A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities From Accelerometer Data

Stephen J. Preece*, John Yannis Goulermas, *Member, IEEE*, Laurence P. J. Kenney, and David Howard

Abstract—Driven by the demands on healthcare resulting from the shift toward more sedentary lifestyles, considerable effort has been devoted to the monitoring and classification of human activity. In previous studies, various classification schemes and feature extraction methods have been used to identify different activities from a range of different datasets. In this paper, we present a comparison of 14 methods to extract classification features from accelerometer signals. These are based on the wavelet transform and other well-known time- and frequency-domain signal characteristics. To allow an objective comparison between the different features, we used two datasets of activities collected from 20 subjects. The first set comprised three commonly used activities, namely, level walking, stair ascent, and stair descent, and the second a total of eight activities. Furthermore, we compared the classification accuracy for each feature set across different combinations of three different accelerometer placements. The classification analysis has been performed with robust subject-based cross-validation methods using a nearest-neighbor classifier. The findings show that, although the wavelet transform approach can be used to characterize nonstationary signals, it does not perform as accurately as frequency-based features when classifying dynamic activities performed by healthy subjects. Overall, the best feature sets achieved over 95% intersubject classification accuracy.

Index Terms—Activity classification, ambulatory monitoring, machine learning, wavelet transform.

I. INTRODUCTION

OVER THE past decade, there has been considerable research effort directed toward the monitoring and classification of physical activity patterns from body-fixed sensor data [1], [2]. This has been motivated by a number of important health-related applications. For example, with the trend toward more sedentary lifestyles, there is growing interest in the link between levels of physical activity and common health problems, such as diabetes, cardiovascular disease, and osteoporosis [3]. As self-reported measures have been shown to be unreliable [4], [5], systems for activity profiling are beginning

*S. J. Preece is with the Centre for Rehabilitation and Human Performance Research, University of Salford, Salford M6 6PU, U.K. (e-mail: s.preece@salford.ac.uk).

J. Y. Goulermas is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3BX, U.K. (e-mail: j.y.goulermas@liverpool.ac.uk).

L. P. J. Kenney and D. Howard are with the Centre for Rehabilitation and Human Performance Research, University of Salford, Salford, M6 6PU, U.K. (e-mail: l.p.j.kenney@salford.ac.uk; d.howard@salford.ac.uk).

to play an important role in large-scale epidemiological studies in this area [6], [7]. Furthermore, such systems can also be used to assess the effectiveness of different interventions aimed at increasing levels of physical activity and for motivating individuals to become more physically active.

The success of a given rehabilitation program is often judged by not only the levels of activity but also the type of activity that an individual can return to after treatment. In addition, as fall risk increases with age, so a better understanding of the factors contributing to fall risk becomes more important. Ambulatory monitoring of various activities, including the time spent in sit–stand transitions, has shown promise as a predictor of fall risk [8]. Furthermore, both type and intensity of individual's activity are of interest to urban designers, and designers, manufacturers, and purchasers of certain medical devices (e.g., advanced responsive pacemakers and orthopedic implants).

In addition to health-related applications, portable systems, which can accurately identify the activity of the user, have the potential to play a fundamental role in a ubiquitous computing scenario [9], [10]. In this field, computing devices use information from a variety of sensors to determine the context of a situation. Different devices can then use the context information to deliver an appropriate service. For example, a mobile phone may detect when a person is driving a vehicle and automatically divert a call.

With recent advances in miniaturized sensing technology, it is now possible to collect and store acceleration data from individual body segments over extended periods of time. Although this technology offers the ideal platform for monitoring daily activity patterns, effective algorithms are also required to interpret the accelerometer data in the context of different activities. Previous studies have shown machine learning or artificial intelligence approaches to be effective for identifying a range of different activities from body-fixed sensor data [11]–[14]. These techniques typically operate via a two-stage process [15]. First, features are derived from windows of accelerometer data. A classifier is then used to identify the activity corresponding to each separate window of data. A range of different approaches has been used to obtain features from accelerometer data, with some researchers deriving features directly from the time-varying acceleration signal [12], [16]–[18] and others from a frequency analysis [11], [13], [19], [20]. More recently, wavelet analysis has been used to derive the so-called time-frequency features [14], [21]–[24].

With wavelet analysis, the original signal is decomposed into a series of coefficients, which carry both spectral and temporal

information about the original signal. From these coefficients, it is possible to identify localized temporal instances at which there is a change in frequency characteristics of the original signal [25]. This concept has been applied successfully to accelerometer signals in order to identify points in the signal at which the subject changes from one activity to another [22], [24]. As well as being used to locate discrete temporal events, wavelet analysis can also be used to derive time-frequency features that characterize the original signal. However, it is not clear whether such time-frequency features lead to more effective activity classification than the more commonly used time-domain or frequency-domain features.

The overall aim of this study was to extensively compare the performance of a number of previously reported and novel wavelet features with a range of time-domain and frequency-domain features for the classification of different activities. Many previous wavelet-based studies have investigated level walking, stair ascent, and stair descent [21]–[23], but have not compared their performance against simpler approaches. Therefore, our first research aim was to compare features for this three-activity classification problem. As a second aim, we sought to compare the same features for a larger set of activities that represents a more challenging problem. Additionally, since the performance of a given set of features can be dependent on the location of the sensor, we compared accuracy for the different features across a number of different lower limb placements. It was felt that this study would underpin the development of an off-the-shelf activity monitor that could be used to classify activity patterns across different subjects.

II. METHOD

A. Data Collection

Accelerometer data were collected using Pegasus activity monitors developed by ETB, U.K. Each of these units contained a triaxial accelerometer, with a dynamic range of ± 5 g, which was sampled with 10-bit resolution. With these devices, it is possible to sample accelerometer data at a user-defined frequency and to store these data for up to 24 h. A sampling frequency of 64 Hz was selected for this study as this is sufficiently higher than the 20 Hz sampling required to assess daily activity [26]. A number of previous activity classification studies have used wavelet analysis to derive features from accelerometer signals collected at relatively high sampling frequencies (>250 Hz). However, for this study, 64 Hz was chosen as this is a realistic sampling frequency that could be implemented by an off-the-shelf activity monitor. No antialiasing filtering was applied to the acceleration data.

For each subject, data were collected with three activity monitors. These were attached to waist (at the sacrum), the thigh (just above the knee), and the ankle (just above the lateral maleollus). To secure each unit in place, specialized bandage (Fabrifoam) was first positioned around each of the body segments and the activity monitors, which were backed with Velcro, adhered to the underwrapped bandage. Once in position, additional bandage was then wrapped over each sensor to ensure no movement could occur from the overlying clothing. This method of

Fig. 1.   Activity monitors attached to the ankle and thigh.

attachment has been illustrated in Fig. 1 for the ankle and thigh placements.

Ten male and ten female subjects participated in the study. As large individual variation has been reported for accelerometer signals corresponding to the same activity [27], subjects with a range of ages and body mass indices were recruited into the study. The mean [standard deviation (SD)] age of the subjects was 31 (7) years, mean (SD) height was 1.71 (0.07) m, and the mean (SD) weight was 68 (10) kg. The subjects covered a wide range of body mass indices from 19 to 30 with mean (SD) 24 (3). Each subject gave informed consent to participate in the trial after approval had been obtained from the Ethical Committee at the University of Salford.

A number of studies have shown that static postures can be differentiated from dynamic activity by applying a single threshold to some measure of acceleration variability [28], [29]. Provided sensors are attached to more than one body segment, it is possible to accurately identify different static postures using a threshold-based approach [30], [31]. However, the situation is more complicated with only a single sensor. In this scenario, more complex signal processing along with an appropriate biomechanical model is required to differentiate between different postures, postural transitions, and continuous dynamic activity [32]. For this study, we chose to investigate the classification of continuous dynamic activities. This choice was motivated by previous studies that have used a range of different features to characterize acceleration signals [11], [14], [16], [17], [21]–[24], [33]–[35].

Subjects completed a total of eight different activities (level walking, walking upstairs and downstairs, jogging, running, hopping on the left and right leg, and jumping). Each of these activities was part of a continuous circuit that started in a building and then followed a route around the university campus. This circuit was described to the subjects before the start of data collection. During the trial, the experimenter recorded the sequence of activities with a portable video camera giving minimal prompting to the subject. With this design, the subjects were free to move at their preferred pace and to transition between different activities when they felt most comfortable.

Fig. 2. GUI used to annotate the accelerometer data from the video record.

To ensure that there were sufficient data to address the first research question, the circuit involved stair walking both inside and outside the building as well as level walking in a number of different environments. In addition to these three everyday activities, both jogging and running were also included in the circuit. For the first of these two activities, subjects were instructed to perform a gentle jog over a 50 m distance, and for the second to perform a fast run over the same distance. Both these activities have been used in previous classification studies [11], [12], [36] and their recognition could prove invaluable in any activity monitoring system for sports rehabilitation. We wanted to collect data across a range of different modes of locomotion and therefore included three additional activities: hopping (on each leg) and jumping. Both hopping [37] and jumping [36] have been used in previous activity monitoring studies and are also used in sports rehabilitation. In order to include each of these activities as part of the circuit, each subject was required to hop (on each leg separately) over a 15 m distance and to jump, moving both legs together, over the same distance.

Just prior to data collection, the three activity monitoring units were synchronized with each other and with the clock of a laptop computer. This procedure was repeated at the end of each experiment to ensure that the units had not drifted relative to each other. Resynchronization was not needed as the units never drifted by more than three to four samples (0.05 s). Custom software was developed in Matlab (The MathWorks, USA) so that the video data could be synchronized with the laptop and thus the accelerometer data. Following data collection, this software was used to annotate the accelerometer data with the transition points between each of the different activities (see Fig. 2). This method allowed for rapid and accurate labeling of the data, which was particularly important for identifying stair ascent and descent as these activities only lasted approximately 10 s. A small pilot study demonstrated minimal (<1 s) intertester variability in the identification of the activity transition points.

Once activity transition points had been identified, features were calculated from 2-s (128-sample) consecutive windows that overlapped by 1 s. The use of a 50% overlap between successive sliding windows has been shown to be effective in previous studies of activity classification [11], [38]. The choice of a 2 s window was motivated by previous studies by Nyan

et al. [24] (2 s) and Wang *et al.* [14] (2.56 s) that had used similar length windows. It was not possible to use shorter windows as signals with less than 128 samples could not be fully decomposed into wavelet coefficients appropriate for comparison with other studies (see Section 2.2). Longer windows limited the amount of data that could be extracted from short-duration activities, such as stair walking. Pilot work also showed minimal differences between classification accuracies calculated from frequency-domain features derived from 2 s or 3 s windows.

If a window corresponded to a transition between two activities, it was excluded from subsequent analysis. Given the continuous nature of the circuit completed by the subjects, there was a disproportionate number of windows of data that corresponded to level walking. Therefore, in order to balance the distribution of the different activities, only a randomly chosen subset of these windows was used in the final analysis.

B. Wavelet Features

A number of previous activity classification studies have derived time-frequency features obtained using the filter bank interpretation of the discrete wavelet transform [22], [24]. With this approach, the original time-domain signal (maximum frequency f) is initially decomposed into a coarse approximation and detail information by low-pass filtering (bandpass $[0, f/2]$) and high-pass filtering (bandpass $[f/2, f]$), respectively [39]. With wavelet decomposition, the half-band filters are designed to enable perfect reconstruction of the original signal and to avoid aliasing effects. In subsequent levels of decomposition, the approximation signal from the previous level is split into a second approximation and a detail coefficient. This process is repeated to the desired decomposition level. For further details, see [40]. Five separate studies were identified, which had previously used wavelet features for classification of accelerometer data [14], [21]–[24]. These studies were then used as a basis for defining seven sets of wavelet features (Table I).

The first set of wavelet features was proposed by Tamura *et al.* [23]. With this approach, the accelerometer signal is decomposed using the wavelet transform and the features defined as signal power measurements, calculated as the sum of the squared detail coefficients at levels 4 and 5. Tamura *et al.* [23] sampled acceleration data at 250 Hz. Given our lower sampling frequency of 64 Hz, we calculated the two features from detail coefficients corresponding to the same frequency bands as those used by Tamura *et al.* [23]. This process of identifying corresponding wavelet coefficients for our lower sampling frequency was performed for other wavelet feature sets where needed.

The second set of features was taken from Nyan *et al.* [24]. These features are calculated in a similar way to Tamura *et al.* [23]; however, rather than treating the scales separately, the summations at levels 4 and 5 are added together. Features suggested by Sekine *et al.* [22] form the basis of the third set of features. Again, there are two features, the first being the total of the summations of the detail signal at levels 6 and 7. This quantity is divided by the number of steps (N), which is obtained by counting the number of times the signal, reconstructed from levels 6 and 7, changes sign. For the second feature, the

TABLE I
SUMMARY OF THE DIFFERENT WAVELET FEATURES

Publication	Wavelet Mother	No of Features	Description of each feature
Tamura et al. [23]	Daubechies 3	6	$\| cD_4 \|^2$ & $\| cD_5 \|^2$ for all three acceleration components
Nyan et al. [24]	Daubechies 5	2	$\sum_{j=4}^{5} \| cD_j^{Vert} \|^2$ & $\sum_{j=4}^{5} \| cD_j^{AP} \|^2$ Where the superscripts *Vert* and *AP* refer to the wavelet coefficients derived from the vertical and anterior-posterior accelerations
Sekine et al. [22]	Daubechies 2	2	$\dfrac{1}{N}\sum_{j=6}^{7} \| cD_j^{Vert} \|^2$ & $\sum_{j=4}^{7} \| cD_j^{AP} \|^2 \Big/ \| x^{AP} \|^2$ *Where x^{AP}* represents the AP accelerometer signal.
Wang et al. [14]	Daubechies 5	33	$\sum_{j=2}^{6} \| cD_j \|^2 + \| dA_3 \|^2 + \| dD_3 \|^2$ Where dA_3 and dD_3 represent the third level wavelet packet approximation and detail coefficient respectively. This feature was calculated for all three components of acceleration along with measures of standard deviation and RMS for wavelet coefficients at different levels.
Sekine et al. [21]	Daubechies 4	3	$FractalDimension = 2 - (\beta - 1)/2$ To calculate fractal dimension, the variance of the detail coefficient is plotted against the decomposition level. The parameter β is the gradient of resulting line. This feature was calculated for each component of acceleration.
Squared coefficients	Daubechies 2	15	$\| cD_1 \|^2$, $\| cD_2 \|^2$, $\| cD_3 \|^2$, $\| cD_4 \|^2$ & $\| cD_5 \|^2$ for each component of acceleration
Magnitude coefficients	Daubechies 2	15	$\| cD_1 \|_1$, $\| cD_2 \|_1$, $\| cD_3 \|_1$, $\| cD_4 \|_1$, $\| cD_5 \|_1$ for each component of acceleration.

The term cDj refers to the detail coefficient at the jth level of decomposition. All other nomenclature is explained within the table.

total of the summations of the detail signal from levels 4 to 7 is normalized against the sum of the squares from the original signal. Although [22] used a Coiflet wavelet mother for wavelet decomposition, our preliminary investigation showed improved classification with a Daubechies wavelet mother. This was therefore used for subsequent analysis. Both [24] and [22] collected data at 256 Hz.

Most previous activity classification studies have used wavelet analysis to derive only a small number of features. In contrast, Wang *et al.* [14] used wavelet packet analysis to derive 33 features from a triaxial accelerometer signal. With wavelet packet analysis, the detail coefficients are split into a further approximation and detail coefficients. This allows additional information to be extracted from the original signal. The features suggested by Wang *et al.* [14] involved summing the squares of the detail and wavelet packet approximation coefficients across different levels. In addition, they calculated standard deviations and rms values of detail and wavelet packet approximation coefficients at a number of different levels. In their study, Wang *et al.* [14] sampled accelerometer data at 50 Hz; therefore, our data were resampled to this frequency.

The fifth set of features is based on the concept of fractal dimension, which was used by Sekine *et al.* [21] to characterize accelerometer signals. The fractal dimension quantifies the variance progression of the detail coefficient over the different wavelet scales and as such gives a measure of the complexity within the original signal [41]. Given the high sampling frequency used by Sekine *et al.* [21] (1024 Hz), they were able to calculate the fractal dimension from the variance of the detail coefficients across seven different levels. Due to our lower sampling frequency of 64 Hz, fractal dimension was estimated from variance progression across three levels. Although this may lead to poorer discriminate ability for this feature set, the use of additional detail coefficients was not possible with our lower sampling frequency.

In addition to the five sets of wavelet features described earlier, we experimented with some alternative wavelet features. Two additional feature sets were then included in this study (Table I). For both of these feature sets, each component of the 64-Hz triaxial acceleration signal was decomposed to five levels using a Daubechies 2 wavelet mother. A sixth wavelet feature set was then defined as the sum of the squared detail coefficients at levels 1–5. These five features were calculated for each component of acceleration, thus giving a total of 15 features. The seventh feature set was obtained in a similar way, but the sums of the absolute values were used to provide a different type of combining norm. All wavelet features in Table I were derived for every window of accelerometer data using Matlab ver.7.4 (The Mathworks, USA).

C. Time- and Frequency-Domain Features

For additional comparison, we also employed three sets of time-domain features and four sets of frequency-domain

TABLE II
SUMMARY OF THE TIME- AND FREQUENCY-DOMAIN FEATURES

No of Features	Description of each feature
6	Mean and SD
15	Mean, SD, median and 25th and 75th percentile
6	Mean low and mean rectified high pass filtered signals
3	Principal frequency
3	Spectral energy
12	Bao and Intille [11]: Mean dc, energy, entropy and correlations between axes
15	Magnitude of first five components of FFT analysis

TABLE III
CLASSIFICATION ACCURACIES (%) FOR THE THREE-ACTIVITY CLASSIFICATION
PROBLEM (LEVEL WALKING, STAIR ASCENT, AND STAIR DESCENT) WITH THE
WAVELET FEATURES (TABLE I)

	Tamura *et al.* [23]	Nyan *et al.* [24]	Sekine *et al.* [22]	Wang *et al.* [14]	Sekine *et al.* [21]	Sum of squares	Absolute Sum
Waist	60	56	67	68	43	69	72
Thigh	62	60	50	65	56	71	74
Ankle	69	66	54	69	42	74	81
Waist and Thigh	71	67	72	83	59	82	88
Waist and Ankle	74	74	72	84	55	79	89
Thigh and Ankle	85	82	61	92	62	91	92
All	85	85	73	95	65	93	95

Accuracies have been reported for each of the different accelerometer combinations.

features (Table II). Within each of these seven sets, the features were derived individually for each of the three components of the triaxial accelerometer signal. Mean and standard deviation (SD) have been used in previous studies [34] to characterize windows of accelerometer data. As an extension of this set, we defined the multiple statistics features set, which additionally included median and 25th and 75th percentile [33]. Low-pass filtering is commonly used to separate the dc and ac components of an accelerometer signal [42]. Previous studies have defined features as the mean dc and the mean of the rectified ac signal [16], [17]. These two statistics were therefore used to define the third set of time-domain features.

In order to derive frequency-domain features, a fast Fourier transform (FFT) was performed on each 2 s window. The principal frequency was defined as the first of the frequency-domain feature sets (fourth in Table II). This has been used previously as an addition to time-domain measures in order to improve classification accuracy [35]. The second frequency-domain feature set was chosen to be spectral energy, which is defined to be the sum of the squared FFT coefficients [11], [43]. A recent study carried out by Bao and Intille [11] obtained high levels of classification accuracy using a mixed set of time and frequency-domain features. Therefore, this was included as the sixth set of features. In addition to spectral energy [11], the sixth set included dc, correlations between axes, and frequency-domain entropy. This latter feature gives a measure of the normalized information entropy of the FFT components and allows for differentiation between activities that have simple acceleration patterns and those with more complex patterns [11]. The final frequency-domain feature set was defined as the magnitude of the first five components of the FFT power spectrum. As with the other feature sets, this set of parameters was derived separately for each of the three components of acceleration. Although it is more common to use the power spectrum of FFT coefficients, preliminary studies showed that the magnitudes gave improved accuracy and were therefore used for the final analysis.

D. Activity Classification

In order to compare the discriminate ability of each of the different features sets, a k-nearest neighbor (kNN) classifier was implemented and its accuracy determined using leave-one-subject-out cross validation. This type of classifier has been shown to be effective in previous activity recognition studies [11], [12] and selects the activity that is closest to the feature under question using the Euclidean distance metric in the multidimensional feature space. We employed kNN as our recognition engine, due to its implementational simplicity and flexibility, and the fact that it can allow analysis of the classification decisions. With leave-one-subject-out cross validation, the classifier is trained with data from all subjects except one and then tested with data from the excluded subject. This process is repeated until each subject has been used once as the testing dataset. With this approach, the overall accuracy is calculated as the average test classification result of each train-test repetition. Cross validation is a popular statistical resampling procedure [44] and we use it here to evaluate the accuracy of the kNN classifier for a given set of features. The mean accuracy of all train-test repetitions can be influenced by a small number of subjects who may bias the overall result. Therefore, in order to compare the performance of two sets of features, the Mann–Whitney U test was used to test for differences in the two distributions of testing accuracies. This test was chosen as it was not possible to guarantee that these distributions were normally distributed. A significance level of $p < 0.01$ was used throughout.

To address our first research aim, only windows of data that corresponded to level walking, stair ascent, and stair descent were included in the analysis. For this three-activity classification problem, accuracy was determined for the waist-mounted accelerometer for each of the seven sets of wavelet features and for each of the seven sets of time-frequency features. This process was then repeated for the thigh and then the ankle-mounted sensor. To establish whether it would be possible to improve classification accuracy using data from more than one sensor, the analysis was performed for all seven possible combinations of the three sensors (as shown in the first column of Table III). Once classification accuracies had been determined for the three-activity problem, the process was repeated with windows of accelerometer data from all eight activities.

III. RESULTS

Table III gives the classification accuracies for the wavelet feature sets and different accelerometer placements for the

TABLE IV
CLASSIFICATION ACCURACIES (%) FOR THE THREE-ACTIVITY CLASSIFICATION
PROBLEM (LEVEL WALKING, STAIR ASCENT, AND STAIR DESCENT) WITH THE
TIME AND FREQUENCY FEATURES (TABLE II)

	Mean and SD	Multiple Statistics	Mean ac and mean dc	Principal Frequency	Energy	Bao and Intille [11]	FFT Magnitude
Waist	69	70	66	77	67	69	77
Thigh	76	80	90	56	63	74	91
Ankle	89	85	93	74	68	88	97
Waist and Thigh	88	85	93	57	85	86	90
Waist and Ankle	95	92	94	83	83	95	98
Thigh and Ankle	95	93	98	73	94	94	98
All	95	93	97	80	95	96	98

Accuracies have been reported for each of the different accelerometer combinations.

TABLE V
CLASSIFICATION ACCURACIES (%) FOR THE EIGHT-ACTIVITY CLASSIFICATION
PROBLEM WITH THE WAVELET FEATURES (SEE TABLE I)

	Tamura et al. [23]	Nyan et al. [24]	Sekine et al. [22]	Wang et al. [14]	Sekine et al. [21]	Sum of squares	Absolute Sum
Waist	53	47	52	60	33	64	70
Thigh	57	51	42	60	30	69	74
Ankle	63	51	33	60	22	73	77
Waist and Thigh	67	63	58	77	46	79	86
Waist and Ankle	69	63	62	78	47	77	87
Thigh and Ankle	77	69	57	82	43	85	88
All	79	73	66	89	55	88	92

Accuracies have been reported for each of the different accelerometer combinations.

TABLE VI
CLASSIFICATION ACCURACIES (%) FOR THE EIGHT-ACTIVITY CLASSIFICATION
PROBLEM WITH THE TIME AND FREQUENCY FEATURES (SEE TABLE II)

	Mean and SD	Multiple Statistics	Mean ac and mean dc	Principal Frequency	Energy	Bao and Intille [11]	FFT Magnitude
Waist	62	65	65	58	58	61	65
Thigh	72	74	84	48	54	68	88
Ankle	83	81	89	60	59	80	92
Waist and Thigh	84	82	88	49	76	78	88
Waist and Ankle	90	89	92	64	75	87	94
Thigh and Ankle	90	90	94	58	83	86	95
All	91	92	95	61	86	90	96

Accuracies have been reported for each of the different accelerometer combinations.

three-activity classification problem. Table IV illustrates the same information but for the time-frequency features. Overall, for the three-activity problem, the highest classification accuracy for a single sensor (97% ± 3%) was obtained using FFT components derived from the ankle-mounted unit. This distribution of accuracies was significantly higher than those obtained from all other feature sets derived from a single sensor ($p < 0.01$). In general, for the wavelet feature sets, the highest performance was obtained using the sum of the absolute values for each sensor configuration (Table III). However, the performance of this feature set was, in some cases, not significantly better than the wavelet feature set proposed by Wang *et al.* [14].

In order to establish whether, in general, the time-frequency features outperformed the wavelet features, a number of statistical tests were performed. First, the performance of the best set of time-frequency features was compared with the best set of wavelet features for each sensor configuration. With the exception of the waist-mounted sensor, the time-frequency feature sets significantly outperformed the wavelet feature sets ($p < 0.01$) in every case (Tables III and IV). Further testing was then carried out by comparing the second-best performing time-frequency feature set with the second-best performing wavelet feature set for every sensor configuration. These tests also showed the time-frequency feature sets to significantly outperform the wavelet feature sets in every case ($p < 0.01$).

The results of the eight-activity classification problem displayed similar trends to the three-activity problem for both the wavelet features (Table V) and the time-frequency features (Table VI). Again, the highest wavelet classification accuracies, for each of the sensor configurations, were obtained using the sum of the absolute values (Table V). However, the performance of this feature set was, in some cases, not significantly better than the sum of the squares feature set. For the time/frequency features, maximal classification accuracy for a single sensor (92% ± 7%) was again obtained when the individual FFT components were derived from the ankle-mounted unit (Table VI). However, this distribution of accuracies was not significantly different to those obtained using FFT coefficients derived from the thigh-mounted sensor ($p = 0.16$). Again, to de-

termine whether differences in accuracy existed between the two types of features, comparisons were made between the best and second-best performing time/frequency and wavelet features. These comparisons showed that, with the exception of features derived from a waist-mounted sensor, the time/frequency features significantly outperformed the wavelet features ($p < 0.01$) (Tables V and VI).

The classification accuracies reported in Tables V and VI represent an average across all of the eight different activities. Although these data would suggest that FFT features give better classification accuracy than wavelet features, it is not clear whether this result is true across all activities. To investigate this further, sensitivity and specificity were calculated separately for each of the eight activities for the ankle-mounted sensor with the best-performing wavelet feature set (sum of the absolute values) and the best performing time/frequency feature set (magnitude of the FFT components). This comparison (Table VII) shows that, for each of the different activities, the FFT feature set outperforms the wavelet feature set.

IV. DISCUSSION

This study was designed to compare the discriminative ability of wavelet features with time/frequency features for two activity classification problems: a simple three-activity problem and

TABLE VII
SENSITIVITY AND SPECIFICITY FOR EACH ACTIVITY FOR THE BEST
PERFORMING TIME/FREQUENCY AND WAVELET FEATURE SETS

	TIME/FREQUENCY FEATURES: FFT MAGNITUDE Sensitivity - Specificity	WAVELET FEATURES: ABSOLUTE SUM Sensitivity - Specificity
Walking	99 – 99	85 – 92
Upstairs	94 -99	67 – 95
Downstairs	96 - 98	88 – 96
Jog	91 -98	78 – 97
Run	91 – 99	87 – 98
Hop(L leg)	83 – 99	77 – 99
Hop(R leg)	74 - 98	69 – 98
Jump	64 – 99	43 - 97

an eight-activity problem. In addition, classification accuracies were compared for three individual lower limb placements, the waist, thigh, and ankle, as well as some of their combinations. In general, similar levels of accuracy were found when data from a waist-mounted sensor were used to obtain either time/frequency or wavelet features. However, for both the ankle- and thigh-mounted sensors, time/frequency features significantly outperformed the wavelet features. For both classification problems, the optimal accelerometer placement for a single sensor was shown to be on the ankle.

Five previous studies were identified, which had used wavelet features to discriminate between level walking, stair ascent, and stair descent. Of these five studies, only Nyan *et al.* [24] and Wang *et al.* [14] reported intersubject classification accuracies. The remaining three studies simply demonstrated significant differences between wavelet parameters corresponding to each of the three activities [21]–[23]. Nyan *et al.* [24] used a simple threshold-based classification scheme that required the manual selection of arbitrary thresholds for both of their features. With this approach, they obtained accuracies ranging from 97% to 99%. The use of thresholds determined by the experimenter reduces the system's ability for fully automatic classification. In our paper, we aimed to build an automated system that can be trained by a set of supervised subjects and activity scenarios. This system can then be applied to new subjects, instrumented with the same sensors, without any further supervision. In their study, Nyan *et al.* [24] collected data using two shoulder-mounted accelerometers so that their results are not directly comparable to those in this study.

Wang *et al.* [14] studied level walking, stair ascent/descent, and walking up/down a slope using data collected from a waist-mounted accelerometer. Using a multilayer perceptron neural network classifier, they obtained classification accuracies ranging from 89% to 92% for these five activities. However, in their study, an individual normalization scheme was used in which the features were divided by those obtained from a 5-s flat walking session. When unadjusted features were used for classification, similar levels of accuracy to those found in this study were obtained.

In order to minimize computational power requirements, activity classification algorithms typically work with relatively short windows of sensor data. As these windows typically cor-

respond to a single activity, the frequency content of the signal varies little with time. Wavelet analysis allows for the analysis of nonstationary signals. However, it is not clear whether parameters derived from wavelet coefficients represent a more effective means of characterizing short windows of data than standard frequency-domain techniques. In this study, data were collected from 20 healthy subjects. Analysis of these data showed that features derived from an FFT analysis outperformed those derived from wavelet coefficients. This may reflect the suitability of standard frequency-domain techniques for characterizing the short-duration stationary signals, which were characteristic of our subject group.

This study found surprisingly good levels of classification accuracy when using simple time-domain features. A number of other studies have reported high levels of classification accuracy using time-domain features. For example, Pirttikangas [34] used means and SDs from a number of body-worn accelerometers to accurately classify (>90%) a wide range of activities. Similarly, Fahrenberg *et al.* [16] used mean dc and mean rectified ac in a hierarchical classification to differentiate between a range of static postures and movements. For the current study, this set normally outperformed the other time-domain features and often gave comparable accuracy to the FFT component feature set.

The highest classification accuracy for a single sensor was obtained for the FFT component feature set and the ankle-mounted sensor. This feature set consistently outperformed both the energy feature and the larger set proposed by Bao and Intille [11]. As they studied a larger range of activities than those of this study, direct comparison of classification accuracies is not possible. However, their reported maximum classification accuracy of 84% using data from five sensors is similar to the maximum accuracy (90%) achieved in our study for the eight-activity problem. Huynh and Schiele [37] also compared the discriminative ability of individual FFT components with simple time-domain features, spectral energy, and spectral entropy for a range of activities including walking, jogging, and hopping. In agreement with this study, they found the FFT component to have higher discriminative ability than the other features. However, they were unable to identify a single component that performed best for each activity. Although, in this study, the first five components were used as input to the classifier, it is possible to use a larger or smaller number of components. Fig. 3 illustrates how the classification accuracy changes as the number of components varies. It can be seen that using the first six components produces maximal accuracy for both the three-activity and eight-activity problems. Although, for the three-activity problem, an almost perfect result is achieved, with the eight-activity problem a maximum accuracy of only 94% is possible. Inspection of the corresponding confusion matrix (Table VIII) showed that jumping was often confused with a number of other activities. When this activity was excluded, the accuracy increased to 97%.

There are a number of limitations of this study. First, subjects performed each of the separate activities while being videoed by the experimenter. Under these conditions, it is possible that individuals may subconsciously modify their habitual movement patterns. However, some method is required for annotating the sensor data. The video method, used in this study, was selected
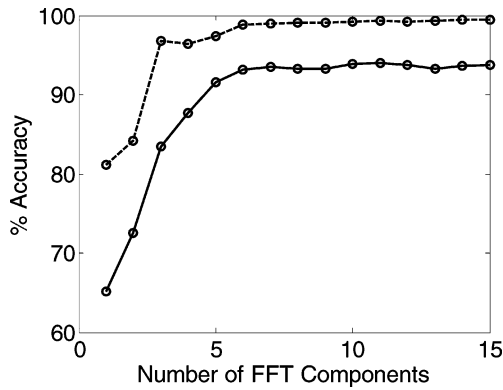
Fig. 3. Plot to show the accuracy of activity recognition as the number of FFT coefficients is increased. The dashed line shows the relationship for the three-activity problem (level walking, stair ascent, and stair descent) and the solid line shows the relationship for the eight-activity problem.

TABLE VIII
CONFUSION MATRIX SHOWING CLASSIFICATION RESULTS FOR THE EIGHT-ACTIVITY PROBLEM USING FEATURES DEFINED AS THE MAGNITUDES OF THE FIRST TEN FFT COMPONENTS OBTAINED FROM THE ANKLE-MOUNTED SENSOR

	Walking	Upstairs	Downstairs	Jog	Run	Hop (L leg)	Hop (R leg)	Jump
Walking	377	1	0	0	0	0	0	0
Upstairs	2	346	2	0	0	0	0	0
Downstairs	2	2	275	0	0	0	0	0
Jog	0	0	0	219	11	0	0	1
Run	0	0	0	11	138	0	0	0
Hop (L leg)	0	1	0	0	0	72	12	2
Hop (R leg)	0	0	0	0	0	14	73	8
Jump	0	2	13	4	1	1	14	59

as it was believed to be more accurate than self-observation by the subject. Another limitation is that only relatively young, healthy subjects were included in the study. Clearly, it is not possible to generalize our findings, that frequency domain features perform better than wavelet features, to other subject groups, such as the elderly or patients with neurological impairment. For such individuals, jerkiness of movement may lead to isolated frequency transients that maybe better characterized using wavelet features. Further research is thus needed to determine the most appropriate features for activity classification for different patient groups.

For this study, a single classifier (kNN) was used to evaluate the discriminatory ability of the different feature sets. Although it is possible to use other methods to identify optimal features, this method was chosen for its simplicity, flexibility, and popularity. In general, different classifiers can have different subsets of optimal features and a larger evaluation study would be needed to perform comparisons between different classifiers.

In conclusion, the findings of this study suggest that future activity monitoring systems, aimed at healthy individuals, should consider using an FFT feature set derived from an ankle-mounted sensor. Further research is required to determine the most appropriate feature sets for other subjects groups, such as the elderly or neurologically impaired.

REFERENCES

[1] K. Aminian and B. Najafi, "Capturing human motion using body-fixed sensors: Outdoor measurement and clinical applications," *Comput. Anim. Virtual Worlds*, vol. 15, pp. 79–94, 2004.
[2] W. Zjlstra and K. Aminian, "Mobility assessment in older people: New possibilities and challenges," *Eur. J. Aging*, vol. 4, pp. 3–12, 2007.
[3] *World Health Organization Global Strategy on Diet, Physical Activity and Health*. Geneva, Switzerland: WHO, 2004.
[4] R. A. Washburn and H. J. Montoye, "The assessment of physical activity by questionnaire," *Amer. J. Epidemiol.*, vol. 123, pp. 563–576, 1986.
[5] B. E. Ainsworth, D. R. Jacobs, Jr., and A. S. Leon, "Validity and reliability of self-reported physical activity status: The lipid research clinics questionnaire," *Med. Sci. Sports Exerc.*, vol. 25, pp. 92–98, 1993.
[6] C. J. Riddoch, C. Mattocks, K. Deere, J. Saunders, J. Kirkby, K. Tilling, S. D. Leary, S. N. Blair, and A. R. Ness, "Objective measurement of levels and patterns of physical activity," *Arch. Dis. Child.*, vol. 92, pp. 963–969, 2007.
[7] R. R. Pate, K. A. Pfeiffer, S. G. Trost, P. Ziegler, and M. Dowda, "Physical activity among children attending preschools," *Pediatrics*, vol. 114, pp. 1258–1263, 2004.
[8] B. Najafi, K. Aminian, F. Loew, Y. Blanc, and P. A. Robert, "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.*, vol. 49, no. 8, pp. 843–851, Aug. 2002.
[9] N. Streitz and P. Nixon, "The disappearing computer," *Commun. ACM*, vol. 48, pp. 32–35, 2005.
[10] J. Coutaz, J. L. Crowley, S. Dobson, and D. Garlan, "Context is key," *Commun. ACM*, vol. 48, pp. 49–53, 2005.
[11] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," *Pers Comput.*, Lecture Notes in Computer Science, vol. 3001, pp. 1–17, 2004.
[12] U. Maurer, A. Rowe, A. Smailagic, and D. Siewiorek, "Location and activity recognition using eWatch: A wearable sensor platform," *Ambient Intell. Everday Life,* Lecture Notes in Computer Science, vol. 3864, pp. 86–102, 2006.
[13] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, "Activity classification using realistic data from wearable sensors," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 119–128, Jan. 2006.
[14] N. Wang, E. Ambikairajah, N. H. Lovell, and B. G. Celler, "Accelerometry based classification of walking patterns using time-frequency analysis," in *Proc. 29th Annu. Conf. IEEE Eng. Med. Biol. Soc.*, Lyon, France, 2007, pp. 4899–4902.
[15] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, K. Meijer, R. H. Crompton, and D. Howard, "Activity identification using body-mounted sensors—a review of classification techniques," *Physiol. Meas.*, 2008, to be published.
[16] J. Fahrenberg, F. Foerster, M. Smeja, and W. Muller, "Assessment of posture and motion by multichannel piezoresistive accelerometer recordings," *Psychophysiology*, vol. 34, pp. 607–612, 1997.
[17] J. Fahrenberg, W. Muller, F. Foerster, and M. Smeja, "A multi-channel investigation of physical activity," *Psychophysiology*, vol. 10, pp. 209–217, 1996.
[18] F. R. Allen, E. Ambikairajah, N. H. Lovell, and B. G. Celler, "Classification of a known sequence of motions and postures from accelerometry data using adapted Gaussian mixture models," *Physiol. Meas.*, vol. 27, pp. 935–951, 2006.
[19] T. Prill and J. Fahrenberg, "Simultaneous assessment of posture and limb movements (e.g., periodic leg movements) with calibrated multiple accelerometry," *Physiol. Meas.*, vol. 27, pp. N47–N53, 2006.
[20] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, and D. Howard, "A comparison of different feature generation methods in activity classification," in *Proc. Int. Conf. Ambul. Monit. Phys. Act. Movement*, Rotterdam, The Netherlands, 2008, p. 87.
[21] M. Sekine, T. Tamura, M. Akay, T. Fujimoto, T. Togawa, and Y. Fukui, "Discrimination of walking patterns using wavelet-based fractal analysis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 10, no. 3, pp. 188–196, Sep. 2002.
[22] M. Sekine, T. Tamura, T. Togawa, and Y. Fukui, "Classification of waist-acceleration signals in a continuous walking record," *Med. Eng. Phys.*, vol. 22, pp. 285–291, 2000.

[23] T. Tamura, M. Sekine, M. Ogawa, T. Togawa, and Y. Fukui, "Classification of acceleration waveforms during walking by wavelet transform," *Methods Inf. Med.*, vol. 36, pp. 356–359, 1997.

[24] M. N. Nyan, F. E. Tay, K. H. Seah, and Y. Y. Sitoh, "Classification of gait patterns in the time-frequency domain," *J. Biomech.*, vol. 39, pp. 2647–2656, 2006.

[25] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 617–643, Mar. 1992.

[26] C. V. C. Bouten, K. T. M. Koekkoek, M. Verduin, R. Kodde, and J. D. Janssen, "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 3, pp. 136–147, Mar. 1997.

[27] E. A. Heinz, K. S. Kunze, S. Sulistyo, H. Junker, P. Lukowicz, and G. Troster, "Experimental evaluation of variations in primary features used for accelerometric context recognition," *Ambient Intell.,* Lecture Notes in Computer Science, vol. 2875, pp. 252–263, 2003.

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

[29] H. B. Bussman, P. J. Reuvekamp, P. H. Veltink, W. L. Martens, and H. J. Stam, "Validity and reliability of measurements obtained with an "activity monitor" in people with and without a transtibial amputation," *Phys. Ther.*, vol. 78, pp. 989–998, 1998.

[30] K. Aminian, P. Robert, E. E. Buchser, B. Rutschmann, D. Hayoz, and M. Depairon, "Physical activity monitoring based on accelerometry: Validation and comparison with video observation," *Med. Biol. Eng. Comput.*, vol. 37, pp. 304–308, 1999.

[31] J. B. J. Bussmann, Y. M. van de Laar, M. P. Neeleman, and H. J. Stam, "Ambulatory accelerometry to quantify motor behaviour in patients after failed back surgery: A validation study," *Pain*, vol. 74, pp. 153–161, 1998.

[32] B. Najafi, K. Aminian, A. Paraschiv-Ionescu, F. Loew, C. J. Bula, and P. Robert, "Ambulatory system for human motion analysis using a kinematic sensor: Monitoring of daily physical activity in the elderly," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp. 711–723, Jun. 2003.

[33] M. Ermes, J. Parkka, J. Mantyjarvi, and I. Korhonen, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 1, pp. 20–26, Jan. 2008.

[34] P. Pirttikangas, K. Fujinami, and T. Nakajima, "Feature selection and activity recognition from wearable sensors," *Ubiquitous Comput. Syst.,* Lecture Notes in Computer Science, vol. 4239, pp. 516–527, 2006.

[35] F. Foerster and J. Fahrenberg, "Motion pattern and posture: Correctly assessed by calibrated accelerometers," *Behav. Res. Methods Instrum. Comput.*, vol. 32, pp. 450–457, 2000.

[36] K. Zhang, P. Werner, M. Sun, F. X. Pi-Sunyer, and C. N. Boozer, "Measurement of human daily physical activity," *Obes. Res.*, vol. 11, pp. 33–40, 2003.

[37] T. Huynh and B. Schiele, "Analyzing features for activity recognition," in *Proc. Conf. Smart Objects Ambient Intell.: Innov. Context-Aware Serv.: Usages Technol.*, Grenoble, France, 2005, pp. 159–164.

[38] K. van Laerhoven and O. Cakmakci, "What shall we teach our pants?," in *Proc. 4th Int. Symp. Wearable Comput.*, 2000, pp. 77–83.

[39] S. G. Mallat, "A thoery for multi-resolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[40] T. Chau, "A review of analytical techniques for gait data. Part 2: Neural network and wavelet methods," *Gait Posture*, vol. 13, pp. 102–120, 2001.

[41] B. B. Mandelbrot, *The Fractal Geometry of Nature*. New York: Freeman, 1983.

[42] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 156–167, Jan. 2006.

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

[44] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001.

Stephen J. Preece received the B.Sc. degree in physics in 1996 and the Ph.D. degree in applied mathematics in 2001, both from the University of Birmingham, Birmingham, U.K.

He has been engaged in research on a range of biomedical engineering projects. He is currently a Postdoctoral Research Fellow in the field of human performance research at the University of Salford, Salford, U.K. His current research interests include algorithm development in activity monitoring and electromyography methodology for characterizing lower limb muscle function during gait.

John Yannis Goulermas (S'98–M'98) received the B.Sc. (Hons, First Class) degree in computation in 1994 from the University of Manchester Institute of Science and Technology (UMIST), Manchester, U.K., and the M.Sc. degree by research in 1996 and the Ph.D. degree in 2000, both from the Department of Electrical Engineering, where he worked in the field of on machine vision.

He has been a Scientific Developer in industry and a Postdoctoral Researcher in virtual reality and biomechanics. He joined the Department of Electrical Engineering and Electronics of the University of Liverpool, Liverpool, U.K., in 2005. His current research interests include pattern recognition, optimization, and image analysis, with applications in biomedical engineering and industrial monitoring.

Laurence P. J. Kenney received the B.Sc. degree in mechanical engineering in 1986 and the Ph.D. degree in engineering design in 1993, both from the University of Salford, Salford, U.K.

From 1998–2000, he was a Postdoctoral Researcher with Roessingh Research and Development in The Netherlands. He rejoined the University of Salford in 2000, where he was engaged in the establishment of the Centre for Rehabilitation and Human Performance Research and where he currently is a Senior Research Fellow. His current research interests include the development and evaluation of functional electrical stimulation and prosthetic devices.

David Howard received the B.E. degree in mechanical and production engineering from Brunel University, Uxbridge, U.K., in 1980, and the Ph.D. degree in mechanical engineering from the University of Bath, Bath, U.K., in 1987.

He is currently with the University of Salford, Salford, U.K., in the Centre for Rehabilitation and Human Performance Research. His current research interests include applying gait simulation to design-assistive devices, controlling functional electrical stimulation, activity monitoring using body-mounted sensors, and real-world monitoring of patients using assistive devices.