TOPICAL REVIEW

Activity identification using body-mounted sensors—a review of classification techniques

To cite this article: Stephen J Preece et al 2009 Physiol. Meas. **30** R1

View the article online for updates and enhancements.

Related content

- Review: Accelerometry
  Merryn J Mathie, Adelle C F Coster, Nigel H Lovell et al.

- Classification of human movement using accelerometry
  Felicity R Allen, Eliathamby Ambikairajah, Nigel H Lovell et al.

- Comparative evaluation of features and techniques for identifying activity type and estimating energy cost from accelerometer data
  Rohit J Kate, Ann M Swartz, Whitney A Welch et al.

Recent citations

- Biradar Shilpa et al

- Fall Detection Using Deep Learning in Range-Doppler Radars
  Branka Jokanovic and Moeness Amin

- A Novel Walking Detection and Step Counting Algorithm Using Unconstrained Smartphones
  Xiaomin Kang et al

TOPICAL REVIEW

Activity identification using body-mounted sensors—a review of classification techniques

Stephen J Preece[1,4], **John Y Goulermas**[2], **Laurence P J Kenney**[1],
Dave Howard[1], **Kenneth Meijer**[3] and **Robin Crompton**[2]

[1] Centre for Rehabilitation and Human Performance Research, University of Salford, Salford,
Greater Manchester, UK
[2] University of Liverpool, Liverpool, UK
[3] University of Maastricht, Maastricht, The Netherlands

E-mail: s.preece@salford.ac.uk, l.p.j.kenney@salford.ac.uk, d.howard@salford.ac.uk,
j.y.goulermas@liverpool.ac.uk, rhcromp@liverpool.ac.uk and Kenneth.Meijer@BW.unimaas.nl

Abstract
With the advent of miniaturized sensing technology, which can be body-worn, it
is now possible to collect and store data on different aspects of human movement
under the conditions of free living. This technology has the potential to be used
in automated activity profiling systems which produce a continuous record of
activity patterns over extended periods of time. Such activity profiling systems
are dependent on classification algorithms which can effectively interpret body-
worn sensor data and identify different activities. This article reviews the
different techniques which have been used to classify normal activities and/or
identify falls from body-worn sensor data. The review is structured according
to the different analytical techniques and illustrates the variety of approaches
which have previously been applied in this field. Although significant progress
has been made in this important area, there is still significant scope for further
work, particularly in the application of advanced classification techniques to
problems involving many different activities.

Keywords: activity monitoring, classification, fall detection, machine
learning

1. Introduction

Physical activity has been defined as 'any bodily movement produced by skeletal muscles
that results in energy expenditure above resting level' (Caspersen *et al* 1985). Activity

---

classification is a recent concept involving the use of technology to automatically recognize different activities and, in some cases, to collate this information into a continuous record. The need for automated activity classification systems has been identified in a number of different fields, from health-related research to pervasive computing, as discussed below.

With the shift towards more sedentary lifestyles in both developed and developing nations, there is a need for research investigating the links between common diseases and levels of physical activity. Conditions such as cardiovascular disease (Barengo *et al* 2004), hypertension (Blair *et al* 1984), diabetes mellitus (Manson *et al* 1991) and depression (Yancey *et al* 2004) have all been linked to physical inactivity. Although some epidemiological studies have used self-reporting (diaries) to quantify activity patterns, these methods have been shown to be unreliable (Ainsworth *et al* 1993, Washburn and Montoye 1986). Instead, fully automated activity classification offers a more objective approach to quantifying levels of physical activity.

Activity classification systems can also be used to investigate the effectiveness of initiatives aimed at increasing physical activity (van Sluijs *et al* 2007). A better understanding of why people choose to exercise and how individuals can be motivated to increase their levels of physical activity is crucial if the current health epidemic resulting from physical inactivity is to be reversed (Dugdill *et al* 2009). Furthermore, activity classification systems could be used to provide feedback to motivate individuals to adhere to daily or weekly physical activity targets (Baker and Mutrie 2005).

Accurate information on daily activity patterns has the potential to improve the treatment and differential diagnosis of neurological, degenerative and respiratory disorders. To date, automated activity classification systems have been used in patients with Parkinson's disease (Dunnewold *et al* 1997, Moore *et al* 2008) and to validate the use of different motor subtypes in delirium (Leonard *et al* 2007). They have also been shown to be valid in the assessment of physical activity levels in patients with multiple sclerosis (Ng and Kent-Braun 1997), osteoarthritis (Brandes *et al* 2008) and chronic pulmonary disease (Pitta *et al* 2006). Furthermore, automated activity classification systems have considerable potential to be used to assess effectiveness of treatments. For example, in stroke, accelerometer-based systems can be used to recognize real-world upper extremity movement which could then be used to derive treatment outcomes (Uswatte *et al* 2000).

With an ageing population the incidence of falls is increasing. As many elderly persons now live alone, falls can go undetected and injured individuals left unaided for lengthy periods of time (Gurley *et al* 1996). Research has shown that the earlier a fall is reported the lower the rate of morbidity and mortality (Gurley *et al* 1996, Wild *et al* 1981). Clearly, any system which can accurately detect a fall and automatically call for help could be of major benefit. A fall is not an intentional movement; however, within the context of activity classification, it can be considered a specific form of activity. As such, the analytical techniques used in activity classification are also applicable to fall detection systems.

In addition to health-related applications, activity-profiling systems could play a fundamental role in ubiquitous computing scenarios (Coutaz *et al* 2005, Streitz and Nixon 2005). In such applications, information from a variety of sensors is used to determine the context of a situation, so that an appropriate service can be provided. For example, a mobile phone may detect when a person is driving or involved in vigorous physical activity and automatically divert a call.

Body-worn sensors are well suited to collecting data on activity patterns over extended periods of time. In contrast to other approaches, such as laboratory-based systems or video analysis, they can be used under conditions of free living with minimal inconvenience to the user. With recent developments in sensor miniaturization, it is now possible to collect data

on many aspects of human motion, such as segmental acceleration, angular velocity and foot pressure. However, for such data to be of value, effective algorithms are required which can interpret the data in the context of a range of different activities. Body-worn sensor data can be used for estimating functional parameters, such as gait speed and energy expenditure, and for activity classification to produce a continuous activity record. The focus of this review is activity classification. Other methods for interpreting body-worn sensor data are reviewed elsewhere, for example Chen and Bassett (2005) and Kavanagh and Menz (2008).

The automated identification of activities using body-worn sensor data is a challenging area of work. Apart from the obvious practical limitations on the number, location and nature of sensors that people will tolerate, there are several issues that directly impact the success of any given algorithm. Problems arise due to the variability in sensor characteristics for the same activity across different subjects and for the same individual. Errors can also arise due to variability in sensor signals caused by differences in sensor positioning and from environmental factors such as sensor temperature sensitivity. Any successful algorithm must overcome all these factors. The ideal activity classification scheme works off-the-shelf, using data from a range of previous subjects to identify activities from an unseen individual. However, sometimes this is not possible and an intra-subject classification scheme is currently all that can be achieved for some problems. With this approach, example data are required for a given individual before classification can be performed.

The aim of this review is to present a conceptual introduction to the different computational techniques that have been applied to activity classification. For this reason, the review is organized by analytical technique rather than by classification problem. The wide variation in choice of activities between previously published studies means it is not possible to identify a single, optimal solution for any given classification problem. Nevertheless, where appropriate we have tried to provide the reader with a degree of guidance as to the advantages of each of the classification techniques.

Most approaches to activity classification, using body-worn sensors, involve a multi-stage process. Firstly, the sensor signal is divided into a number of small time segments, referred to as windows, each of which is considered sequentially. For each window, one or more features are derived to characterize the signal. These features are then used as input to a classification algorithm which associates each window with an activity. This review article has been structured according to this multi-stage process. After body-worn sensors are briefly discussed, the different techniques used to define data windows are outlined in section 3. In section 4, we present different approaches to generating features from sensor data, and in section 5, the classification techniques are described.

2. Body-worn sensors

2.1. Inertial sensors

The vast majority of activity classification systems have used inertial sensors, notably accelerometers and rate gyros. Most accelerometers respond to gravity as well as to their true acceleration and, therefore, if the acceleration of the body segment is small with respect to g (9.81 m s^{-2}), as is the case when measuring body sway or static posture, these devices can be used to estimate the inclination of a body segment from the vertical. When the acceleration component becomes large, more sophisticated approaches to separating the orientation component (g) from body segment acceleration are necessary. For a detailed description of accelerometry, the reader is directed to a comprehensive review by Mathie *et al* (2004b).

Rate gyro-based measurements of angular velocity are subject to significant calibration errors, electronic noise and temperature effects (Woodman 2007). This means that if the output is simply integrated to estimate a change in orientation, then the gyro error will be integrated leading to a continuously increasing error. Therefore, integration of rate gyro outputs to estimate orientation changes can only be done over short time periods. A number of approaches have been proposed to overcome this problem including the use of a state estimator with input from a tri-axial gyroscope and a tri-axial accelerometer (Luinge and Veltink 2005) and the use of wavelet analysis (section 4.4) to remove both drift and high frequency noise from gyroscope signals (Najafi *et al* 2002) before integrating to obtain orientation.

Inertial sensors can be complemented by a magnetic compass or magnetometer (Parkka *et al* 2006), which can enable more accurate orientation measurement about the vertical axis (Sabatini 2006) and by GPS (Murakami and Makikawa 1997) to enable location tracking.

2.2. Other sensors

Although inertial sensors have been used in the vast majority of activity classification studies, other sensors which may be considered include devices for measuring segment angles, such as inclinometers (Dai *et al* 1996) and goniometers (Kostov *et al* 1995); skin temperature (Krause *et al* 2003); galvanic skin resistance (Dolgov and Zane 2006); heart rate (Bussmann *et al* 1998b); humidity (Lester *et al* 2006) or respiratory rate (Parkka *et al* 2006).

Recently the concept of 'smart textiles' has been proposed, in which miniature sensors are distributed and integrated into clothing (Wijesiriwardana *et al* 2003). The strict limitation on the size of such integrated sensors means it may not be currently possible to use accelerometers or rate gyroscopes. As an alternative, smaller binary sensors, such as tilt switches, have been proposed (van Laerhoven and Gellersen 2004, van Laerhoven *et al* 2006).

Simple foot pressure switches can be used to identify gait events such as heel strike and toe off (Mansfield and Lyons 2003). However, as they do not give additional information on limb movement, it is not possible to classify different activities from these signals alone. More detailed information, such as net ground reaction force, can be obtained from pressure sensitive insoles which can assess the pressure distribution across the planter aspect of the foot (Veltink *et al* 2005, Zhang *et al* 2005).

3. Windowing techniques

Most activity classification methods use windowing techniques to divide the sensor signal into smaller time segments (windows). Activity classification algorithms are then applied separately to each window. In real-time applications, windows are defined concurrently with data collection and a continuous real-time activity profile is produced. When the sensor data are processed off-line, the windows are defined first and classification algorithms applied sequentially to each window. This information is then combined to give an activity profile along the entire signal.

Three different windowing techniques have been used in activity monitoring, sliding windows, event-defined windows and activity-defined windows. With the sliding window method, the signal is divided into windows of fixed length with no inter-window gaps (figure 1(a)). A range of window sizes have been used in previous studies from 0.25 s (Huynh and Schiele 2005) to 6.7 s (Bao and Intille 2004), with some studies including a degree of overlap between adjacent windows (Bao and Intille 2004, Preece *et al* 2008b). The sliding window approach does not require pre-processing of the sensor signal and is therefore ideally suited to real-time applications. Due to its implementational simplicity, most activity classification studies have employed this approach.

In order to use event-defined windows, pre-processing is required to locate specific events, such as heel strike or toe-off. These events are then used to define successive windows (figure 1(b)). Given that such events may not be uniformly spaced in time, the size of these windows is not fixed. A number of different approaches have been proposed for identifying heel strike and toe-off from body-worn sensor signals. For example, it is possible to define search windows from either a low pass filtered version of the original signal (Aminian *et al* 1999a, Selles *et al* 2005) or segmental angles (Jasiewicz *et al* 2006), within which maxima or minima correspond to gait events. Another approach is to identify the times at which the anterio-posterior component of the trunk acceleration changes sign. Heel strike is then located at a given time offset from these points (Mansfield and Lyons 2003, Zijlstra 2004, Zijlstra and Hof 2003).

The use of activity-defined windows is dependent on determining the times at which the activity changes. These points are then used to define windows of sensor data, each of which correspond to a different activity (figure 1(c)). A number of methods have been proposed to identify activity-transition points prior to explicitly identifying the specific activities. For example, wavelet analysis can be used to identify localized changes in frequency characteristics (Nyan *et al* 2006a, Sekine *et al* 2000a) which correspond to a change between activities. Once defined, classification is performed for each window, sometimes using only a subset of the data contained within the window (Nyan *et al* 2006a).

4. Feature generation

Previous activity classification studies have used a wide range of approaches to generate features which characterize windows of body-fixed sensor data. These features are then used as inputs to classification schemes (section 5). In this section, the different feature generation techniques are presented within a number of different sub-categories. Firstly, heuristic features, for both the recognition of everyday activities and falls, are discussed. In this review, we use the term heuristic to refer to features which have been derived from a fundamental and often intuitive understanding of how a specific movement or posture will produce a characteristic body-worn sensor signal. In sections 4.2–4.4, time-domain, frequency-domain and time–frequency (wavelet) features are described. In contrast to the heuristic approach, these features are not typically related to specific aspects of individual movements or postures. Instead they simply represent different ways of characterizing the information within the time-varying signal. For a given classification problem it is often difficult to identify optimal time- and frequency-domain features. Therefore, methods for selecting optimal features from a larger set and methods for reducing dimensionality of features can be used for pre-processing before advanced classification algorithms are applied. These two techniques are described in sections 4.5 and 4.6.

4.1. Heuristic features

The signal measured by an accelerometer comprises two components. The first 'static acceleration' is due to the effect of gravity and gives a measure of the inclination of the sensor to the vertical. The second 'dynamic acceleration' results from the acceleration of the body segment to which the unit is attached. In the absence of motion, the measured acceleration (in units of g) is equal to the cosine of the sensor orientation angle relative to the vertical. This angle is often used as an input to classification algorithms, particularly those designed to differentiate between static postures (Aminian *et al* 1999b, Maxwell 2002) and identify postural transitions (Najafi *et al* 2003).
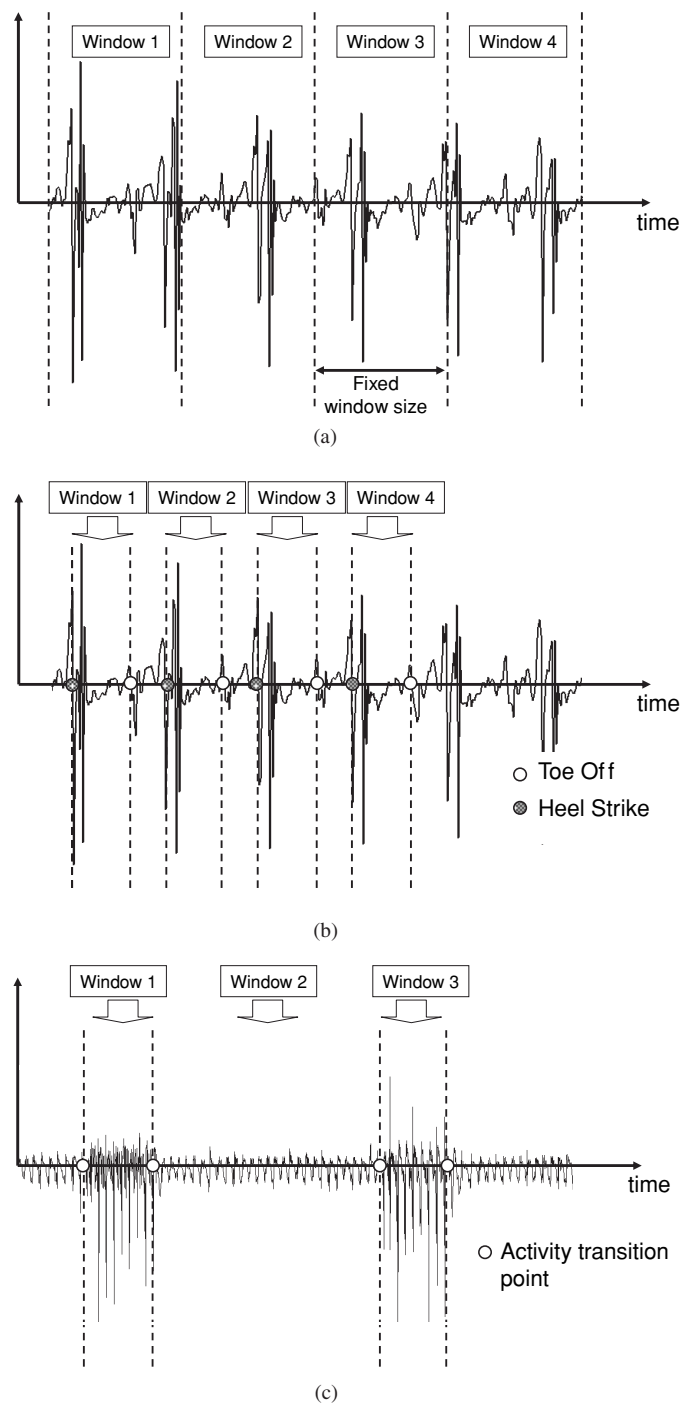
Figure 1. Defining (a) sliding windows, (b) event-based windows and (c) activity-defined windows along a continuous body-worn sensor signal.

Change in segmental orientation can be obtained by integrating a gyroscope signal, provided some method is used to eliminate drift. This feature can then be used to identify different postures and postural transitions (Najafi *et al* 2002, 2003). A gyroscope utilizes the Coriolis force to quantify segmental angular velocity. Activities which exhibit unique patterns of angular velocity can thus be identified from a gyroscope signal with a simple classification algorithm, such as a threshold-based classifier. This idea was exploited by Coley *et al* (2005) who demonstrated that the peak shank angular velocity in the anterior–posterior direction at midstance was positive during stair ascent but negative during level walking and stair descent.

All movement patterns result in time varying segmental accelerations. A number of different methods have been used to derive features which quantify the amplitude of these accelerations. Before these features are derived, the signal is first high pass filtered (typically 0.2–0.5 Hz) to remove any baseline offset. The range of different features includes the signal magnitude area (SMA—the area under the high pass filtered acceleration curve) (Mathie *et al* 2003, Preece *et al* 2008c), peak-to-peak acceleration (Makikawa and Iizumi 1995), mean rectified value (Bussmann *et al* 1998a, 1998b) and root mean square (Veltink *et al* 1996). This type of feature is often used to differentiate between static and dynamic activity (Mathie *et al* 2003).

As well as being used as input to classification algorithms, the SMA can be used to quantify the level of intensity of physical activity. This measure is normally expressed in units known as activity counts. It is now well established that, for a given activity, a linear model can be used to relate metabolic energy expenditure to the number of activity counts (Bouten *et al* 1997, Hendelman *et al* 2000, Terrier *et al* 2001). For an excellent review of this area the reader is directed to Chen and Bassett (2005).

The development of robust algorithms which can accurately differentiate between everyday activities and falls using body-worn sensor data is a rapidly growing area of study. The vast majority of previously developed algorithms have utilized heuristic features exploiting one or more intrinsic characteristics of a fall, such as velocity (Bourke *et al* 2008), acceleration (Bourke *et al* 2007), orientation (Hwang *et al* 2004) or even sound (Doukas and Maglogiannis 2008). During a fall, there is an initial period of free fall, during which velocity increases rapidly. This is then followed by a rapid deceleration, as contact is made with the ground. In addition, there is also a measurable change in the orientation of a number of body segments. Using accelerometers or gyroscopes mounted at the wrist (Degen *et al* 2003), waist (Karantonis *et al* 2006), chest (Bourke and Lyons 2008, Bourke *et al* 2007) or head (Lindemann *et al* 2005), it is possible to characterize these different events. Typically, a threshold-based classifier (section 5.1) is then applied to differentiate between falls and everyday activities.

4.2. Time-domain features

Time-domain features are derived directly from a window of sensor data and are typically statistical measures. Example time-domain features used in activity monitoring include the mean, median, variance, skewness, kurtosis (Baek *et al* 2004, Herren *et al* 1999) and inter-quartile range (Maurer *et al* 2006). Other studies have used high and low pass filters to separate accelerometer signals on a frequency basis. Separate means for the low frequency and rectified high frequency components are then used as inputs to the classification schemes (Fahrenberg *et al* 1997, Foerster and Fahrenberg 2000, Lee *et al* 2003).

As an alternative to the above, Veltink *et al* (1996) developed a classification scheme based on measures of signal morphology. With this approach, a cross-correlation coefficient was used to quantify the similarity of an event-defined window of data to a previously obtained template

signal for each activity. In general, it has been suggested that measures of correlation between different accelerometer axes may improve activity recognition (Aminian *et al* 1995, Herren *et al* 1999). Following this idea, Bao and Intille (2004) used cross-correlation coefficients to quantify the similarity between acceleration signals from different axes on the same body segment and across different segments.

4.3. Frequency-domain features

In order to derive frequency-domain features, the window of sensor data must first be transformed into the frequency domain, normally using a fast Fourier transform (FFT). The output of a FFT typically gives a set of basis coefficients which represent the amplitudes of the frequency components of the signal and the distribution of the signal energy. Different methods can then be used to characterize the spectral distribution from these coefficients. For example, median frequency (Foerster and Fahrenberg 2000) or a subset of the different FFT coefficients can be used (Preece *et al* 2008a, 2008b). Alternatively, information from a number of coefficients can be combined to give a single feature. Examples include spectral energy, which is the sum of the squared FFT coefficients (Huynh and Schiele 2005, Sugimoto *et al* 1997), and frequency-domain entropy, which is the normalized information entropy of the FFT components (Bao and Intille 2004). This latter feature allows for differentiation between activities which have simple acceleration patterns and those with more complex patterns. For example, as cycling involves a uniform movement of the legs, a frequency-domain analysis of thigh acceleration shows a single dominant frequency. In contrast, running may result in more complex acceleration pattern and often displays many major FFT components. This difference leads to a much higher frequency-domain entropy for running in comparison to cycling (Bao and Intille, 2004).

4.4. Wavelet analysis (time–frequency features)

Unlike Fourier analysis which can only be used to extract information on the frequency content of a signal, wavelet analysis can be used to investigate both time and frequency characteristics. Like Fourier analysis, wavelet analysis can be formulated via a continuous or discrete wavelet transform. Previous work on activity monitoring has employed the discrete wavelet transform (DWT), therefore our discussion will focus on this method. The discrete wavelet transform is normally implemented using the filter bank interpretation. In this approach, the original signal is successively decomposed into separate low and high pass filtered signals, referred to as approximation and detail coefficients respectively.

If we consider the original signal (S) with maximum frequency f_{max}, then the first approximation coefficient (cA_1) is obtained by passing the original signal through a low pass filter with passband $[0, f_{max}/2]$. Similarly, to obtain the first detail coefficient (cD_1), the original signal is filtered using a high pass filter with passband $[f_{max}/2, f_{max}]$. The wavelet coefficients cA_1 and cD_1 represent the first level of wavelet decomposition. Subsequent levels of decomposition are obtained by high and low pass filtering the approximation coefficient from the previous level. This process is illustrated schematically in figure 2 and on an example signal in figure 3. At each level of wavelet decomposition, the filtered signal is downsampled by a factor of 2 in order to produce the approximation and detail coefficients. Thus, if the original signal contains N samples, then the first approximation and detail coefficients will be of length $N/2$. Similarly, the length of the approximation and detail coefficients at the second level of decomposition will be $N/4$. This process of subsampling reduces the number of time samples and effectively decreases time resolution. This successive halving of the frequency band
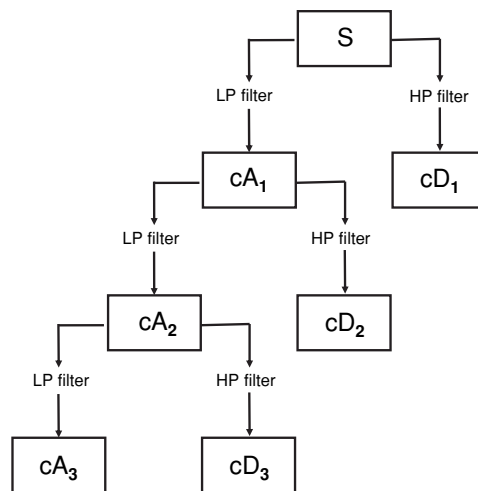
Figure 2. Wavelet decomposition tree. *S* refers to the original signal and cA₁ and cD₁ to the approximation and detail coefficients at the first level of decomposition. These two coefficients are obtained by low pass and high pass filtering of the original signal respectively. Subsequent levels of wavelet decomposition are obtained by filtering the approximation coefficient from the previous level.
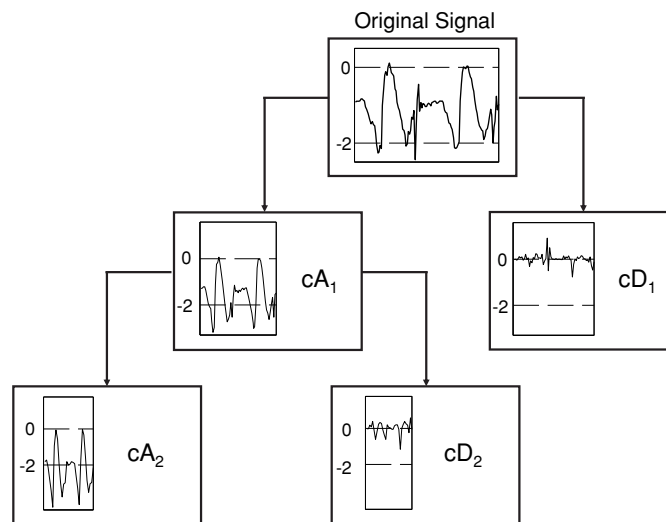

Figure 3. An example of the accelerometer signal along with the approximation and detail coefficients at the first and second levels of decomposition. It can be seen that the high pass filtering associated with the detail coefficient results in the removal of the non-zero offset present in the original signal.

with each level of wavelet decomposition increases frequency resolution. This compromise between time and frequency resolution allows the wavelet transform to provide not only good frequency resolution at low frequencies (higher levels of decomposition) but also better time resolution at higher frequencies (lower levels of decomposition).

Table 1. Different applications of wavelet analysis in activity classification.

Signal enhancement	Najafi *et al* (2002, 2003), Paraschiv-Ionescu *et al* (2004)
Identification of transition points	Nyan *et al* (2006a), Sekine *et al* (2000a)
Generation of time–frequency features	Nyan *et al* (2006a), Sekine *et al* (2000a, 2000b, 2002), Tamura *et al* (1997), Wang *et al* (2007), Preece *et al* (2008b)

Wavelet analysis allows a body-worn sensor signal to be decomposed into a number of individual coefficients, each of which contains data on a specific frequency band. As these coefficients characterize the original signal along its entire length, they contain information on temporal changes in frequency content. Thus, unlike Fourier analysis, wavelet techniques can be used to analyse and characterize non-stationary signals (those in which frequency context changes over time). There are a number of different types of DWT, such as the Haar, Daubachies and Coiflets transform, the difference between these different transforms being in the filters used for decomposition. For a more complete description of the fundamental principles underlying wavelets, the reader is directed to Rioul and Vetterli (1991), Walker (1999) and Graps (1995).

Wavelet analysis has been applied to three different types of problem within activity monitoring. These are signal enhancement, identification of activity transition points and generation of time–frequency features subsequently used for classification (table 1). Each of these applications is now discussed in detail. For signal enhancement, multi-resolution analysis is used to reconstruct the original signal from the decomposed approximation and detail signals. By performing this reconstruction and omitting/modifying specific detail and approximation signals, it is possible to enhance and focus on the frequency band of interest. In activity monitoring applications, this has been used to remove both high frequency noise and low frequency offsets, such as drift, from the original signal (Najafi *et al* 2002). Specific parameters are then derived which characterize the reconstructed signal. This method has primarily been used in threshold-based classification schemes (section 5.1) which use heuristic features to characterize some aspect of movement or posture (Najafi *et al* 2003, Paraschiv-Ionescu *et al* 2004).

Wavelet analysis can be used to identify the points in a body-worn sensor signal at which there is a change in the frequency content. Recent work by Sekine *et al* (2000a) and Nyan *et al* (2006a) demonstrated that, by identifying such points, it was possible to determine the transition times between three different types of gait. Sekine *et al* (2000a) used wavelet packet analysis to decompose the signal and then reconstructed a low frequency version of the original signal. With wavelet packet analysis, the detail coefficients (figure 2) are also split into approximation and detail signals (Mallat 1999). Manually set thresholds were then applied to the original signal to identify changes in frequency content and thus walking pattern. Rather than reconstructing the original signal, Nyan *et al* (2006a) determined transitions from a correlation signal. This was obtained by multiplying the wavelet approximation signals at the two highest levels of decomposition. The technique of multiplying wavelet coefficients, known as direct spatial correlation, can be used to sharpen major signal edges while suppressing noise (Xu *et al* 1994). By comparing a rescaled version of the correlation signal with the approximation signal at the largest scale, they were able to determine the points at which the walking pattern changed. Both Sekine *et al* (2000a) and Nyan *et al* (2006a) specified activity-defined windows from the previously determined transition points. Classification was then performed separately for each window.

Several studies have compared time–frequency features between different activities (Sekine *et al* 2000a, 2000b, Tamura *et al* 1997). These parameters are typically obtained

by calculating one or more statistical measures from the approximation or detail coefficients at specific decomposition levels. For example, Tamura *et al* (1997) calculated the sum of the squares of the approximation coefficients at two different decomposition levels for accelerometer signals collected during level walking and ascending/descending stairs. Using the same activity set, Sekine *et al* (2002) derived a fractal dimension parameter, describing variance progression, from the detail coefficients of a seven-level wavelet decomposition. In both these studies, the wavelet parameters were shown to be significantly different between the three activities.

A number of studies have demonstrated the possibility of using wavelet analysis in activity classification. For example, Sekine *et al* (2000a, 2000b) and Nyan *et al* (2006a) used wavelet parameters based on the sum of the squares or RMS of specific detail coefficients as inputs to threshold-based classification algorithms. In a similar spirit Wang *et al* (2007) derived wavelet parameters using simple statistical measures, such as SD and RMS, of specific approximation and detail coefficients which were subsequently used as input to an artificial neural network. In a recent paper by Preece *et al* (2008b), the performance of a number of wavelet-based sets was compared to previously used time- and frequency-domain features for the classification of eight different activities. In general, the wavelet features tended to be outperformed by the time- and frequency-domain features. This result suggests that, although wavelet analysis can be used to analyse non-stationary signals, it may not be the most effective method for characterizing short windows of sensor data over which there is minimal variation in frequency content.

4.5. Feature selection methods

Different individuals may perform the same movement in a variety of different ways. This can lead to substantial variability in the features derived from body-fixed sensor data (Heinz *et al* 2003). Hence, for effective classification, it is important to identify a set of features which have high discriminative ability (Kiani *et al* 1997). A good feature set should show little variation between repetitions of the same movements and across different subjects but should vary considerably between different activities. Furthermore, it is important to minimize any redundancy between features as this can result in unnecessarily increased computational demands and, also, reduced accuracy with some classification methods (Duda *et al* 2001, Theodoridis and Koutroumbas 2006).

A number of different techniques, of varying complexity, have been used to select appropriate features for activity classification. For example, Parkka *et al* (2006) used simple visual and statistical analysis to assess the distribution of a given feature for different activities. Features which changed markedly between activities and showed little overlap were selected for subsequent analysis. In their study of six daily activities, Maurer *et al* (2006) used correlation-based feature selection. With this approach optimal features are defined as those which exhibit high within-class but low between-class correlations. Another method for feature selection is a forward–backward search in which features are sequentially added and removed from a larger set. Optimal features are identified depending on the resulting classification accuracies for each feature subset. This approach was used by Pirttikangas *et al* (2006) to identify the best sensors/features for the classification of 17 different activities

Huynh and Schiele (2005) compared a range of different acceleration-derived features, including mean, variance, spectral energy and FFT coefficients. The acquired data patterns were subjected to *k*-means clustering in the feature space. Clustering is a method to locate concentrations of data points well separated from each other, and to extract the cluster centres that represent those patterns. Huynh and Schiele (2005) measured the cluster homogeneity

to assess whether individual activities tended to cluster together. In order to automatically recognize activities, they labelled the centroid of each cluster with the dominating activity, and they used a nearest neighbour rule to assign a class to a new pattern (section 5.4). Their analysis showed that, in general, FFT coefficients were best for differentiating between dynamic activities, but they were unable to identify single FFT coefficients which performed best for all activities.

4.6. Dimensionality reduction methods

As an alternative to selecting a subset of the existing features, it is often possible to combine the original features to define a new set of variables. There are two benefits associated with such a procedure. Firstly, the often unnecessarily large numbers of features, resulting from many sensors, can be reduced. Secondly, the new reduced set of variables frequently has better discriminative ability for classification problems. One of the most common techniques for reduction is principal component analysis (PCA) (Chau 2001a, Duda *et al* 2001, Webb 2002). PCA locates the directions of maximal variance and the data are projected onto those directions. This achieves decorrelation of the resulting variables and removal of likely statistical interactions. Additionally, all the directions of smaller variance can be ignored, so a dramatic reduction is often achieved, without sacrificing useful information content. Independent component analysis (ICA) (Duda *et al* 2001, Theodoridis and Koutroumbas 2006) extends PCA to non-Gaussian data, where the sought directions produce variables that are statistically independent from each other. The variables are still reduced as in PCA, but a general linear transformation, as opposed to the rotation of PCA, is performed and often enhances the classification ability of many algorithms.

Previous authors have applied dimensionality reduction methods to different aspects of activity classification problems. For example, Mantyjarvi *et al* (2001) preprocessed data from two tri-axial accelerometers using PCA and ICA, and used this as input to a wavelet-based feature generation technique. For a five-activity problem, high levels of classification accuracy (83–90%) were obtained using a neural network classifier. Following similar principles, Krause *et al* (2003) used PCA to reduce the high dimensionality of a 128 FFT feature set (section 5.11). Using a slightly different approach Huynh and Schiele (2006b) developed an algorithm based on multiple eigenspaces. This technique extends PCA in the sense that it uses multiple spaces spanned by subsets of the PCA eigenvectors. Using this method, Huynh and Schiele (2006b) were able to decipher structure in accelerometer data without any user annotation or information on the activities involved. In subsequent work, Huynh and Schiele (2006a) applied this approach to data collected by Kern *et al* (2003) using tri-axial accelerometers distributed across the body. The multiple eigenspaces algorithm provided a low-dimensional description of original sensor data which was then used as input to an SVM (section 5.6) classification algorithm.

5. Classification schemes

Once features have been derived to characterize a window of sensor data, they are used as input to a classification algorithm. The degree of complexity of these different classification schemes varies from simple threshold-based schemes to more advanced algorithms, such as artificial neural networks or hidden Markov models. With these advanced classification algorithms, appropriately implemented software learns to recognize and associate patterns in the input features with each activity. As such, this field of study is often referred to as machine learning. Machine learning techniques are generally considered to fall within one of two

categories, either supervised or unsupervised (Duda *et al* 2001, Theodoridis and Koutroumbas 2006, Webb 2002). With supervised learning, a significant amount of fully labelled activity data is required in order to 'train' the classification algorithm. Once the training phase is complete, the classifier is able to assign an activity label to an unknown window of sensor data. With unsupervised approaches no activity labels are required for the training dataset. Instead, all the sensor data are passed to the algorithm which automatically identifies a number of states or data clusters, each of which may correspond to a particular activity.

Within the field of activity classification, the classical cross-validation (CV) (Duda *et al* 2001) can be adapted to evaluate the accuracy of the system in two ways: between-subject and within-subject evaluation. In the former case, the classifier is first trained with data from all subjects except a few and then tested with data from the excluded subjects. The accuracy is then calculated as the proportion of correctly classified windows of data across all activities. The process of excluding some subjects and performing a train–test cycle is repeated until all subjects have participated in the testing datasets. The finally overall accuracy is then calculated as the average accuracy across all train–test cycles. When one subject is used for the testing, for a number of cycles equal to the number of subjects, this is called leave-one-subject-out CV. For within-subject evaluation, training is performed using a portion of windows for a specific subject, while testing takes place with the remaining samples of the same subject. This process is then repeated, each time using a different portion of the subject samples for testing. The overall accuracy is determined from the average of all the cycles for all available subjects.

Although an overall accuracy is often provided, more detailed views of the classifier's performance can be given through sensitivity and specificity. These are calculated separately for each activity by determining whether each data window in the test dataset has been identified as the correct activity or not. Sensitivity represents the ability of the classifier to select instances of a certain activity class, whereas specificity represents the true negative rates of an activity. These measures are based on the analysis of the confusion matrix, which summarizes the predicted and actual instances for each class.

This section begins with a brief discussion of threshold-based classification. Following this, in sections 5.2–5.10, the different supervised learning approaches are described. There has only been a very small amount of work applying unsupervised techniques to activity classification, therefore this work is reviewed in a single section, 5.11. Finally, in section 5.12, we present an overview of the different classification techniques. Table 2 lists the different classification methods along with corresponding published studies.

5.1. Threshold-based classification

With threshold-based classification, a derived feature is simply compared to a predetermined threshold to determine whether a particular activity is being performed. This approach has been used successfully to differentiate between static postures, such as standing, sitting and lying, using angles derived from accelerometers placed on combinations of the pelvis/trunk (Boyle *et al* 2006, Culhane *et al* 2004, Uiterwaal *et al* 1998), lower limb segments (Busser *et al* 1997, Bussmann *et al* 1998c, Culhane *et al* 2004, Makikawa and Iizumi 1995) and chest (Aminian *et al* 1999b, Najafi *et al* 2003). Threshold-based classification has also been used successfully to identify postural transitions using data on the change in segmental angles derived from either accelerometers (Najafi *et al* 2003) or gyroscopes (Najafi *et al* 2002, 2003). These algorithms are typically sensitive to the exact choice of threshold angle (Najafi *et al* 2003). Therefore, as an alternative, Najafi *et al* (2003) proposed a simple kinematic model in which vertical displacement of the chest sensor was estimated from double integration of the acceleration signal.

Table 2. Different machine learning techniques used in activity classification.

Threshold-based classification[a]	Culhane *et al* (2004), Uiterwaal *et al* (1998), Boyle *et al* (2006), Bussman *et al* (1998c), Makikawa and Iizumi (1995), Busser *et al* (1997), Najafi *et al* (2002, 2003), Aminian *et al* (1999b), Coley *et al* (2005), Sekine *et al* (2000a, 2000b), Nyan *et al* (2006a)
Hierarchical methods	Fahrenberg *et al* (1996), Fahrenberg *et al* (1997), Lee *et al* (2003), Mathie *et al* (2004a), Karantonis *et al* (2006), Parkka *et al* (2006), Ermes *et al* (2008)
Decision trees	Bao and Intille (2004), Maurer *et al* (2006), Parkka *et al* (2006), Ravi *et al* (2005), Ermes *et al* (2008)
k-nearest neighbour	Foerster *et al* (1999), Foerster and Fahrenberg (2000), Bao and Intille (2004), Maurer *et al* (2006), Bussmann *et al* (2001), Huynh and Schiele (2005), Preece *et al* (2008b), Zhang *et al* (2006c), Zhang *et al* (2006a)
Artificial neural networks	Zhang *et al* (2005), Parkka *et al* (2006), Pirttikangas *et al* (2006), Kiani *et al* (1998), van Laerhoven and Gellersen (2004), Mantyjarvi *et al* (2001), Ermes *et al* (2008), Baek *et al* (2004), Wang *et al* (2007)
Support vector machines	Doukas and Maglogiannis (2008), Krause *et al* (2005), Ravi *et al* (2005), Huynh and Schiele (2006a), Zhang *et al* (2006b), Zhang *et al* (2006a)
Naive Bayes and Gaussian mixture models	Bao and Intille (2004), Maurer *et al* (2006), Kern *et al* (2003), Ravi *et al* (2005), Huynh and Schiele (2006a), Wu *et al* (2007), Allen *et al* (2006)
Fuzzy logic	Salarian *et al* (2007), Lee and Mase (2002), Boissy *et al* (2007)
Markov models	Pober *et al* (2006), Krause *et al* (2003), Krause *et al* (2005), Lester *et al* (2005), (2006), Lukowicz *et al* (2004), Ward *et al* (2006), Van Laerhoven and Cakmakci (2000)
Combining classifiers	Lester *et al* (2005), (2006), Ravi *et al* (2005), van Laerhoven and Gellersen (2004)
Unsupervised learning	Nguyen *et al* (2007), Van Laerhoven and Cakmakci (2000), Krause *et al* (2003)

[a] See table 3 for studies using threshold-based classification for fall detection.

It is common to differentiate between static postures and dynamic activity by using a feature which quantifies variation in the acceleration signal (section 4.1) (Mathie *et al* 2003, Maxwell 2002, Veltink *et al* 1996). Although advanced classification schemes are normally required to recognize different dynamic activities, a small number of researchers have successfully applied threshold-based algorithms to this problem. For example, Coley *et al* (2005) were able to differentiate stair ascent from level walking or stair descent using an algorithm based on the peak angular velocity of the shank. This feature was calculated after a wavelet-based algorithm had been used to identify stance phase (Aminian *et al* 2002). Similarly, using wavelet-based features (section 4.4), both Sekine *et al* (2000a, 2000b) and Nyan *et al* (2006a) were able to differentiate between three different gaits.

Threshold-based classification has been successfully applied to the detection of falls. A fall can be considered an extreme instance of a postural transition. As explained in section 4.1, a range of different characteristics have been used to develop heuristic features which are then used in threshold-based classification schemes. This range of different characteristics has been summarized in table 3.

Both Nyan *et al* (2006b) and Bourke and Lyons (2008) studied angular velocity and angular acceleration characteristics during a fall. They found significantly larger values during a fall than in everyday activities, demonstrating the potential for accurate fall identification. Wu (2000) studied the horizontal and vertical velocity characteristics of falls using an optical motion capture system and showed that trunk vertical velocities associated with falls were two to three times those of everyday activities. More recently, Bourke *et al* (2008) used an inertial measurement unit (accelerometer and gyroscope) to measure vertical velocity and then applied a threshold of 1.3 ms^{-1} to identify falls with 100% accuracy. Other researchers have obtained

Table 3. Different fall characteristics used in threshold-based fall detection.

Angular velocity during free fall	Nyan *et al* (2006b), Bourke and Lyons (2008)
Angular acceleration during free fall	Bourke and Lyons (2008)
Linear velocity during free fall	Wu (2000), Bourke *et al* (2008), Degen *et al* (2003), Lindemann *et al* (2005)
Impact deceleration	Hwang *et al* (2004), Lindemann *et al* (2005), Bourke *et al* (2007), Chen *et al* (2005), Karantonis *et al* (2006), Degen *et al* (2003), Doughty *et al* (2000)
Linear acceleration during free fall	Bourke *et al* (2007)
Change in orientation	Hwang *et al* (2004), Bourke and Lyons (2008), Chen *et al* (2005)

an estimate of linear velocity by directly integrating the signal from an accelerometer (Degen *et al* 2003, Lindemann *et al* 2005) and again applied simple thresholds to identify falls.

The most common characteristic used to identify the presence of a fall is the rapid deceleration which occurs as the faller contacts the ground (Chen *et al* 2005). Different thresholds have been reported for different accelerometer placements (Doughty *et al* 2000). Thresholds of 6 g, 3.5 g and 2.7 g have been reported for accelerometers mounted at the ear, trunk and thigh, respectively (Lindemann *et al* 2005, Bourke *et al* 2007). Bourke *et al* (2007) further compared the accuracy of fall detection using two threshold rules, one applied to the impact deceleration and the other to the acceleration during free fall. Their results showed the impact deceleration to be a more effective means of identifying falls from everyday activities, with 100% specificity compared to 91% specificity for free fall acceleration.

A number of studies have demonstrated improved fall detection accuracy when a number of different threshold rules are combined together. For example, Lindemann *et al* (2005) used a tri-axial accelerometer located in the ear and combined two acceleration-based and one velocity-based threshold. In a similar spirit Bourke and Lyons (2008) combined three threshold-based rules using angular velocity, angular acceleration and orientation and demonstrated that falls could be differentiated from everyday activities with 100% accuracy. Other studies have combined acceleration thresholds with a measure of change in orientation (Chen *et al* 2005, Hwang *et al* 2004), again reporting high levels of accuracy. After detection using threshold-based methods, the occurrence of a fall is often confirmed by checking for a period of inactivity (Doughty *et al* 2000). For example Hwang *et al* (2004) suggested a period of 10 s and Karantonis *et al* (2006) a period of 60 s.

5.2. Hierarchical methods

To implement a hierarchical classification scheme, a binary decision structure is constructed which consists of a number of consecutive nodes. At each node, a binary decision is made depending on the input features. This decision results in either a definite classification being made or in a transition to another node, where further differentiation between activities is performed. The exact nature and parameters of the decision made at each node is obtained via manual inspection and analysis of the training data, which means that this approach is very time consuming. An example decision structure is illustrated in figure 4. Although this example uses only simple threshold-based rules, it is possible to base the node decision on any mathematical operation.

Fahrenberg *et al* (1996, 1997) classified four activities using a hierarchical approach. Their classification scheme used threshold rules which were applied to time-domain features obtained from accelerometers mounted on the chest, wrist, shank and thigh. For a within-subject design, they were able to identify every activity with almost 100% accuracy (Fahrenberg *et al* 1997). Although similarly high accuracy (97%) was obtained for the between-subject
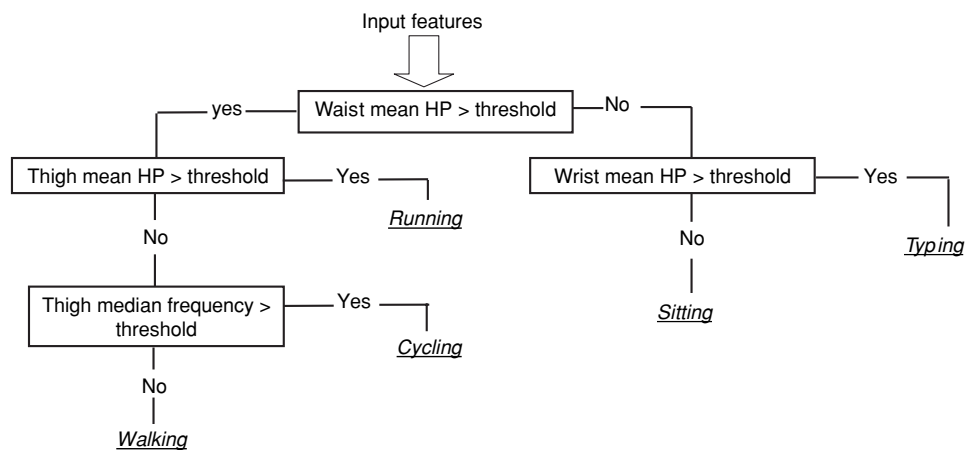
Figure 4. An example hierarchical decision structure. Classification is based on simple threshold rules for each of the four input parameters. These are (1) waist HP mean, (2) wrist HP mean, (3) thigh HP mean AC and (4) thigh median frequency (HP refers to a high pass filtered signal).

design, they were unable to differentiate between level walking and stair walking. A similar approach was used by Lee *et al* (2003) to differentiate between five static and four dynamic activities using a single waist-mounted accelerometer.

More recently, Parkka *et al* (2006) applied a threshold-based hierarchical classification scheme to differentiate between eight different dynamic activities (table 4). In a follow-on study, Ermes *et al* (2008) investigated not only the same set of activities but also included football and compared the performance of the hierarchical approach to other standard classification schemes (table 4). They also developed a hybrid classification scheme in which each node of the hierarchical structure consisted of an artificial neural network. When all data were used for evaluation purposes, this hybrid model was shown to outperform both an artificial neural network and the hierarchical classifier.

The hierarchical approach was also used by Mathie *et al* (2004a) in a study using a single tri-axial accelerometer. In addition to threshold-based rules, they used probabilistic methods and signal morphology techniques to make the classification decision at each node. They demonstrated that this approach could be used to differentiate between a large range of postures, activities and postural transitions across 26 healthy subjects. Furthermore, by including an additional node which identified abnormal peaks in the accelerometer signal, they could identify possible falls. A simplified and computationally efficient version of this algorithm was later developed by Karantonis *et al* (2006) which used only simple threshold-based decisions at each node and demonstrated the potential of their hierarchical approach for real-time fall detection.

5.3. Decision trees

The decision tree approach is similar to hierarchical classification. However, rather than the decision structure being constructed manually by the user, rigorous algorithms exist to automate the process and create a compact set of rules (Duda *et al* 2001, Webb 2002). These algorithms work by examining the discriminatory ability of the features one at a time to create a set of rules which ultimately leads to a complete classification system. For further details of

Table 4. Studies comparing different classifiers.

Publication (number of subjects)	Activities (number of activities)	Accelerometer placements	Inter-subject classification accuracy
Bao and Intille (2004) (20 subjects)	Walking, sitting, cycling, running, vacuuming, folding laundry and more (20)	Shank, thigh, upper arm, wrist and hip	Decision tree (84%) kNN (83%) Naive Bayes (52%)
Parkka et al (2006) (16 subjects)	Lying, sitting, walking, Nordic walking, rowing, cycling and more (8)	Chest and wrist	Decision tree (86%) Hierarchical (82%) Neural network (82%)
Maurer et al (2006) (6 subjects)	Sitting, standing, walking, ascending/descending stairs and running (6)	Wrist	Decision tree (87%) Naive Bayes[a] (<87%) kNN[a] (<87%)
Pirttikangas et al (2006) (13 subjects)	Typing, watching TV, drinking, walking upstairs, cycling and more (17)	Both wrists, thigh and necklace	Neural network (93%) kNN (90%)
Ermes et al (2008) (12 subjects)	Lying, sitting, walking, Nordic walking, rowing, playing football and more (9)	Hip and wrist	Neural network (87%) Hierarchical (83%) Decision tree (60%)
Ravi et al (2005) (2 subjects)	Standing, running, sit-ups, vacuuming, brushing teeth, walking and more (8)	Waist	Naive Bayes (64%) SVM (63%) Decision trees (57%) kNN (50%)
Lester et al (2005) (2 subjects)	Walking, driving, jogging, ascending/descending in an escalator and more (10)	Shoulder	Naive Bayes (67%) HMM (47%) HMM and binary classifiers (95%)
Allen et al (2006) (6 subjects)	Sitting, standing, lying, walking and four postural transitions (8)	Waist	Gaussian mixture model[b] (91%) Hierarchical (71%)

[a] No data were presented on classification accuracy.
[b] Some subject-specific training was used for this classifier.

the different types of decision tree algorithms, the reader is directed to Godfrey et al (2008), Quinlan (1996) and Duda et al (2001).

Decision trees have been applied to a wide range of classification problems (Ermes et al 2008, Parkka et al 2006, Ravi et al 2005). One of the most comprehensive studies was carried out by Bao and Intille (2004) who used both time and frequency features to differentiate between 20 activities (table 4). Using five sensors, they obtained an accuracy of 86%. However, additional analysis showed an accuracy reduction of only 3% if only data from a thigh and wrist sensor was used. Maurer et al (2006) investigated the performance of different features and classifiers in the recognition of six different activities (table 4). The long-term goal of their research was to develop a real-time classification algorithm using data from only one wrist-mounted sensor. Ultimately, they used time-domain features which can be calculated with less computational power than frequency-domain features, as input to their decision tree classifier.

5.4. k-nearest neighbour

With a k-nearest neighbour (kNN) classification scheme (Duda et al 2001, Theodoridis and Koutroumbas 2006), a multi-dimensional feature space is constructed, in which each dimension corresponds to a different feature. The feature space is first populated with all training data points, each of which corresponds to a particular activity. Unknown windows of sensor data are represented in the feature space and the k-nearest points (or neighbours) of training data identified. Classification is then determined by the majority of the k-nearest neighbours which correspond to a given activity. The value of k typically varies from 1 to
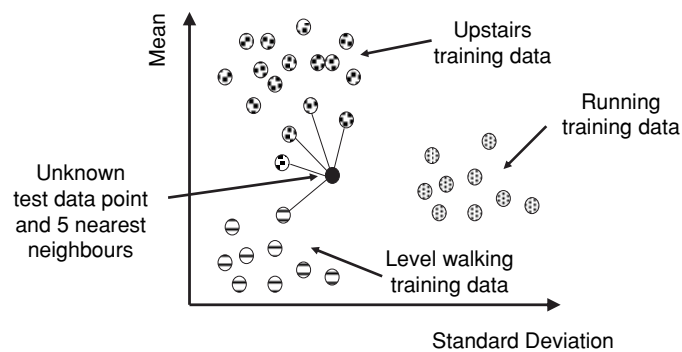
Figure 5. A simple illustration of the *k*-nearest neighbour approach using a 2D feature space. Training data from three separate activities, level walking, walking upstairs and running, have been shown along with an unknown data point which is classified as walking upstairs.

a small percentage of the training data and is selected using trial and error, or ideally using cross-validation procedures. Figure 5 illustrates the *k*-nearest neighbour approach, where a 2D feature space has been constructed. In general, the kNN approach can be applied to any number of dimensions.

Foerster *et al* (1999) were the first to apply the kNN approach to activity classification. They derived simple time-domain features from three uni-axial accelerometers. Using a within-subject design they were able to differentiate between nine common activities. In subsequent work Foerster and Fahrenberg (2000) used a reduced sensor set but extended their original approach, combining a kNN classifier with a hierarchical decision structure and including a frequency-domain feature. At each node of their hierarchical decision structure, they constructed an appropriate feature space using a subset of features. With this approach they were able to accurately classify a wider range of activities than in their previous work (Foerster *et al* 1999).

A similar approach was used by Bussmann *et al* (2001) who defined a 21-dimensional feature space using data derived features from three different accelerometers. Rather than applying the standard kNN approach, they used training data for each activity to specify a maximum and minimum value along each axis. This effectively defined a volume corresponding to each activity within the feature space. For an unknown window of activity data, classification was determined by the closest activity volume within the feature space. With this approach, they were able to identify a wide range of movements and postures with good levels of accuracy (89–93%). More recently, the kNN approach has been compared to other classification schemes (Bao and Intille 2004, Maurer *et al* 2006) (table 4) and used as part of an algorithm for comparing different features for activity classification (Huynh and Schiele 2005, Preece *et al* 2008b).

Zhang *et al* (2006c) used the kNN approach to differentiate between falls and everyday activities. With their classification scheme, windows of accelerometer data were identified immediately before any period during which no sensor motion was detected. Non-negative matrix factorization was then used to extract features from the sensor data which were used as input to the classifier. This factorization is used to decompose the data matrix into a vector basis matrix and a coefficient matrix, under certain constraints, so that new features can be obtained. The results showed that, in most scenarios, it was possible to differentiate between falls and common activities with >95% accuracy.

5.5. Artificial neural networks

An artificial neural network (ANN) can be likened to a flexible mathematical function configured to represent complex relationships between its inputs (independent variables) and outputs (dependent variables). The ANN is initially presented with a set of training data and some form of optimization process is employed to enable known outputs to be predicted for a given set of inputs. Once trained, the ANN can then be used to obtain the outputs for any set of inputs. In the field of activity classification, the inputs are normally features derived from sensor data with the outputs being the different classes of activities. As well as being used for classification problems, ANNs can also be used to estimate continuously varying outputs from a set of input variables (Aminian *et al* 1995, Goulermas *et al* 2005, 2008, Herren *et al* 1999). ANNs have been widely used within the field of human movement research (Chau 2001b, Ohno-Machado and Rowland 1999). For further background information, the reader is directed to Haykin (1999) and Bishop (1999).

One of the most common ANNs is referred to as a multi-layer feedforward neural network or multilayer perceptron (MLP) (Bishop 1999, Haykin 1999). This consists of inputs and outputs which are interconnected via special nodes, distributed in so-called hidden layers. The flow of information through the network is controlled by the weighting of the links between the nodes and the transfer function within each node. This type of network is trained by iteratively optimizing the weights in order to accurately produce the desired training outputs from the corresponding inputs. Zhang *et al* (2005) used this approach in a four-activity problem using data from pressure sensitive insoles. By using features derived via parameterizing the ground reaction force as input to the ANN, Zhang *et al* (2005) were able to accurately ($>97\%$) identify the type of activity as well as predict the speed of walking and running. Other studies which have used an MLP include Baek *et al* (2004), Mantyjarvi *et al* (2001) and Wang *et al* (2007) with a further three studies (Ermes *et al* 2008, Parkka *et al* 2006, Pirttikangas *et al* 2006) comparing the accuracies obtained using an MLP to those obtained with other classification approaches (table 4).

An alterative to the feedforward ANN is the probabilistic neural network (Specht 1990). Unlike most ANNs which require an extensive training period, this type of network enables classification to be rapidly performed using example patterns stored in memory (Specht 1990). Kiani *et al* (1998) employed this approach, training their ANN using template waveform patterns for each activity, rather than using features derived from sensor signals. Although their classification scheme was straightforward to implement, an individually designed network was required for each subject.

Spiking (Gerstner and Kistler 2002) or pulsed neural networks (PNN) (Maass and Bishop 2001) are another class of ANNs which have been applied to activity classification problems. Whereas most ANNs accept a relatively small number of continuously varying inputs, PNNs work with a much larger number of binary inputs and can be implemented very efficiently in hardware. Van Laerhoven and Gellersen (2004) compared the performance of a PNN, using inputs from tilt switches, with an approach which used standard accelerometer data and a self-organizing map (van Laerhoven and Cakmakci 2000) (section 5.11). Although tilt switches did not perform as well as the accelerometer-based approach, relatively good classification accuracy was demonstrated for a range of daily activities, especially static postures.

5.6. Support vector machines

Support vector machines (SVMs) (Cristianini and Shawe-Taylor 2000, Vapink 1998) constitute a popular machine learning method which is based on finding optimal separating decision

hyperplanes between classes with the maximum margin between patterns of each class. Additionally, by using the so-called kernel functions, they can project the data from the original feature space they lie in, to another higher dimensional space. In this way, a linear separation in the new space becomes equivalent to a non-linear classification in the original space. An optimization technique is used to find the optimal separating hyperplanes that perform the required classifications.

SVMs have only been applied in a small number of activity classification studies. Huynh and Schiele (2006a) combined a multiple eigenspaces approach (section 4.6) with SVM and were able to consistently outperform a naive Bayes approach even with very small numbers of training data. In another study, Krause *et al* (2005) used an SVM and showed better performance of frequency-domain over time-domain features for the recognition of eight daily activities.

Three studies have used SVM techniques to differentiate between simulated falls and other activities. Doukas and Maglogiannis (2008) collected data from a microphone and tri-axial accelerometer and were able to accurately differentiate between falls and walking and running in two subjects. Zhang *et al* (2006b) also used features derived from a waist-mounted accelerometer and demonstrated a fall recognition accuracy of 96% across 12 subjects. In another study, Zhang *et al* (2006a) collected data from a tri-axial accelerometer embedded in a mobile phone which was either carried in the subject's pocket or hung around their neck. Classification was performed by first identifying potential falls using a SVM algorithm. True falls were then confirmed by using features, extracted with a dimensionality reduction approach (section 4.6), as input to a kNN classifier. With this approach, they demonstrated high recognition accuracy (>92%), although this dropped to 84% when they attempted to identify falls from other high intensity activities, such as running and jumping.

5.7. Naive Bayes and Gaussian mixture models

The Bayesian classifier is based on the estimated conditional probabilities or likelihoods of the signal patterns available from each activity class. Given such likelihoods, the probability of a new unknown pattern having been generated by a specific activity can be estimated directly. With a naive Bayes classifier, the input features are assumed to be independent of each other. With this assumption, it is possible to express the likelihood function for each activity as the product of n simple probability density functions, where n is the number of features. These functions are typically expressed as one-dimensional normal distributions. Although the assumption of feature independence is often violated, the Bayesian approach is popular due to its simplicity and ease of implementation. A more general version of the naive Bayesian is discriminant analysis, where cross-correlations between features are taken into account. For further details the reader is directed to Duda *et al* (2001) and Theodoridis and Koutroumbas (2006).

Mixed results have been reported when the Bayesian approach to activity classification has been compared to other methods (table 4). For example, Maurer *et al* (2006) and Ravi *et al* (2005) found this approach to either outperform or match the classification accuracy of other methods, whereas Bao and Intille (2004) found low levels of classification accuracy. Bao and Intille (2004) suggested that the reason for this poor performance may have been the questionable assumptions that acceleration features can be considered conditionally independent and modelled by a normal distribution. Other studies which have used the Bayesian approach are Huynh and Schiele (2006a), Kern *et al* (2003) and Wu *et al* (2007). In this latter study, Wu *et al* (2007) developed a generic classification method which could discern when to use available sensors to achieve a specified level of certainty. They demonstrated
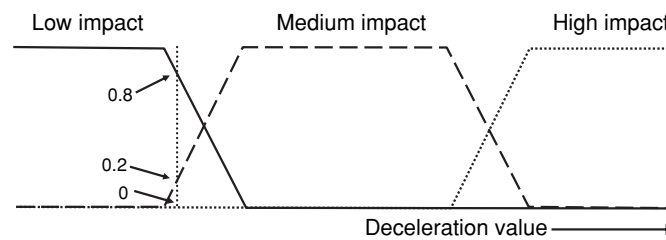
Figure 6. Three example membership functions used to specify the input to a fuzzy classification scheme. The vertical line represents a particular value of deceleration and corresponds to a separate fuzzy truth value for each of the three functions (0 for high impact, 0.2 for medium impact and 0.8 for low impact).

their approach through a case study in which they distinguished between different types of limp using accelerometers and knee-angle sensors.

A Gaussian mixture model (GMM) (Haykin 1999) operates along similar principles to a Bayesian classifier. However, the likelihood function is not assumed to be a single Gaussian probability density. Instead, it is assumed to be of unknown shape and functional form and thus approximated by a weighted mixture of Gaussian functions. The weights and the parameters (centres and covariances) of the mixture components are calculated using the expectation-maximization (EM) algorithm. Allen *et al* (2006) employed this approach using time-domain features to construct separate GMMs for a number of movements/postures. To train the GMMs and calculate the parameters, they used an approach similar to EM but which employed a statistical estimate proposed in the field speech recognition. Classification of test data was achieved by selecting the GMM (activity) with the highest probability of having produced that particular set of features. Allen *et al* (2006) showed that, provided subject-specific training was used, the GMM outperformed a hierarchical classifier (table 4).

5.8. Fuzzy logic

Fuzzy logic is derived from fuzzy set theory and uses reasoning which is approximate rather than precisely defined. It allows mapping from a set of inputs to one or more outputs via a set of if–then statements called rules. For an activity classification problem, features derived from body-worn sensor signals constitute the inputs, with the outputs being fuzzy truths corresponding to each class of activity. Information flows through a fuzzy system via a number of steps. Firstly, the inputs (or features) are assigned membership to fuzzy sets via appropriate membership functions. In contrast to classical set theory in which a data point's membership is either in or out, by allowing the membership function to range between 0 and 1, fuzzy set theory permits partial membership in multiple sets. Figure 6 illustrates example membership functions which could be used to describe the size of an impact deceleration in terms of three sets: low, medium or high impact. As an example, a dotted vertical line has been used to specify a deceleration value which has the membership function values of 0, 0.2 and 0.8 for high, medium and low, respectively. Once each input has been assigned membership of a fuzzy class, the rules can be applied to produce a corresponding output. For an activity classification problem, this output is a membership value, or fuzzy truth, ranging from 0 to 1 for each class of activity. The classification result is then normally taken to be the activity with the maximum fuzzy truth.

Using fuzzy logic, it is possible to reason with imprecise concepts. As such, fuzzy logic is sometimes better suited for dealing with real-world problems than conventional logic which is

normally used in hierarchical or decision tree classification schemes. Despite this, fuzzy logic has only been applied to a limited number of activity classification problems. Lee and Mase (2002) applied this approach, first using simple heuristic features to identify different static postures, and then using the fuzzy classifier to differentiate between different movements. They defined membership functions in terms of the standard deviations of the sensor signals and the short-term changes in orientations, calculated from the gyroscope signal. By using a set of rules based around the min operation (the fuzzy equivalent of AND), Lee and Mase (2002) were able to distinguish between different gaits with good accuracy (>90%).

The Mamdani fuzzy inference method is one of the most common techniques for developing a fuzzy logic classifier. With this approach, it is possible to specify certain membership functions and then to develop a set of rules which allow the training inputs (features) to be mapped to the training outputs (activity classes). Salarian *et al* (2007) used this method as part of a three-stage activity classification scheme. This scheme first used a statistical classifier to identify sit-to-stand and stand-to-sit transitions, and then employed a threshold-based approach to identify periods of walking and lying. Finally, a fuzzy classifier was used to identify periods of sitting and standing. This classifier was developed using membership functions constructed from a knowledge of activity states before and after the period of interest. Classification accuracies obtained using this approach were shown to be better than those obtained using simple threshold rules (Najafi *et al* 2002).

Boissy *et al* (2007) used Mamdani's fuzzy inference to identify falls. Data from a tri-axial accelerometer were used as input to a fuzzy classifier and the amplitude of each acceleration component was used to determine membership values for the classes: low, medium and high (figure 6). A total of 27 rules were used to produce the output, which was expressed in terms of a three-class membership function ('no', 'maybe' and 'yes') representing the occurrence of a fall. The value of this output function was then combined with the knowledge of body orientation using conventional Boolean logic to determine whether a fall had occurred. By collecting a large dataset of fall and non-fall events from 10 subjects, they were able to demonstrate average fall detection accuracies ranging between 86 and 93%, depending on sensor location.

5.9. Markov chains and hidden Markov models

For certain classification problems, some transitions between activities are more likely to occur than others. For example, it is highly unlikely that an individual would sit down directly after descending stairs, but would be likely to start walking. A Markov chain is a discrete time stochastic process in which each activity is represented as a different state. Markov chains can be used to represent the likelihood of transitions between different activities.

A HMM is similar to the Markov chain, but the state of the model at any given time is unknown (or hidden) and can only be determined from observable parameters which depend on the state. In contrast to the Markov chain, the HMM can be used directly for activity classification problems. The observable parameters are the features derived from body-worn sensor data, with the states corresponding to the different activities. Unlike a Markov chain, states in a HMM can correspond to more than one activity. As with previous classification techniques, a HMM is first trained using example data. Once trained, it can then be used to determine the most likely sequence of state transitions (and thus activities) which could have resulted from an observed sequence of features. HMMs are trained by determining state transitions along with the probabilities that each possible set of observations (features) will be observed for a given state. These probabilities are obtained using the Baum–Welch algorithm (McLachilan and Peel 2000). In activity classification studies, HMMs have been used as a

single classifier (Pober *et al* 2006) and as part of a two-stage classification scheme (Lester *et al* 2005, 2006, Ward *et al* 2006). They have the advantage over other classifiers that they can be used to model any constraints that are imposed on the sequence in which activities can occur.

Pober *et al* (2006) used a HMM to recognize four different activities using only one computationally simple input feature (SMA (section 4.1)). In their model, they used three states per activity and achieved an overall classification accuracy of 80%, with most misclassifications arising due to confusion between level and uphill walking. Lukowicz *et al* (2004) and Ward *et al* (2006) used HMMs to classify a range of workshop activities, such as using a saw or screw driver, from body-worn accelerometers and microphones. Features derived from the accelerometer data were used as input to an HMM and the output combined with that of a separate classifier for the sound data. Using this approach, they were able to identify the different workshop activities with an accuracy of 74–78% with subject-specific training (Ward *et al* 2006).

Lester *et al* (2005, 2006) used the HMM formulation as part of a two-layer classification for differentiating between a range of daily activities. The output probabilities from a large number of static binary classifiers (section 5.10) were used as input to the HMM. The addition of the HMM layer allowed the classifier to account for sequence constraints thereby increasing the accuracy of activity recognition by as much as 10–15% (Lester *et al* 2006). It also had the effect of smoothing out sporadic errors which occurred when the simple static classifiers were used alone, ensuring temporal smoothness in the final activity profile.

For applications in which the aim is not to determine a continuous activity profile from a set of observed features, but simply to know transition probabilities to subsequent activities, a simple Markov chain can be used. Krause *et al* (2005) used Markov chains to determine the optimal strategy for selectively sampling sensor data, demonstrating the potential to reduce power consumption. Specifically, for activities which were known to have short duration, a short sampling interval was selected, whereas for longer duration activities, the sampling interval was increased. Markov chains have also been used as part of unsupervised learning algorithms (Krause *et al* 2003, van Laerhoven and Cakmakci 2000); for details see section 5.11.

5.10. Combining different classifiers

Meta-level classification schemes have recently gained popularity within the biomedical community. They improve the performance of individual classifiers by combining their output using different techniques. These include majority voting (where the majority class is accepted), stacked generalization (which trains the base classifiers and then uses their predictions as data to a new learning stage) or boosting (which assigns weights to the training patterns to combine the performance of weak classifiers) (Theodoridis and Koutroumbas 2006, Webb 2002). Ravi *et al* (2005) used a meta-level classification scheme in a pilot study with two subjects who performed eight common activities. Five base-level classifiers were used in their study, including SVMs, decision trees, kNN and naive Bayes. In general, when an inter-subject design was used, the boosted SVM was shown to outperform other meta-level classification schemes.

AdaBoost is a type of adaptive boosting that incrementally trains classifiers by suitably increasing the pattern weights to favour the misclassified data. Thus, it combines multiple weak classifiers to create a single more powerful one and has been used by Lester *et al* (2005, 2006) and van Laerhoven and Gellersen (2004). Lester *et al* (2005, 2006) studied ten common daily activities deriving a large number of statistical and frequency-domain features from

a range of sensors. They then constructed a set of weak binary classifiers, each of which accepted only a single feature as input and obtained a classification result from a weighted combination of the weak classifiers. They compared the performance of two different weak classifiers: a discriminative decision-stump (a binary decision tree classifier constrained to the use of a single feature) and a generative naive Bayes model (section 5.7) and found the Bayesian approach to perform best. Classification accuracy was then improved by using the output from the weak classifiers as input to a HMM (section 5.9).

5.11. Unsupervised learning

Unsupervised learning techniques (Duda *et al* 2001, Theodoridis and Koutroumbas 2006) can be used for the analysis and interpretation of body-worn sensor data without the need of activity labels for each data window. In contrast to supervised methods, where the classifier is trained to identify unseen windows of sensor data, unsupervised approaches are used to identify clusters of related patterns in the feature space. Ideally each of these clusters will correspond to a different activity or subclass. Such techniques have the advantage that they allow exploratory data analysis and investigation of the importance of individual features. Also, once the cluster structure has been determined, a process of labelling of the clusters, followed by a supervised learning layer is often adopted. This can largely reduce the cost of labelling large datasets. By combining unsupervised with supervised approaches, it is possible to develop off-the-shelf systems which can be trained by the user with only occasional input. This allows for considerable flexibility and adaptation to new scenarios inevitably encountered by the real-world user. In addition, unsupervised learning has the potential to be used as the first stage of a system for detecting adverse events, such as falls, which may differ significantly from typical daily activity patterns. Despite their potential usefulness in the field of activity monitoring, their application has been limited to only a few studies.

Van Laerhoven and Cakmakci (2000) were the first to demonstrate the potential of unsupervised learning techniques in activity monitoring. They defined a feature space from a number of simple time-domain features obtained from two thigh-mounted accelerometers. A Kohonen self-organizing feature map (SOM) was then used to identify localized patterns within the feature space. A SOM can be considered an array of discrete nodes or neurons, used to store projections of the original data to a much lower dimensional feature space. It does so by recognizing and maintaining the groupings and proximity characteristics of the data in the original space. During unsupervised learning, each original pattern is projected onto the network topology and the strongest pattern activation is used to update the weights in the corresponding neighbourhoods. Through this process, the clusters and patterns of points within the original high dimensional feature space are identified and mapped to well-defined regions of the two-dimensional SOM. Once the original data have been grouped in this way, they can be labelled with minimal user input and then a supervised classification layer added to recognize unknown windows of sensor data.

After associating specific regions of their SOM with one of seven activities, Van Laerhoven and Cakmakci (2000) added a supervised layer which comprised a kNN classifier and a Markov chain. Unknown windows of sensor data were assigned a potential label via projection onto the SOM after which the kNN method was used to locate the closest activity cluster. Finally the transition probabilities, modelled by the Markov chain, were used to determine whether a particular transition was likely and the classification result modified accordingly.

Krause *et al* (2003) also used a SOM to develop an unsupervised learning algorithm for interpreting data collected from five different types of sensor located within a single arm-mounted unit. Due to the high dimensionality of their feature space, they employed an

initial PCA stage (section 4.6) to reduce the dimensionality of their data before applying the SOM. Once potential clusters had been identified within the SOM, a Markov chain was used to represent transition probabilities between clusters and a graph reduction strategy used to eliminate transient states, i.e. those with low transition probabilities. With this approach, Krause *et al* (2003) were able to automatically identify clusters corresponding to a range of different activities.

Unsupervised approaches can be used to identify unusual events from sensor data containing a range of repeated activities. This could be of value for a fall detection system, where there are typically no available training data (sensor outputs during a fall). Recently, Nguyen *et al* (2007) presented a preliminary study demonstrating the potential of unsupervised clustering for the recognition of both usual and unusual events. For their study, they used data from a waist-mounted accelerometer as input to an algorithm which combined hidden Markov models and Gaussian mixture models to perform data segmentation and clustering without prior knowledge. Each particular activity was represented by a HMM whose density function was estimated using a Gaussian mixture model. They investigated the degree to which similar activities clustered together and showed that optimal results were obtained using 'raw features' in comparison to other time-domain features. Although their study was carried out on a single subject, they demonstrated the potential for an unsupervised fall detection system. Future work is required to better understand the potential of this approach.

5.12. Overview of machine learning classification techniques

Almost all previously published activity classification studies differ in the type and number of activities and in the location, type and number of body-fixed sensors. Furthermore, there is considerable variation in the number and type of features which are derived from the sensor signals. This is the case both for studies investigating a range of normal activities and those concerned with the identification of falls from everyday activities. The variability in activities, sensors and features means that it is not possible to directly compare classification accuracies between different studies. However, a number of studies have compared the performance of two or more classifiers using exactly the same set of input features and therefore allow us to gain some insight into the relative performance of individual classifiers. These studies are summarized in table 4.

Initial inspection of table 4 suggests that either decision trees or neural networks may give the highest levels of classification accuracy for a number of representative activity classification problems. However, in some studies, differences in accuracy between classifiers are often as low as 1–4%, which may not be statistically significant. Furthermore, some studies report differing findings on the relative accuracy of different classifiers. For example, Parkka *et al* (2006) studied eight activities and found that maximal classification accuracy could be obtained with a decision tree classifier. In a subsequent study of a similar set of activities, performance of the decision tree classifier was considerably lower, with the best performance from an artificial neural network. Similarly, although two studies (Lester *et al* 2005, Ravi *et al* 2005) obtained relatively good performance with a naive Bayesian classifier, Bao and Intille (2004) found poor performance using this approach.

Taken together, the studies summarized in table 4 may suggest that there is no classifier which performs optimally for a given activity classification problem. However, many of the different techniques have been evaluated using small numbers of subjects. Therefore, there is a need for further studies investigating the relative performance of the range of different classifiers for different activities and sensor features and with large numbers of subjects. For example, techniques such as SVM and Gaussian mixture models show considerable promise

but have not been applied to large datasets. Fuzzy logic and Markov models also have the potential to be of value in future algorithms as they can be implemented as either a single classifier or as part of a hybrid classification algorithm.

The choice of classifier for any given problem will be determined by a number of considerations. As well as accuracy, factors such as ease of development and speed of real-time execution will influence the final choice. The following paragraphs briefly summarize the different techniques, giving a simple overview of the potential advantages and disadvantages of each method.

The hierarchical approach has been widely used as it is an intuitive approach which can be modified after development to include additional activities (Mathie *et al* 2004a). Although it can take a long time to develop, it is normally executed with minimal computational power and is therefore well suited to real-time applications. Unlike the hierarchical approach, decision trees use automated algorithms, and are thus faster to develop and considerably less user intervention is required. Again, once developed, this classifier can be used effectively in real time (Maurer *et al* 2006). With both the hierarchical and decision tree approach, the classification scheme can be represented graphically and so the underlying rules are easy to understand and interpret.

Classification schemes using the kNN approach can be developed rapidly, are highly versatile and can be used to classify a large range of different activities. However, on-line execution may be slower than decision trees due to the distance evaluation requirements. Similar to the kNN approach, artificial neural networks are a very flexible, powerful approach which have the potential to be used for a range of different classification problems as well as for predicting functional parameters. Although they have demonstrated high levels of accuracy for a number of classification problems (table 4), they can be slow to train and some types of networks difficult to implement. SVMs are also a very powerful and popular method and, although they have shown significant potential, they have not been applied to many activity monitoring problems. With this approach, it is possible to work reliably with difficult and noisy classification datasets, but they may be very slow to train with large datasets and difficult to set their kernel type and kernel parameters.

Naive Bayes classifiers are simple to develop and can be executed rapidly. However, they are based on the weak assumption of feature independence. Although they have been shown to work well on studies with small numbers of subjects, they tend to be outperformed by other classifiers in larger studies (table 4). Gaussian mixture models are more powerful than the naive Bayes method. However, it is often difficult to set the number of mixtures to obtain optimal density functions. Promising results have been obtained using this approach in one study (Allen *et al* 2006) and further work is required to establish whether it is applicable to other activity classification problems.

Fuzzy logic holds considerable promise for activity classification problems as it enables reasoning with imprecise concepts. Potential disadvantages of this approach are the difficulties in construction of appropriate membership functions and uniquely interpreting and combining fuzzy rules. A small number of previous studies have demonstrated good classification accuracies using fuzzy logic, particularly in fall detection, but further work is required to determine the full potential of this approach.

Markov chains are graphical models which contain information on the probability of transition between different activity states. Although a simple chain cannot be directly used for classification, it can often lead to improved accuracy when combined with other classification techniques. Hidden Markov models are a development of the simple Markov chain and represent a powerful approach for identifying a sequence of activities from a sequence of measured features. With this technique, classification of a particular window of sensor data

depends not only on the observed features but also on the likelihood of a transition from a previous activity. Hidden Markov models have been shown to be effective both as a single classifier and for improving the performance of other classifiers.

Combining classifiers is another very promising approach. As multiple classifiers applied to the same dataset create different decision boundaries, they can exhibit different sensitivity to different patterns. Therefore, combining them can provide complementary decisions and improve the overall accuracy. Although the overall algorithm can be relatively complex, the advantage is that much simpler existing algorithms can be combined to create a strong classifier, and hence its application to activity monitoring seems very promising.

Unsupervised learning operates along different principles to the previously discussed techniques. This approach can be used to explore the data and, hence, provide insight into the structure of activity data within the feature space. It can therefore play a valuable role in the development of supervised classification schemes. When unsupervised methods are combined with a supervised layer, it is possible to rapidly develop individual specific classification algorithms which can be readily adapted to include new activities. However, this approach will always require some input from the user. With the limited number of studies utilizing this approach, there is a considerable need for further work in this area.

6. Conclusion

This review has presented an overview of the different techniques which have been used for activity classification from body-worn sensor data. Information has been organized into two principal sections, the first dealing with feature generation and simple threshold-based classification and the second dealing with more advanced classification techniques. Within this framework, features were categorized as heuristic, time-domain, frequency-domain or time–frequency (wavelet). Heuristic features are derived from a fundamental understanding of how a specific movement or posture will produce a characteristic body-worn sensor signal. By using such features in simple threshold-based classification schemes, it is possible to accurately differentiate between static postures and dynamic activity and to identify falls with high levels of accuracy (Bourke and Lyons 2008, Bourke *et al* 2008). In order to differentiate between large numbers of dynamic movements and postures, it is necessary to use advanced classification schemes which accept one or more features as input. In section 5, a range of different classification techniques has been reviewed. Although a small number of studies, comparing the performance of different classifiers, suggest that either decision trees or artificial neural networks may give the highest classification accuracy, differences are often small. Furthermore, there are many other methods such as support vector machines, fuzzy logic and hidden Markov models which have shown promise in small pilot studies but have yet to be tested in larger-scale studies. Therefore considerable further work is required to establish the suitability of the different techniques for a range of classification problems.

Most previously published activity monitoring studies vary considerably in the choice of sensor placements and in the range of activities analysed, which means that comparisons of the results from different studies should be treated with caution. Future work in activity monitoring may therefore benefit from the use of a standard predefined set of activities and sensor placements (Ward *et al* 2005). In line with this idea, Noury *et al* (2007) recently proposed a standard set of potential fall scenarios which could be used as a common framework to evaluate fall detection algorithms. The identification of similar sets of activities for other problems, such as epidemiological studies, rehabilitation and ubiquitous computing, may lead towards an improved understanding of the relative effectiveness of the different classification algorithms presented in this review.

Acknowledgments

The authors gratefully acknowledge the support for this work from EU IST Framework VI project Healthy Aims and from the DTI/Vodafone sponsored sensor-phone project.

References

Ainsworth B E, Jacobs D R Jr and Leon A S 1993 Validity and reliability of self-reported physical activity status: the Lipid Research Clinics questionnaire *Med. Sci. Sports Exerc.* **25** 92–8

Allen F R, Ambikairajah E, Lovell N H and Celler B G 2006 Classification of a known sequence of motions and postures from accelerometry data using adapted Gaussian mixture models *Physiol. Meas.* **27** 935–51

Aminian K, Najafi B, Bula C, Leyvraz P F and Robert P 2002 Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes *J. Biomech.* **35** 689–99

Aminian K, Rezakhanlou K, De Andres E, Fritsch C, Leyvraz P F and Robert P 1999a Temporal feature estimation during walking using miniature accelerometers: an analysis of gait improvement after hip arthroplasty *Med. Biol. Eng. Comput.* **37** 686–91

Aminian K, Robert P, Buchser E E, Rutschmann B, Hayoz D and Depairon M 1999b Physical activity monitoring based on accelerometry: validation and comparison with video observation *Med. Biol. Eng. Comput.* **37** 304–8

Aminian K, Robert P, Jequier E and Schutz Y 1995 Incline, speed, and distance assessment during unconstrained walking *Med. Sci. Sports Exerc.* **27** 226–34

Baek J, Lee G, Park W and Yun B J 2004 Accelerometer signal processing for user activity detection *Knowledge-Based Intelligent Information and Engineering Systems, Pt 3, Proceedings* pp 610–7

Baker G and Mutrie N 2005 Are pedometers useful motivational tools for increasing walking in sedentary adults? *6th Int. Conf. on Walking in the 21st Century (Zurich)*

Bao L and Intille S S 2004 Activity recognition from user-annotated acceleration data *Pervasive Computing (Lecture Notes in Computer Science* vol 3001*)* (Berlin: Springer) pp 1–17

Barengo N C, Hu G, Lakka T A, Pekkarinen H, Nissinen A and Tuomilehto J 2004 Low physical activity as a predictor for total and cardiovascular disease mortality in middle-aged men and women in Finland *Eur. Heart J.* **25** 2204–11

Bishop C M 1999 *Neural Networks for Pattern Recognition* (Oxford: Oxford University Press)

Blair S N, Collingwood T R, Reynolds R, Smith M, Hagan R D and Sterling C L 1984 Health promotion for educators: impact on health behaviors, satisfaction, and general well-being *Am. J. Public Health* **74** 147–9

Boissy P, Choquette S, Hamel M and Noury N 2007 User-based motion sensing and fuzzy logic for automated fall detection in older adults *Telemed J. E. Health* **13** 683–93

Bourke A K and Lyons G M 2008 A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor *Med. Eng. Phys.* **30** 84–90

Bourke A K, O'Brien J V and Lyons G M 2007 Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm *Gait Posture* **26** 194–9

Bourke A K, O'Donovan K J and Olaighin G 2008 The identification of vertical velocity profiles using an inertial sensor to investigate pre-impact detection of falls *Med. Eng. Phys.* **30** 937–46

Bouten C V C, Koekkoek K T M, Verduin M, Kodde R and Janssen J D 1997 A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity *IEEE Trans. Biomed. Eng.* **44** 136–47

Boyle J, Karunanithi T, Wark T, Chan W and Colavitti C 2006 Quantifying functional mobility progress for chronic disease management *28th Annual Conf. of the IEEE Engineering in Medicine and Biology Society (New York)* pp 5916–9

Brandes M, Schomaker R, Mollenhoff G and Rosenbaum D 2008 Quantity versus quality of gait and quality of life in patients with osteoarthritis *Gait Posture* **28** 74–9

Busser H J, Ott J, van Lummel R C, Uiterwaal M and Blank R 1997 Ambulatory monitoring of children's activity *Med. Eng. Phys.* **19** 440–5

Bussmann H B, Reuvekamp P J, Veltink P H, Martens W L and Stam H J 1998a Validity and reliability of measurements obtained with an 'activity monitor' in people with and without a transtibial amputation *Phys. Ther.* **78** 989–98

Bussmann J B, Martens W L, Tulen J H, Schasfoort F C, Van Den Berg-Emons H J and Stam H J 2001 Measuring daily behavior using ambulatory accelerometry: the activity monitor *Behav. Res. Methods Instrum. Comput.* **33** 349–56

Bussmann J B, Tulen J H, van Herel E C and Stam H J 1998b Quantification of physical activities by means of ambulatory accelerometry: a validation study *Psychophysiology* **35** 488–96

Bussmann J B J, van de Laar Y M, Neeleman M P and Stam H J 1998c Ambulatory accelerometry to quantify motor behaviour in patients after failed back surgery: a validation study *Pain* **74** 153–61

Caspersen C J, Powell K E and Christensen G M 1985 Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research *Public Health Rep.* **100** 126–31

Chau T 2001a A review of analytical techniques for gait data: part 1. Fuzzy, statistical and fractal methods *Gait Posture* **13** 49–66

Chau T 2001b A review of analytical techniques for gait data: part 2. Neural network and wavelet methods *Gait Posture* **13** 102–20

Chen J, Kwong K, Chang D, Luk J and Bajcsy R 2005 Wearable sensors for reliable fall detection *27th Annual Conf. of the IEEE Engineering in Medicine and Biology Society (Shanghai)* pp 3551–4

Chen K Y and Bassett D R Jr 2005 The technology of accelerometry-based activity monitors: current and future *Med. Sci. Sports Exerc.* **37** S490–500

Coley B, Najafi B, Paraschiv-Ionescu A and Aminian K 2005 Stair climbing detection during daily physical activity using a miniature gyroscope *Gait Posture* **22** 287–94

Coutaz J, Crowley J L, Dobson S and Garlan D 2005 Context is key *Commun. ACM* **48** 49–53

Cristianini N and Shawe-Taylor J 2000 *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods* (Cambridge: Cambridge University Press)

Culhane K M, Lyons G M, Hilton D, Grace P A and Lyons D 2004 Long-term mobility monitoring of older adults using accelerometers in a clinical environment *Clin. Rehabil.* **18** 335–43

Dai R, Stein R B, Andrews B J, James K B and Wieler M 1996 Application of tilt sensors in functional electrical stimulation *IEEE Trans. Rehabil. Eng.* **4** 63–72

Degen T, Jaeckel H, Rufer M and Wyss S 2003 SPEEDY: a fall detector in a wrist watch *Proc. 7th International Symposium on Wearable Computers (White Plains, NY)* pp 184–7

Dolgov A B and Zane R 2006 Low-power wireless medical sensor platform *28th Annual Conf. of the IEEE Engineering in Medicine and Biology Society (New York)* pp 2067–70

Doughty K, Lewis R and McIntosh A 2000 The design of a practical and reliable fall detector for community and institutional telecare *J. Telemed. Telecare* Suppl 1 **6** S150–4

Doukas C and Maglogiannis I 2008 Advanced patient or elder fall detection based on movement and sound data *IEEE Pervasive Health Conference and Workshops 2008 (Tampere, Finland)*

Duda R O, Hart P E and Stork D G 2001 *Pattern Classification* 2nd edn (New York: Wiley)

Dugdill L, Crone D and Graham R 2009 *Physical Activity and Health Promotion: Evidence-Based Approaches to Practice* (Oxford: Wiley-Blackwell)

Dunnewold R J, Jacobi C E and van Hilten J J 1997 Quantitative assessment of bradykinesia in patients with Parkinson's disease *J. Neurosci. Methods* **74** 107–12

Ermes M, Parkka J, Mantyjarvi J and Korhonen I 2008 Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions *IEEE Trans. Inf. Tech. Biomed.* **12** 20–6

Fahrenberg J, Foerster F, Smeja M and Muller W 1997 Assessment of posture and motion by multichannel piezoresistive accelerometer recordings *Psychophysiology* **34** 607–12

Fahrenberg J, Muller W, Foerster F and Smeja M 1996 A multi-channel investigation of physical activity *Psychophysiology* **10** 209–17

Foerster F and Fahrenberg J 2000 Motion pattern and posture: correctly assessed by calibrated accelerometers *Behav. Res. Methods Instrum. Comput.* **32** 450–7

Foerster F, Smeja M and Fahrenberg J 1999 Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring *Comp. Hum. Behav.* **15** 571–83

Gerstner W and Kistler W 2002 *Spiking Neuron Models* (Cambridge: Cambridge University Press)

Godfrey A, Conway R, Meagher D and Olaighin G 2008 Direct measurement of human movement by accelerometry *Med. Eng. Phys.* **30** 1364–86

Goulermas J Y, Findlow A H, Nester C J, Liatsis P, Zeng X J, Kenney L P J, Tresadern P, Thies S B and Howard D 2008 An instance-based algorithm with auxiliary similarly information for the estimation of gait kinematics from wearable sensors *IEEE Trans. Neural Networks* **19** 1574–82

Goulermas J Y, Howard D, Nester C J, Jones R K and Ren L 2005 Regression techniques for the prediction of lower limb kinematics *J. Biomech. Eng.* **127** 1020–4

Graps A 1995 An Introduction to wavelets *IEEE Comput. Sci. Eng.* **2** 50–61

Gurley R J, Lum N, Sande M, Lo B and Katz M H 1996 Persons found in their homes helpless or dead *N. Engl. J. Med.* **334** 1710–6

Haykin S 1999 *Neural Networks; a Comprehensive Foundation* (Englewood Cliffs, NJ: Prentice-Hall)

Heinz E A, Kunze K S, Sulistyo S, Junker H, Lukowicz P and Troster G 2003 Experimental evaluation of variations in primary features used for accelerometric context recognition *Ambient Intelligence (Lecture Notes in Computer Science* vol 2875*)* (Berlin: Springer) pp 252–63

Hendelman D, Miller K, Baggett C, Debold E and Freedson P 2000 Validity of accelerometry for the assessment of moderate intensity physical activity in the field *Med. Sci. Sports Exerc.* **32** S442–9

Herren R, Sparti A, Aminian K and Schutz Y 1999 The prediction of speed and incline in outdoor running in humans using accelerometry *Med. Sci. Sports Exerc.* **31** 1053–9

Huynh T and Schiele B 2005 Analyzing features for activity recognition *Proc. Conf. Smart Objects and Ambient Intelligence: Innovative Context-Aware Services: Usages and Technologies (Grenoble)* pp 159–64

Huynh T and Schiele B 2006a Towards less supervision in activity recognition from wearable sensors *Proc. of the 10th IEEE International Symposium on Wearable Computers (Montreaux)* pp 3–10

Huynh T and Schiele B 2006b Unsupervised discovery of structure in activity data using multiple eigenspaces *Second International Workshop on Location- and Context-Awareness (Dublin)* pp 151–67

Hwang J Y, Kang J M, Jang Y W and Kim H C 2004 Development of novel algorithm and real-time monitoring ambulatory system using bluetooth module for fall detection in the elderly *26th Annual Conf. of the IEEE Engineering in Medicine and Biology Society (San Francisco)* pp 2204–7

Jasiewicz J M, Allum J H, Middleton J W, Barriskill A, Condie P, Purcell B and Li R C 2006 Gait event detection using linear accelerometers or angular velocity transducers in able-bodied and spinal-cord injured individuals *Gait Posture* **24** 502–9

Karantonis D M, Narayanan M R, Mathie M, Lovell N H and Celler B G 2006 Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring *IEEE Trans. Inf. Technol. Biomed.* **10** 156–67

Kavanagh J J and Menz H B 2008 Accelerometry: a technique for quantifying movement patterns during walking *Gait Posture* **28** 1–15

Kern N, Schiele B and Schmidt A 2003 Multi-sensor activity context detection for wearable computing *European Symposium on Ambient Intelligence (Eindhoven)* pp 220–32

Kiani K, Snijders C J and Gelsema E S 1997 Computerized analysis of daily life motor activity for ambulatory monitoring *Technol. Health Care* **5** 307–18

Kiani K, Snijders C J and Gelsema E S 1998 Recognition of daily life motor activity classes using an artificial neural network *Arch. Phys. Med. Rehabil.* **79** 147–54

Kostov A, Andrews B J, Popovic D B, Stein R B and Armstrong W W 1995 Machine learning in control of functional electrical stimulation systems for locomotion *IEEE Trans. Biomed. Eng.* **42** 541–51

Krause A, Ihmig M, Rankin E, Leong D, Gupta S, Siewiorek D, Smailagic A, Deisher M and Sengupta U 2005 Trading off prediction accuracy and power consumption for context-aware wearable computing *Proc. of the 9th IEEE International Symposium on Wearable Computers* pp 20–6

Krause A, Siewiorek D P, Smailagic A and Farringdon J 2003 Unsupervised, dynamic identification of physiological and activity context in wearable computing *Proc. 7th International Symposium on Wearable Computers (White Plains, NY)* pp 88–97

Lee S and Mase K 2002 Activity and location recognition using wearable sensors *IEEE Perv. Comp.* **1** 25–32

Lee S H, Park H D, Hong S Y, Lee K J and Kim Y H 2003 A study on the activity classification using a triaxial accelerometer *25th Annual Conf. of the IEEE Engineering in Medicine and Biology Society (Cancun)* pp 2941–3

Leonard M, Godfrey A, Silberhorn M, Conroy M, Donnelly S, Meagher D and Olaighin G 2007 Motion analysis in delirium: a novel method of clarifying motoric subtypes *Neurocase* **13** 272–7

Lester J, Choudhury T and Borriello G 2006 A practical approach to recognizing physical activities *Pervasive Computing (Lecture Notes in Computer Science* vol 3968*)* (Berlin: Springer) pp 1–16

Lester J, Choudhury T, Kern N, Borriello G and Hannaford B 2005 A hybrid discriminative/generative approach for modeling human activities *19th International Joint Conference on Artificial Intelligence (Edinburgh)* pp 766–72

Lindemann U, Hock A, Stuber M, Keck W and Becker C 2005 Evaluation of a fall detector based on accelerometers: a pilot study *Med. Biol. Eng. Comput.* **43** 548–51

Luinge H J and Veltink P H 2005 Measuring orientation of human body segments using miniature gyroscopes and accelerometers *Med. Biol. Eng. Comput.* **43** 273–82

Lukowicz P, Ward J, Junker H, Stager M, Troster G, Atrash A and Starner T 2004 Reconizing workshop activity using body worn microphones and accelerometers *Pervasive Computing (Lecture Notes in Computer Science* vol 3001*)* (Berlin: Springer) pp 18–32

Maass W and Bishop C M 2001 *Pulsed Neural Networks* (Cambridge, MA: MIT Press)

Makikawa M and Iizumi H 1995 Development of an ambulatory physical activity memory device and its application for the categorization of actions in daily life *Medinfo* **8** 747–50

Mallat S 1999 *A Wavelet Tour of Signal Processing* 2nd edn (San Diego: Elsevier)

Mansfield A and Lyons G M 2003 The use of accelerometry to detect heel contact events for use as a sensor in FES assisted walking *Med. Eng. Phys.* **25** 879–85

Manson J E, Rimm E B, Stampfer M J, Colditz G A, Willett W C, Krolewski A S, Rosner B, Hennekens C H and Speizer F E 1991 Physical activity and incidence of non-insulin-dependent diabetes mellitus in women *Lancet* **338** 774–8

Mantyjarvi J, Himberg J and Seppanen T 2001 Recognizing human motion with multiple acceleration sensors *IEEE International Conference on Systems, Man, and Cybernetics (Tucson)* pp 747–52

Mathie M J, Celler B G, Lovell N H and Coster A C 2004a Classification of basic daily movements using a triaxial accelerometer *Med. Biol. Eng. Comput.* **42** 679–87

Mathie M J, Coster A C, Lovell N H and Celler B G 2003 Detection of daily physical activities using a triaxial accelerometer *Med. Biol. Eng. Comput.* **41** 296–301

Mathie M J, Coster A C F, Lovell N H and Celler B G 2004b Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement *Physiol. Meas.* **25** R1–20

Maurer U, Rowe A, Smailagic A and Siewiorek D 2006 Location and activity recognition using eWatch: a wearable sensor platform *Ambient Intelligence in Everday Life (Lecture Notes in Computer Science* vol 3864*)* (Berlin: Springer) pp 86–102

Maxwell D 2002 Addressing the challenge of quantifying free-living activity—the activPAL^TM professional *Conf. on Recent Advances in Assistive Technology and Engineering (RAATE) (Birmingham)* p 23

McLachilan G J and Peel P 2000 *Finite Mixture Models* (New York: Wiley)

Moore S T, MacDougall H G and Ondo W G 2008 Ambulatory monitoring of freezing of gait in Parkinson's disease *J Neurosci. Methods* **167** 340–8

Murakami D and Makikawa M 1997 Ambulatory behavior map, physical activity and biosignal monitoring system *Methods Inf. Med.* **36** 360–3

Najafi B, Aminian K, Loew F, Blanc Y and Robert P A 2002 Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly *IEEE Trans. Biomed. Eng.* **49** 843–51

Najafi B, Aminian K, Paraschiv-Ionescu A, Loew F, Bula C J and Robert P 2003 Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly *IEEE Trans. Biomed. Eng.* **50** 711–23

Ng A V and Kent-Braun J A 1997 Quantitation of lower physical activity in persons with multiple sclerosis *Med. Sci. Sports Exerc.* **29** 517–23

Nguyen A, Moore D and McCowan I 2007 Unsupervised clustering of free-living human activities using ambulatory accelerometry *29th Annual Conf. of the IEEE Engineering in Medicine and Biology Society (Lyon)* pp 4895–8

Noury N, Fleury A, Rumeau P, Bourke A K, Laighin G, Rialle V and Lundy J E 2007 Fall detection—principles and methods *29th Annual Conf. of the IEEE Engineering in Medicine and Biology Society (Lyon)* pp 1663–6

Nyan M N, Tay F E, Seah K H and Sitoh Y Y 2006a Classification of gait patterns in the time-frequency domain *J. Biomech.* **39** 2647–56

Nyan M N, Tay F E H, Tan A W Y and Seah K H W 2006b Distinguishing fall activities from normal activities by angular rate characteristics and high-speed camera characterization *Med. Eng. Phys.* **28** 842–9

Ohno-Machado L and Rowland T 1999 Neural network applications in physical medicine and rehabilitation *Am. J. Phys. Med. Rehabil.* **78** 392–8

Paraschiv-Ionescu A, Buchser E E, Rutschmann B, Najafi B and Aminian K 2004 Ambulatory system for the quantitative and qualitative analysis of gait and posture in chronic pain patients treated with spinal cord stimulation *Gait Posture* **20** 113–25

Parkka J, Ermes M, Korpipaa P, Mantyjarvi J, Peltola J and Korhonen I 2006 Activity classification using realistic data from wearable sensors *IEEE Trans. Inf. Technol. Biomed.* **10** 119–28

Pirttikangas P, Fujinami K and Nakajima T 2006 Feature selection and activity recognition from wearable sensors *Ubiquitous Computing Systems (Lecture Notes in Computer Science* vol 4239*)* (Berlin: Springer) pp 516–27

Pitta F, Troosters T, Probst V S, Spruit M A, Decramer M and Gosselink R 2006 Quantifying physical activity in daily life with questionnaires and motion sensors in COPD *Eur. Respir. J.* **27** 1040–55

Pober D M, Staudenmayer J, Raphael C and Freedson P S 2006 Development of novel techniques to classify physical activity mode using accelerometers *Med. Sci. Sports Exerc.* **38** 1626–34

Preece S J, Goulermas J Y, Kenney L P J and Howard D 2008a A comparison of different feature generation methods in activity classification *Proc. ICAMPAM (Rotterdam)* p 87

Preece S J, Goulermas J Y, Kenney L P J and Howard D 2008b A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data *IEEE Trans. Biomed. Eng.* at press

Preece S J, Goulermas J Y, Kenney L P J and Howard D 2008c Does accelerometer placement affect metabolic energy expenditure estimation in normal weight and obese subjects? *Proc. ICAMPAM (Rotterdam)* p 193

Quinlan J R 1996 Improved use of continuous attributes in c4.5 *J. Artif. Intell. Res.* **4** 77–90

Ravi N, Dandekar N, Mysore P and Littman M 2005 Activity recognition from accelerometer data *Proceedings of the 7th Innovative Applications of Artificial Intelligence Conference (CA)* pp 11–8

Rioul O and Vetterli M 1991 Wavelets and signal processing *IEEE Signal Process. Mag.* **8** 14–38

Sabatini A M 2006 Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing *IEEE Trans. Biomed. Eng.* **53** 1346–56

Salarian A, Russmann H, Vingerhoets F J G, Burkhard P R and Aminian K 2007 Ambulatory monitoring of physical activities in patients with Parkinson's disease *IEEE Trans. Biomed. Eng.* **54** 2296–9

Sekine M, Tamura T, Akay M, Fujimoto T, Togawa T and Fukui Y 2002 Discrimination of walking patterns using wavelet-based fractal analysis *IEEE Trans. Neural Syst. Rehabil. Eng.* **10** 188–96

Sekine M, Tamura T, Togawa T and Fukui Y 2000a Classification of waist-acceleration signals in a continuous walking record *Med. Eng. Phys.* **22** 285–91

Sekine M, Toshiyo T, Toshiro F and Yasuhiro F 2000b Classification of walking pattern using acceleration waveform in elderly people *22nd Annual Conf. of the IEEE Engineering in Medicine and Biology Society (Chicago)* pp 1356–9

Selles R W, Formanoy M A G, Bussmann J B J, Janssen P J and Stam H J 2005 Automated estimation of initial and terminal contact timing using accelerometers; development and validation in transtibial amputees and controls *IEEE Trans. Neural Syst. Rehabil. Eng.* **13** 81–8

Specht D F 1990 Probabilisitic neural networks *Neural Netw.* **3** 109–18

Streitz N and Nixon P 2005 The disappearing computer *Commun. ACM* **48** 32–5

Sugimoto A, Hara Y, Findley T W and Yoncmoto K 1997 A useful method for measuring daily physical activity by a three-direction monitor *Scand. J. Rehabil. Med.* **29** 37–42

Tamura T, Sekine M, Ogawa M, Togawa T and Fukui Y 1997 Classification of acceleration waveforms during walking by wavelet transform *Methods Inf. Med.* **36** 356–9

Terrier P, Aminian K and Schutz Y 2001 Can accelerometry accurately predict the energy cost of uphill/downhill walking? *Ergonomics* **44** 48–62

Theodoridis S and Koutroumbas K 2006 *Pattern Recognition* 3rd edn (San Diego: Academic)

Toreyin B U, Soyer E B, Onaran I and Cetin A E 2008 Falling person detection using multisensor signal processing *EURASIP J. Adv. Signal Process.* article no 149304

Uiterwaal M, Glerum E B, Busser H J and van Lummel R C 1998 Ambulatory monitoring of physical activity in working situations, a validation study *J. Med. Eng. Technol.* **22** 168–72

Uswatte G, Miltner W H, Foo B, Varma M, Moran S and Taub E 2000 Objective measurement of functional upper-extremity movement using accelerometer recordings transformed with a threshold filter *Stroke* **31** 662–7

van Laerhoven K and Cakmakci O 2000 What shall we teach our pants? *Proc 4th Int. Symposium on Wearable Computers* pp 77–83

van Laerhoven K and Gellersen H W 2004 Spine versus porcupine: a study in distributed wearable activity recognition *Proc. 8th Int. Symposium on Wearable Computers (Arlington)* pp 142–50

van Laerhoven K, Gellersen H W and Malliaris Y G 2006 Long-term activity monitoring with a wearable sensor node *International Workshop on Wearable and Implantable Body Sensor Networks* pp 171–4

van Sluijs E M, McMinn A M and Griffin S J 2007 Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials *Br. Med. J.* **335** 703

Vapink V N 1998 *Statistical Learning Theory* (NJ: Wiley)

Veltink P H, Bussmann H B, de Vries W, Martens W L and Van Lummel R C 1996 Detection of static and dynamic activities using uniaxial accelerometers *IEEE Trans. Rehabil. Eng.* **4** 375–85

Veltink P H, Liedtke C, Droog E and Van Der Kooij H 2005 Ambulatory measurement of ground reaction forces *IEEE Trans. Neural Syst. Rehabil. Eng.* **13** 423–7

Walker J S A 1999 *Primer on Wavelets and their Scientific Applications* (New York: Chapman and Hall–CRC)

Wang N, Ambikairajah E, Lovell N H and Celler B G 2007 Accelerometry based classification of walking patterns using time-frequency analysis *29th Annual Conf. of the IEEE Engineering in Medicine and Biology Society (Lyon)* pp 4899–902

Ward D S, Evenson K R, Vaughn A, Rodgers A B and Troiano R P 2005 Accelerometer use in physical activity: best practices and research recommendations *Med. Sci. Sports Exerc.* **37** S582–8

Ward J A, Lukowicz P, Troster G and Starner T E 2006 Activity recognition of assembly tasks using body-worn microphones and accelerometers *IEEE Trans. Pattern Anal.* **28** 1553–67

Washburn R A and Montoye H J 1986 The assessment of physical activity by questionnaire *Am. J. Epidemiol.* **123** 563–76

Webb A 2002 *Statistical Pattern Recognition* (Chichester: Wiley)

Wijesiriwardana R, Dias T and Mukhopadhyay S 2003 Resistive fibre-meshed transducers *Proc. 7th International Symposium on Wearable Computers (White Plains, NY)* pp 200–9

Wild D, Nayak U S and Isaacs B 1981 How dangerous are falls in old people at home? *Br. Med. J. (Clin. Res. Ed.)* **282** 266–8

Woodman O 2007 An introduction to inertial navigation http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-696.html

Wu G 2000 Distinguishing fall activities from normal activities by velocity characteristics *J. Biomech.* **33** 1497–500

Wu W H, Bui A A, Batalin M A, Liu D and Kaiser W J 2007 Incremental diagnosis method for intelligent wearable sensor systems *IEEE Trans. Inf. Technol. Biomed.* **11** 553–62

Xu Y S, Weaver J B, Healy D M and Lu J 1994 Wavelet transform—a spatially selective noise filtration technique *IEEE Trans. Image Process.* **3** 747–58

Yancey A K, Wold C M, McCarthy W J, Weber M D, Lee B, Simon P A and Fielding J E 2004 Physical inactivity and overweight among Los Angeles County adults *Am. J. Prev. Med.* **27** 146–52

Zhang K, Sun M, Lester D K, Pi-Sunyer F X, Boozer C N and Longman R W 2005 Assessment of human locomotion by using an insole measurement system and artificial neural networks *J. Biomech.* **38** 2276–87

Zhang T, Wang J, Liu P and Hou J 2006a Fall detection by embedding an accelerometer in cellphone and using KFD algorithm *Int. J. Comput. Sci. Netw. Secur.* **6** 277–84

Zhang T, Wang J, Xu L and Liu P 2006b Fall detection by wearable sensor and one-class SVM algorithm *Intell. Comput. Signal Process. Pattern Recognit.* **345** 858–63

Zhang T, Wang J, Xu L and Liu P 2006c Using wearable sensor and NMF algorithm to realize ambulatory fall detection *Advances in Natural Computation: Part 2 (Lecture Notes in Computer Science* vol 4222*)* (Berlin: Springer) pp 488–91

Zijlstra W 2004 Assessment of spatio-temporal parameters during unconstrained walking *Eur. J. Appl. Physiol.* **92** 39–44

Zijlstra W and Hof A L 2003 Assessment of spatio-temporal gait parameters from trunk accelerations during human walking *Gait Posture* **18** 1–10