Evaluating Human Activity Recognition Systems for AAL Environments

Conference Paper · January 2013

DOI: 10.1007/978-3-642-37419-7_10

1 author:

Juan Antonio Alvarez-Garcia
Universidad de Sevilla

Evaluating Human Activity Recognition Systems for AAL Environments

Juan Antonio Álvarez-García *

Computer Languages and Systems Department, University of Seville, Spain
jaalvarez@us.es

Abstract. EvAAL Activity Recognition track's main goal is to evaluate one of the pillars of Ambient Assisted Living (AAL): human activity recognition (AR). In this edition 4 teams from United States, Ireland, Spain and Japan participated in the competition. Results show that accelerometer based solutions are promising due to their small size and their integration in complex devices such as mobile phones or elastics wearable straps.

Keywords: AAL, activity recognition, tracking

1 Introduction

Activity Recognition (AR) is a research area where the objective is to recognize human activities. The automatic and unobtrusive identification of users activities is one of the challenging goals of context-aware computing [1] and is expected to be a practical solution to monitor aged people: According to UNFPA[2] although currently only Japan has an older population of more than 30 per cent, by 2050, 64 countries are expected to join Japan.

Our objective is to measure AR hardware and software performance through a competition. Since competitions in computer science have a long-established tradition, the design of EvAAL has been inspired by other successful competitions. Some past and current competitions have been analyzed in order to identify successful practices, specifically: HARL [3],OPPORTUNITY [4], HASC [5] and BSN Contest [6]. The main difference with the preceding ones is that this track addresses both software and hardware so that whole state of the art of activity recognition can be examined.

The main objective of this competition is to implement an activity recognition system (ARS) that recognizes the following activities: lie, sit, stand, walk, bend, fall and cycle (using a stationary bike).

The competition took place from 9 to 13 July, 2012 at the CIAmI Living Lab [7] in Valencia (Spain). In this track there is no limitation to the number of devices that can be used and competing solutions can be based on a variety

* This work was supported in part by the European Commission in the framework of the FP7 project universAAL under Contract 247950.

of sensors and technologies, including: accelerometers, gyroscopes, magnetometers, pressure sensors, microphones, sensor networks, mobile phones, cameras, etc.. Other technologies or combinations of them are also considered acceptable provided they are compatible with the constraints of the hosting Living Lab. Figure 1 shows the CIAmI map.

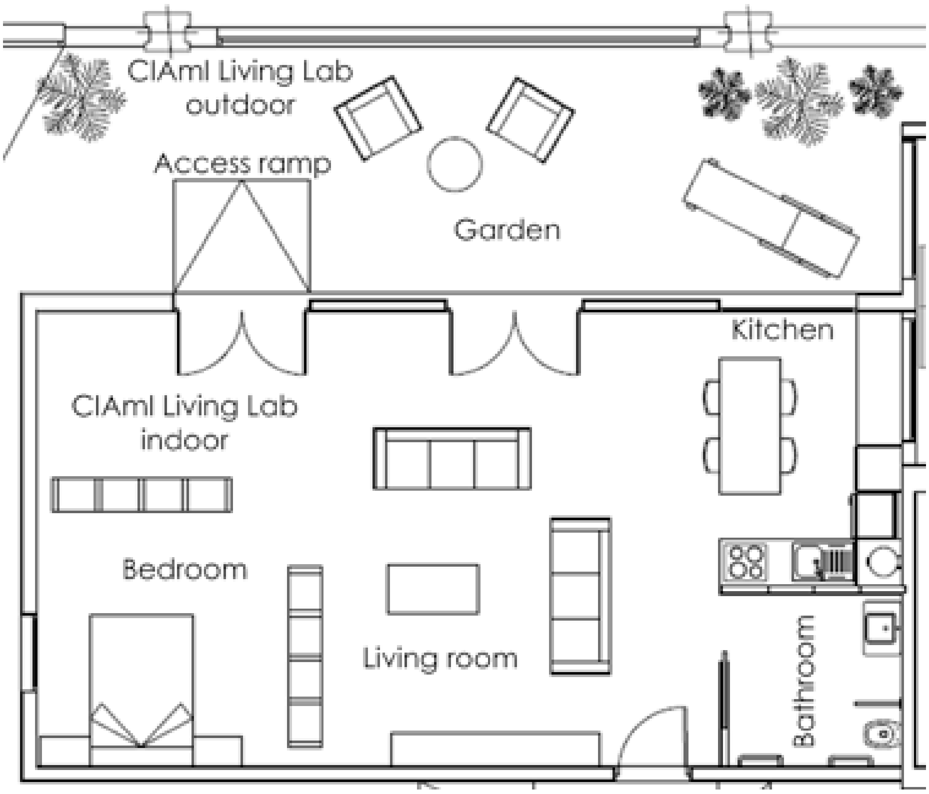

Fig. 1. Map from CIAmI Living Lab.

2 Benchmarks

Competitors are invited to install and run their ARS during a time slot, in this case for two hours and a half, divided in three subslots: Installation, Benchmark and Removal phases.

During the second phase, the ARS are evaluated. An actor (an evaluation committee member) performs a predefined physical activity trip across the smart home. Audio signals synchronize the actor movements in each performance (twice per team) in order to get the same ground truth for all the participants. The path followed by the actor and the activities are the same for each performance, and they were not disclosed to competitors before the application of the benchmarks. Similarly, the position of the stationary bike and the place of the fall are not revealed either.

Once the two performances are executed, the one with better overall mark is used to compare each team.

A critical issue (that was communicated in advance to the competitors) is the age of the actor, that is required in order to train and prepare their algorithms properly. The actor was also trained to repeat the activities in the benchmarks always in the same way (following the mp3 file explained in Section 3). The fall is also critical, because many different kind of falls are possible [8]. For this reason we published in advance a video of the fall that had been executed by the actor during the experiment. A recorded performance can be seen in this video[1].

3 Reference Localization System

The reference ARS is used to obtain the ground truth data. In order to get approximately the same ground truth for all the contestants, audio signals were used to synchronize the actor movements. An mp3 file indicates the next activity that the actor must perform and a countdown ("three, two, one, now") to perform it. When the actor hears the word "now", he begins the transition to the next activity. In some cases, such as BSN Contest, researchers identify transitions to recognize the next activity. In our case with 7 activities, the number of possible transitions (some of them not very probable) is 42. Since the number of transitions is high and it is not trivial to evaluate them, we decided not to evaluate the transitions but only the activities.

To retrieve competitor and ground truth data, a local server accepts sockets with the activity code and the time when it is identified. A competitor system only need to send this information to the server. A local NTP server is also available to synchronize the time. To obtain ground truth data, an evaluator uses an Android application to mark the activities (activity code) or transitions (-1 code) through his mobile phone and to send this information to the socket server.

To this purpose, the evaluator follows the actor to see exactly when he starts the activities or transitions. For instance, if the actor is standing still and hears "Cycling 3, 2, 1, now!" he begins to hold the handlebars and move his leg to go up the bike. When the word "now" is spelled, the evaluator pushes the button in the application to identify the end of previous activity and the beginning of a transition. The evaluator later on pushes again the button to mark the end of the activity.

[1] http://vimeo.com/52843550

4 Evaluation Criteria

The following criteria and weights were used to evaluate the system of every team:

Accuracy (0.25) It evaluates the recognized activity instances (500 ms) using F-measure $\frac{2*precision*recall}{precision+recall}$ to compute it.

Recognition delay (0.2) It refers to the elapsed time between the instant in which the user begins an activity and the time in which the system recognizes it.

Installation complexity (0,15) It is a measure of the effort required to install the ARS in a flat, measured by the evaluation committee as a function of the person-minutes of work needed to complete the installation.

User Acceptance (0.25) It captures how much invasive the ARS is in the users daily life and thereby the impact perceived by the user; this parameter is evaluated by the Evaluation Committee using a questionnaire.

Interoperability with AAL systems (0,15) It is evaluated using a questionnaire that measure the use of open source solutions, use of standards, availability of libraries for development and integration with standard protocols.

5 Contestants and results

After peer review, five teams were accepted but one of them withdrew due to financial cutbacks of its institution. Hence only four competitors participated in the challenge, CUJ (from Chiba University, Japan), CMU (from Carnegie Mellon and Utah Universities, USA), DCU (from Dublin City University, Ireland) and USS (from University of Seville, Spain). The technology used by every team is described here:

5.1 DCU Team

Title of paper Visual Experience for Recognising Human Activities

List of Authors and affiliations Na Li, Martin Crane, Heather Ruskin from Centre for Scientific Computing and Complex Systems Modelling, School of Computing, Dublin City University.

Brief description DCU Team uses a SenseCam[2] hanging from the actor's neck to evaluate off-line the activities. The DCU system worked offline and therefore no accuracy and delay score was assigned.

5.2 CUJ Team-Bronze

Title of paper Human Behavior Recognition by a Mobile Robot Following Human Subjects.

[2] http://research.microsoft.com/en-us/um/cambridge/projects/sensecam/

List of Authors and affiliations Myagmarbayar Nergui[1], Nevrez Imamoglu[1], Yuki Yoshida[1], Jose Gonzalez[2], Masashi Sekine[2], Kazuya Kawamura[1] and Wenwei Yu[1].

1. Medical System Engineering Department, Graduate School of Engineering, Chiba University.
2. Research Center for Frontier Medical Engineering, Chiba University.

Brief description CUJ Team uses a robot (initially a Pioneer 3-AT[3] but due to difficulties in transport from Japan, a Roomba[4] was used instead) with two kinects[5]. The first Kinect is used to avoid obstacles when following the actor and the other to recognize their activities. After the Evaal experience CUJ decided to use only one Kinect.

5.3 CMU Team-Silver

Title of paper An Activity Recognition System for Ambient Assisted Living Environments

List of Authors and affiliations Jin-Hyuk Hong, Julian Ramos, Choonsung Shin, Anind Dey from Human-Computer Interaction Institute Carnegie Mellon University.

Brief description CMU Team proposes a solution composed by three subsystems: A chest wearable elastic strap[6] capable of measuring several physiological signals, an Android mobile phone and a system for indoor localization based on Radio tomographic imaging[7].

5.4 USS Team-Gold

Title of paper Activity recognition system using AMEVA method

List of Authors and affiliations Luis Miguel Soria[1], Luis Gonzalez-Abril[2], Miguel Ángel Álvarez de la Concepción[1],, Juan Antonio Ortega Ramírez[1]

1. Computer Languages and Systems Department, University of Seville.
2. Applied Economics I Department, University of Seville.

Brief description USS Team uses an android mobile phone placed on the right hip. The user activities are recognized by means of the accelerometer embedded in the mobile phone.

5.5 Results

Table 1 shows the results for each team and criteria for the best of both performances. The winner (USS team) obtained acceptable results in performance, but its simplicity (although it uses multiple mathematical methods it only rely on accelerometers) and interoperability give good marks in all the evaluated criteria. Gold and Silver teams used accelerometer-based solutions.

[3] http://www.mobilerobots.com/researchrobots/p3at.aspx
[4] http://store.irobot.com/family/index.jsp
[5] http://en.wikipedia.org/wiki/Kinect
[6] http://www.zephyr-technology.com/bioharness-bt
[7] http://span.ece.utah.edu/radio-tomographic-imaging

Table 1. Best performance result

Team	Accuracy	Delay	Installation	User Acceptance	Interoperability	Final Score
USS	4,33	9	10	7,47	7,63	7,3945
CMU	7,17	9	0	7,93	6,15	6,4975
CUJ	2,59	2	0	5,6	5,09	3,5235
DCU	0	0	10	5,2	1,25	2,9875

6 Conclusions

Next edition some improvements may regard the kind of activities (some more complex activities can be taken into account) and the possibility to simulate an aged actor using specifics kits to avoid rapid movements. Results suggest that there is still space for other editions of this competition in the future, and encourage us to continue in this track for the next year.

References

1. Brush, A.B., Krumm, J., Scott, J., Saponas, S.: Recognizing activities from mobile sensor data: Challenges and opportunities. In: Ubicomp' 11. (2011)
2. United Nations Population Fund and HelpAge International, ed.: Ageing in the Twenty-First Century: A Celebration and A Challenge. United Nations Population Fund (December 2012)
3. Wolf, C., Mille, J., Lombardi, L., Celiktutan, O., Jiu, M., Baccouche, M., Dellandrea, E., Bichot, C.E., Garcia, C., Sankur, B.: The liris human activities dataset and the icpr 2012 human activities recognition and localization competition. Technical Report RR-LIRIS-2012-004, LIRIS Laboratory (March 2012)
4. Sagha, H., Digumarti, S.T., del R. Millán, J., Chavarriaga, R., Calatroni, A., Roggen, D., Tröster, G.: Benchmarking classification techniques using the opportunity human activity dataset. In: Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on. (October 2011) 36 –40
5. Kawaguchi, N., Ogawa, N., Iwasaki, Y., Kaji, K., Terada, T., Murao, K., Inoue, S., Kawahara, Y., Sumi, Y., Nishio, N.: Hasc challenge: Gathering large scale human activity corpus for the real-world activity understandings. In: ACM AH 2011. (2011)
6. Vitali, L., Roozbeh, J.: Power aware wireless data collection for bsn data repositories. In: International Conference on Body Sensor Networks. (2011)
7. Martínez, A., Llorente, M.A., Lázaro, J.P.: Ciami living lab: an economically sustainable technological tool for open innovation. In: I ENoLL Living Lab Summer School. (2010)
8. Noury, N., Fleury, A., Rumeau, P., Bourke, A.K., ÓLaighin, G., Rialle, V., Lundy, J.: Fall detection - principles and methods. In: Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2007). (2007) 1663–1666