Evaluating Sensor Placement and Modality for Activity Recognition in Active Games

Natalie Jablonsky
School of Information Technology Deakin University, Geelong, Australia
njablons@deakin.edu.au

Sophie McKenzie
School of Information Technology Deakin University, Geelong, Australia
sophie.mckenzie@deakin.edu.au

Shaun Bangay
School of Information Technology Deakin University, Geelong, Australia
shaun.bangay@deakin.edu.au

Tim Wilkin
School of Information Technology Deakin University, Geelong, Australia
tim.wilkin@deakin.edu.au

ABSTRACT

Active games augment physical activity with the immersive elements of computer games running on mobile devices to encourage healthy behaviours in players. Such games rely on automated human activity recognition systems to obtain meaningful input for gameplay. The nature, number and placement of the sensors used to measure game activities affects the quality of activity recognition and the resulting value that this has for game play. This study investigates the recognition performance impact of using parallel sensing strategies from multiple body locations, as well as multiple sensor modalities without data fusion. C4.5 decision trees are trained on both raw sensor data and extracted features and classification performance is evaluated with tenfold cross-validation and 80/20 training/test methods. It was found that recognition accuracy depends on location and sensor types. Best results are achieved at locations closer to the core of the body. Classifiers derived from other sensor data achieve comparable performance to triaxial accelerometers. This study suggests that exergame hardware would benefit from incorporating multiple sensor modalities into a single device.

Keywords

pervasive gaming, activity recognition, sensor placement, sensor modality, parallel sensing

CCS Concepts

•**Computing methodologies** → *Mixed / augmented reality;* Feature selection; •**Software and its engineering** → **Interactive games;** •**Hardware** → *Sensor devices and platforms;*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ACSW '17, January 31-February 03, 2017, Geelong, Australia

© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ISBN 978-1-4503-4768-6/17/01…$15.00

DOI: http://dx.doi.org/10.1145/3014812.3014875

1. INTRODUCTION

A variety of gaming technologies implementing automated gesture and activity recognition have been created over the past two decades, to facilitate active gaming products. Using detected physical activity as an input to gameplay processes enables the relaxation of the boundaries between the virtual and physical world, enhancing both the gaming experience and the meaning of the physical activity. In-game reward systems that promote appropriate physical activities thus provide extrinsic motivation for exercise.

Coincident with the rapid uptake of mobile devices such as tablets and smartphones, active games have moved outside of the loungeroom and into school yards and playgrounds, providing broader and richer opportunities for engaging children in augmented reality exergames [19] [20], [21]. Such games have been proposed as a novel means of re-engaging children with physical activity and in encouraging a healthy balance of exercise and entertainment [12].

Smartphones are an ideal platform for delivering exergames, given their small size, low energy consumption, high resolution displays and growing computational power. They are also a useful platform for real-time activity recognition - an essential component of active games software - given their array of onboard sensors and onboard processing capabilities [9, 10, 3].

While there is a growing body of literature covering embedded activity recognition systems using smartphones and other wearable sensors, much of this research has been conducted in the context of aged care. In such contexts researchers are interested in automated monitoring of daily living activities of subjects, in order to identify abnormal events such as falls. While it may be reasonable to expect that existing techniques for activity recognition on smartphones will apply to sporting and exercise contexts, there remain several open questions that are considered in this study. In particular, the question as to the most suitable location for collecting data during activities such as running, jumping, and throwing, cannot easily be answered by results obtained for tasks such as eating with a spoon, drinking, or sitting down. Furthermore, the predominant sensor modality used in on-body activity recognition systems is the accelerometer. Smartphones offer a variety of sensors and thus an assessment of the capability of these sensors to produce useful data in the active games context is called for.

To this end, this article presents our study of sensor modality and placement within the context of children's playground activities. Parallel sensing using a set of smartphones is conducted to generate activity dependent data, which is then used to train decision tree classifiers. We consider several data protocols, including training per-sensor classifiers with raw (temporally unordered) data, features extracted from raw data and combined raw and feature data. We also propose a novel approach whereby a classifier is trained using all sensor data at one location, but without prior data fusion. Our findings suggest that batch processing of data from multiple sensors in one location offers better results than using a single sensor type (such as accelerometer) placed at several bodily locations. This is a new finding with regards to previous studies into the location and number of wearable sensors used for activity recognition.

The remainder this article is structured as follows. We consider related studies that have investigated aspects of the current problem but within other contexts and then present our experimental methodology, including our data collection and pre-processing protocols, as well as a justification for our classifier algorithm selection. We present results from performance evaluation of C4.5 decision trees as human activity classifiers, considering parallel sensing and diverse sensor modalities. We highlight our significant findings and discuss briefly their implication. Finally, we present our conclusions and future work.

2. RELATED WORK

Methods for activity recognition using wearable sensors and smartphones have focused predominantly on data obtained from accelerometers [14, 9, 10, 7, 15, 18], following the trend set out by early biomechanics studies [24]. Given the noisy data obtained from MEMS accelerometers and the subsequent difficulty in reconstructing user trajectories, classification techniques based on both supervised and unsupervised learning from features extracted from acceleration data have become the norm in this field. Chen and Nugent suggest that recognition algorithms can be divided into two broad approaches: machine learning techniques based on probabilistic and statistical reasoning; and, logical modelling and reasoning [8]. The majority of recent literature considers robust classification techniques using supervised learning (although unsupervised learning approaches have been proven viable).

A variety of supervised learning techniques have been applied, including Hidden Markov Models [17, 16, 6], naïve Bayes [22, 3], decision trees [2, 22] and k-Nearest Neighbours clustering [22, 3, 15]. Whereas historically most classifiers have been trained using windowed data, recently Bruno [4, 6] proposed training of an activity classifier using raw, temporally unordered accelerometer data. With increasing computational power available in mobile devices, the real-time batch processing of large quantities of raw sensor data is possible and thus further consideration of this approach and its performance compared to traditional feature-based methods is warranted.

Schmidt and Van Laerhoven identified sensor placement as a critical factor in activity recognition success and related this to the increased computational burden of training induced by poor sensing location [23]. Bao and Intille evaluated performance of recognition derived from accelerometers placed at the hip, wrist, arm, ankle and thigh [2]. They found that for upper body movements, the upper arm and wrist offered the best results. They also found that when using two accelerometers, the thigh and wrist provided the best combination of locations. Lester suggested that placing multiple sensors on a subject can be cumbersome and obstructive when collecting regular data [16].

However, considering the use of multiple sensors, Bao and Intille also showed that recognition accuracy was greatly improved using two sensors over one (by around 30%). Although they still focused on one sensor modality (i.e., acceleromaters), they also found that using two locations (hip/wrist, thigh/wrist) affected accuracy only marginally (less than 5%) when compared to five sensors used simultaneously, suggesting that careful selection of a small number of sensing sites could maintain accuracy, while diminishing the burden on the user. Of note though was that they were using only biaxial sensors in their study and thus it remains to be seen whether these results apply to smartphones incorporating triaxial sensors.

A common theme of these studies was that the sensor location was constant for both training and testing. Henpraserttae abandoned this assumption and presented a study of activity recognition regardless of device orientation and location on the body [13]. Not surprisingly they found that for accurate recognition different models were required according to the sensor location.

Further to the question of accuracy and its reliance on sensor modality, Ravi found that for activities such as climbing stairs, using only accelerometers made the activities indistinguishable regardless of the direction of motion (up versus down) [22]. Lester though found that the loss of accuracy using a single sensing location could be compensated for through the use of complementary sensors [17]. That study considered accelerometer, audio, IR/visible light, high frequency light, pressure, humidity, temperature and compass data. Three sensors were found to be the most significant (accelerometer, microphone and barometer) accounting for nearly 75% of the fifty features most significant in the classification of ten distinct activities.

Most studies of human activity recognition cite applications in aged health care and tele-care, considering tasks such as ambulatory monitoring for fall detection, metabolic energy expenditure, and recognition of daily living activities. In most cases these are low energy activities (for instance, during even gait, 99% of accelerometer signal energy lies below 15 Hz [1]). While it is reasonable to surmise that modern approaches to activity recognition can be applied to higher energy locomotor activities (such as running, skipping, hopping and jumping), it remains to be seen as to what effect sensor modality and placement has in this context and whether previous results hold true.

3. METHODOLOGY

3.1 Data Collection

In this study five LG Nexus 4 smartphones running Android were used to collect data from onboard hardware sen-

Figure 1: Smartphone placements. The 5 phones are placed on dominant lower arm (outside), dominant upper arm (outside), thigh (outside), lower leg (outside) and non-dominant side hip.

sors: triaxial accelerometer and gyroscope, (visible) light, barometer, and magnetometer. Additionally data was collected from software-defined sensors that computed linear acceleration, gravitational acceleration and the instantaneous rotation vector.

Two subjects (1 female and 1 male, ages 21 and 22) were used in the study to collect data simultaneously from five body locations:

- Dominant lower arm (outside).
- Dominant upper arm (outside).
- Thigh (outside).
- Lower leg (outside).
- Non-dominant side hip.

Figure 1 shows one of the two subjects in this study wearing phones affixed with commercially available bands. Some issues were noted during data collection, where some records were not captured. This is not expected to significantly affect the outcome but those cases have been tagged where they affect the results presented.

All phones were connected to an isolated wireless network, which permitted synchronisation of data collection using a common network time signal. A custom application was developed to use the Android SensorManager to poll sensors at a frequency of 30Hz, which was the limiting frequency above which data writes to the SD card interfered with sensor read requests on the CPU. While some variation in network event processing of the time trigger signal was noted across phones, it was found to be less than 5 milliseconds and thus not a significant effect.

Five activities from the gameplay of an exergame [19, 20] developed for our ongoing studies with school children were undertaken by each subject. These were: running, walking, jumping, hopping and throwing. To ensure sufficient data for each activity trial from each subject, the total duration of data for each activity was set as: walking for 120 seconds, running for 100 seconds, jumping for 80 seconds, hopping for 45 seconds and throwing for 130 seconds.

Crashes on individual phones during the activities resulted in some data loss. This is not believed to impact on the later analysis, but annotations have been included in the results where this may affect direct comparability. The thigh phone did not record the walking activity and lost 80 seconds of the running activity. A review of the data collected from the phones on hip and upper arm indicated that their sampling rate was lower than the other phones, resulting in producing approximately 40-50% and 25-30% respectively of the samples returned by the other devices.

3.2 Algorithm Selection

Selection of an appropriate algorithm for supervised learning of an activity classifier was undertaken by an evaluation of several algorithms available through WEKA [11]. Using data sets provided by Bruno [5] and Kwapisz [15], the following algorithms were used to train classifiers: J48 (C4.5 decision tree), Logistic (logistic regression), IBk (k-Nearest Neighbours), and multilayer perceptron (using backpropogation learning).

Bruno's dataset contains data of 16 subjects wearing a triaxial accelerometer attached to their right wrist. The activities performed were: sit down on chair, stand up from chair, get out of bed, eat soup, eat meat, pour water, ascend stairs, descend stairs, brush teeth, lie down, drink from a glass, comb hair, and walk.

Kwapisz's dataset consists of triaxial accelerometer data gathered from 29 subjects wearing an Android phone (Nexus One, HTC Hero, or Motorola Backflip) positioned on the upper leg area. The activities performed were: walking, jogging, ascending stairs, descending stairs, sitting, and standing.

Two approaches to data pre-processing consistent with those taken by each team were used. As per Bruno's methodology [6] classifiers were trained using raw, temporally unordered accelerometer data. Kwapisz [15] on the other hand constructed feature vectors from the per channel arithmetic mean, standard deviation, root mean square and time between peaks, as well as the average resultant acceleration and binned distribution.

A ten-fold cross validation study of the algorithms was conducted on each data set and accuracy and precision results consistent with those published in [6] and [15] were obtained. Table 1 shows the results of our replication study of Kwapisz's work against their published results. The J48 algorithm also produced results consistent with Bruno's work but are not directly comparable since the original results were obtained using a Hidden Markov Model. On this basis the J48 algorithm was selected for our present study.

3.2.1 Data Processing

Data from all subjects was combined to produce a single data set which was then used for all subsequent analysis. Three feature extraction and training protocols were used in this study. The first, in line with Bruno's recent work [6] was to train classifiers using raw, temporally unordered data. Data from each sensor, for each instance of an activity, was used as exemplars for training and testing using a ten-fold cross-validation methodology.

The second protocol involved the extraction of features from non-overlapping consecutive windows of 10 data samples, for each sensor channel. The features used were those reported by Kwapisz [15], being: per-channel arithmetic mean, standard deviation, minimum and maximum, as well as the RMS magnitude for each sensor and the binned his-

Activity	J48	Logistic Regression	Multilayer Perceptron
Replication	89.4611%	84.9575%	92.1189%
Original	85.1%	78.1%	91.7%

Table 1: Evaluating the reproducibility of results using 10-fold cross validation using data set from [15]. The classifier chosen for this study is the J48 (C4.5 decision tree) given both its consistency with this study and with results produced using a Hidden Markov Model from [6].

togram. Classification using only extracted features, and using extracted features plus raw data was evaluated.

Finally, in a break from previous studies, complete log files obtained for each activity at each location were used to train individual activity classifiers for that location. Within the log file all sensors report at each poll event, providing a fully observed data set for all hardware and software sensors. Each log file is then used as an exemplar for training and testing.

4. RESULTS & DISCUSSION

4.1 Raw Data Models

Typical properties of the data collected is summarized in Table 2.

The results presented represent the analysis performed on data collected from two subjects. Models were built using J48 decision tree using default settings provided by WEKA (unless specified) and assessed using 10-fold cross validation. Some data loss occurred during collection meaning that not all activities are recorded for all devices but we believe the impact of this is minimal. Affected records have been tagged.

The accuracy (true positive rate) achieved for each sensor and for each location is as shown in Table 3.

The individual raw sensor values are used for creating these models, analogous to strategies used elsewhere [4, 6, 5]. While a more meaningful and accurate assessment of accuracy would be achieved through cross validation that removed entire subjects or activity records rather than individual measurements, the focus in this study is on the merits of alternative sensor placements and modalities.

These results show a number of unexpected effects. The accuracy achieved for each case is surprisingly high given that each involves identifying activities with a single sensor reading. This does suggest that individual sensor values are often specific to a specific activity. While this may be due to the small data set used it does imply that low latency activity recognition could be feasible using short sequences of sensor values.

The accuracy achieved by other sensors in comparison to the accelerometer is unexpected, particularly given the prevalence of accelerometers for activity measurement and recognition. There are several sensors that outperform the accelerometer in specific locations. Particularly surprising is the accuracy achieved with light and pressure sensors which are low dimension limited fidelity sensors. It is anticipated that these will be of less benefit once larger data sets are collected under a range of environmental settings.

Alternative motion related hardware sensors such as the magnetic compass and gyroscope provide better or equivalent performance to the hardware accelerometer. Software enhanced values such as gravity and rotation perform better still. We hypothesize that filtering operations involved in producing these values reduce sensor noise and result in improved classification.

Accuracy does vary with location. As all locations are measured simultaneously for exactly the same locations, the effect of different locations can be compared directly. Highest values are associated with the upper arm, hip and thigh. These locations are closest to the core of the body and may have more limited movement or benefit from the implicit smoothing resulting from their fewer degrees of freedom. These positions are also common locations for attaching accelerometers for clinical trials or jogging apps. This may also suggest that activity recognition for trendy wrist based devices may be a harder problem.

4.2 Transformed Data Models

Two approaches were investigated to explore the impact of transformations on the data using just the acceleration values. Consecutive sequences of 10 elements from the raw data sets are used to produce transformed data. In the first approach, a small set of features are extracted (means of each accelerometer channel, and mean, min and max of the magnitude). The second approach extends this feature vector to include the acceleration values for each element.

The accuracy values achieved by J48 decision trees combined with a partition membership filter are shown in Table 4.

Despite the limited complexity of the feature vector this approach performs at levels comparable to the raw data. We anticipate that this approach will be more robust as the range of features improves. This is supported by the results when these features are augmented with the raw data values which, as would be expected, gives accuracy superior to anything achieved by individual raw sensor values regardless of the sensor type. These results could also be indicative of over-fitting. We are, however, able to reduce the proportion of data used for training as low as 20% of the data set size while still achieving superior performance.

The effect of location mirrors that for the raw data set, thus supporting the importance of appropriate sensor placement for activity recognition.

4.3 Multiple Sensors

The benefits of this investigation include the ability to use multiple sensors concurrently and to use sensors at multiple locations to enhance activity recognition. The latter involves synchronizing readings from multiple devices and that analysis will be reported in a later paper. The effect of heterogeneous sensors on recognition is discussed in this section.

Since all the sensors on an individual device report values at different rates and asynchronously there are several ways of producing individual data records. We elected to create a new record every time a single sensor updates, recording all other fields as the last reported value of that sensor. As may

Activity	Total Duration/[s]
Walking	120
Running	104
Jumping	75
Hopping	34
Throwing	135

Sensor	#Samples	Minimum	Maximum	Mean	Standard Deviation
Accelerometer	6913	-15.283	24.315	1.663	3.159
Gravity	4626	-4.854	9.704	1.935	1.431
Gyro	4341	-1.2	1.238	-0.021	0.357
Light	708	1253	9926	2970.768	1280.585
Linear Acceleration*	4317	-9.017	6.899	-0.136	2.071
Magnetic Field	6567	-46.259	60.419	7.723	23.031
Pressure	3055	1004.598	1005.397	1005.187	0.064
Rotation*	4217	-0.235	0.7	0.514	0.145

Table 2: Summary of data collected. Data from 5 activities was collected, using 8 physical and virtual (through fusion and internal filtering) from the devices available. The number of samples per sensor is related to the rate at which new data points are available from that sensor.

Sensor\Location	Upper Arm*	Lower Arm	Hip*	Thigh*	Lower Leg
Accelerometer	0.773	0.636	0.739	0.704	0.618
Gravity	0.905	0.633	0.779	0.771	0.690
Gyro	0.959	0.480	0.755	0.650	0.593
Light	0.866	0.447	0.736	0.681	0.446
Linear Acceleration	0.965	0.593	0.870	0.742	0.577
Magnetic Field	0.870	0.804	0.894	0.883	0.847
Pressure	0.921	0.642	0.717	0.644	0.650
Rotation	0.990	0.808	0.971	0.919	0.872

Table 3: Accuracy achieved for J48 decision trees using 10-fold cross validation for each combination of sensor and location. Records for which some samples were lost during collection are marked with *.

Evaluation Location	Features only 10-fold CV	Features + raw 10-fold CV	Features + raw 20%/80% training/test
Upper Arm	0.775	0.969	0.853
Lower Arm	0.682	0.911	0.780
Hip	0.757	0.940	0.852
Thigh	0.756	0.934	0.860
Lower Leg	0.692	0.927	0.831

Table 4: Accuracy achieved for J48 decision trees with partition membership filter on transformed accelerometer data. A 10-fold cross validation is used for the first 2 columns while a 20%/80% train/test split is used for the third. Summary features produce results comparable to those achieved with individual readings while including raw acceleration values improves accuracy further. Consistent performance even with smaller training sets suggests that over fitting may not be the cause in this case.

Assessment Location	10-fold CV	2%/98% training/test
Upper Arm	1.000	0.953
Lower Arm	1.000	0.953
Hip	1.000	0.984
Lower Leg	1.000	0.996

Table 5: Accuracy achieved for J48 decision trees on records combining all sensor readings. The thigh readings are omitted because of partial data loss for some sensors. The 10-fold CV results are believed to result from over-fitting. However using only 2% of data for training still indicates that combining readings from sensors at multiple locations would add value to activity recognition.

be expected, this does lead to subsequent records having many elements in common with their predecessors which can assist with clustering. The several sensors with high update rate such as the accelerometer derived values, gyro and compass ensure that the specific fields changed on each update will typically vary between adjacent records. The similarity between records does increase the likelihood of very similar records in training and test sets providing over-fitting effects.

The accuracy values calculated using data records containing all sensor values are reported in Table 5.

Evaluation with 10-fold cross validation returns unbelievable results corresponding to the expected over-fitting. We would expect this effect to be less significant when much smaller proportions of the data are used for creating the model. Table 5 shows that the size of the training set can be reduced down to 2% of the data set size while still achieving very credible values. This does suggest that there is still merit in using combined sensors and worth pursuing this direction once larger data sets are available. It would similarly be interesting to explore if the improved performance at extremities is significant in this case. Over-fitting issues could also be investigated by collecting multiple records for the same activities and dividing these complete records between training and test data sets.

4.4 Discussion

It is quite apparent that phones placed closer to the core of the body generate better models than the phones placed on the extremities. This may be caused by the more vigorous movement that the lower arm and lower leg generate and partly by these two phone placements being prone to loosening through the collection phase. These optimal positions are also non-intrusive upon the subject's movement and can generate natural results. If we were to consider the best placement for phones the best candidates would be upper arm, thigh and hip.

This result is consistent with the previous research [2] using multiple simultaneous accelerometers for activity recognition. Thigh and hip were identified as providing the best recognition results overall, consistent with the body-code hypothesis identified here. Results for pairs of accelerometers were able to achieve results comparable to accuracy achieved using all five, suggesting that additional sensors add value as observed in this study, but that the extent of improvement with additional sensors decreases as homogeneous sensors are added at other body positions.

Magnetic Field and Rotation sensors are on average the most accurate sensors in this device when operating on raw data. The unprocessed accelerometer sensor is one of the worst performing sensors. This suggests that different sensors and a preliminary filter would benefit activity recognition processes.

The small size of the data set and use of only two subjects may have introduce bias within the data. However initial results indicate potential benefits for using and combining different classes of sensors. This approach achieved this through the machine learning mechanisms employed here rather than through any deliberate data fusion mechanism. A clear direction for future research is to extend these approaches and validate them with more diverse data.

5. CONCLUSION

This paper examines the placement of smartphone devices for activity recognition. Considering the literature on the subject we looked at the methods and tools for classifying accelerometer data. A number of factors affected (and improved) recognition relative to single accelerometer based approaches strapped to thigh/upper arm. Through the use of multiple smartphones placed across the body we were able to determine relationships between sensor placements and accuracy of activity recognition. In general, positions closer to the core of the body are more desirable. Activity recognition can be also achieved with smartphone sensors other than the accelerometer. Both the magnetic field and (software defined) rotation vector sensors achieve good results. Furthermore, it is an indication that virtual sensors (those that involve filtering within the Android SDK) give better results than raw sensor data, suggesting noise removal may be the causative factor.

The significance of these results is limited by the relatively small sample size. Significant accuracy levels are being reported regardless of the type or placement of the sensor and we believe this is indicative of limitations of typical activity recognition. Many related studies have taken a similar approach but limited to just a single data source although often collecting more data. It is clear from these results that sparse sampling of the sensor space can result in misleading levels of accuracy and we recommend our strategy of including additional sensors or senor locations to serve as a reference level. Recognition levels would have to be significantly better than the reference standard to be considered meaningful.

5.1 Contributions

Our investigation has identified the effect of smartphone placement on accuracy of activity recognition. This has implications for future data collection procedures but also impacts on the placement of wearable devices for activity related ubiquitous computing. Applications include development of sensor bracelets and similar devices for monitoring activity and exercise levels, and placement for systems using activity sensing to provide feedback for training and physical therapies.

Different and multiple sensors play a significant role in recognition outcomes. The accelerometer should not be automatically regarded as the de facto activity measurement and recognition sensor. Other sensors commonly found in mobile and wearable platforms could work as well or better.

Combinations of sensors would provide benefits for activity monitoring in wearable systems.

While the small size of the data set does limit the significance of the findings it does also suggest that systems could be trained during start-up stages of a mobile exergame using relatively short periods of activity to accommodate the characteristics of each individual player.

5.2 Future Investigation

This investigation has identified some surprising opportunities for improving activity recognition outcomes using permutations of different sensors and placements. These effects need to be validated using more comprehensive data sets representing a greater range of activities and conditions. A further direction for investigation within the scope of this work and which is being pursued is an investigation into the concurrent use of sensor data from multiple locations. For our context of exergames we want to know how quickly, accurately and robustly can we identify a particular activity and if having more sensors means that we can use shorter sequences of sensor readings, effectively trading sequential values for parallel.

6. REFERENCES

[1] E. K. Antonsson and R. W. Mann. The frequency content of gait. *Journal of Biomechanics*, 18(1):39–47, 1985.

[2] L. Bao and S. Intille. Activity recognition from user-annotated acceleration data. In *Pervasive 2004*, volume 3001 of *LNCS*, pages 1–17. Springer, 2004.

[3] T. Brezmes, J.-L. Gorricho, and J. Cotrina. Activity recognition from accelerometer data on a mobile phone. In *Proceedings of the 10th International Conference on Artificial Neural Networks: Part II*, pages 796–799. Springer, 2009.

[4] B. Bruno, F. Mastrogiovanni, A. Saffiotti, and A. Sgorbissa. *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, volume 443, chapter Using Fuzzy Logic to Enhance Classification of Human Motion Primitives, pages 596–605. Springer, 2014.

[5] B. Bruno, F. Mastrogiovanni, and A. Sgorbissa. A public domain dataset for adl recognition using wrist-placed accelerometers. In *Robot and Human Interactive Communication*, pages 738–743, 2014.

[6] B. Bruno, F. Mastrogiovanni, and A. Sgorbissa. Recognition of human activities through wearable accelerometers. In *Robot and Human Interactive Communication*, 2014.

[7] P. Casale, O. Pujol, and P. Radeva. Human activity recognition from accelerometer data using a wearable device. In *Proceedings of the 5th Iberian Conference on Pattern Recognition and Image Analysis*, volume 6669 of *LNCS*, pages 289–296. Springer, 2011.

[8] L. Chen and C. Nugent. Ontology-based activity recogntion in intelligent pervasive environments. *Information Systems*, 5(4):410–430, 2009.

[9] T. Choudhury, G. Borriello, S. Consolvo, B. Harrison, J. Hightower, A. Lamarca, L. Legrand, A. Rahimi, A. Rea, G. Bodello, B. Hemingway, P. Klasnja, K. Hoscher, J. A. Landay, D. Wyatt, and D. Haehnel. The mobile sensing platform: an embedded activity recognition system. *Pervasive Computing*, 7(2):32–41, April 2008.

[10] N. Györbíró, Á. Fábián, and G. Hományi. An activity recognition system for mobile phones. *Mobile Networks and Applications*, 14(1):82–91, February 2009.

[11] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten. The weka data mining software: An update. *SIGKDD Explorations*, 11(1):596–605, 2009.

[12] L. W. Hansen and S. W. Saunders. Active gaming: A new paradigm in childhood physical activity. *Digital Culture & Education*, 3, 2011.

[13] A. Hennpraserttae, S. Thiemjarus, and S. Marukatat. Accurate activity recognition using a mobile phone regardless of device orientation and location. In *Proceedings of the 2011 International Conference on Body Sensor Networks*, 2011.

[14] D. M. Karontonis, M. R. Narayanan, M. Mathie, H. H. Lovell, and B. G. Celler. Implementphone of a real-time human movement classifier using a triaxial accelerometer for amulatory monitoring. *IEEE Transactions on Information Technology in Biomedicine*, 10(1):156–167, 2006.

[15] J. Kwapisz, G. Weiss, and S. Moore. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, 12(2):74–82, 2011.

[16] J. Lester, T. Choudhury, and G. Borriello. A practical approach to recognizing physical activities. In *Lecture Notes in Computer Sciences*, volume 3968, pages 1–16. Springer, 2006.

[17] J. Lester, T. Choudhury, N. Kern, G. Borriello, and B. Hannaford. A hybrid discriminative/generative approach for modelling human activities. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 766–772, 2005.

[18] N. Li, Y. Hou, and Z. Huang. Implementation of a real-time human activity classifier using a triaxial accelerometer and smartphone. *International Journal of Advancements in Computing Technology*, 5(4):234–242, 2013.

[19] S. McKenzie, S. Bangay, L. Barnett, and N. Ridgers. Active gaming as a mechanism to promote physical activity and fundamental movement skill in children. *Frontiers in Public Health*, 1(74), December 2013.

[20] S. McKenzie, S. Bangay, L. Barnett, N. Ridgers, and J. Salmon. Design elements and feasibility of an organized multiplayer mobile active videogame for primary school-aged children. *Games for Health Journal*, 3(6):379–387, September 2014.

[21] B. Mortazavi, S. Nyamathi, S. Lee, T. Wilkerson, H. Ghasemzadeh, and M. Sarrafzadeh. Near-realistic mobile exergames with wireless wearable sensors. *Journal of Biomedical and Health Informatics*, 18(2):449–456, March 2014.

[22] N. Ravi, N. Dandekar, P. Mysore, and M. Littman. Activity recognition from accelerometer data. In *Proceedings of the National Conference on Artificial Intelligence*, volume 20, pages 1541–1546. MIT Press, 2005.

[23] A. Schmidt and K. van Laerhoven. How to build smart appliances. *IEEE Personal Communications*, 8(4):66–71, 2001.

[24] P. H. Veltnik, H. Bussman, W. W. de Vries, W. Martens, and R. C. van Lummel. Detection of static and dynamic activities using uniaxial accelerometers. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 4(4):375–385, 1996.