

# Physical Activity Recognition from Accelerometer Data using a Wearable Device.

Pierluigi Casale, Oriol Pujol, and Petia Radeva

Computer Vision Center, Bellaterra, Barcelona, Spain  
Dept. Applied Mathematics and Analysis, University of Barcelona, Barcelona, Spain  
{pierluigi, oriol, petia}@cvc.uab.es

**Abstract.** Activity Recognition is an emerging field of research, born from the larger fields of ubiquitous computing, context-aware computing and multimedia. Recognizing everyday life activities will be in the short future a fundamental application in pervasive computing. In our work, we developed a wearable system easy to use and comfortable to bring. We obtain very good classification accuracy for each activity, using one stage of classification without any further sequential classification stage.

**Keywords:** Physical Activity Recognition, Wearable Computing, Pervasive Computing-

## 1 Introduction

Activity Recognition is an emerging field of research, born from the larger fields of ubiquitous computing, context-aware computing and multimedia. Recognizing everyday life activities will be in the short future a fundamental application in pervasive computing, with a lot of interesting developments in the health care domain, the human behavior modeling domain and the human-machine interaction domain. Even if first works about activity recognition used high dimensional and densely sampled audio and video streams [7], in many recent works activity recognition is based on classifying sensory data using one or many accelerometers. Accelerometers have been widely accepted due to their miniaturization, their low-power requirement and for their capacity to provide data directly related to motion.

In recent years, a great number of papers have been published where accelerometer data analysis has been applied and investigated for physical activity recognition. Nevertheless, few of them override the difficulty to perform experiments out-of-the-lab. The necessity to perform experiments out-of-the-lab brings to the necessity to build easy to use and easy to bring systems and to free the testers from the expensive task of labeling the activities they perform. In addition, many different kind of features have been widely used, each of them good for discriminating different typologies of activities.

In our work, we developed a wearable system easy to use and comfortable to bring. From 14 volunteers, we collected data using an accelerometer and a cheap webcam. Experimenters performed five basic every-day life activities. Activities

are walking, walking up/down stairs, staying standing, talking with people and working at computer. Experimenters were free to perform the activities where they want and for the time they want. We free them from the task of labeling activities and we just ask them to annotate the sequential order they perform the activities and to restart the system each time they start performing an activity. With the data collected, we select which are the best features from accelerometer data for classifying the five activities we select. Using these results, we reach very good classification performances, using just one classification stage without any further sequential classification stage.

This paper is structured as follow. After discussing related work in Section 2, we will describe in Section 3 how we create the dataset using our wearable device. In Section 4 we will provide the technical details about how we select the best features for classify our activities and in Section 5 we will present the results of the classification of the activities. Finally, in Section 6, we will conclude.

## 2 Related Works

In [2], Mannini and Sabatini give a complete review about the state of the art of activity classification using data from one or more accelerometers. In their review, the best classification approaches are based on wavelet features using threshold classifiers. Furthermore, they classify 7 basic activities and transactions between activities, from 5 biaxial accelerometer placed in different part of the body, using a 17th-dimensional feature vector and a HMM-based sequential classifiers, achieving 98.4% of accuracy. In their work, they separate high-frequency ( AC ) components, related to the dynamic motion the subject is performing from low-frequency ( DC ) components of the acceleration signal related to the influence of gravity and able to identify static postures. They finally extract features from the DC components. In our work, we will use the same separation and we try to understand if both AC and DC components can be useful for classifying activities.

Lester, Choudhury and Borriello in [1] resume their experience in developing an automatic physical activities recognition system. In their work, they answer some important questions about where sensors have to be placed in a person, if variation across users helps to improved the accuracy in activity classification and which are the best modalities for recognizing activities. Their principal results can be summarized in the following:

- it does not matter where the users place the sensors ;
- variation across users do help improving accuracy classification ;
- the best modalities for physical activities recognition are accelerometers and microphones.

In the developing phase of our wearable system and the dataset, we take into account all these points. The authors use two classification stages. On the first stage, they use a modified version of AdaBoost and, in the second stage, they also share the philosophy of using an HMM classifiers. They achieve 87% of accuracy in classifying 8 activities using 50 features.

Hermes, Parkka, Mantyjarvy and Korhonen in [3] used an hybrid model composed by a Decision Tree and an Artificial Neural Network (ANN) for classifying 9 not strictly basic activities with 90% of accuracy. They suggest that embedding a priori knowledge into a classifier can improve the results of classification and thus they build a decision tree which split nodes are ANNs. An important issue of their system is due to the noise and inaccuracy of data, that can prevent the convergence of the system if a not adequate learning rate in the ANNs is chosen. Noisy data are mostly outliers and outliers affect both features selection process and classification. For that reason, in our experiments we use Random Forest, more robust respect to AdaBoost and ANN on noisy data.

### 3 The Dataset

The wearable system is based on Beagleboard, a low-price board built around the TI OMAP system on chip. We use Linux as operating system on the board. A low-cost usb webcam and a bluetooth accelerometer are connected with the board. The system is powered using a portable lithium battery able to power up to four hours the system. Users can wear the system as in Figure 1, where we print upon the picture the directions of the acceleration axis. The system



**Fig. 1.** The Wearable System.

works with three modalities, video, audio and accelerometer data. It takes photos, grabs audio continuously applying a filter for voice removal and it receives via bluetooth data from the accelerometer. Audio and video are acquired using the Gstreamer framework [8]. Bluetooth data are acquired using the Bluez interface [9]. Accelerometer data are sampled at 52Hz with a resolution of  $\pm 4g$ . Audio is sampled at 44100bps. Video is sampled at 1fps. All the sensors are localized in the same part of the body. In particular, we select to put the sensors on sternum because the sternum can be seen as the center of mass of a person.

Data have been collected from fourteen testers, three women and eleven men with age between 27 and 35 performing as many activities as they want. Activities are the following:

- walking up/down stair ;
- walking ;
- talking with people ;
- staying standing ;
- working at PC.

For labeling activities, people were asked to annotate the sequential order of the activities they performed and restart the system. Every time the system starts, data are named with a serial number. Once the user presses the starting button, she/he can start to perform the activity. The system boots in less than 2 minutes and the acquisition automatically starts while the user is already performing the activity. In this way, there are no “border effects“ due to starting. The user can stop the acquisition in every moment pressing again the start button.

The dataset collected is composed by:

- 33 minutes of walking up/down stairs ;
- 82 minutes of walking ;
- 115 minutes of talking ;
- 44 minutes of staying standing ;
- 86 minutes of working at computer.

## 4 Features Selection for Motion Data.

In this section, we discuss the features we extract for the motion modality and the quantitative results about the best features for classification.

In our setting, we take into account four separated accelerometer data time series, one time series for acceleration on each axis  $A_x, A_y, A_z$  and, in addition, one more time series,  $A_m$ , obtained computing the magnitude of the acceleration as in Formula 1.

$$A_m = \sqrt{A_x^2 + A_y^2 + A_z^2} \quad (1)$$

From these four time series, we create four new time series  $J_x, J_y, J_z, J_m$  obtained computing the derivative of the acceleration signals. In Physics, the derivative of acceleration is called Jerk. It represents the variation of the force acting on a body. The derivative has been approximated by finite differences.

### 4.1 Data Preprocessing and Features Extraction

Each time series  $T_i$ , with  $T = A, J$  and  $i = x, y, z, m$  has been filtered with a digital filter in order to separate low frequencies (*DC*) components and high frequencies (*AC*) components as suggested in [2]. The cut-off frequency has been set to 1Hz, arbitrarily. In this way, we obtain for each time series, three more time series  $T_{jf}$  with  $f = b, dc, ac$ , where  $b, dc, ac$  represent respectively the time series without filtering, the time series resulting from the low pass filtering and the time series resulting from the high pass filtering. Finally, we extract features from each one of the time series. A successful technique for extracting features from sequential motion data has been demonstrated to be windowing with overlapping. We extract features from data using windows of 52 samples, corresponding to 1 second of accelerometer data, with 50% of overlapping between windows. From each window, we extract the following features :

- mean value ;
- standard deviation ;
- skewness ;
- kurtosis ;
- correlation between each pairwise of accelerometer axis ( not including magnitude ) ;
- energy of coefficients of seven level wavelet decomposition ;
- root mean squared value of integration of acceleration in a window ;
- mean value of Minmax sums.

In this way, we obtain a 319-dimensional feature vector. While many of these features have been widely used in previous work on accelerometer data analysis and physical activity recognition, the last two features has been not yet exploited. In next session, we well show that many between the most important features are of these two typologies. The integration of acceleration is the Velocity. For each window, we compute the integration of signal and the RMS value of the series obtained. We approximate the integral using running sums with step equals to 10 samples. The Minmax sums are computed adding all the samples in the windowed time series sorted in crescent order with all the samples in the windowed time series sorted in decrescent order. From the series obtained, we compute the half of the mean value. The mean value of Minmax sums can be considered as a naive version of standard deviation.

## 4.2 Importance Measurement

For measuring attribute importance, we use Random Forest. The obvious advantage in using Random Forest instead of AdaBoost is that Random Forest performs much more better on noisy data as accelerometers data are. We will give here a short explanation about how Random Forest can be used for measuring attribute importance. For more details, see [4]. Random Forest build many classification trees. Each tree votes for a class and the forest choose the classification having the most votes over all the trees. Each tree is built as follows:

- if the number of cases in the training set is  $N$ ,  $N$  cases are sampled at random with replacement. This sample is the training set.
- if there are  $M$  input variables, a number  $m \ll M$  of variables are selected at random and the best split on these  $m$  is used to split the node. The value of  $m$  is held constant during the construction of the forest.
- Trees are not pruned.

When the training set for the current tree is drawn with replacement, about one-third of the cases are left out of the sample. This out-of-bag ( oob ) data is used to get an unbiased estimate of the classification error as trees are added to the forest. It is also used to get estimates of attribute importance. Measuring the importance of attributes on random forest is based on the idea that randomly changing an important attribute between the  $m$  selected for building a tree affects classification, while changing an unimportant attribute does not affect it

much. In this work, we use the Random Forest implementation provided by the Orange framework [10].

Importance of all attributes for a single tree are computed as: correctly classified oob examples minus correctly classified oob examples when an attributes is randomly shuffled. The importance measure is obtained dividing the accumulated attribute by the number of used trees and multiplying the result by 100. In Table 1 we report the best 20 features and their respective importance measures.

**Table 1.** List of Features selected by Random Forest.

Feature	Importance	Feature	Importance
Mean Value $A_{zdc}$	4.64	Mean Value $A_{ydc}$	3.86
MinMax $A_{zdc}$	4.61	Rms Velocity $A_{ydc}$	3.67
RMS Velocity $A_{zdc}$	4.23	Mean Value $A_{zb}$	3.59
RMS Velocity $A_{mdc}$	4.2	Mean Value $A_{xdc}$	3.57
RMS Velocity $A_{xac}$	4.14	MinMax $A_{xdc}$	3.52
Mean Value $A_{mdc}$	4.07	MinMax $A_{zb}$	3.51
MinMax $A_{ydc}$	3.92	Mean Value $A_{yb}$	3.33
Standard Deviation $A_{xb}$	3.9	Rms Velocity $A_{xdc}$	3.22
MinMax $A_{mdc}$	3.89	Rms Velocity $A_{zb}$	3.2
Standard Deviation $J_{xdc}$	3.87	MinMax $A_{yb}$	2.96

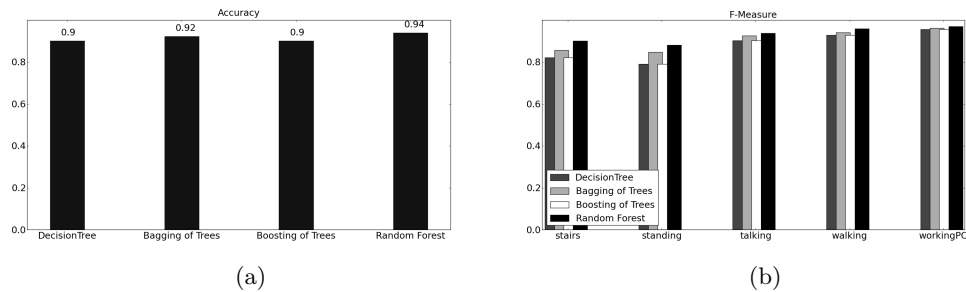
### 4.3 Discussion

Random Forest selects really meaningful features for classifying activities. The most important features selected are related to the  $Z$  axis that is, the direction of the movements. The majority of the features are relative to the  $DC$  components of movements and only the RMS velocity feature relative to the  $X$  axis from the  $AC$  components has been selected. Random Forest selects also the standard deviation of the *Jerk* relative to  $X$  axis. The information relative to the variation of movements on the  $X$  axis can help to discriminate between activities like staying standing, talking and working at pc. On the other side, features relative to the variation of movements on  $Y$  axis, can help to discriminate between activities like walking and walking up/down stairs. Mean value, minmax features and RMS velocity are selected for all the  $DC$  components of all the time series. Random Forest selects the best features but it is not able to discriminate between features bringing the same information. For example, all the features selected that have been extracted from the time series without filtering are also selected from the  $DC$  time series and, in all the cases, the features selected from the  $DC$  time series have an importance value bigger than the corresponding value from the series without filtering. Features deriving from higher level statistics ( skewness and kurtosis ) and features relative to the correlation between axis are features with the lowest importance. It has to be notes that the features selected are conceptually simple, significative and really easy to compute.

## 5 Classification Results.

In order to verify if the features selected are really informative, we use different classification methods for classifying our five activities. In the following section we will show the classification results obtained using Decision Trees, Bagging of 10 Decision Trees, AdaBoost using Decision Trees as base classifiers and a Random Forest of 10 Decision Trees. All the results are validated by 5-fold cross validation. The dataset  $D_m$  has been create using the 20 features selected from Random Forest.

In Figure 2(a) we show the classification accuracy of the classifiers trained on  $D_m$ . In Figure 2(b) we show the F-Measure of each activity for every classifier. The best classification accuracy is obtained using Random Forest. The confusion



**Fig. 2.** (a) Classification Accuracy for Different Classifiers.(b) F-Measure for each Activity on the Motion Dataset.

matrix is reported in Table 2.

**Table 2.** Confusion Matrix of Random Forest trained on  $D_m$ .

	stairs	standing	talking	walking	workingPC
stairs	<b>3535</b>	10	67	292	12
standing	24	<b>4592</b>	562	16	68
talking	61	481	<b>13175</b>	78	128
walking	296	11	89	<b>9471</b>	20
working	18	76	284	18	<b>9929</b>

All the classifiers have accuracy above the 90% using only the motion modality. The Random Forest classifier trained on  $D_m$  shows many very consistent confusions between similar activities like walking and walking up/down stairs, and between talking and standing. The F-Measure does not present significant differences between the classes. That means that the five activities can be recognized with high confidence.

## 6 Conclusions.

In this work, we developed a wearable system easy to use and comfortable to bring. We collected data using an accelerometer and a cheap webcam. Experimenters performed five basic every-day life activities like walking, walking up/down stairs, staying standing, talking with people and working at computer. The wearable device we made up allows experimenters to perform the activities where they want and for the time they want, freeing them from the expensive task of labeling activities.

Using the accelerometer data, we select the best features for classifying activities. Using Random Forest instead of AdaBoost for features selection allows us to override the limitation due to the noise, characteristic of accelerometer data. Random Forest selects very significative features for classifying activity. Those features are strictly related to the type of activity people are performing and result to be very simple to compute. Using the features selected, we reach very good classification performances.

Based on these results obtained using only the motion sensor, future works plan to add the other sensors to increase the classification performances. We expect that adding further information from the camera and the microphone can help considerably in discriminating between activities like “standing”, “talking” and “workingPC” or “walking” and “walking up/down stairs”, activities where the biggest confusions are present.

**Acknowledgments.** This work is partially supported by a research grant from projects TIN2009-14404-C02 and CONSOLIDER (CSD2007-00018).

## References

1. J. Lester, T. Choudhury, G. Borriello: A Practical Approach to Recognizing Physical Activities, **13** (1916), 114–119.
2. A. Mannini, A.M. Sabatini: Machine Learning Methods for Classifying Human Physical Activities from on-body sensors, **13** (1916), 114–119.
3. M. Ermes, J. Parkka, J. Mantyjarvi, I. Korhonen: Detection of Daily Activities and Sports with wearable Sensors in Controlled and Uncontrolled Conditions, **13** (1916), 114–119.
4. L. Breiman: Random Forests. Machine Learning 45 (1) pp: 532 (2001)
5. A.Krause, D. Siewiorek, A. Smailagic, J. Farrigdon: Unsupervised, dynamic identification of Physiological and Activity Context in Wearable Computing, **13** (1916), 114–119.
6. T. Huynh, M. Fritz, B. Schiele: Discovery of Activity Patterns using Topic Models,
7. B. Clarkson, A. Pentland: Unsupervised Clustering of ambulatory audio and video,
8. GStreamer: open source multimedia framework : [www.gstreamer.com](http://www.gstreamer.com)
9. BlueZ, Official Linux Bluetooth protocol stack : [www.bluez.org](http://www.bluez.org)
10. Orange - Data Mining Fruitful & Fun : [www.aillab.si/orange/](http://www.aillab.si/orange/)