster, "Combining motion sensors and ultrasonic hands tracking for continuous activity recognition in a maintenance scenario," in *Wearable Computers, 2006 10th IEEE International Symposium on.* IEEE, 2006, pp. 97–104.

[114] N. Bidargaddi, A. Sarela, L. Klingbeil, and M. Karunanithi, "Detecting walking activity in cardiac rehabilitation by using accelerometer," in *Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007. 3rd International Conference on.* IEEE, 2007, pp. 555–560.

[115] N. C. Krishnan, D. Colbry, C. Juillard, and S. Panchanathan, "Real time human activity recognition using tri-axial accelerometers," in *Sensors, signals and information processing workshop*, 2008.

[116] J.-Y. Yang, J.-S. Wang, and Y.-P. Chen, "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers," *Pattern recognition letters*, vol. 29, no. 16, pp. 2213–2220, 2008.

[117] C. Ó Conaire, D. Connaghan, P. Kelly, N. E. O'Connor, M. Gaffney, and J. Buckley, "Combining inertial and visual sensing for human action

recognition in tennis," in *Proceedings of the first ACM international workshop on Analysis and retrieval of tracked events and motion in imagery streams.* ACM, 2010, pp. 51–56.

[118] Z. He, L. Jin, L. Zhen, and J. Huang, "Gesture recognition based on 3d accelerometer for cell phones interaction," in *Circuits and Systems, 2008. APCCAS 2008. IEEE Asia Pacific Conference on.* IEEE, 2008, pp. 217–220.

[119] J. Wu, G. Pan, D. Zhang, G. Qi, and S. Li, "Gesture recognition with a 3-d accelerometer," in *Ubiquitous intelligence and computing.* Springer, 2009, pp. 25–38.

[120] J. A. Ward, P. Lukowicz, and G. Tröster, "Gesture spotting using wrist worn microphone and 3-axis accelerometer," in *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies.* ACM, 2005, pp. 99–104.

[121] S. Pirttikangas, K. Fujinami, and T. Nakajima, "Feature selection and activity recognition from wearable sensors," in *Ubiquitous Computing Systems.* Springer, 2006, pp. 516–527.

[122] C. Lombriser, N. B. Bharatula, D. Roggen, and G. Tröster, "On-body activity recognition in a dynamic sensor network," in *Proceedings of the ICST 2nd international conference on Body area networks.* ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2007, p. 17.

[123] S. Lee, H. Park, S. Hong, K. Lee, and Y. Kim, "A study on the activity classification using a triaxial accelerometer," in *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, vol. 3. IEEE, 2003, pp. 2941–2943.

[124] F. Carpi and D. De Rossi, "Electroactive polymer-based devices for e-textiles in biomedicine," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 9, no. 3, pp. 295–318, 2005.

[125] D. Bradley, "Patterns of breathing dysfunction in hyperventilation syndrome and breathing pattern disorders," *Multidisciplinary approaches to breathing pattern disorders. London: Churchill Livingstone*, 2002.

[126] L. Garza and S. J. Ford, "Implementing a breathing technique to manage performance anxiety in softball." *Sport Journal*, vol. 12, no. 2, 2009.

[127] S. C. FitzSimmons, "The changing epidemiology of cystic fibrosis," *The Journal of pediatrics*, vol. 122, no. 1, pp. 1–9, 1993.

[128] A. Hristara-Papadopoulou and J. Tsanakas, "Results of active cycle of breathing techniques and conventional physiotherapy in mucociliary clearance in children with cystic fibrosis," *Hippokratia*, vol. 11, no. 4, p. 202, 2007.

[129] S. Enright, K. Chatham, A. A. Ionescu, V. B. Unnithan, and D. J. Shale, "Inspiratory muscle training improves lung function and exercise capacity in adults with cystic fibrosis," *CHEST Journal*, vol. 126, no. 2, pp. 405–411, 2004.

[130] S. Coyle, D. Morris, K.-T. Lau, D. Diamond, and N. Moyna, "Textile-based wearable sensors for assisting sports performance," in *Wearable and Implantable Body Sensor Networks, 2009. BSN 2009. Sixth International Workshop on.* IEEE, 2009, pp. 307–311.

[131] C. E. Association, "14th annual ce ownership and market potential study," Tech. Rep., 2012.

[132] ABIresearch, "Mobile device user interfaces - touchscreens, software, accelerometers, haptics, and sensors," Electronic, ABIresearch, http://www.abiresearch.com/research/1003813, 3Q 2010.

[133] S. L. Lau, I. Konig, K. David, B. Parandian, C. Carius-Dussel, and M. Schultz, "Supporting patient monitoring using activity recognition with a smartphone," in *Wireless Communication Systems (ISWCS), 2010 7th International Symposium on.* IEEE, 2010, pp. 810–814.

[134] Y.-S. Lee and S.-B. Cho, "Activity recognition using hierarchical hidden markov models on a smartphone with 3d accelerometer," in *Hybrid Artificial Intelligent Systems.* Springer, 2011, pp. 460–467.

[135] S. Zhang, P. McCullagh, C. Nugent, and H. Zheng, "Activity monitoring using a smart phone's accelerometer with hierarchical classification," in *Intelligent Environments (IE), 2010 Sixth International Conference on.* IEEE, 2010, pp. 158–163.

[136] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, "Activity classification using realistic data from wearable sensors," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 10, no. 1, pp. 119–128, 2006.

[137] N. V. Chawla, N. Japkowicz, and A. Kotcz, "Editorial: special issue on learning from imbalanced data sets," *ACM Sigkdd Explorations Newsletter*, vol. 6, no. 1, pp. 1–6, 2004.

[138] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SigKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.

[139] Y. Xue and L. Jin, “A naturalistic 3d acceleration-based activity dataset & benchmark evaluations,” in *Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 4081–4085.

[140] C.-W. Hsu, C.-C. Chang, C.-J. Lin *et al.*, “A practical guide to support vector classification,” 2003.

[141] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, “Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine,” in *Ambient assisted living and home care*. Springer, 2012, pp. 216–223.

[142] A. Fleury, M. Vacher, and N. Noury, “Svm-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results,” *Information Technology in Biomedicine, IEEE Transactions on*, vol. 14, no. 2, pp. 274–283, 2010.

[143] I. Rish, “An empirical study of the naive bayes classifier,” in *IJCAI 2001 workshop on empirical methods in artificial intelligence*, vol. 3, no. 22, 2001, pp. 41–46.

[144] C. Zhu and W. Sheng, “Human daily activity recognition in robot-assisted living using multi-sensor fusion,” in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 2154–2159.

[145] L. K. Au, M. A. Batalin, T. Stathopoulos, A. A. Bui, and W. J. Kaiser, “Episodic sampling: Towards energy-efficient patient monitoring with wearable sensors,” in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*. IEEE, 2009, pp. 6901–6905.

[146] L. K. Au, W. H. Wu, M. A. Batalin, D. Mclntire, and W. J. Kaiser, "Microleap: Energy-aware wireless sensor platform for biomedical sensing applications," in *Biomedical Circuits and Systems Conference, 2007. BIOCAS 2007. IEEE.* IEEE, 2007, pp. 158–162.

[147] C. G. Snoek, M. Worring, and A. W. Smeulders, "Early versus late fusion in semantic video analysis," in *Proceedings of the 13th annual ACM international conference on Multimedia.* ACM, 2005, pp. 399–402.

[148] T. Westerveld, A. P. De Vries, A. van Ballegooij, F. de Jong, and D. Hiemstra, "A probabilistic multimedia retrieval model and its evaluation," *EURASIP Journal on Applied Signal Processing*, vol. 2003, pp. 186–198, 2003.

[149] D. E. Warburton, C. W. Nicol, and S. S. Bredin, "Health benefits of physical activity: the evidence," *Canadian medical association journal*, vol. 174, no. 6, pp. 801–809, 2006.

[150] J. L. Dragoo, H. J. Braun, J. L. Durham, M. R. Chen, and A. H. Harris, "Incidence and risk factors for injuries to the anterior cruciate ligament in national collegiate athletic association football data from the 2004-2005 through 2008-2009 national collegiate athletic association injury surveillance system," *The American journal of sports medicine*, vol. 40, no. 5, pp. 990–995, 2012.

[151] A. Hreljac, "Impact and overuse injuries in runners," *Medicine and science in sports and exercise*, vol. 36, no. 5, pp. 845–849, 2004.

[152] B. R. van Gent, D. D. Siem, M. van Middelkoop, T. A. van Os, S. S. Bierma-Zeinstra, and B. B. Koes, "Incidence and determinants of lower extremity running injuries in long distance runners: a systematic review," *British journal of sports medicine*, 2007.

[153] B. T. Zazulak, T. E. Hewett, N. P. Reeves, B. Goldberg, and J. Cholewicki, “Deficits in neuromuscular control of the trunk predict knee injury risk a prospective biomechanical-epidemiologic study,” *The American journal of sports medicine*, vol. 35, no. 7, pp. 1123–1130, 2007.

[154] K. A. Moran and B. M. Marshall, “Effect of fatigue on tibial impact accelerations and knee kinematics in drop jumps.” *Medicine and science in sports and exercise*, vol. 38, no. 10, pp. 1836–1842, 2006.

[155] J. Iguchi, H. Tateuchi, M. Taniguchi, and N. Ichihashi, “The effect of sex and fatigue on lower limb kinematics, kinetics, and muscle activity during unanticipated side-step cutting,” *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 22, no. 1, pp. 41–48, 2014.

[156] D. A. Winter, *Biomechanics and motor control of human movement.* John Wiley & Sons, 2009.

[157] S. O. Madgwick, A. J. Harrison, and R. Vaidyanathan, “Estimation of imu and marg orientation using a gradient descent algorithm,” in *Rehabilitation Robotics (ICORR), 2011 IEEE International Conference on.* IEEE, 2011, pp. 1–7.

[158] A. M. Sabatini, “Quaternion-based extended kalman filter for determining orientation by inertial and magnetic sensing,” *Biomedical Engineering, IEEE Transactions on*, vol. 53, no. 7, pp. 1346–1356, 2006.

[159] A. K. Stuifbergen and S. Rogers, “Health promotion: An essential component of rehabilitation for persons with chronic disabling conditions,” *Advances in Nursing Science*, vol. 19, no. 4, pp. 1–20, 1997.

[160] F.-C. Yen, K. Behbehani, E. A. Lucas, J. R. Burk, and J. R. Axe, “A noninvasive technique for detecting obstructive and central sleep

apnea," *Biomedical Engineering, IEEE Transactions on*, vol. 44, no. 12, pp. 1262–1268, 1997.

[161] S. Horstmann, C. Hess, C. Bassetti, M. Gugger, and J. Mathis, "Sleepiness-related accidents in sleep apnea patients," *SLEEP-NEW YORK-*, vol. 23, no. 3, pp. 383–392, 2000.

[162] J. Teran-Santos, A. Jimenez-Gomez, and J. Cordero-Guevara, "The association between sleep apnea and the risk of traffic accidents," *New England Journal of Medicine*, vol. 340, no. 11, pp. 847–851, 1999.

[163] T. D. Bradley and J. S. Floras, "Sleep apnea and heart failure part i: obstructive sleep apnea," *Circulation*, vol. 107, no. 12, pp. 1671–1678, 2003.

[164] U. D. O. HEALTH, H. SERVICES *et al.*, *Wake Up America National Sleep Alert*, 1993.

[165] W. W. Flemons, M. R. Littner, J. A. Rowley, P. Gay, W. M. Anderson, D. W. Hudgel, R. D. McEvoy, and D. I. Loube, "Home diagnosis of sleep apnea: A systematic review of the literaturean evidence review cosponsored by the american academy of sleep medicine, the american college of chest physicians, and the american thoracic society," *CHEST Journal*, vol. 124, no. 4, pp. 1543–1579, 2003.

[166] S. D. Ross, I. Sheinhait, K. J. Harrison, M. Kvasz, J. E. Connelly, S. A. Shea, and I. E. Allen, "Systematic review and meta-analysis of the literature regarding the diagnosis of sleep apnea." *Sleep*, vol. 23, no. 4, pp. 519–532, 2000.

[167] J. Gubbi, A. Khandoker, and M. Palaniswami, "Classification of sleep apnea types using wavelet packet analysis of short-term ecg signals,"

Journal of clinical monitoring and computing, vol. 26, no. 1, pp. 1–11, 2012.

[168] B. Yιlmaz, M. H. Asyalι, E. Arιkan, S. Yetkin, and F. Özgen, "Sleep stage and obstructive apneaic epoch classification using single-lead ecg," *Biomedical engineering online*, vol. 9, p. 39, 2010.

[169] D. S. Morillo, J. L. R. Ojeda, L. F. C. Foix, and A. L. Jiménez, "An accelerometer-based device for sleep apnea screening," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 14, no. 2, pp. 491–499, 2010.

[170] Wearable wellness system (wws). [Accessed: 1 Feb 2013]. [Online]. Available: http://www.smartex.it/index.php/en/products/wearable-wellness-system

[171] R. Paradiso, G. Loriga, and N. Taccini, "A wearable health care system based on knitted integrated sensors," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 9, no. 3, pp. 337–344, 2005.

[172] K.-P. Hoffmann and R. Ruff, "Flexible dry surface-electrodes for ecg long-term monitoring," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE.* IEEE, 2007, pp. 5739–5742.

[173] L. Hejjel and L. Kellenyi, "The corner frequencies of the ecg amplifier for heart rate variability analysis," *Physiological measurement*, vol. 26, no. 1, p. 39, 2005.

[174] A. B. Sankar, D. Kumar, and K. Seethalakshmi, "Enhanced method for extracting features of respiratory signals and detection of obstructive sleep apnea using threshold based automatic classification algorithm,"

International Journal of Computer Science & Emerging Technologies, vol. 1, no. 4, pp. 38–43, 2010.

[175] A. J. Block, P. G. Boysen, J. W. Wynne, and L. A. Hunt, "Sleep apnea, hypopnea and oxygen desaturation in normal subjects: a strong male predominance," *New England Journal of Medicine*, vol. 300, no. 10, pp. 513–517, 1979.

[176] I. Daubechies *et al.*, *Ten lectures on wavelets.* SIAM, 1992, vol. 61.

[177] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE transactions on systems, man, and cybernetics*, vol. 21, no. 3, pp. 660–674, 1991.

[178] F. Duman, A. Erdamar, O. Erogul, Z. Telatar, and S. Yetkin, "Efficient sleep spindle detection algorithm with decision tree," *Expert Systems with Applications*, vol. 36, no. 6, pp. 9980–9985, 2009.

[179] M. Wysocki, C. Cracco, A. Teixeira, A. Mercat, J.-L. Diehl, Y. Lefort, J.-P. Derenne, and T. Similowski, "Reduced breathing variability as a predictor of unsuccessful patient separation from mechanical ventilation*," *Critical care medicine*, vol. 34, no. 8, pp. 2076–2083, 2006.

[180] P. De Chazal, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. O'Malley, "Automatic classification of sleep apnea epochs using the electrocardiogram," in *Computers in Cardiology 2000.* IEEE, 2000, pp. 745–748.

[181] R. Leardi, "Genetic algorithms in chemometrics and chemistry: a review," *Journal of chemometrics*, vol. 15, no. 7, pp. 559–569, 2001.

[182] M. Tiwari and N. Vidyarthi, "Solving machine loading problems in a flexible manufacturing system using a genetic algorithm based heuris-

tic approach," *International Journal of Production Research*, vol. 38, no. 14, pp. 3357–3384, 2000.

[183] M. Gen and R. Cheng, "Genetic algorithms and engineering optimization," 2000.

[184] K. Safarzyńska and J. C. van den Bergh, *Evolutionary models in economics: a survey of methods and building blocks.* Springer, 2010, vol. 20, no. 3.

[185] A. Chaturvedula, M. E. Sale, and H. Lee, *Genetic algorithm guided population pharmacokinetic model development for simvastatin, concurrently or non-concurrently co-administered with amlodipine.* Wiley Online Library, 2014, vol. 54, no. 2.

[186] U. Maulik, S. Bandyopadhyay, and A. Mukhopadhyay, "Multiobjective genetic algorithms for clustering: Applications in data mining and bioinformatics," 2011.

[187] T. Vidal, T. G. Crainic, M. Gendreau, N. Lahrichi, and W. Rei, "A hybrid genetic algorithm for multidepot and periodic vehicle routing problems," 2012.

[188] R. Helaers and M. C. Milinkovitch, "Metapiga v2. 0: maximum likelihood large phylogeny estimation using the metapopulation genetic algorithm and other stochastic heuristics," *BMC bioinformatics*, vol. 11, no. 1, p. 379, 2010.

[189] A. Gorbenko and V. Popov, "The force law design of artificial physics optimization for robot anticipation of motion," *Advanced Studies in Theoretical Physics*, vol. 6, no. 13, pp. 625–628, 2012.

[190] J.-F. Li and J. Peng, *Task scheduling algorithm based on improved genetic algorithm in cloud computing environment.* Chengdu Institute of

Computer Applications,— a 16 Donghuangchenggen North Street— c Beijing— z 100717— e bibomputerapplications. com. c— u www. computerapplications. com. cn, 2011, vol. 31, no. 1.

[191] M. Mitchell, *An introduction to genetic algorithms.* MIT press, 1998.

[192] B. Chakraborty and P. Chaudhuri, "On the use of genetic algorithm with elitism in robust and nonparametric multivariate analysis," *Austrian Journal of statistics*, vol. 32, no. 1-2, pp. 13–27, 2003.

[193] W.-Y. Lin, W.-Y. Lee, and T.-P. Hong, "Adapting crossover and mutation rates in genetic algorithms," *J. Inf. Sci. Eng.*, vol. 19, no. 5, pp. 889–903, 2003.