A Multi-modal Sensing Framework for Human Activity Recognition

Barbara Bruno[1], Jasmin Grosinger[2], Fulvio Mastrogiovanni[1], Federico Pecora[2], Alessandro Saffiotti[2], Subhash Sathyakeerthy[2], and Antonio Sgorbissa[1]

[1] University of Genova, Dept. DIBRIS,
Via Opera Pia 13, 16145 Genova, Italy
{barbara.bruno,fulvio.mastrogiovanni,antonio.sgorbissa}@unige.it
[2] Örebro University, AASS Cognitive Robotic Systems Lab,
Fakultetsgatan 1, S-70182 Örebro, Sweden
{jasmin.grosinger,fpa,asaffio,subhash.sathyakeerthy}@aass.oru.se

Abstract. Robots for the elderly are a particular category of home assistive robots, helping people in the execution of daily life tasks to extend their independent life. Such robots should be able to determine the level of independence of the user and track its evolution over time, to adapt the assistance to the person capabilities and needs. We present an heterogeneous information management framework, allowing for the description of a wide variety of human activities in terms of multi-modal environmental and wearable sensing data and providing accurate knowledge about the user activity to any assistive robot.

1 Introduction

Home assistive robotics addresses the design of robots to be deployed in domestic environments, to assist the residents in the execution of daily life tasks. Robots for the elderly are a particular category of home assistive robots, which rely on social interaction with the person and aim at extending the elderly independent life [1]. To properly and effectively perform the assistive duties, robots for the elderly should be *context-aware*, i.e., able to assess the status of the environment they are in, *user-aware*, i.e., able to assess the status of the person they are working for, and also *ageing-aware*, i.e., able to perform a long term analysis of the person cognitive and physical evolution, to adapt to their current capabilities.

Human Activity Recognition (HAR) systems for elderly-care are devoted to the identification, among all actions executed by a person during a day, of specific activities of interest, the Activities of Daily Living (ADL), which require the use of different cognitive and physical abilities and are used by gerontologists to estimate the level of autonomy of a person [2]. ADL cover a wide variety of human activities: consequently, a number of sensing strategies have been developed for their automatic recognition. ADL occurring at home, in particular, are usually monitored with smart environments [3] and wearable sensing systems [4]. Unfortunately, wearable sensing systems are prone to ambiguity, while smart environments may reach erroneous conclusions due to incomplete information.

We address the problem of endowing robots for the elderly with the ability of monitoring the Activities of Daily Living, by designing a HAR system which allows for a seamless integration with the robot planning system. We propose the integration of multiple sensing strategies in a single framework, to compensate the weaknesses of each modality and increase the recognition reliability.

The abstract is organized as follows. Section 2 details the system architecture. Preliminary experimental results are analysed in Section 3. Conclusions follow.

2   System Architecture

We set up a test bed in an apartment located in the city of Örebro (SWE), in the elderly care facility Ängen. The apartment, shown in Figure 1, is composed of fully furnished living-room, bathroom, bedroom and kitchen.
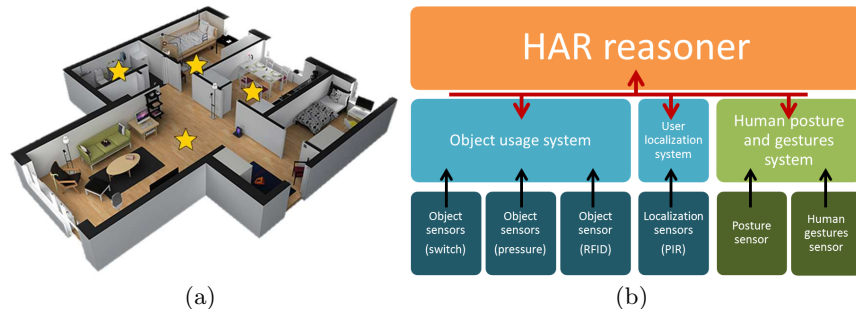

(a)   (b)

Fig. 1. (a) The test bed apartment in the elderly care facility Ängen, in Örebro (SWE). (b) WearAmI system architecture: dark boxes at the bottom denote the adopted sensors. Shades of blue represent environmental sensing components, while shades of green represent wearable sensing components.

We propose the multi-modal monitoring system with the architecture shown in Figure 1 for the reliable detection of the activities: *transferring* (denoting the motions of sitting down, standing up, lying down, getting up); *feeding* (eating, drinking); *food preparation*; *indoor transportation* (climbing stairs, descending stairs, walking). The system makes use of: (i) a wrist-placed inertial sensor; (ii) a waist-placed inertial sensor; (iii) a network of Passive Infra-Red sensors; (iv) RFID tags, pressure sensors and switches; and (v) a temporal reasoner.

The data extracted by the wrist sensor are used to detect occurrences of gestures [5, 6], such as *walking, picking up* or *sitting.* The data provided by the waist sensor, instead, are used to estimate the person posture on the basis of the angle between the torso and the gravity force. The combined analysis of wrist and waist acceleration data also allows for detecting falls with high accuracy [7]. Person localization is achieved via a network of Passive Infra-Red (PIR) sensors. We identified three categories of objects to monitor: *cutlery and dishes*, which are

assumed to be in use when located on the kitchen table and that we detect via an RFID network; *furniture*, such as chairs, armchairs and bed, for which pressure sensors detect whether and which is in use; *household appliances*, such as the fridge and the oven, whose usage can be inferred by checking the status of their doors with switches. All elements in the architecture are envisioned as Physically Embedded Intelligent Systems (PEIS) [8], i.e., devices incorporating computational, communication, sensing and/or actuating resources, connected with each other by a uniform communication model. The analysis systems (focusing on objects usage, user localization and user posture & gestures, respectively) share information among each other and with a reasoning system which is responsible for the recognition of all occurrences of activities of interest. The adopted temporal reasoner uses and extends Allen's interval algebra to model the activities as sets of temporal constraints [9].

3   Experimental Evaluation

Listing 1.1 reports the models of *sitting* and *standing*. The field `Head` defines the entity it refers to and the name of the model, separated by a `::`. As an example, `Head Human::SitDown()` indicates that whenever the reported constraints are satisfied, the reasoner should infer that the activity of *sitting* has been executed by the person. The field `RequiredState` defines the sensor values which correspond to the execution of the motion. The field `Constraint` defines the temporal relation between each sensor value of interest and the activity.

Listing 1.1. DDL models for the activities *sit down* and *stand up*.

```
( SimpleOperator
 ( Head  Human : : SitDown ( ) )
 ( RequiredState  req1  Gesture : : Sit ( ) )
 ( RequiredState  req2  Posture : : Sitting ( ) )
 ( RequiredState  req3  Chair : : On ( ) )
 ( Constraint  OverlappedBy ( Head , req1 ) )
 ( Constraint  During ( Head , req2 ) )
 ( Constraint  EndEnd ( Head , req3 ) )
)

( SimpleOperator
 ( Head  Human : : SitDown ( ) )
 ( RequiredState  req1  Gesture : : Sit ( ) )
 ( RequiredState  req2  Posture : : Sitting ( ) )
 ( RequiredState  req3  Armchair : : On ( ) )
 ( Constraint  OverlappedBy ( Head , req1 ) )
 ( Constraint  During ( Head , req2 ) )
 ( Constraint  EndEnd ( Head , req3 ) )
)
```

```
( SimpleOperator
  ( Head  Human : : StandUp ( ) )
  ( RequiredState  req1  Gesture : : Stand ( ) )
  ( RequiredState  req2  Posture : : Standing ( ) )
  ( RequiredState  req3  Human : : SitDown ( ) )
  ( Constraint  MetByOrOverlappedBy ( Head , req1 ) )
  ( Constraint  Starts ( Head , req2 ) )
  ( Constraint  MetByOrAfter ( Head , req3 ) )
)
```
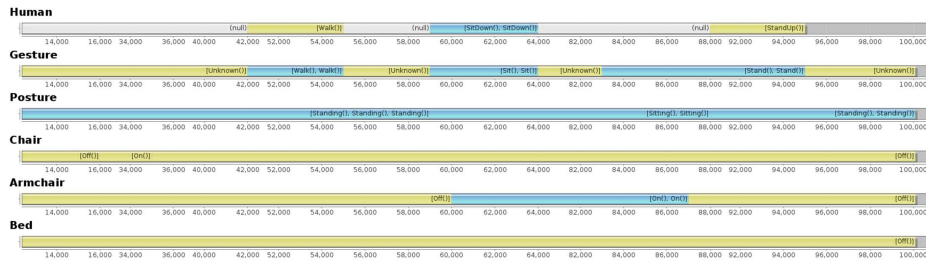

Fig. 2. Validation of the models of *sit down* and *stand up*.

We performed preliminary tests defining sequences of sensor values to analyse the reasoner inferences they trigger. In Figure 2, the timeline of the context variable Human is computed by the reasoner and list all corresponding recognized activities, as indicated by the Head fields. The other timelines report the sensor values (i.e., gesture, posture, location and objects sensors). At each time instant, the reasoner samples the sensors, keeping track of all modelled activities which are consistent with the sensors readings up to that instant (i.e., those that *could be* the one currently being executed). As time passes, the number of possible activities progressively reduces, until it converges to the one effectively performed, if it is among the modelled ones, or to none. All sensors or context variables statuses supporting an inferred activity are marked with a blue filling.

Figure 2 reports the simulated sensor readings related to a person who drops a heavy bag on the kitchen chair, then walks to the living-room and sits on the armchair. Although the chair pressure sensor is activated by the bag (for $t = [16; 35]$), the wearable gesture and posture sensors do not signal any sitting motion, therefore preventing the reasoner from making an erroneous inference. Later on, when the person sits on the armchair, environmental and wearable sensors agree on indicating that the person sat down, therefore triggering the correct recognition of the *sitting* motion. The example also highlights one advantage deriving from overloading rules. Since the two modelled sitting actions (i.e., sitting down on the kitchen chair and sitting down on the armchair) are defined as SitDown, it is possible to define a single model for the standing up motion, constrained by the previous occurrence of *any* sitting action.

4   Conclusions

In this abstract, we introduce the idea of a multi-modal monitoring system, which combines information retrieved via different monitoring approaches and we prove, in simulation, that the integration of wearable and environmental information is beneficial for the purposes of human activity monitoring. Future work will focus on the set up of a test bed apartment in an elderly care facility in Örebro, Sweden, to test the performance of the system under realistic conditions.

References

1. Johnson, D.O., Cuijpers, R.H., Juola, J.F., Torta, E., Simonov, M., Frisiello, A., Bazzani, M., Yan, W., Weber, C., Wermter, S., Meins, N., Oberzaucher, J., Panek, P., Edelmayer, G., Mayer, P., Beck, C.: Socially Assistive Robots: A Comprehensive Approach to Extending Independent Living. International Journal of Social Robotics 6,2, 195–211 (2014)
2. Katz, S., Chinn, A., Cordrey, L.: Multidisciplinary studies of illness in aged persons: a new classification of functional status in activities of daily living. J. Chron. Dis. 9,1, 55–62 (1959)
3. Alam, M.R., Reaz, M.B.I., Ali M.A.M.: A Review of Smart HomesPast, Present, and Future. IEEE Trans on Systems, Man, and Cybernetics, Part C: Applications and Reviews 42,6, 1190–1203 (2012)
4. Lara O.D., Labrador, M.A.: A Survey on Human Activity Recognition using Wearable Sensors. IEEE Communications Surveys and Tutorials, 15,3, 1192–1209 (2013)
5. Bruno, B., Mastrogiovanni, F., Sgorbissa, A., Vernazza, T., Zaccaria, R.: Human motion modelling and recognition: A computational approach. In: IEEE Int Conf on Automation Science and Engineering (CASE), 156–161 (2012)
6. Bruno, B., Mastrogiovanni, F., Sgorbissa, A., Vernazza, T., Zaccaria, R.: Analysis of human behavior recognition algorithms based on acceleration data. In: IEEE Int Conf on Robotics and Automation (ICRA), 1602–1607 (2013)
7. Kangas, M., Konttila, A., Lindgren, P., Winblad, I., Jamsa, T.: Comparison of low-complexity fall detection algorithms for body attached accelerometers. Gait & Posture 28, 285–291 (2008)
8. Saffiotti, A., Broxvall, M., Gritti, M., LeBlanc, K., Lundh, R., Rashid, J., Seo, B.S., Cho, Y.J.: The PEIS-ecology project: vision and results. In: IEEE/RSJ Int Conf on Intelligent Robots and Systems (IROS), 2329–2335 (2008)
9. Pecora, F., Cirillo, M., Dell'Osa, F., Ullberg, J., Saffiotti, A.: A constraint-based approach for proactive, context-aware human support. Journal of Ambient Intelligence and Smart Environments 4, 347–367 (2012)