INSTITUTE OF PHYSICS PUBLISHING PHYSIOLOGICAL MEASUREMENT

Physiol. Meas. **27** (2006) 1–17 UNCORRECTED PROOF

Classification of a known sequence of motions and postures from accelerometry data using adapted Gaussian mixture models

Felicity Allen[1], Eliathamby Ambikairajah[2], Nigel H Lovell[1,3] and Branko G Celler[2]

[1] Graduate School of Biomedical Engineering, University of New South Wales, Sydney, NSW, 2052, Australia
[2] School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW, 2052, Australia
[3] National Information and Communications Technology Australia (NICTA), Eveleigh NSW 1308, Australia

E-mail: N.Lovell@unsw.edu.au

Received 12 February 2006, accepted for publication 27 June 2006
Published DD MMM 2006
Online at stacks.iop.org/PM/27/1

Abstract

Accelerometry shows promise in providing an inexpensive but effective means of long-term ambulatory monitoring of elderly patients. The accurate classification of everyday movements should allow such a monitoring system to exhibit greater 'intelligence', improving its ability to detect and predict falls by forming a more specific picture of the activities of a person and thereby allowing more accurate tracking of the health parameters associated with those activities. With this in mind, this study aims to develop more robust and effective methods for the classification of postures and motions from data obtained using a single, waist-mounted, triaxial accelerometer; in particular, aiming to improve the flexibility and generality of the monitoring system, making it better able to detect and identify short-duration movements and more adaptable to a specific person or device. Two movement classification methods were investigated: a rule-based Heuristic system and a Gaussian mixture model (GMM) based system. A novel time-domain feature extraction method is proposed for the GMM system to allow better detection of short-duration movements. A method for adapting the GMMs to compensate for the problem of limited user-specific training data is also proposed and investigated. Classification performance was considered in relation to data gathered in an unsupervised, directed routine conducted in a three-month field trial involving six elderly subjects. The GMM system was found to achieve a mean accuracy of 91.3%, distinguishing between three postures (sitting, standing and lying) and five movements (sit-to-stand, stand-to-sit, lie-to-stand, stand-to-lie and walking), compared to 71.1% achieved by the Heuristic system. The adaptation method was found to offer a mean accuracy of 92.2%; a relative improvement of 20.2% over tests without subject-specific data and 4.5% over tests using only a limited amount of subject-specific data. While limited to a restricted subset of possible motions and postures, these results provide a significant step in the search for a more robust and accurate ambulatory classification system.

Keywords: accelerometer, ambulatory monitoring, falls, human movement, Gaussian mixture models

(Some figures in this article are in colour only in the electronic version)

1. Introduction

The population of the world is aging. The United Nations predicts that by 2100, 28.1% of the world population will be aged 65 years or older compared to 10.0% in 2000 and 6.9% in 1900 (United Nations 2003). With this aging population comes an increased demand for ongoing health monitoring and support of elderly patients.

Recent advances in information communications and sensor technologies have meant that telemedicine systems for automated monitoring of elderly subjects either living at home or in institutionalized care have become a near reality. Some of the primary aims of any such system are to monitor the functional ability of patients and to predict and detect falls.

Accelerometry has been proposed as a practical, inexpensive and reliable method for monitoring motion in free-living subjects (Mathie *et al* 2004). In this paper, a system is described which uses a single triaxial accelerometer for ambulatory monitoring of elderly people living alone. The initial scope of the study is to monitor for potential falls only within the patient's residence. While monitoring outside the home is likely to provide further useful information in terms of functional well-being and energy expenditure, most falls which go undiscovered occur within the home and thus such a limited scope is still likely to be fundamentally useful for the elderly community.

In order to accurately detect and predict falls, it is beneficial not just to examine data for acceleration spikes, but to make some semantic sense of the accelerometry signal in terms of the sequences of motions and postures that occur. If an automated system is able to accurately and reliably classify such sequences of movements, then it is expected that the system will be better able to detect and predict falls. Robust classification of motions and postures should also allow more intelligent monitoring of long-term change in physiological indicators such as parameters of gait and balance, energy expenditure and functional well-being. Hence this study examines several methods for classification of postures and motions and searches for methods to improve performance and robustness of accelerometry-based movement classification systems. In particular, we consider the advantages of moving away from a static rule-based classification approach and towards a more general Gaussian mixture model (GMM) based approach. We also aim to address existing shortcomings in the ability of such generalized systems to identify and distinguish between short-duration movements and to adapt to the parameters of a specific person or device.

Existing literature on movement classification using accelerometry data has been widely varied in approach, intention and outcome. Individual researchers or research groups have each investigated their own particular set of movements using their own device/s and their own data collection method and have applied a wide variety of algorithms and methods. Consequently, it is difficult to make significant comparisons or draw meaningful conclusions from the existing literature beyond noting that accelerometry shows promise in ambulatory monitoring.

Many studies use multiple accelerometers fixed to specific places on the body, usually a subset of the thighs, wrists, arms, sternum, waist and lower legs (Bao and Intille 2004, Bussmann *et al* 2001, Foerster *et al* 1999). A smaller number of studies have investigated the use of a single accelerometry device attached at the waist, sternum or back (Lee *et al* 2003, Mathie *et al* 2004, Sung *et al* 2005, Ravi *et al* 2005). Whereas the use of a larger number of accelerometers is likely to provide a higher accuracy in terms of classifying motions and postures, such a system is also likely to be too cumbersome and inconvenient to be truly feasible for long-term ambulatory monitoring.

A variety of classification methods have been trialed. Many studies have focused on the development of simple threshold-based methods or Heuristic classifiers specific to a small subset of possible motions and postures. The rule-based Heuristic methods described in Mathie *et al* (2004), Mathie (2003) are one such example which will be used as a baseline for comparative tests in this paper. Other simple methods assessing similar motion and posture subsets include, but are by no means limited to, Bussmann *et al* (2001), Foerster *et al* (1999), who classified motions based on a distance measure to an expected range of parameter values and Lee *et al* (2003), who used a simple Heuristic approach.

Other researches employing more generic, automatic methods from the machine learning literature include applications of decision trees, nearest neighbour and Naïve Bayes (Ravi *et al* 2005, Bao and Intille 2004), support vector machines (SVM) (Ravi *et al* 2005), neural networks (Kiani *et al* 1997), Gaussian mixture models and Markov chains (Sung *et al* 2005, Sung and Pentland 2005). It is currently unclear which method is most effective. Minimal experimentation has been performed in relation to finding the optimal front-end features for these methods. Most studies have used frequency-derived features employing an FFT or parameters such as averages or correlations calculated over long time-windows (Sung and Pentland 2005, Ravi *et al* 2005), potentially reducing the ability of the systems to detect short-duration movements, for example, transitions between sitting and standing or taking a couple of steps.

In this paper, two methods for distinguishing between three postures (sitting, standing and lying) and five motions (stand-to-sit transition, sit-to-stand transition, stand-to-lie transition, lie-to-stand transition and walking) using a single, waist-mounted, triaxial accelerometer are implemented and compared. The first is a rule-based Heuristic method, which will be described further in section 2.3. The second uses GMMs as described in section 2.4.

For the GMM system, we propose that time-domain features incorporating temporal information via delta and shifted delta coefficients are sufficient for accurate movement classification. Based on a time-derivative estimate of the gravity and body acceleration waveforms, these have the potential to reduce computation time and better detect short-duration movements. A method for adapting the GMMs to a particular person or device is also proposed. This allows models pre-trained on data from multiple subjects, to be adjusted to better fit a particular person or device, using only a very limited amount of training data from that person or device. If this kind of device is to be deployed clinically in long-term ambulatory monitoring situations, this kind of adaptation method is likely to prove very useful.

2. Methods

2.1. The device

The device used in these studies is a single, waist-mounted triaxial accelerometer identical to that described in Salleh *et al* (2000). Two ADXL210 biaxial accelerometers were mounted orthogonally within a pager case measuring $71 \times 50 \times 18$ mm^3. The whole device is

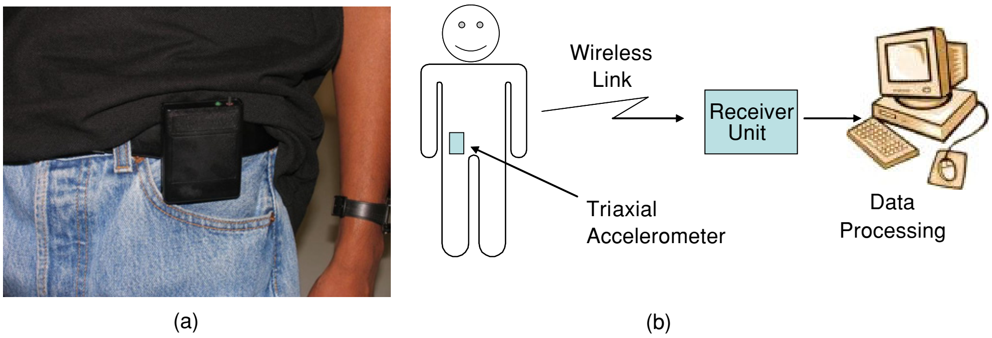

Figure 1. The device (a) being worn by a subject and (b) a schematic diagram.

light weight, weighing approximately 50 g including a single AA 1.5 V battery, a wireless transmitter and an emergency push button (figure 1(a)). The data from the waist-mounted device is transmitted over a 433.92 MHz wireless link to a personal computer where the data are stored and further processing is carried out (see figure 1(b)). The device was designed only for use within a single residence and consequently there is minimal onboard data storage and the range of the wireless link of the device covers only the vicinity of a small house. The resultant sampling rate after transmission is approximately 45 Hz and the data resolution is better than 25×10^{-3} g.

The device is capable of measuring both static and dynamic accelerations and consequently the signal is the net result of the body acceleration due to movement of the subject, acceleration due to gravity, acceleration due to other external forces and noise. Given that this device was designed for use by elderly patients in the home, acceleration due to external forces is assumed to be negligible and the noise level is fairly low (median filtering is applied to reduce noise spikes in the signal and the device output when completely still is very low (Mathie 2003)). Thus the main components of the device signal which need to be considered are the body acceleration and the acceleration due to gravity. These two components overlap in the frequency domain so cannot be completely separated by filtering, however most of the gravity component is found below 0.2–0.5 Hz and thus a reasonable estimate can be separated out using a low-pass filter (Mathie 2003). The gravity component of the signal allows one to determine the orientation or inclination of the subject.

Most commonly the device is placed at the waist above the iliac spine where it is easy to attach without assistance and has been found to be comfortable and unnoticeable (Mathie *et al* 2004). However, for further convenience and to lower the incidence of bruising or discomfort the algorithms are intended to handle any placement around the waist.

The use of a single device has some obvious limitations in terms of its ability to detect the angle of the legs with respect to the torso and consequently to distinguish between basic postures such as sitting and standing. However, it is hoped that such a single device located close to the centre of mass of the person still provides sufficient information that an automated system might be able to infer whether the person is sitting or standing from the sequence of movements which led the person to be in that posture. In a field trial with six elderly recipients conducted over a period of 2–3 months, subject compliance with this single device was found to be 88% (Mathie *et al* 2004). It is expected that the attachment of additional devices would further limit this compliance level and would thus limit the overall benefit that such a system might offer.

2.2. Data collection

The data for these experiments were gathered in an unsupervised pilot study in which six healthy, elderly subjects (4 women, 2 men aged 80–86 years) living independently at home wore the device each day for a period of 2–3 months. Further details of the study are given in Mathie *et al* (2004), Mathie (2003). A desktop computer was set up within each home to collect the data from the device via the wireless connection as well as to provide instructions during a daily routine. Data were collected and stored on the computer automatically whenever the subjects were within range (i.e. whenever they were in the house). The trial was primarily conducted as a feasibility study. Most of the data collected during the days were not annotated.

The data used to assess classification performance in this paper were collected during a short directed routine performed daily by the subjects. The routine was performed according to a set of instructions, presented both as text and orally by the computer. Periods of activity were interspersed with periods of rest and were followed by a button press from the user to allow easy annotation by the researchers after the trial was complete. The full set of instructions are given in Mathie *et al* (2004), Mathie (2003) but can be summarized as follows:

(1) Press the button to start the testing.
(2) Remain standing (30 s).
(3) Sit down.
(4) Once you are seated, press the button on the device.
(5) Remain seated (10 s).
(6) Stand up.
(7) Once you are standing, press the button on the device.
(8) Remain standing (10 s).
(9) Walk around.
(10) Once you are standing beside your bed, press the button on the device.
(11) Remain standing (10 s).
(12) Lie down.
(13) Once you are lying down, press the button on the device.
(14) Remain lying (10 s).
(15) Stand up.
(16) Once you are standing, press the button on the device.
(17) Remain standing (10 s).

Because the routines were performed on successive days and the device was removed overnight, the placement of the device varied with every iteration of the routine. In some cases, the device would be placed in roughly the same position as on previous days. On other days, some participants chose to completely change the position of the device in order to prevent bruising.

After completion of the trial, the routine data were manually checked to ensure that the expected movements were performed, and annotated with each of the eight classifications: standing, sitting, lying, stand-to-sit transition, sit-to-stand transition, stand-to-lie transition, lie-to-stand transition and walking. Each separate instance of each movement type and classification was segmented out into a separate file so that each training or test file contained the complete duration of one particular movement or posture occurrence only. Between 40 and 100 instances of each classification type were collected for each subject resulting in over 340 instances of each movement and posture. The movement and posture data for each subject were then divided randomly into training and test sets in a roughly 60–40% split.

A representative set of routine data is shown in figure 2.

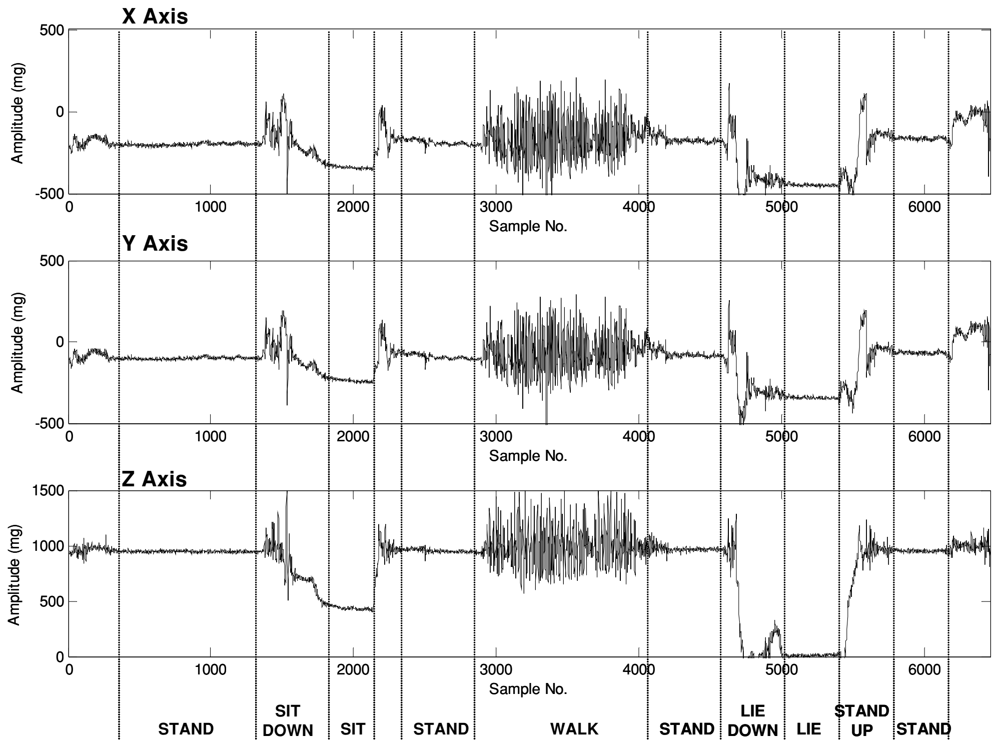

Figure 2. Representative data from the daily routine for each of the three axes of the triaxial device.

While the movements and postures contained within the routine are by no means a complete set of all possible activities that a given person might perform, they do form a basic set of simple activities which form an underlying structure to a person's daily life, and are likely to provide a great deal of information in terms of the person's balance, gait and activity levels if they can be accurately identified.

2.3. Rule-based Heuristic system

The rule-based Heuristic classification system is a modified version of that described in Mathie *et al* (2004), Mathie (2003). It is based on a hierarchical decision tree in which general decisions are made about the signal in the higher nodes of the tree, e.g. classifying activity versus rest, and then the classification is further refined in the lower branches, e.g. standing versus sitting versus lying. Most of the nodes contain Heuristic tests whereby a parameter of the signal is compared to some form of pre-set threshold; for example, is the tilt angle less than 20°? The full classification scheme is shown in the block diagram in figure 3.

2.4. Gaussian mixture model system

2.4.1. Overview. The GMM-based approach is a more general and arguably sophisticated approach to classification. GMMs are a parametric representation of a probability density function, based on a weighted sum of multi-variate Gaussian distributions. A separate GMM is trained for each movement and posture classification using the expectation maximization

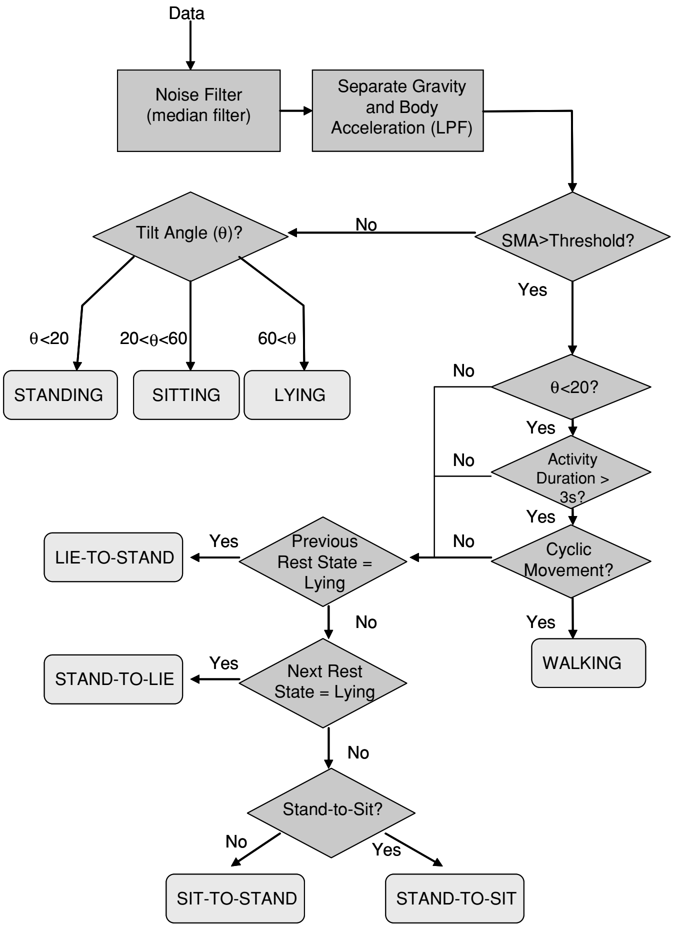

Figure 3. Block diagram of the rule-based Heuristic system used to classify various movements.

(EM) algorithm applied to data initialized with the *k*-means clustering algorithm; see Reynolds and Rose (1995) for further details. Test data are classified by selecting the GMM that gives the highest likelihood value of having produced those data. GMMs with 32 mixtures were found to give the best compromise between performance and computational load in the experiments described in section 3.

GMMs have been employed for a long time in the pattern classification literature and are commonly used in a range of applications including speaker and language identification (Reynolds and Rose 1995, Torres-Carrasquillo *et al* 2002, Allen *et al* 2005). The application of GMMs to this field is not new (Sung and Pentland 2005). The main distinction between our system and the use of GMMs in previous accelerometry systems is the use of a novel set of time-domain features (see section 2.4.2) and a proposed subject adaptation method (see section 2.4.3).

2.4.2. Feature extraction. A range of front-end parameters have been used in movement classification systems to date. Minimal comparison of the various parameterization methods has been performed and thus the optimal front-end processing methods for such applications remain unknown.

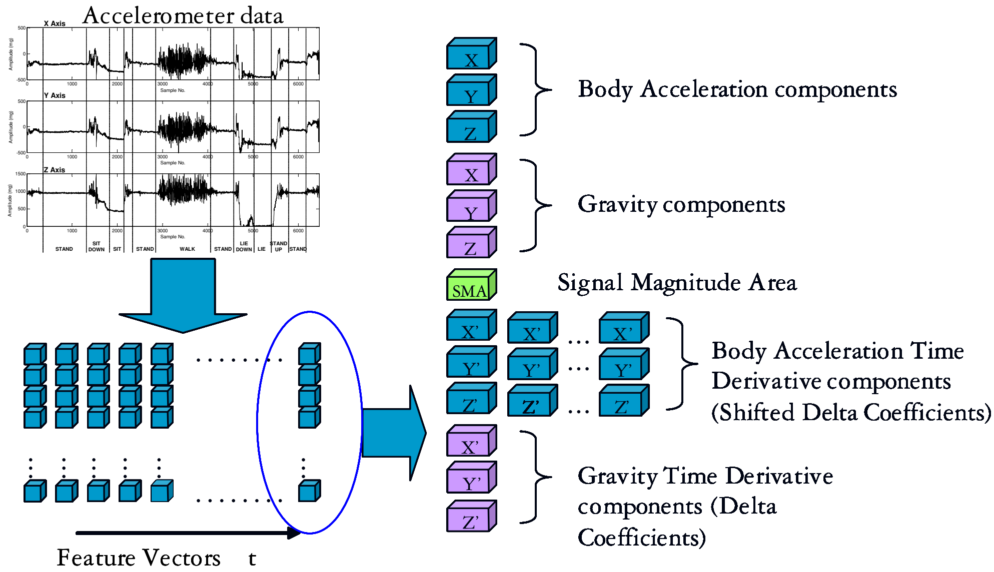

Figure 4. Block diagram showing the component features of the proposed feature extraction method used in the Gaussian mixture model system.

Most studies have used frequency-derived features employing an FFT or parameters such as averages or correlations calculated over long time-windows (Sung and Pentland 2005, Ravi *et al* 2005). While these features may be good for long duration, quasi-periodic signals like walking, cycling or brushing teeth, they are likely to be less effective in identifying shorter duration, non-periodic activities such as transitions between sitting and standing or taking a couple of steps. Since a trained eye can identify the signals fairly easily from their time-domain representation, time-domain features were selected rather than FFT-based features. These have the potential to improve both the ability of the system to detect short-duration movements and the computational complexity of the feature calculations.

Following extensive testing, the final proposed front-end uses 25-dimensional feature vectors composed of

- separated body acceleration (three features) and gravity components of the signal (three features),
- first-order delta coefficients for the gravity component (three features),
- a 3–3–3–5 shifted delta configuration for the body acceleration component (15 features),
- signal magnitude area (SMA) (one feature).

These are shown in figure 4 and will be described in the following subsections. Given the low sampling rate of 45 Hz, it was found to be computationally feasible to calculate a feature vector for every sample.

Gravity and body acceleration components. As previously described in section 2.1, the accelerometer signal is composed of two parts: acceleration due to movement of the wearer, i.e. body acceleration, and acceleration due to gravity. Both signals play an important, though distinct, role in characterizing a particular movement or posture. Consequently, these components are included separately in our feature vectors. The two components were separated by filtering. An eighth-order elliptic low-pass filter with a cut-off at approximately 0.25 Hz

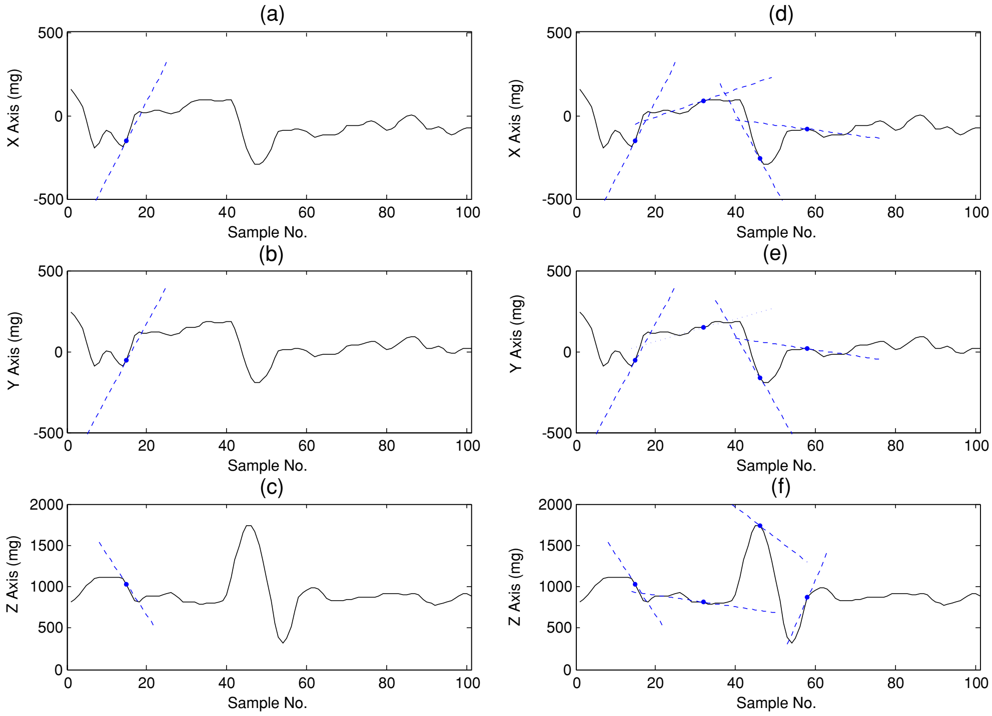

Figure 5. Diagrams (a)–(c) show a representation of the first-order delta coefficients, which approximate the slope of the time-domain waveforms. Diagrams (d)–(f) show a set of corresponding shifted delta coefficients, which are a sampling of the delta coefficients in the temporal vicinity of the sample.

was used to extract the gravity component. The body acceleration component was then found by subtracting the gravity component from the original signal. The final feature vectors include the signals from all three axes of both the gravity and body acceleration components (i.e. six features in total).

Delta coefficients. Delta coefficients are a method commonly used for representing temporal information in speech signal processing. Originally proposed by Furui (1986) they form a regression-based estimate of the first-order time derivative of the features on which they are based. For this application, first-order delta coefficients were calculated for each of the three axis components of the acceleration due to gravity according to

$$\Delta g(t) = \frac{\sum_{d=-D}^{D} d g(t+d)}{\sum_{d=-D}^{D} d^2} \tag{1}$$

where $g(t)$ is the gravity component at time t. See figures 5(a)–(c) for a graphical representation of these features.

Shifted delta coefficients (SDC). The SDC are a simple extension of the delta coefficients and have recently been used to better represent temporal information in speech signals for automatic language identification (Torres-Carrasquillo *et al* 2002, Allen *et al* 2005). They are calculated by taking a sampling of the delta coefficients from past or future frames and concatenating them in the current feature vector. This forms a numerical representation of the way the feature slope changes over time in the vicinity of the current feature (figures 5(d)–(f)). Given that the gravity component does not undergo rapid changes, little is gained by adding an extensive SDC component for gravity. However, the body acceleration component does undergo rapid changes in slope and thus useful information can be gained by including an SDC calculated from the body acceleration component.

The SDC are specified by four parameters: N, D, P and K. N is the number of base features from which they are calculated, D is the same D as in the delta calculations, P is the distance between samples and K is the number of samples taken. So the final shifted delta feature vector is given by the concatenation of $\Delta b(t + iP)$ for all $0 \leqslant i < K$, where

$$\Delta b(t + iP) = \frac{\sum_{d=-D}^{D} db(t + iP + d)}{\sum_{d=-D}^{D} d^2} \tag{2}$$

and $b(t)$ is the gravity component at time t.

After extensive experimentation, a parameter configuration of 3-3–3-5 was chosen to represent the body acceleration components of the signal (a total of 15 features). Given the 45 Hz sampling rate, this means that each feature incorporates detailed temporal information from the surrounding 0.4 s.

Signal magnitude area (SMA). The SMA has been found to be a suitable measure for distinguishing between activity and rest using triaxial accelerometer signals (Bouten *et al* 1997, Mathie 2003). It was used for that purpose in the Heuristic system described in section 2.3.

It is calculated according to

$$\mathrm{SMA} = \frac{1}{t} \times \left(\int_0^t |x(t)|\,\mathrm{d}t + \int_0^t |y(t)|\,\mathrm{d}t + \int_0^t |z(t)|\,\mathrm{d}t \right) \tag{3}$$

where $x[n]$, $y[n]$ and $z[n]$ are the discrete samples of the body components of the x-, y- and z-axis samples, respectively, and N is the window length over which the SMA value is calculated.

2.4.3. Adaptation methods to compensate for limited or mismatched training data. One of the difficulties associated with the practical deployment of an ambulatory monitoring system based on accelerometry data is the loss in accuracy caused by either training the system on data from other people or training the system on limited data from the person for whom it is intended.

If the system is trained on data from that person only, it is likely to be undertrained and therefore insufficiently robust to variation in the movements (since gathering sufficiently large training data sets from just that person is impractical). Whereas, if the system is pre-trained on a larger set of data, taken from multiple people, the models may not be specific enough to that person for accurate classification. For an algorithm like the EM algorithm (which can take a substantial amount of time to run), it is likely to be impractical to gather data from the intended subject and retrain the whole system using those data as well as other previously gathered data from multiple people.

As a possible solution to this problem, it is proposed that the GMMs should first be trained on a large data set from multiple subjects using the EM algorithm. Those models can then be adapted to the intended subject using Bayesian adaptation. A diagram of the proposed method is given in figure 6. A variation of this method is currently widely accepted as a solution to

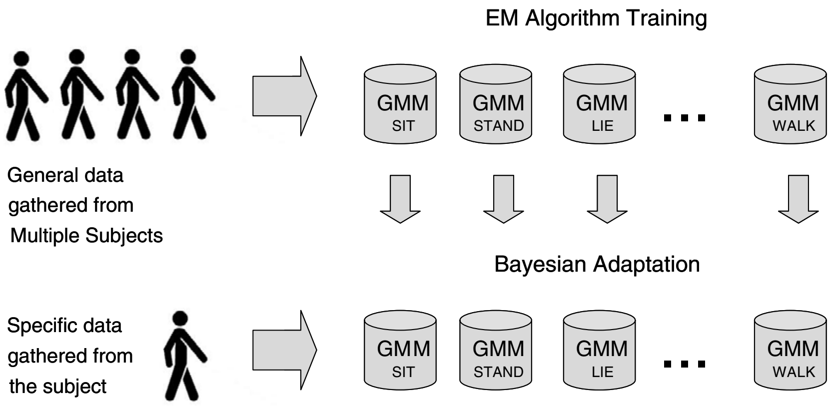

Figure 6. Proposed method of training movement GMMs. The GMMs are first trained on data from multiple subjects using the EM algorithm and then adapted to a specific subject using Bayesian adaptation.

the problem of limited training data in speaker verification and identification tasks (Reynolds *et al* 2000, Lucey and Chen 2003).

Bayesian adaptation works by adjusting Gaussian mixtures which are similar to values seen in the training data to take on values closer to those data whilst leaving mixtures which are further from the training data at roughly their original values. Thus in our proposed application for movement classification it is anticipated that the GMM mixtures similar to the subject-specific training data will become very close to those subject-specific data, thus incorporating the subject-specific information. Whereas, the GMM mixtures which are dissimilar to the subject-specific data will remain close to their original values, thus retaining the robustness and generality offered by the larger training set.

In this application, the adaptation of the GMMs is performed on the mixture means over successive iterations as follows (Reynolds *et al* 2000, Lucey and Chen 2003).

For the ith Gaussian mixture component, the probabilistic count of training observations is defined as

$$n_i = \sum_{t=1}^{T} P(i \mid \vec{x}_t, \lambda) \tag{4}$$

where λ is the parameterization of the GMM on the previous iteration and $P(i|\vec{x}_t, \lambda)$ is given by Reynolds and Rose (1995)

$$P(i \mid \vec{x}_t, \lambda) = \frac{w_i b_i(\vec{x}_t)}{\sum_{k=1}^{K} w_k b_k(\vec{x}_t)}. \tag{5}$$

The adaptation coefficient for the ith mixture component is then defined as

$$\alpha_i = \frac{n_i}{n_i + r} \tag{6}$$

where r is a fixed relevance parameter which determines how much similar new data must be observed before the new parameters will begin to replace the old ones ($r = 2$ or 4 was found to be best for accelerometry data).

The sufficient statistic is estimated according to

$$E_i(X) = \frac{1}{n_i} \sum_{t=1}^{T} P(i \mid x_t, \lambda) x_t. \tag{7}$$

Table 1. Comparison of accuracy levels obtained using the Heuristic system with that of the Gaussian mixture model system showing breakdown across movement categories.

		Heuristic system		GMM system	
Classification		Mean accuracy (%)	SD %)	Mean accuracy (%)	SD %)
Postures	Sitting	53.4	39.8	79.2	10.0
	Standing	85.9	17.5	77.3	20.3
	Lying	90.5	10.7	91.4	10.9
Motions	Sit-to-stand	28.9	20.9	93.1	6.2
	Stand-to-sit	53.4	14.1	88.3	6.6
	Stand-to-lie	85.8	11.9	98.5	2.3
	Lie-to-stand	90.5	9.2	98.9	2.7
	Walking	83.4	12.0	97.8	2.5
Total		71.1	5.7	91.3	4.2

The new GMM means, $\vec{\hat{\mu}}_i$, are then adapted from their previous values, $\vec{\mu}_i$, according to

$$\vec{\hat{\mu}}_i = \alpha_i E_i(X) + (1 - \alpha_i)\vec{\mu}_i \tag{8}$$

where γ is a scaling factor used to ensure that the adapted weights sum to unity.

3. Experimentation and results

3.1. The Heuristic system versus the GMM system

The GMM system was trained using the training data sets from all of the six subjects. The separate test data sets from each of the six subjects were then used to assess the classification performance of both the Heuristic and GMM systems. As described in section 2.2, the data were separated into individual movement and posture instances. A single classification was given for each movement in its entirety. Note that this method of classification is not possible in a real situation where the movement boundaries are unknown. However, it offers a reliable and effective means of assessing basic system accuracy.

Results are given as an average and standard deviation of the accuracy taken across the six subjects where the accuracy is the percentage of data segments in that category which were correctly classified. The results are displayed in table 1.

3.2. GMM system feature extraction experimentation

A number of configurations of the front-end feature vectors were trialed before arriving at the set detailed in section 2.4.2. Initial experimentation-tested system performance using a feature vector containing only the three noise-filtered (median filter) triaxial signals directly without further processing or separation into parts (i.e. a three-dimensional feature vector). This was then compared to the same system using features containing the separated gravity and body acceleration components (i.e. a six-dimensional feature vector). Further experimentation then considered the benefit obtained by the addition of the SMA value to the features above. Finally, the benefit of the addition of delta and SDC coefficients to the system was considered in two stages, first considering the use of first-order delta coefficients for both the gravity and body acceleration components of the signal and then adding a more extensive 3-3–3-5 configuration of the shifted delta coefficients for the body acceleration component. The results are given in table 2.

Table 2. Benefit of each separate feature component in terms of overall accuracy.

Features	Accuracy (%)
Unseparated triax signals (three features)	67.0
Triax signals separated into gravity and body acceleration components (six features)	82.9
Separated triax signals with SMA (seven features)	83.6
Separated triax signals with SMA plus delta coefficients for both gravity and body acceleration components (13 features)	88.2
Separated triax signals with SMA plus delta coefficients for the gravity component and 3-3–3-5 SDC coefficients for the body acceleration component (25 features)	91.5

Table 3. Results of experiments demonstrating the effectiveness of the proposed adaptation technique.

	Accuracy (%)							
	Subject							
Training scheme	1	2	3	4	5	6	Mean	SD
Same subject training	90.0	89.4	87.5	96.8	80.7	84.6	88.2	5.0
Other subject training	67.7	75.6	76.3	83.1	76.4	80.5	76.6	4.8
Other subject training + same subject adaptation	94.3	92.9	92.6	98.7	87.0	87.8	92.2	4.0

3.3. GMM subject adaptation experimentation

In these tests, the performance of the GMM system was considered when faced with limited training data or data trained on other subjects, and the proposed adaptation method was tested. For each of the six subjects, three tests were run. All three tests used the same test data for evaluation—the test data set aside for that particular subject. In the first tests, the GMMs were trained on the designated training sets from that subject only. In the second tests, the GMMs were trained on both the designated training and test sets from the other five subjects. In the final tests, the GMMs were first trained on all data from the other five subjects (as in the second tests) but were then adapted using the designated training data of that subject. The results are displayed in table 3.

For all six subjects, the adapted GMM method outperformed both of the other training methods.

Further experimentation was conducted to determine how much user-specific training data was required by the adaptation algorithm for the performance to improve. This was measured by considering the number of routine iterations required to be performed by each subject before improved performance was observed. If the routine is performed once per day, then the time taken for the system to adapt to the new user can be measured in days. The results are displayed in table 4.

In all cases, reduced performance was obtained by adapting the GMMs using only one routine iteration; however, adapting the GMMs using just five routine iterations was sufficient to improve system performance for all subjects except subject 2. Adaptation using ten routine iterations was sufficient to increase accuracy for all six subjects and after 21 routine iterations, significant performance improvements were obtained for all subjects with only minor improvements at larger data quantities.

Table 4. Considering the number of routine iterations required for successful adaptation. Grey values indicate where insufficient data are available for further adaptation to take place.

Number of routine iterations for adaptation	Subject 1	2	3	4	5	6	Mean
0 (no adaptation)	67.7	75.6	76.3	83.1	76.4	80.5	76.6
1	59.4	43.7	62.9	91.3	68.2	62.8	64.7
5	81.2	62.2	79.1	95.9	83.1	83.3	80.8
10	87.9	85.8	89.4	96.1	85.4	86.4	88.5
21 (3 weeks)	92.9	89.4	93.7	98.7	87.0	87.0	91.5
28 (4 weeks)	93.8	91.2	92.6			87.8	91.8
35 (5 weeks)	94.4	92.8					92.2
42 (6 weeks)	94.3	92.9					92.2

4. Discussion

These results show that the Heuristic system achieves some success in correctly identifying movements. In particular, it is able to distinguish lying from upright positions fairly effectively and consequently is able to identify transitions to or from lying with some success. It is also reasonably competent at identifying walking. However, it has great difficulty distinguishing between sit-to-stand and stand-to-sit transitions and between sitting and standing. This is demonstrated by the lower mean accuracy levels and the larger variability of results between subjects for those classifications.

The decision to use this particular set of heuristics is based on a previous work in which a large amount of careful study was conducted on the differentiating aspects of the movements and postures considered. However, this particular set of rules is still a fairly arbitrary set and the accuracies found are by no means a representation of the limits obtainable with other heuristic systems. In fact, given a particularly refined and accurate set of Heuristic rules and an optimal arrangement of the nodes within the tree, such a system has the potential to make very accurate and computationally efficient decisions regarding the classification of a given movement. The primary drawback of such a system, however, is that finding such a set of rules is an incredibly time-consuming process involving what is largely a manual search through the data for similarities and differences between the signal classifications.

The addition of other motions or postures to the system also presents significant problems. In the best case, it may be a simple case of adding an extra leaf node to one of the branches. In the worst case, it may require a complete restructuring of the tree in order to arrive at the best classification results. Consequently, more automated methodologies such as the GMM are desirable.

The GMM system obtained a higher mean accuracy for all motions except standing and correspondingly lower variances, than the Heuristic system, indicating a significantly better overall performance. The Heuristic system obtained a higher mean accuracy for identifying standing. However, this came at the cost of a significantly reduced ability to identify sitting.

Unsurprisingly, given the use of only a single accelerometer attached at the waist, both the Heuristic and the GMM system have some difficulty distinguishing standing from sitting, commonly confusing the two. The results are not wholly discouraging however, and by no means invalidate the efficacy of using only a single accelerometer. Firstly, because the GMM system still has a far from trivial capability for distinguishing between sitting and standing, with mean accuracies of around 78% for each. And secondly, because the GMM system identifies transitions between sitting and standing with a high degree of accuracy (close to 90% for both) and very rarely confuses the two. In fact, the loss of accuracy in identifying sit-to-stand and stand-to-sit transitions is more often due to confusion with lie-to-stand and stand-to-lie transitions, respectively, rather than confusion between the two. This is not surprising given that the lie-to-stand will often consist of a lie-to-sit and sit-to-stand transition and similarly for the stand-to-lie transition. This improved knowledge of transitional movements between sitting and standing should allow better estimation of whether the person is standing or sitting through the application of Markov chains or other forms of sequence modelling, i.e. if the system recognizes that the person is currently either sitting or standing and they have just undergone a sit-to-stand transition, then it can infer that the person is now standing. Other indications such as duration might also improve system performance since an elderly person is unlikely to remain standing still for long periods of time. The use of such information would be entirely meaningless in analysing data from a directed routine however, and consequently the incorporation of such methods could not be conducted as part of this study and must be left to future work.

It is also interesting to note that the inclusion of short-term temporal information in the signal via the SDC is sufficient to model and identify the walking signals without resorting to a direct measure of periodicity such as the FFT. This might potentially allow us to better identify short-duration periods of walking, such as taking only a few steps, where the cyclic nature of the gait is not yet apparent. Further tests are required, however, to determine whether these features are also sufficient for distinguishing walking from other periodic activities such as brushing teeth, sweeping or walking up stairs.

The proposed adaptation method was found to give improved classification performance for all subjects tested, overcoming the problem of limited training data (as encountered in the first tests) and the problem of mismatched training data (as encountered in the second tests). It shows great promise for allowing robust, pre-trained models to be generated prior to system deployment and then trained for the particular person or device once they are installed. The experiments in section 3.3 indicate that only a small amount of annotated user-specific data, around 21 iterations of the daily routine, are required to achieve significant improvements in classification performance for a given subject.

The use of six different subjects in six different places of residence, the extended duration of the study such that all routine iterations were at least 24 h apart providing a wide variety of device placements, and subject health statuses lead us to believe that the modelling capability of the system is fairly robust to variation. However, like all previous similar studies, this study used only a small subset of possible motions and postures and the classification decisions are based on closed-set choices where no ‘other’ option was made available. It is likely that given sufficient training data, this system could be simply extended to apply to a variety of other movements, e.g. walking up stairs, sweeping, bending over etc, and would retain similar levels of robustness and variation in doing so. The primary unknown, however, is the extensibility of this system to general unconstrained movements which cannot be modelled by training data, since it is impossible to gain training data for all possible movements. In fact, the movement type which is most conspicuously absent from this study, and also one of the most difficult to gather representative data for, is ‘falls’. Ultimately, for a system such as this one to be successful in real-world applications there must be a capability for the system to identify ‘unusual’ or ‘unknown’ movement categories, where a fall will be one such possibility. This remains a significant challenge in this area which future work will hope to address. The current study, however, offers a significant step towards a more robust and realizable system for human movement classification from accelerometry data.

5. Conclusion

Gaussian mixture models show significant potential in their ability to accurately and robustly model ambulatory data from a single triaxial accelerometer. They significantly outperform the Heuristic approach previously employed by this research group and are simpler to implement and optimize.

Using a single, waist-mounted triaxial accelerometer limits the ability of the system to directly distinguish between sitting and standing postures. However, given that the GMMs can accurately identify stand-to-sit transitions and sit-to-stand transitions, a single accelerometer may still prove sufficient to identify these postures once sequence modelling (e.g. Markov chains) are implemented. Thus we postulate that a single triaxial accelerometer will be sufficient for accurate movement classification in long-term ambulatory monitoring.

GMMs are likely to be one of a number of possible automated classification methodologies which achieve a high degree of accuracy in movement classification tasks. Other methodologies likely to perform equally well include support vector machines, and others. In implementing any of these systems, we propose that time-domain features such as those described here may be sufficient to accurately model and identify ambulatory motions. These have the potential to better identify short-duration movements such as stand-to-sit transitions and taking of only a few steps.

The method proposed here for adapting the GMMs to more accurately model a particular person or device via Bayesian adaptation also shows great promise. Similar methods are likely to be beneficial to any such classification system. This method is a very simple one by which good results can be achieved for Gaussian mixture models.

References

Allen F R *et al* 2005 Language identification using warping and the shifted delta cepstrum *IEEE Int. Workshop on Multimedia Signal Processing (Shanghai, China)*

Bao L and Intille S S 2004 Activity recognition from user-annotated acceleration data *PERVASIVE 2004 (Berlin Heidelberg)*

Bouten C V *et al* 1997 A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity *IEEE Trans. Biomed. Eng.* **44** 136–47

Bussmann J B *et al* 2001 Measuring daily behaviour using ambulatory accelerometry: the activity monitor *Behav. Res. Methods Instrum. Comput.* **33** 349–56

Foerster F *et al* 1999 Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring *Comput. Hum. Behav.* **15** 571–83

Furui S 1986 Speaker-independent isolated word recognition using dynamic features of speech spectrum *IEEE Trans. Acoust. Speech Signal Process.* **34** 52–9

Kiani K *et al* 1997 Computerized analysis of daily life motor activity for ambulatory monitoring *Technol. Health Care* **5** 307–18

Lee S H *et al* 2003 A study of the activity classification using a triaxial accelerometer *IEEE EMBS*

Lucey S and Chen T 2003 An investigation into subspace rapid speaker adaptation for verification *ICME: IEEE Conf. Multimedia and Expo (Baltimore MD, USA)*

Mathie M 2003 Monitoring and interpreting human movement patterns using a triaxial accelerometer *PhD Thesis* Faculty of Engineering. Sydney, Australia, University of New South Wales

Mathie M J *et al* 2004a Classification of basic daily movements using a triaxial accelerometer *Med. Biol. Eng. Comput.* **42** 670–87

Mathie M J *et al* 2004b Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement *Physiol. Meas.* **25** R1–R20

Mathie M J *et al* 2004c A pilot study of long term monitoring of human movements in the home using accelerometry *J. Telemed. Telecare* **10** 144–51

Ravi N *et al* 2005 Activity recognition from accelerometer data *12th National Conf. on Artificial Intelligence*

Reynolds D A *et al* 2000 Speaker verification using adapted Gaussian mixture models *Digit. Signal Process.* **10** 19–41

Reynolds D A and Rose R C 1995 Robust text-independent speaker identification using Gaussian mixture speaker models *IEEE Trans. Speech Audio Process.* **3** 72–83

Salleh R *et al* 2000 Low power tri-axial ambulatory falls monitor *10th Int. Conf. on Biomedical Engineering (Singapore)*

Sung M *et al* 2005 Wearable feedback systems for rehabilitation *J. Neuroeng. Rehabil.* **2** **See endnote 1**

Sung M and Pentland A 2005 Minimally-invasive physiological sensing for human-aware interfaces *HCI International*

Torres-Carrasquillo P A *et al* 2002 Approaches to language identification using Gaussian mixture models and the shifted delta cepstrum *ICSLP*

United Nations 2003 Long-range world population projections: based on the 1998 revision. The Population Division, Department of Economic and Social Affairs, United Nations Secretariat

Endnotes

(1) Author: Please provide the page number in reference 'Sung *et al* 2005'.
(2) Author: Please be aware that the colour figures in this article will only appear in colour in the Web version. If you require colour in the printed journal and have not previously arranged it, please contact the Production Editor now.
