Identification of Activities of Daily Living using Sensors Available in Off-the-shelf Mobile Devices: Research and Hypothesis

Ivan Miguel Pires[1,2,3], Nuno M. Garcia[1,3,4], Nuno Pombo[1,3] and Francisco Flórez-Revuelta[5]

[1]Instituto de Telecomunicações, University of Beira Interior, Covilhã, Portugal
[2]Altranportugal, Lisbon, Portugal
[3]ALLab - Assisted Living Computing and Telecommunications Laboratory, Department of Informatics, University of Beira Interior, Covilhã, Portugal
[4]ECATI, Universidade Lusófona de Humanidades e Tecnologias, Lisbon, Portugal
[5]Faculty of Science, Engineering and Computing, Kingston University, Kingston upon Thames, United Kingdom

[1]impires@it.ubi.pt, [2]ngarcia@di.ubi.pt, [3]ngpombo@di.ubi.pt, [4]F.Florez@kingston.ac.uk

Abstract. This paper presents a PhD project related to the identification of a set of Activities of Daily Living (ADLs) using different techniques applied to the sensors available in off-the-shelf mobile devices. This project consists on the creation of new methodologies, to identify ADLs, and to present some concepts, such as definition of the set of ADLs relevant to be identified, the mobile device as a multi-sensor system, review of the best techniques for data acquisition, data processing, data validation, data imputation, and data fusion processes, and creation of the methods for the identification of ADLs with data mining, pattern recognition and/or machine learning techniques. However, mobile devices present several limitations, therefore techniques at each stage have to be adapted. As result of this study, we presented a brief review of the state-of-the-art related to the several parts of a mobile-system for the identification of the ADLs. Currently, the main focus consists on the study for the creation of a new method based on the analysis of audio fingerprinting samples in some Ambient Assisted Living (AAL) scenarios.

Keywords: Sensors; data fusion; mobile devices; activities of daily living; data acquisition; data processing; data imputation; audio fingerprinting; pattern recognition.

1 Introduction

Mobile devices, *e.g.,* smartphones, tablets, smartwatches, and other specific devices, are commonly used while performing Activities of Daily Living (ADLs). Mobile devices may be considered in Ambient Assisted Living (AAL) systems, because they are able to collect data related to the user's environment and user activities, which can be used for different purposes, including the identification of ADLs, what is important to support the independent living or the management of chronic condition of elderly people and/or individuals with some type of disability.

This study is part of a PhD project that aims to create a new method to recognize ADLs with good accuracy, and to reduce the constraints of mobile devices for these tasks. This project started with the definition of two different sets: ADLs that may be reliably identified with mobile devices, and, valid sensors available in mobile devices. This work continues the research presented in [1], which included the calculation of a jump flight time that makes the use of pattern recognition techniques to identify patterns in vertical jumps [2]. Another work where similar analysis has been carried out is related to the Heel-Rise test [3], based on a test used mainly by physiotherapists that allows the detection of fatigue and/or specific diseases. A preliminary study for the identification of ADLs using mobile devices was already presented in [4], which will pave the way for the development of a personal digital life coach [5].

The proposed method includes three different stages, as follows: acquisition and processing, data fusion, and identification of ADLs based on the fused data. The first stage includes the activities of data acquisition, data processing, data validation, data cleaning, and data imputation. The second stage is focused on data fusion, which consists on the consolidation of the data acquired from all sensors, whereas the last stage consists on several techniques, including data mining, pattern recognition and machine learning techniques combined with the users' feedback in order to identify the ADLs. The main goal is that the method could provide an user profile based on events inferred from a regular activity day, making use of multi-sensor data fusion technologies and context awareness approaches, combined with other information available from the user context.

The remaining sections of this paper are organized as follows. Section 2 presents a review of the state of the art, focusing in the main concepts of this work. Section 3 introduces the proposed method to be developed during the PhD project, presenting a small overview of the mobile multi-sensor system, methods that can be applied to sensors available on off-the-shelf devices, the current approaches obtained and the work in progress. Section 4 presents the discussion and conclusions.

2 Related Work

This research involves several topics, such as the identification of ADLs, the sensors available in off-the-shelf mobile devices and the different methods for acquisition, processing, validation, imputation, cleaning, and fusion of the data acquired. To conclude, a summary of the review about data mining, pattern recognition and machine learning techniques are presented.

2.1 Identification of ADLs

People's self-care activities are commonly named as ADLs, which involve a desire to achieve a degree of physical comfort, self-care, and autonomy, that may promote feelings of independence and personal control. Examples of ADLs include activities related to personal appearance and hygiene, domestic skills, household management, family and child care, conversational and social skills, and leisure, education, training and work activities [6].

Several sensors, including the accelerometer, gyroscope, magnetometer, communication, proximity, light, gravity, RFID (Radio-Frequency IDentification) sensors, GPS (Global Positioning System) receiver, microphone and camera, may be used to detect when a person is having/preparing a meal, washing up, bathing, waking up, sleeping, standing, sitting, watching TV, using the phone, doing chores, cycling, jogging or perform other activities [7-10].

2.2 Sensors

Sensors are small hardware components that capture different types of signals, which are widely available in several mobile devices, including smartphones, smartwatches, and tablets, which may be used to collect data in a plethora of situations, including the identification of ADLs [11].

Sensors may be used for different purposes and working environments, collecting different types of data [12]. On the one hand, these environments can be classified as controlled, uncontrolled, static, dynamic, uncertain and undefined. On the other hand, sensors, can be categorized in mechanical, environmental, motion, imaging, proximity, acoustic, medical, chemical, optical, and force sensors.

Mobile devices may be considered as a multi-sensor platform, because they are equipped with several types of sensors whose combination may increase the reliability in the detection and identification of ADLs. However, the most important stage in multi-sensor systems is the signal classification with pattern recognition or machine learning methods [13]. One example is a wearable multi-sensor ensemble classifier for physical activity pattern recognition, developed in [14], which combines several classifiers based on different sensor feature sets to improve the accuracy of physical activities identification and recognition.

2.3 Data Acquisition

Data acquisition is supported with the sensors available in off-the-shelf mobile devices, which allows anywhere and at anytime the capture of data and conversion of the electrical signals received by each sensor into a readable format [15]. Several challenges are associated with the data acquisition process when recognizing ADLs, including the positioning of the mobile device, the data sampling rate and the number of sensors to be managed [16]. This process can be enhanced with several frameworks, including the *Acquisition Cost-Aware QUery Adaptation (ACQUA)*, consisting in a query processing engine implemented for mobile devices that dynamically modifies both the order and the segments of data streams requested from different sources [17].

2.4 Data Processing

Data processing may be implemented by two different architectures, such as the Device Data Processing Architecture and the Server Data Processing Architecture [18]. The first one, is designed to acquire the data from the sensors embedded in a mobile device and process the data locally. This architecture is useful when the processing methods require low resources, such as processing the accelerometer data, proximity sensor data, and others. On the contrary, Server Data Processing Architecture consists in the dispatch of the data collected to a remote server allowing the computation of a larger amount of data as well as computations of complex nature. This architecture is useful to work with data acquired by a GPS receiver or imaging sensors.

2.5 **Data Validation**

Data validation is the process executed simultaneously with data acquisition and data processing in order to validate if the data are correctly acquired or processed so as to obtain valid, accurate and reliable results. During the Data Validation stage, some tests are performed depending on the characteristics of the application, *e.g.* if it depends on continuous data collected during a time interval, if any value at a random instant is not correctly acquired or missing, the value should be inferred using data imputation techniques. Aiming at reducing noise artifacts the data captured or processed should be cleaned, treated or imputed, and therefore more prone to offer better and reliable results.

Data validation methods are mainly implemented with machine learning algorithms, such as principal component analysis (PCA), relevance vector machine (RVM) and artificial neural networks (ANN) [19, 20].

2.6 **Data Imputation**

Data imputation is applied when the acquired data is faulty. The missing data failures can be classified as Missing Completely At Random (MCAR), Missing At Random (MAR) and Missing Not At Random (MNAR) [21]. MCAR happens when missing values are randomly distributed across all observations. MAR is observed when missing values are randomly distributed within one or more subsamples instead of the whole data set. MNAR is the type of missingness that arises when missing values are not randomly distributed across observations. Several methods for data imputation are presented in [22], being the K-nearest neighbors (k-NN) the most used method [23].

2.7 **Data Cleaning**

Data cleaning is process similar to data imputation and is executed when the data contains incorrect values, also considered as noise. Data cleaning methods may be statistical and probabilistic methods, which are used to apply filters to the collected data and adjust the values into correct measurements [24]. The data cleaning process may be supported on spatial and temporal characteristics of the collected data.

2.8 **Data Fusion**

Data fusion is the most important step for the integration of the data collected by several sources, aiming at to increase the reliability of the algorithms for the identification of ADLs with mobile devices, to reduce the effects of the incorrect data captured by the sensors, or to support the data processing [25].

Data fusion methods may be executed either on the mobile application, as a background process, or on a remote system. These methods may be categorized as probabilistic, statistic, knowledge base theory and evidence reasoning methods [26]. Although, the most used method for data fusion is the Kalman filter, which is a dynamically weighted recursive least-squares algorithm [27].

2.9 Data Mining, Pattern Recognition and Machine Learning techniques

This is the last step in the recognition of ADLs, comprising the identification of the main features of the collected data, which may include the peaks of acceleration, the standard deviation of the acceleration, the maximum and minimum values of acceleration, the amplitude of the audio data captured, and others explored during this PhD project.

The use of data mining, pattern recognition and machine learning techniques allows the classification of the sensors' data. Several studies have been carried out to identify ADLs [28-35].

3 Proposed Solution

The proposed solution in this PhD project consists on a multi-sensor mobile system in which several techniques are applied aiming to create new methods to identify ADLs with high accuracy.

3.1 Method Description

The proposed method consists on an architecture to support the development of applications for the monitoring and identification of ADLs. Mobile devices combined with mobile applications are able to acquire and process the sensors' data, and finally, show the obtained results.

This method includes different techniques for the accurately identification of ADLs. First, it makes use of the sensors in mobile devices to enable the capture of several physical and physiological parameters optimized with sensor data acquisition methods. Second, the data may be processed with lightweight methods implemented on a mobile application. However, some of the data captured are too complex to be processed locally, therefore they must be sent to a remote server, which carries out the data processing and returns the results back to the mobile device. However, these methods need a constant connection to the network. Besides, during the data acquisition and/or data processing stages, some failures may occur and data imputation and data cleaning techniques should be applied. Depending on the types of data captured, different data processing techniques must be applied. Currently, the case study in this PhD project is focused on the processing of audio for the identification of ADLs, applying audio fingerprinting methods. After that, data fusion techniques should be applied in order to merge all the features obtained by all the sensors available in the mobile device. After the data is fused and taking into account user's context, data mining, pattern recognition and/or machine learning techniques may be applied to create models for the identification of ADLs, which should be validated with the user's feedback.

3.2 Work in Progress

Until now, the state of the art about the roadmap of the proposed method has been already researched and published at [43], discovering the most and the best models used in other research studies.

At the same time of the capture of the data from the accelerometer, gyroscope, magnetometer, strength of the Wi-Fi signal and basic service set identification (BSSID) of the currently connected network, this project is aiming at creating a set of audio fingerprints for the identification of ADLs using the acoustic signal. The audio fingerprints will be based on the main features of the collected audio data using a mobile application specifically developed. This mobile application captures audio data, which is later stored in text files with signed or unsigned amplitude values, the accelerometer data, the values of the location based on a GPS receiver and the data related to the Wi-Fi connection, *e.g.,* the BSSID and the strength of the signal, which are also stored in text files for further analysis. At the end of this project, the data from all sensors are fused and the ADLs may be recognized with more accuracy.

The analysis of audio data for the identification of ADLs is now focused on four scenarios: watching TV, working at a software development company, sleeping and walking. We are collecting several samples for further analysis and creation of the metrics for audio fingerprints of each activity. In addition, the accelerometer, gyroscope, magnetometer, GPS, BSSID of Wi-Fi connected network and the strength of the Wi-Fi signal are captured for further analysis, in order to complement the recognition of the ADLs using the audio fingerprinting.

After the acquisition of a sample of audio data by the mobile device, it is processed to obtain the correspondent audio fingerprint that is then compared with the learned models based on the scenarios previously defined. However, each audio fingerprint obtained can match on several activities, thus distance metrics are applied. In addition, the use of multiple sensors for comparison are desirable, because it is expected to produce better results.

At the learning stage, the data should be captured by a smartphone and/or a smartwatch combined with the user's feedback and the users' agenda to create a set of signals' patterns and audio fingerprints with the characteristics of the data captured related to the different ADLs. After that, when a sample of the sensors' data is collected, it should be compared with the data stored in the collection, identifying the ADLs more accurately. Generally, the sensors' data related to the ADLs performed by a sedentary individual working at a software development company, includes the several ADLs, these are:

- between 0h00m and 8h00m, the activity is sleeping (during the working days);
- between 9h00m and 19h00m, the activity is working (during the working days, stopping the capture during the breaks);
- between 19h00m and 0h00m, the activity is watching TV (during the working days);
- between 15h00m and 19h00m, the activity is walking (during the weekend).

More samples should be acquired to compare and to create a large set of identifiable ADLs, in order to create the different methods to the development of a personal digital life coacher.

4 Discussion and Conclusions

The review of the state of the art related to part of the proposed method for this PhD project has been already published. It was focused on several concepts, such as the identification of ADLs, the existence of several mobile platforms, the diversity of the sensors available in off-the-shelf mobile devices, the concept of multi-sensor systems, the acquisition, processing, validation, cleaning, imputation, and fusion of the sensors' data, data analysis with pattern recognition techniques, and user's feedback.

The diversity of sensors available in mobile devices and the constraints related to the low resources of the mobile devices and the position of the sensors during the acquisition of the data are some challenges that this PhD project needs to address in order to create a robust method for the identification of the ADLs. The method that we expect to create during the project should be based on a subjective validation provided by the user combined with adjustable models in order to obtain better results, and statistically validated. Thus, several models should be implemented and tested for the different user's daily activities, such as: the k-NN for data imputation, and the Kalman filter for data fusion.

The research is now focused on the capture and process of audio data in order to create a set of consistent audio fingerprints for the identification of ADLs.

Acknowledgments

This work was supported by FCT project **UID/EEA/50008/2013** (*Este trabalho foi suportado pelo projecto FCT UID/EEA/50008/2013*).

The authors would also like to acknowledge the contribution of the COST Action IC1303 – AAPELE – Architectures, Algorithms and Protocols for Enhanced Living Environments.

References

1. Pires, I.M.S., *Aplicação móvel e plataforma Web para suporte à estimação do gasto energético em atividade física*, in *Informatics Department*. 2012, University of Beira Interior: Covilhã.
2. Pires, I.M., N.M. Garcia, and M.C. Canavarro Teixeira. *Calculation of Jump Flight Time using a Mobile Device*. in *Proceedings of the HEALTHINF 2015 8th International Conference on Health Informatics*. 2015. Lisbon, Portugal.
3. Pires, I.M., et al., *Measurement of Heel-Rise Test Results using a Mobile Device*, in *Proceedings of PhyCS 2015, 2nd International Conference on Physiological Computing Systems*. 2015: Angers, France.
4. Pires, I.M., N.M. Garcia, and F. Flórez-Revuelta. *Multi-sensor data fusion techniques for the identification of activities of daily living using mobile devices*. in *Proceedings of the ECMLPKDD 2015 Doctoral Consortium, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*. 2015. Porto, Portugal.
5. Garcia, N.M., *A Roadmap to the Design of a Personal Digital Life Coach*. 2016.
6. Morgan, S., *Activities of daily living*, in *Community Mental Health*. 1993, Springer Berlin Heidelberg. p. 141-158.
7. Kuspa, K. and T. Pratkanis, *Classification of Mobile Device Accelerometer Data for Unique Activity Identification*. 2013.
8. Kwapisz, J.R., G.M. Weiss, and S.A. Moore, *Activity recognition using cell phone accelerometers.* ACM SIGKDD Explorations Newsletter, 2011. **12**(2): p. 74.
9. Siirtola, P. and J. Röning, *Recognizing Human Activities User-independently on Smartphones Based on Accelerometer Data.* International Journal of Interactive Multimedia and Artificial Intelligence, 2012. **1**(5): p. 38.
10. Dernbach, S., et al. *Simple and Complex Activity Recognition through Smart Phones*. in *Intelligent Environments (IE), 2012 8th International Conference on*. 2012. Guanajuato, Mexico: IEEE.
11. Salazar, L.H.A., et al., *A Systematic Literature Review on Usability Heuristics for Mobile Phones.* International Journal of Mobile Human Computer Interaction, 2013. **5**(2): p. 50-61.
12. White, R.M., *A Sensor Classification Scheme.* Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on, 1987. **34**(2): p. 124-126.
13. Xu, R. and L. He, *GACEM: Genetic Algorithm Based Classifier Ensemble in a Multi-sensor System.* Sensors, 2008. **8**(10): p. 6203-6224.

14. Mo, L., et al., *Multi-Sensor Ensemble Classifier for Activity Recognition.* Journal of Software Engineering and Applications, 2012. **05**(12): p. 113-116.
15. Scalvini, S., et al., *Information and communication technology in chronic diseases: a patient's opportunity.* Journal of Medicine and the Person, 2013. **12**(3): p. 91-95.
16. Bersch, S.D., et al., *Sensor data acquisition and processing parameters for human activity classification.* Sensors (Basel), 2014. **14**(3): p. 4239-70.
17. Lim, L., A. Misra, and T. Mo, *Adaptive data acquisition strategies for energy-efficient, smartphone-based, continuous processing of sensor streams.* Distributed and Parallel Databases, 2012. **31**(2): p. 321-351.
18. Dolui, K., S. Mukherjee, and S.K. Datta, *Smart Device Sensing Architectures and Applications.* 2013 International Computer Science and Engineering Conference (Icsec), 2013: p. 91-96.
19. Sun, S., et al., *Literature review for data validation methods.* Science and Technology, 2011. **47**(2): p. 95-102.
20. Garza-Ulloa, J., H. Yu, and T. Sarkodie-Gyan, *A mathematical model for the validation of the ground reaction force sensor in human gait analysis.* Measurement, 2012. **45**(4): p. 755-762.
21. Vateekul, P. and K. Sarinnapakorn, *Tree-Based Approach to Missing Data Imputation*, in *Data Mining Workshops, 2009. ICDMW '09. IEEE International Conference on* 2009, IEEE: Miami, FL. p. 70-75.
22. Patrician, P.A., *Multiple imputation for missing data.* Res Nurs Health, 2002. **25**(1): p. 76-84.
23. Ling, W. and F. Dong-Mei, *Estimation of Missing Values Using a Weighted K-Nearest Neighbors Algorithm*, in *Environmental Science and Information Application Technology, 2009. ESIAT 2009. International Conference on*. 2009, IEEE: Wuhan. p. 660-663.
24. Jeffery, S.R., et al., *Declarative Support for Sensor Data Cleaning.* 2006. **3968**: p. 83-100.
25. Banos, O., et al., *On the use of sensor fusion to reduce the impact of rotational and additive noise in human activity recognition.* Sensors (Basel), 2012. **12**(6): p. 8039-54.
26. Pombo, N., et al., *Medical decision-making inspired from aerospace multisensor data fusion concepts.* Inform Health Soc Care, 2015. **40**(3): p. 185-97.
27. Tanveer, F., O.T. Waheed, and Atiq-ur-Rehman, *Design and Development of a Sensor Fusion based Low Cost Attitude Estimator.* Journal of Space Technology,, 2011. **1**(1): p. 45-50.
28. Bujari, A., B. Licar, and C.E. Palazzi. *Movement pattern recognition through smartphone's accelerometer*. in *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*. 2012. Las Vegas, NV: IEEE.
29. Chikhaoui, B., S. Wang, and H. Pigot. *A Frequent Pattern Mining Approach for ADLs Recognition in Smart Environments*. in *Advanced Information Networking and Applications (AINA), 2011 IEEE International Conference on*. 2011. Biopolis: IEEE.
30. Kouris, I. and D. Koutsouris, *A comparative study of pattern recognition classifiers to predict physical activities using smartphones and wearable body sensors.* Technol Health Care, 2012. **20**(4): p. 263-75.
31. Zhu, C. and W. Sheng, *Realtime recognition of complex human daily activities using human motion and location data.* IEEE Trans Biomed Eng, 2012. **59**(9): p. 2422-30.
32. Farrahi, K. and D. Gatica-Perez, *Daily Routine Classification from Mobile Phone Data*, in *Machine Learning for Multimodal Interaction*. 2008, Springer Berlin Heidelberg. p. 173-184.
33. Hong, J.-H., et al., *An Activity Recognition System for Ambient Assisted Living Environments*, in *Evaluating AAL Systems Through Competitive Benchmarking*. 2013, Springer Berlin Heidelberg. p. 148-158.
34. Phithakkitnukoon, S., et al., *Activity-Aware Map: Identifying Human Daily Activity Pattern Using Mobile Phone Data*, in *Human Behavior Understanding*. 2010, Springer Berlin Heidelberg. p. 14-25.
35. Nam, Y., S. Rho, and C. Lee, *Physical activity recognition using multiple sensors embedded in a wearable device.* ACM Transactions on Embedded Computing Systems, 2013. **12**(2): p. 1-14.