Activity recognition from body worn accelerometers - toward real-time event detection

Ali K. Mohamed Ali[1,2], **Rachel King**[1], **Balazs Janko**[1], **Emma Sack**[3,4], **Ann Ashburn**[3], **Malcolm Burnett**[3], **Ian Craddock**[5] **and William Harwin**[1]

Abstract Inadequate physical activity is leading risk factor in public health and inactive people are more vulnerable to have chronic diseases. In addition levels of physical activity may be an indicator of health problems in elderly individuals, a particular problem in many societies where there is a growing ratio of elderly people. Identifying levels of physical activity may have a significant effect on reducing healthcare costs in the future. As a consequence, finding approaches for measuring the individuals' activities is a persistent need, in order to provide a view about their quality life and to observe their current health status. This may be best achieved by using low-cost wearable technology such as accelerometer based inertial sensors. In this work, the angle between the posture of the individual's trunk and the gravity trajectory was used together with the velocity to extract significant features from accelerometer data. Two datasets were used, the first had been collected for SPHERE project from an individuals with Parkinson's disease in their home and the second is a public domain benchmark data set. Decision Trees and Naïve Bayes classifiers have been used on both datasets. The classification results of a small set of activities of single individual from first dataset show that Naïve Bayes have high overall accuracy rate of 85%. The second dataset of daily household activities is used to provide a comparison with one state-of-the-art approach in the literature. The result shows that Decision Trees with the proposed features outperform the literature approach by having overall accuracy rate exceeded 91%.

[1] Biomedical Engineering group, School of Biological Sciences, University of Reading, Reading RG1 5AQ, UK

[2] Department of Computer Systems, Babel Technical Institute, Al-Furat Al-Awsat Technical University, Babylon, Iraq

[3] Faculty of Health Sciences, University of Southampton, Southampton SO17 1BJ, UK

[4] NIHR CLAHRC Wessex, University of Southampton, Southampton SO17 1BJ, UK

[5] Faculty of Engineering, University of Bristol, Bristol, BS8 1UB, UK

1 Introduction

According to the World Health Organization, inadequate physical activity is a leading risk factor in public health resulting in more than three million deaths each year [1].In addition, a considerable number of studies showed that physically inactive people are more susceptible to have chronic illnesses and have poorer levels of health-related fitness when compared to physically active people [2]–[5]. On the other hand, the United Nations reports for world population show that there is an increasing rate of elderly people around the world [6]. For example, in 2015, the reports estimates that people aged 60 or over comprise a 12% of world population and this sector of the population has grown steadily by 3.26% per year[6]. These factors cause a substantial impact on healthcare costs [7].As a consequence, there is a persistent need for methods to assess the individuals' activities of daily life (ADL) in order to present a view about their quality life and to monitor their current health status. A sufficient and effective intervention might then be considered if early detection of specific health threat factors can be made [8]. New intelligent technologies might enable transparent detection and evaluation of the activities of daily life [9]. Recently the stage for a noticeable modify in health monitoring has been set by the latest advances in wearable technology [8], [9]. This technology can use as a way of assessing health and could enable elderly people to live independently and in safety at home [8].

One of the sensors that widely investigated in most new technology research for monitoring of human movement and for assessing healthcare is accelerometers [8], [10]. Unfortunately although most of the methods in literature that deal with human activity recognition using the accelerometer, they have differences in many aspects [10]. These aspects include the number of measured axes in each accelerometer sensor, the number and placement of individual sensors, the sampling rate, the number of participants, the type and number of activities, computed features, and the type (sliding or non-sliding), the design and the size of the filter window[10]. These differences, along with the lack of good benchmark dataset, make it difficult to compare the new approaches to existing methods.

The purpose of this paper is to propose an approach for human activity recognition, apply it to a dataset collected from a small cohort of individuals with Parkinson's disease and to compare the proposed approach with a dataset collected from unimpaired individuals [10].

2 Background

Interest in collecting and analysing data from individuals in their natural surroundings to assist with management and diagnosis of health and healthcare problems is growing [10]. To investigate this concept further the SPHERE research project (http://irc-sphere.ac.uk) aims to deploy, collect and analyse data from a range of sensors in a 100 residential homes in the Bristol area and consider the data analysis and data mining techniques that can be employed to enable this data to be used by the individual, their carers, and researchers to monitor healthcare related problems [9], [11]. The project is nonspecific but healthcare issues are likely to range from COPD (chronic obstructive pulmonary disease), Parkinson's disease, stroke, frailty, depression, sleep disorders, and obesity [9].

Data collection will include body-worn sensors that can operate for up to a month without recharging while transmitting key information to the house infrastructure for data storage and analysis [11]. A key hypothesis is that changes in the data characteristics over time will be indicative of health related concerns [9], [11]. Thus short-term change may be an indication of a medical emergency such as a fall [9], [11]. Long-term changes in the data would be indicative of an ongoing chronic medical problem that may require a change in management, for example, a change in the gait pattern of an individual with Parkinson's disease [9], [11].

This paper aims to recognise different postures and ambulation for individual using the angle between the gravity vector and the posture of the individual's body and using the velocity of data. The successful classification models could help in detection of activities of daily life (ADL) in real-time.

2.1 Related work

There is a considerable body of work in the literature on human activity recognition using inertial sensors. Most publications are based on their own dataset which is collected using a variety of different methods. Some of this literature are summarised in Table 1 which shows for each method, the conditions of collecting data, the number of participant subjects, the number of performed activities, the features and the algorithms used to classify the collected data. Because of the different approaches in the literature, it is difficult to make direct comparisons between different studies. However, the recent availability of movement datasets allows us the opportunity for limited comparison of our method across a wider set of experimental data.

Key questions are the number of sensors used and type of classification algorithm. The study was done by [12] showed that using one sensor placed on the human trunk can be efficient in detecting various types of fall event, with results showing up to 96.7% accuracy. However, using more than one sensor improve accuracy in recognising a number of activities of daily life, as shown by [13]. In addition, [10] explained that every classification algorithm can obtain substantial accuracy for specific activities and consequently, for one method, a number of classifiers can be exploited for different activity groups' recognition.

Table1. Summary of previous work for activity recognition.

Authors	No. Of Sensors	Sensor Placement	Sampling Rate (Hz)/ No. Of Subject/ No. Of Activities	Window Size	Extracted Features	Used Classifier
[10]	4 triaxial accelero meter & gyroscope	right wrist, chest, right hip and left ankle.	204.8/19/13	5 s with 50% overlap	minimum amplitude, maximum amplitude, mean amplitude, variance of amplitude, spectral centroid, bandwidth, energy and gravitational component of the acceleration signal.	Hierarchial classification using ADA, kNN and SVM
[14]	5 biaxial accelero meter	right hip, dominant wrist, non-dominant upper arm, dominant ankle and non-dominant thigh	76.25/20/20	6.7 s with 50% overlap	mean, energy, frequency domain entropy, correlation of the acceleration signals	Decision Tree
[15]	3 triaxial accelero meter	waist, thigh and ankle.	64/20/8	2 s with 50% overlap	magnitude of first five components of FFT analysis	kNN
[12]	1 triaxial accelero meter, gyroscope & tilt	chest	-/3/5	50 & 100 s	raw data and mean	Suggested Algorithm for real-time fall detection
[16]	1 iPhone triaxial accelerometer &Nike +iPod	thigh & foot	200/8/4	1 s	magnitude, mean, standard deviation, minimum value, maximum value, min minus max, max minus min and energy	Naive Bayseian Network

	Sport kit					
[17]	1 triaxial accelero meter	waist	45/6/12	1 s	median filtering and low pass filtering	Suggested Algorithm

3 Data source

In preparation for deployment to residential houses in the Bristol area the Sphere project collected data from a small cohort of individuals with Parkinson's disease [18]. Data was collected in the individual's home over a period of approximately 1-2 hours during which time the individual was asked to continue his or her daily activities as normal [18]. During this data collection period the individuals demonstrated typical activities [18]. A suite of five accelerometers with sampling rate set to 50 Hz was used for data collection, and ground truth captured by video recording the activities [18]. The five accelerometers were worn on the waist at the lower back, left wrist, right wrist, left ankle and right ankle [18]. For the work presented in this paper, only the waist worn accelerometer data was used. The accelerometer data was labeled using the ELAN 4.9.3 package to categories eight key activities [18], [19]. These were sitting in a chair, standing, sit to stand, stand to sit, walking, stair-up, stair-down and turning [18]. For the comparison purpose, the proposed method has also been applied on a data called "Benchmark dataset" that collected by [10] which available on (http://www.activitynet.org). This dataset collected from 19 subjects for 13 activities. These activities are sitting, lying, standing, washing dishes, vacuuming, sweeping, walking outside, ascending stairs, descending stairs, treadmill running (8.3 km/h), bicycling (50 watt), bicycling (100 watt) and rope jumping [10].

4 Data Analysis

The data analysis recognizes that the principal component from accelerometer data is the omnipresent 1g field. Thus features relating to angle and magnitude with respect to this field dominate, and movement activities are effectively imposed on this data. Data from waist sensor of a single individual with Parkinson's disease was used for preliminary analysis. The data was initially processed by considering the principal movements of the individual to be in the 'sagittal plane' (i.e. forwards and backwards). The angle between the gravity vector and the pose of the individual's trunk was computed using the 'atan2' function. The axes that used as arguments for 'atan2' function are the vertical axis and the axis that point to forward or backward direction. For feature extraction, a non-sliding of 24, 48, 72, 96, 144, 192 and 240 sample window sizes with 50% overlap and same sizes sliding windows was used to determine the best window size and window type. Three features were extracted for classification which are mean, standard deviation and energy. The energy is computed by adding together the sum of the squared values for each axis, divided the addition result by three then divided by the number of samples. Classification was performed using DTs (Decision trees) and NB (Naïve Bayes) separately, and validated using 10 fold Cross Validation method.

5 Results

The obtained results from both of the classification methods show that the 48 sample sliding window has better classification accuracy than the other windows, with 79% accuracy for DTs and 85% for NB. Table 1 shows that NB outperform DTs in recognition of five activities which are the ambulation acts. While the best results for stationary acts gained by TDs.

Table 2. Classification accuracy (in per cent) for DTs and NB using Parkinson's data set.

Activities	DTs	NB
Sitting	**99.63**	96.08
Standing	**98.35**	97.69
Sit to Stand	68.75	**75.00**
Stand to Sit	66.67	66.67
Walking	90.91	**95.04**
Stairs-up	87.50	**100.00**
Stairs-down	66.67	**77.78**
Turning	50.00	**75.00**
Overall Accuracy	78.56	**85.41**

As a result of the success of the 'atan2' function, addition analysis was considered to allow increase the veracity of the data. This was done by considering the projection of the gravity vector onto a sphere radius (g). Thus, static and slow movements will be characterised by points on or near the surface of this sphere, and more dynamic movements will be characterised by signature trajectories above or below the surface of the sphere. Figure (1. a, b & c) shows the projection of three of the labeled activities sitting, standing, and stand-to-sit.

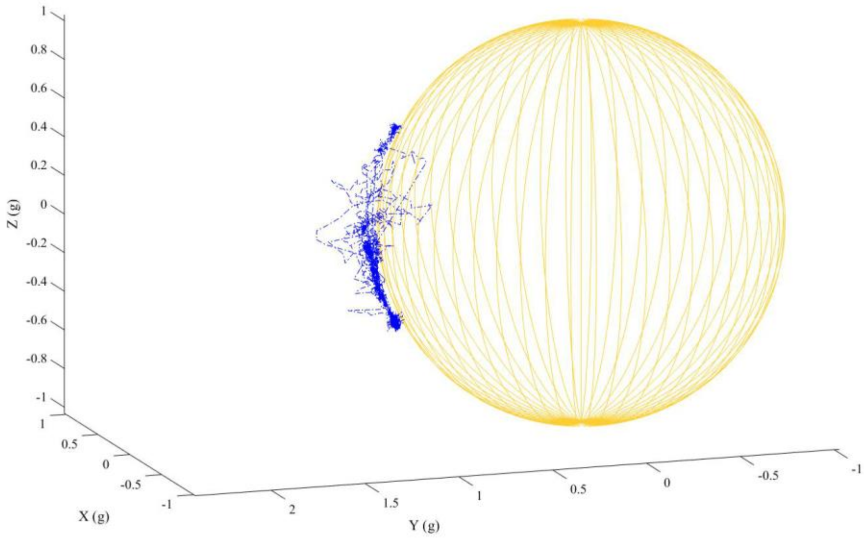

Figure 1.a projection of the waist accelerometer sensor data for sitting activity onto the sphere.

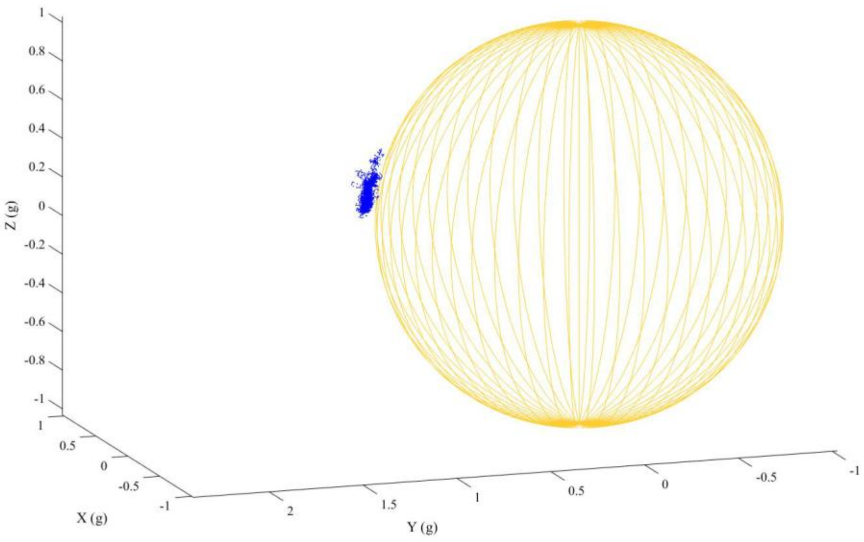

Figure 1.b projection of the waist accelerometer sensor data for standing activity onto the sphere.

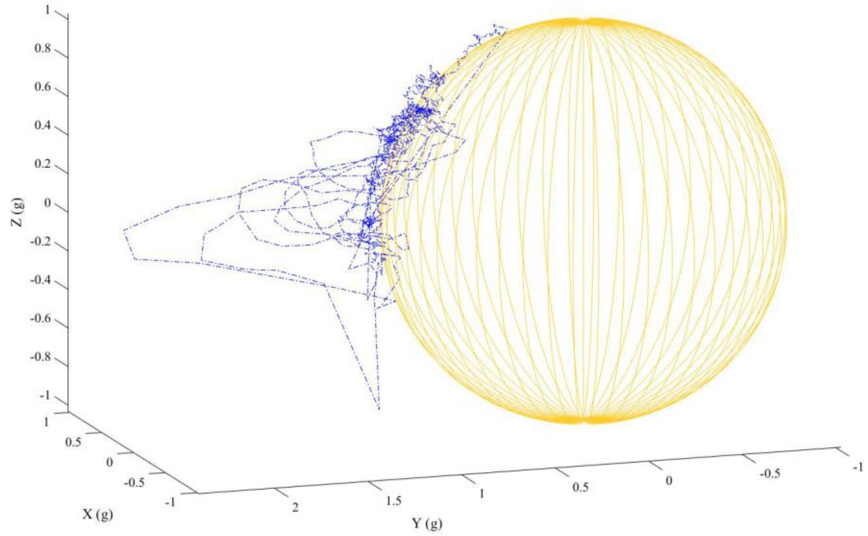

Figure 1.c projection of the waist accelerometer sensor data for stand-to-sit activity onto the sphere.

According to this, the use of data velocity will be significantly useful in increasing the accuracy of classification algorithm. This can be achieved by using the dot product operation. To examine the features gained using both 'atan2' function and dot product operation and for comparison purpose, a benchmark dataset collected by [10] has been used.

The benchmark dataset was collected using four sensors with 204.8 Hz accelerometer and gyroscope from 19 subjects achieved 13 activities [10]. These sensors were placed on right wrist, right hip, left ankle and chest. A sliding window of 5 second size has been used in [10] approach. For each sliding window, the total number of features extracted and used by [10] is 152 features for dynamic activities and 12 features for static activities. All of the four sensors have used in our work. Both of 'atan2' function and dot product operation have applied on the data of the four accelerometer sensors. Because of using sensors placed on limbs and performing activities which have moving in different directions, all accelerometer axes are used with 'atan2' function, which result in three values for every accelerometer. For feature extraction, a sliding windows of size 204, 408, 612, 816, 1020 and 1224 sample, which nearly correspond to 1, 2, 3, 4, 5 & 6 seconds respectively, with 50% overlap to find the best window size. For classification, three features where extracted for each accelerometer which are mean, standard deviation and energy; and mean for every gyroscope axis and energy for each gyroscope. The total number of features used in this work for all the four sensors is 52 features. Classification was performed using DTs (Decision trees) and NB (Naïve Bayes) separately, and validated using 10 fold Cross Validation method.

The results obtained from both DTs and NB algorithms with the results of [10] approach are shown in Table 3. The results show that for DTs the 612 sample (nearly 3s) and for NB the 1020 sample (nearly 5s) sliding windows have better classification accuracy than the other windows, with overall accuracy 91.05% for DTs and 79.26% for NB. The results shown in this table and in Table 2 will be discussed in the discussion section.

Table 3. Classification accuracy for 13 activities and overall accuracy for DTs, NB and [10], using the benchmark dataset.

Activities	Proposed DTs	Proposed NB	[10]
Sitting	**95.26**	13.19	88.9
Lying	97.65	95.43	**100.0**
Standing	**95.10**	90.11	89.8
Washing Dishes	96.88	97.02	**98.1**
Vacuuming	77.72	**90.61**	85.4
Sweeping	81.80	69.35	**89.9**
Walking	97.13	94.38	**99.0**
Ascending Stairs	87.59	93.75	**95.5**
Descending Stairs	86.30	88.09	**95.2**
Treadmill Running	98.62	98.03	**100.0**
Bicycling on Ergometer (50 W)	**86.51**	7.35	69.1
Bicycling on Ergometer (100 W)	87.36	**95.37**	53.5
Rope Jumping	95.69	97.71	**100.0**
Overall Accuracy	**91.05**	79.26	89.6

6 Discussion

In this work, two classification algorithms (DTs and NB) have been applied separately to examine the usefulness of the proposed method to classify activities for an individual with PD; and to compare their results with one state-of-the-art approach in the literature [10] using a benchmark dataset. The application of DTs and NB on the waist sensor data of single individual with Parkinson's disease have shown the advantage of the NB algorithm over the DTs algorithm, with 85.41% and 78.56% overall accuracy respectively. As outlined in Table 2, from the eight activities in the dataset, DTs had the best classification accuracy when compared to NB for sitting and standing activities. Whereas, NB outperformed DTs in recognition of success for the five dynamic activities. However, performing these two classification algorithms on a benchmark dataset using four sensors (placed on the right wrist, right hip, left ankle and chest) for 19 individuals result in outperforming of DTs algorithm over NB algorithm. As shown in Table 3, the failure of the NB algorithm was in recognition of sitting, sweeping and bicycling ergometer (50) activities, with classification accuracy 13.19%, 69.35% and 7.35% for each of them respectively. The reason for these poor results is that the misclassification particularly of SI (sitting), SW (sweeping), and BC50 (bicycle ergometer at 50) by the NB algorithm as shown in Table 4. The uncertainty of NB appears to be between two stationary acts, [sitting and standing] and between the matched repetitive high-velocity acts, [vacuuming and sweeping, bicycling ergometer (50) and bicycling ergometer (100)].

There was a high classification accuracy of DTs so the confusion matrix has not been shown. It is possible that this classification accuracy may be due to the correlations in features values that are picked up by the rule-based activity recognition of DTs. For instance, the DTs classified sitting and standing as activities having different angles between individual's trunk and gravity vector at the hip and low velocity at the hip, wrist and ankle sensors. It distinguishes bicycling (50) activity from bicycling (100) activity because each one of them involves different levels of velocity (moderate and high) at the ankle sensor. Also, it differentiates between sweeping and vacuuming, even though both activities show high energy in wrist acceleration because the first activity involves different angles in wrist sensor. The weaker performance of NB approach may be due to its inability to adequately model such rules.

Comparing these results with the approach of [10] it can be noted that although the DTs method has an overall classification accuracy that is higher than the results shown in [10] the latter outperform the DT method in specific recognition of eight of the thirteen activities. The better individual recognition results of [10] are probably due to the hierarchical approach used. This approach divided the thirteen activities into groups and use a SVM for the initial group classification and different algorithms to do the subgroup classification. However, for this dataset, this approach had a low classification accuracy for sitting, standing, bicycling (50) and bicycling (100). In addition, this approach used 152 features for each sliding window of dynamic activities. While our proposed method uses just 52 features for each sliding window. The high number of features could result in more computational complexity in real-time systems.

Nevertheless, this method has been applied on two different datasets, one acquired from an individual with Parkinson's disease and the other from 19 young healthy individuals. The results show that the use of a number of sensor with both accelerometer and gyroscope have a substantial impact on the classification accuracy for all activities rather than the use of single accelerometer sensor. An improvement on the proposed method could be achieved by exploiting of number of classification algorithms to gather in this method. The use of the combination of different classification algorithms applied on the proposed features with different window sizes could result in high classification accuracy real-time recognition system for activities of daily live (ADL). The difference between window sizes of the DTs and NB best results from using both the Parkinson's and the benchmark datasets possibly because Parkinson's data collected in a home environment while the benchmark data collected in a lab.

Table 4. Confusion matrix for NB algorithm of the proposed method. Boxed numbers highlight the misclassification of key activities.

	SI	LY	ST	WD	VC	SW	WK	AS	DS	RU	BC 50	BC 100	RJ
Sitting (SI)	**60**	7	6	0	1	2	0	0	0	0	2	0	2
Lying (LY)	45	**439**	2	0	0	0	0	0	0	0	0	0	0
Standing (ST)	$\boxed{316}$	14	**410**	7	6	0	0	0	0	0	0	0	0
Washing-Dishes (WD)	7	0	27	**912**	5	3	0	0	0	0	0	0	0
Vacuuming (VC)	2	0	7	10	**415**	$\boxed{196}$	6	0	0	0	2	0	0
Sweeping(SW)	3	0	3	11	28	**516**	18	0	0	1	5	8	0
Walking (WK)	0	0	0	0	0	1	**1933**	16	3	2	3	1	0
Ascending Stairs (AS)	0	0	0	0	1	10	32	**300**	26	0	4	1	0
Descending Stairs (DS)	0	0	0	0	2	15	57	4	**244**	9	1	0	0
Treadmill Running (RU)	0	0	0	0	0	0	2	0	1	**897**	25	0	0
Bicycling 50 (BC 50)	0	0	0	0	0	0	0	0	0	0	**68**	33	3
Bicycling 100 (BC 100)	0	0	0	0	0	1	0	0	0	0	$\boxed{814}$	**886**	1
Rope Jumping (RJ)	22	0	0	0	0	0	0	0	3	6	1	0	**256**

7 Conclusion

This paper considers the recognition of a small set of activities based on accelerometer data from an individual with Parkinson's disease. A novel feature set is considered and a benchmark dataset of daily household activities is used to provide a comparison. Since accelerometers are primarily in a 1g environment it is relatively easy to compute the angle between a sensor worn on the individual's trunk and the gravity trajectory, and this together with the velocity are significant features for recognising person activities using accelerometry. A decision tree classification method with sliding window of size nearly 3 second was shown to be significantly better than a Naïve Bayes approach. This study also shows that the inclusion of data from a number of sensors placed on different part of subject body, especially wrist and ankle, in classification process would lead to enhancing the results, particularly for ambulation acts. A multi-level classification system that uses more than one classification algorithm with different window sizes, will be exploited in future work to increase recognition accuracy.

Knowledge of the underlying cause of the data should lead to the higher veracity of information transmitted at a lower rate. This is particularly important in this application where data must be transmitted from the individual to a base station through a low energy channel at a low bit rate.

Acknowledgements

This work was performed under the SPHERE IRC funded by the UK Engineering and Physical Sciences Research Council (EPSRC), Grant EP/K031910/1. The first author is sponsored by a studentship from the Al-Furat Al-Awsat Technical University, Iraq.

References

[1] World Health Organization (http://www.who.int/beat-ncds/take-action/en/). "Global Ncd Target, Reduce Physical Inactivity". Accessed on Nov. 2nd 2016.

[2] Physical Activity Guidelines Advisory Committee, "Physical Activity Guidelines Advisory Committee Report," *Washingt. DC US*, vol. 67, no. 2, p. 683, 2008.

[3] S. N. Blair, Y. Cheng, and J. S. Holder, "Is physical activity or physical fitness more important in defining health benefits?," *Med. Sci. Sports Exerc.*, vol. 33, no. 6 Suppl, pp. S379-99–20, 2001.

[4] Y. Oguma, H. D. Sesso, R. S. Paffenbarger, and I. M. Lee, "Physical activity and all cause

mortality in women: a review of the evidence," *Br. J. Sports Med.*, vol. 36, no. 3, pp. 162–172, 2002.

[5] D. E. R. Warburton, C. W. Nicol, and S. S. D. Bredin, "Health benefits of physical activity: the evidence.," *CMAJ*, vol. 174, no. 6, pp. 801–9, 2006.

[6] P. D. United Nations, Department of Economic and Social Affairs, "World Population Prospects: The 2015 Revision, Key Findings and Advance Tables," 2015.

[7] 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, 2002.

[8] R. C. King, E. Villeneuve, R. J. White, R. S. Sherratt, W. Holderbauma, and W. S. Harwin, "A Survey of Data Fusion Techniques and Technologies for Wearable Health Monitoring," *Medical Engineering & Physics* (Accepted).

[9] N. Twomey, T. Diethe, M. Kull, H. Song, M. Camplani, S. Hannuna, X. Fafoutis, N. Zhu, P. Woznowski, P. Flach, and I. Craddock, "The SPHERE Challenge: Activity Recognition with Multimodal Sensor Data," *arXiv preprint arXiv:1603.00797,* March, 2016.

[10] H. Leutheuser, D. Schuldhaus, and B. M. Eskofier, "Hierarchical , Multi-Sensor Based Classification of Daily Life Activities : Comparison with State-of-the-Art Algorithms Using a Benchmark Dataset," *PloS one,* vol. 8, no. 10, 2013.

[11] P. Woznowski, X. Fafoutis, T. Song, S. Hannuna, M. Camplani, L. Tao, A. Paiement, E. Mellios, M. Haghighi, N. Zhu, G. Hilton, D. Damen, T. Burghardt, M. Mirmehdi, R. Piechocki, D. Kaleshi, and I. Craddock, "A multi-modal sensor infrastructure for healthcare in a residential environment," *2015 IEEE Int. Conf. Commun. Work. ICCW 2015*, pp. 271–277, 2015.

[12] J. Y. Hwang, J. M. Kang, Y. W. Jang, and H. C. Kim, "Development of Novel Algorithm and Real-time Monitoring Ambulatory System Using Bluetooth Module for Fall Detection in the Elderly," vol. 1, pp. 2204–2207, 2004.

[13] S. Liu, R. X. Gao, D. John, J. W. Staudenmayer, and P. S. Freedson, "Multisensor Data Fusion for Physical Activity Assessment," *Biomed. Eng. IEEE Trans.*, vol. 59, no. 3, pp. 687–696, 2012.

[14] L. Bao and S. S. Intille, "Activity Recognition from User-Annotated Acceleration Data."

[15] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, and D. Howard, "A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data," *IEEE Trans. Biomed. Eng.*, 2009.

[16] T. S. Saponas, J. Lester, J. Froehlich, J. Fogarty, and J. Landay, "iLearn on the iPhone : Real-Time Human Activity Classification on Commodity Mobile Phones," in *University of Washington CSE Tech Report*, 2008, no. UW-CSE-08-04-02.

[17] D. M. Karantonis, S. Member, M. R. Narayanan, M. Mathie, N. H. Lovell, S. Member, and B. G. Celler, "Implementation of a Real-Time Human Movement Classifier Using a Triaxial Accelerometer for Ambulatory Monitoring," *IEEE transactions on information technology in biomedicine,* vol. 10, no. 1, pp. 156–167, 2006.

[18] E. Stack, R. King, B. Janko, M. Burnett, N. Hammersley, V. Agarwal, S. Hannuna, A. Burrows, and A. Ashburn, "Could In-Home Sensors Surpass Human Observation of People with Parkinson's at High Risk of Falling? An Ethnographic Study," *Biomed Res. Int.*, vol. 2016, 2016.

[19] H. Brugman and A. Russel, "Annotating multi-media/multi-modal resources with ELAN," in *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*, 2004, pp. 2065–2068.