Assessing computer systems for monitoring elderly people living at home

Nadia ZOUBA[1], François BREMOND[1], Monique THONNAT[1], Alain ANFOSSO[2], Éric PASCUAL[2], Patrick MALLEA[3], Véronique MAILLAND[3], and Olivier GUERIN[3]

[1] INRIA Sophia-Antipolis, 2004 route des Lucioles, 06902 Sophia Antipolis Cedex, France.
[2] CSTB Sophia Antipolis, 290 route des Lucioles, 06904 Sophia Antipolis Cedex, France.
[3] CHU Nice, Hôpital Cimiez, 4 avenue Reine Victoria, 06003 Nice Cedex 1, France.

Abstract: The elderly population is expected to grow dramatically over the next 20 years. The number of people requiring care will grow accordingly, while the number of people able to provide this care will decrease. Without receiving sufficient care, elderly are at risk of loosing their independence. Thus a system permitting the elderly person to live safely at home is more than needed.
In this context, we propose an automatic monitoring system which consists in analysing human behaviours and looking for changes in their activities. The system includes detecting people, tracking people as they move, recognizing events of interest based on multi-sensor analysis, and human activity recognition.

A laboratory (called GERHOME) equipped with many sensors and composed of four rooms (a kitchen, a living-room, a bedroom, and a bathroom) has been built to evaluate the performance of the monitoring system and to explore the ADLs (Activities of Daily Living) that can be recognized by such a computer system. While evolving in the laboratory, fourteen volunteers (aged from 60 to 85 years) have been observed during 4 hours to measure the accuracy of the detected events recognized by the monitoring system. This paper describes the experimentation and tries to highlight what are the capabilities of a computer system to assess automatically the frailty of elderly people in order for them to have a better and more secure life at home.

Keywords: *video event, monitoring system, ADLs, activity recognition.*

Introduction

In order to improve elderly life conditions at home and to reduce the costs of long hospitalizations, the medical world is more and more interested in telemonitoring techniques. These techniques will allow elderly people to stay safely at home, to benefit from an automated medical supervision and will delay their entrance in nursing homes. Medical professionals believe that one of the best ways to detect emerging physical and mental health problems, before it becomes critical (particularly for the elderly), is analysing the human behaviour and looking for changes in the activities of daily living (ADLs).

For this purpose, the proposed monitoring system aims at building an innovative framework for modelling activities of daily living (ADLs) at home. In particular, the goal is to collect and combine multi-sensor information to detect activities and assess behavioural trends to provide user services at different levels: (i) the elderly at home will be supported in handling and preventing critical incidents, and (ii) the care provider will be able to provide better support via the monitoring of health trends. To reach this goal, we have designed a video understanding platform described in [1], which includes an event recognition algorithm extended from [2].

Previous work

Healthcare technology for the elderly is a popular area of research. This technology represents a sub-discipline of “gerontechnology” [3]. Automatic monitoring of ADLs has been a common focus in gerontechnology. Over the last several years much effort has been put into developing and employing a variety of sensors to monitoring activities at home, including camera networks for people tracking [4], cameras and microphones for activity recognition [5, 6], and embedded sensors (e.g., with switch sensors [7, 8]). Other research work detects activities using sensors that are placed on the body (e.g. [9]). Posture, for example, can be detected automatically from accelerometers, as can some types of physical activity such as walking, sitting, standing, lying down, and bending. Recently, Jesse Hoey et al. [10] successfully used only cameras to assist person with dementia during hand-washing. The system uses only video inputs, and combines a Bayesian sequential estimation framework for tracking hands and towel, with a decision using a partially observable Markov decision process. Further, recent work suggests that many sensors placed throughout a home environment in combination with a few sensors worn on the body (e.g. [11]) may permit a system to automatically recognize everyday activities.

Approach overview

The approach consists in collecting multi-sensor data of the elderly person at home in order to build up a "normal" profile of his/her daily activity patterns (e.g. use the refrigerator, prepare a meal, sitting on a chair, go to bed). Large deviations from this profile should alert a human operator. To detect activity in a house, a number of sensing devices have been studied: video cameras and environmental sensors embedded in the home infrastructure (e.g. contact sensors, presence sensors, pressure sensors, water consumption sensors).

As described in figure1, the proposed monitoring system processes in real time three types of input: (1) video stream(s) acquired by video camera(s), (2) data resulting from environmental sensors, and (3) a priori knowledge containing a 3D model of a person (e.g. 3D size), event models and a description of the static environment including its 3D geometric information (e.g. position and size of furniture, zones of interest) and semantic information (e.g. people are expected to sit on a chair). The output of the system is the set of recognized events at each instant.

The monitoring system is composed of three components:

(1) A video analysis component: detects and tracks people evolving in the scene, recognizes both the posture of the person and a set of simple video events (e.g. the person leaves the kitchen).

(2) An environmental sensor analysis component: collects information about interactions between people and the contextual objects and recognizes a set of simple environmental events (e.g. the fridge is open).

(3) A multimodal activity recognition component: recognizes complex activities by combining video events and environmental events (e.g. person prepares a meal).

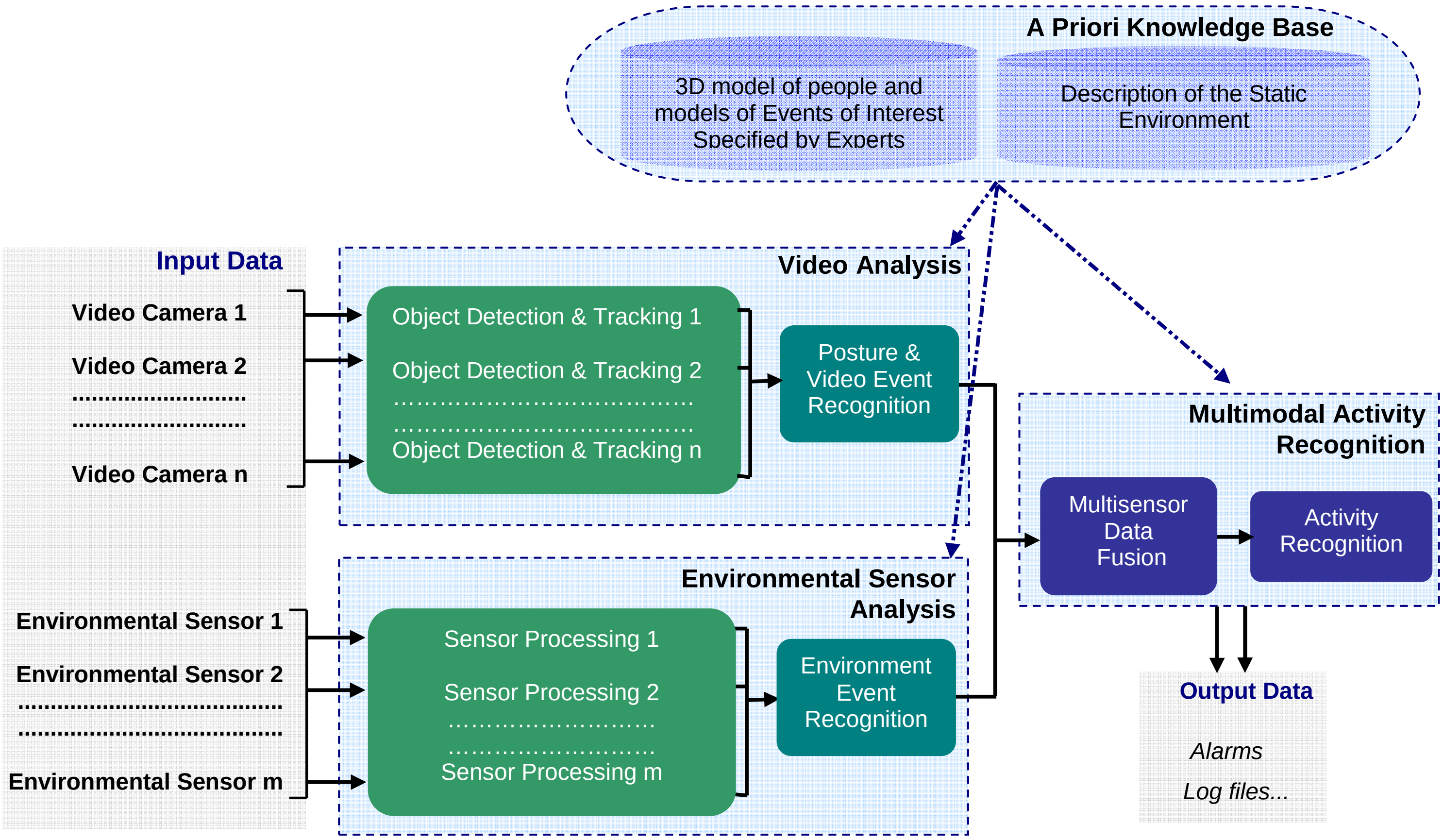

Fig.1. Architecture for the monitoring system

Data analysis

- **Person detection and tracking**

For detecting and tracking person we use a set of vision algorithms coming from a video interpretation platform described in [1]. A first algorithm segments moving pixels in the video into a binary image by subtracting the current image with the reference image. The reference image is updated along the time to take into account changes in the scene (light, object displacement, shadows). The moving pixels are then grouped into connected regions, called blobs. A set of 3D features such as 3D position, width and height are computed for each blob. Then the blobs are classified into predefined classes (e.g. person). After that the tracking task associates to each new classified blob a unique identifier and maintains it globally throughout the whole video. Once a person evolving in the scene is correctly detected by the video camera, a human posture recognition algorithm [12] recognizes in real time a set of human postures by using a 3D geometrical human model.

- **Activity recognition**

Activities at home (corresponding to ADLs) which are studied in this experimentation include using the fridge, turning on a microwave, preparing meal and so on. Critical situations are limited for the time being to falling and fainting.

The activity models are defined using a generic event description language [2]. Four types of video event have been defined: primitive states, composite states, primitive events and composite events. A state describes a stable situation in time characterizing one or several physical objects. A primitive state (e.g. a person is located inside a zone) corresponds to a perceptual property directly computed by the vision/environmental components. A composite state is a combination of primitive states. An event is an activity containing at least a change of state values between two consecutive times. A primitive event corresponds to a significant change of primitive state values (e.g. a person enters a zone). A composite event is a combination of primitive states and/or primitive events (e.g. preparing meal). An event model is composed of five parts:

- The "physical objects" involved in the event (e.g. person, equipment, zones of interest).
- The event "components" corresponding to the sub-events composing the event.
- The "forbidden components" corresponding to the events which should not occur during the main event.
- The "constraints" which are conditions between the physical objects and/or the components (constraints can be logical, spatial or temporal including the temporal Allen's algebra [13]).
- The "alerts" describing the actions to be taken when the event is recognized.

Figure 2 shows the video and environmental sensors analysis.

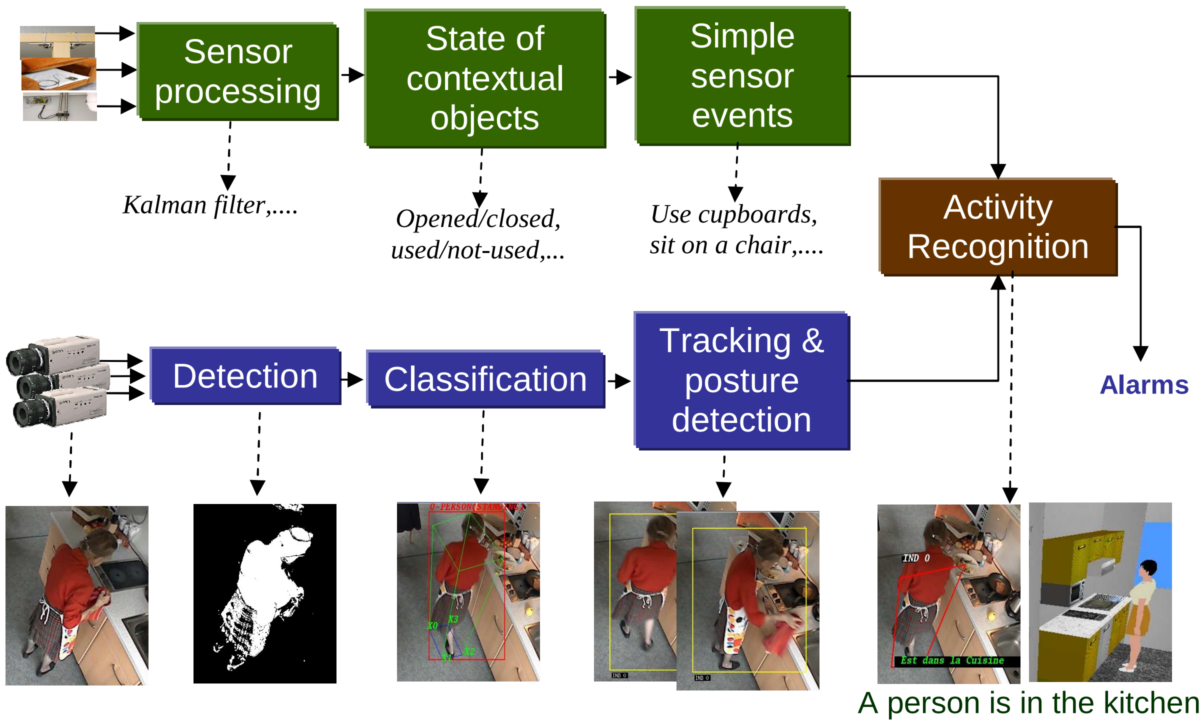

Fig.2. Sensors analysis

For homecare applications and in collaboration with gerontologists and geriatrics from Nice hospital, we have modelled forty activities. In particular, we have defined fourteen primitives and composite states, four of them are related to the location of the person in the scene (e.g. inside the kitchen, inside the living-room) and the ten remaining are related to the human postures (e.g. standing, sitting). We have defined also four primitive events related to the combination of these primitive states: "standing up" which represents a change state from sitting or slumping to standing, "sitting down" which represents a change state from standing, or bending to sitting on a chair, "sitting up" represents a change state from lying to sitting on the floor, and "lying down" which represent a change state from standing or sitting on the floor to lying. We have defined also twelve primitive events such as: stay in kitchen, stay in living-room, person use the stove. These primitive states and events are used to model ten composite events such as: "preparing meal", "taking meal", "feeling faint", "falling down".

Examples of modelled activities

1. "Falling down" model

```
CompositeEvent (PersonFallingDown,
    PhysicalObjects ((p: Person))
        Components ((pStand: PrimitiveState Standing(p))
                    (pSit: PrimitiveState Sitting_Flexed_Legs(p))
                    (pLay: PrimitiveState Lying_Outstretched_Legs(p)))
        Constraints ((pSit before_meet pLay)
                    (pLay's Duration >=50))
        Alert (AText ("Person is Falling Down")
              AType ("VERYURGENT")) )
```

2. "Taking a meal" model:

```
CompositeEvent (TakingMeal,
    PhysicalObjects ((p : Person), (z : Zone), (eq1 : Equipment), (eq2 : Equipment))
        Components ((s_inz : PrimitiveState inside_Livingroom(p, z))
                    (s_close : PrimitiveState close_to_table(p, eq1))
                    (c : CompositeState chair_used(p, eq2))
                    (s_sit : CompositeState person_sitting_inLivingroom(p, z)))
        Constraints ((z->Name = Livingroom)
                    (eq1->Name = table)
                    (eq2->Name = chair1)
                    (s_close's Duration >= 120)
                    (c's Duration >= 120)
                    (s_sit's Duration >= 120))
        Alert (AText ("Person Takes a Meal")
              AType("NOTURGENT") )
```

Results and evaluation

- ### Experimental site

An experimental laboratory called GERHOME has been built in the CSTB (Scientific Centre of Building Techniques) at Sophia Antipolis to assess the monitoring system and to explore the ADLs (Activities of Daily Living) that can be recognized by a computer system. This laboratory looks like a typical apartment of an elderly person: $41m^2$ with entrance, living-room, bedroom, bathroom, and kitchen. The kitchen includes an electric stove, a microwave oven, a fridge, cupboards, and drawers.

This laboratory is equipped with different sensors to evaluate the scenarios predefined by the domain specialists (gerontologists). Four video cameras are installed in the kitchen, in the living-room and in the bedroom. Twelve contact sensors are mounted on many devices in the kitchen (e.g. fridge, cupboards, drawers) to determine the state (opened or closed) of the various devices. Two power supply sensors are installed on electrical outlets to detect the utilisation (used or not used) of the electrical appliances (e.g. microwave). Three presence sensors are installed in front of the sink, the cooking stove, and the washbowl to detect the presence of people nearby. Three water sensors are installed in the kitchen and in the bathroom to detect hot/cold water consumption (e.g. used or not used). Four pressure sensors are installed under the chairs, the armchair and the bed. The selected sensors can easily and quickly be installed in home environments and are removable without damage to the cabinets or furniture. Figure 3 represents some pictures, and a 3D visualization of Gerhome laboratory.

Fig.3. Gerhome laboratory

- ### Experimental results

To validate our models, we have tested a set of human activities in the Gerhome laboratory. Two experiments have been performed. In the first experiment, we have acquired ten video sequences with only one human actor. The duration of each video is about ten minutes and each video contains about 4800 frames (about eight frames per second). In this experiment, we have tested some normal activities such as: open and close kitchen cupboards, use microwave and warm up a meal. We have also tested two abnormal activities: fainting and falling down. Figure 4 and figure 5 show respectively the 2D and the 3D visualization of the recognition of “fainting” and “falling down” situations.

The preliminary results of the recognition of the different activities are encouraging. The errors in the recognition of some postures for example occur when the system detects a wrong posture (e.g. the bending posture instead of the sitting one). These errors are due to the segmentation errors (i.e. shadow, light change) and to object occlusions.

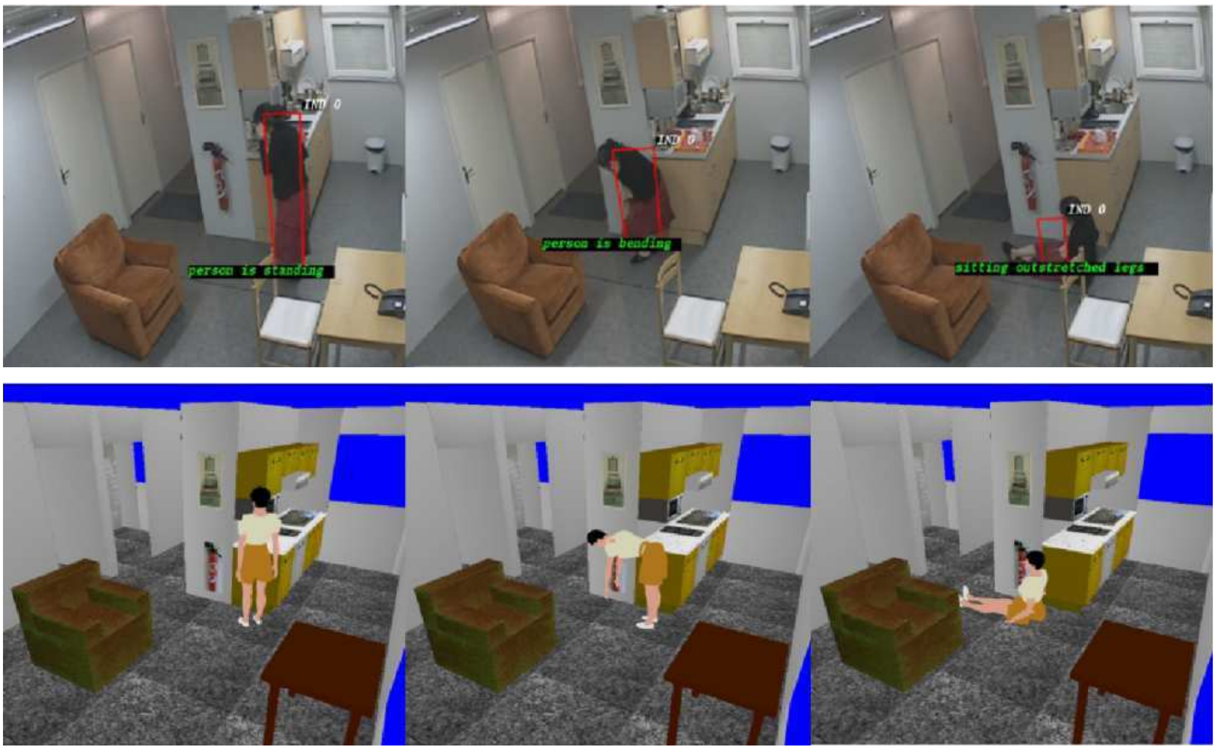

Fig.4. 2D and 3D visualization of the recognition of "fainting" situation

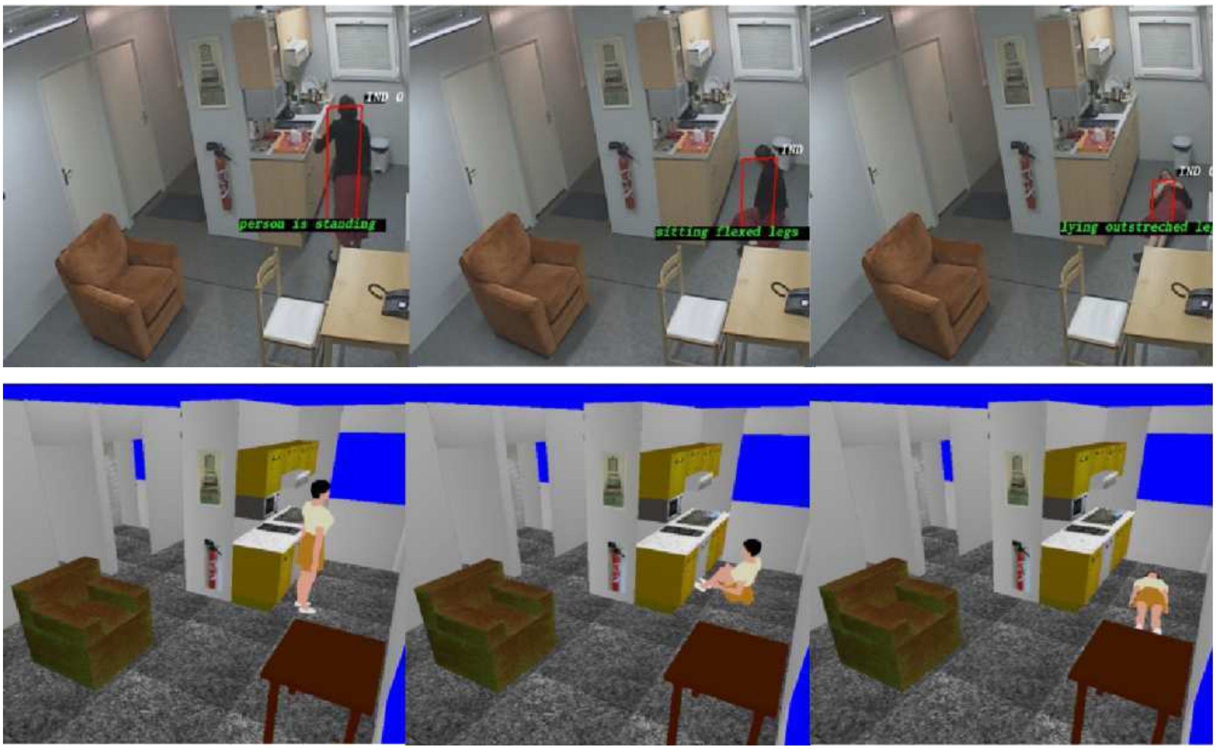

Fig.5. 2D and 3D visualization of the recognition of "falling down" situation

In the second experiment, while evolving in the laboratory, fourteen volunteers (aged from 60 to 85 years) have been observed during 4 hours and 14 video scenes have been acquired by 4 video cameras (about ten frames per second).

In this experiment the fourteen volunteers could behave freely, but were asked to perform a set of household activities, such as preparing meal, taking meal, washing dishes, cleaning the kitchen, and watching TV. Each volunteer was alone in the laboratory during the experiment.
Figure 6 and figure 7 show respectively the 2D and the 3D visualization of the recognition of “bending in the kitchen” and “taking meal” activities.

Fig.6. 2D and 3D visualization of the recognition of "bending in the kitchen" activity

Fig.7. 2D and 3D visualization of the recognition of "taking meal" activity

Among all analysed data, preliminary results for two different people (2 volunteers) observed during 4 hours are shown in Table 1. This table summarizes the number of instances, the mean duration and the total duration of each monitored activity. Time units are in minutes.

Volunteers	Volunteer 1 (< 65 years)			Volunteer 2 (> 84 years)		
Activity	**Mean duration (min)**	**Total duration (min)**	Number of instance	**Mean duration (min)**	**Total duration (min)**	Number of instance
- Use fridge	**0 :12**	**2 :50**	14	**0 :13**	**1:09**	5
- Use stove	**0 :08**	**4 :52**	35	**0 :16**	**27:57**	102
- Use microwave	**0**	**0**	0	**0**	**0**	0
- Use hot-water (in the kitchen)	**0 :19**	**12:40**	40	**0 :09**	**3:31**	22
- Use cold-water (in the kitchen)	**0 :28**	**9:36**	20	**0 :03**	**0:58**	19
- Use the upper-cupboard	**0 :51**	**21:34**	25	**4 :42**	**42:24**	9
- Use the lower-cupboard	**0 :21**	**3:09**	9	**1 :50**	**7:23**	4
- Use the middle-cupboard	**0 :10**	**1:51**	11	**0 :10**	**0 :52**	5
- Sitting on a chair 1	**0**	**0**	0	**92 :42**	**185:25**	2
- Sitting on a chair 2	**6 :07**	**73:27**	12	**0**	**0**	0
- Sitting on armchair	**0**	**0**	0	**0 :01**	**0:06**	6
- Use the toilet	**0**	**0**	0	**0 :56**	**0 :56**	1
- Use the bathroom-cupboard	**0 :03**	**0:07**	2	**0 :03**	**0:07**	2
- Use hot-water (in the bathroom)	**0 :10**	**0:21**	2	**0 :12**	**0:36**	3
- Use cold-water (in the bathroom)	**0 :12**	**0:24**	2	**0 :07**	**0:07**	1
- Use TV	**42 :18**	**169:12**	4	**16 :19**	**65:18**	4
- Enter in the kitchen	**2 :45**	**12 :00**	5	**2 :66**	**8 :00**	3
- Enter in the living-room	**1 :25**	**25 :00**	20	**2 :68**	**35 :00**	13
- Enter in the entrance	**2 :00**	**8 :00**	4	**1 :83**	**11 :00**	6
- Enter in the bedroom	**1 :25**	**5 :00**	4	**3 :28**	**23 :00**	7
- Enter in the bathroom	**1 :00**	**2 :00**	2	**1 :00**	**5 :00**	5
- Standing	**0:09**	**30 :00**	200	**0:16**	**12 :00**	45
- Bending	**0:04**	**2 :00**	30	**0:20**	**5 :00**	15

The volunteer 1 was detected to change more often its zone of activity (for entering in the living-room 20 vs. 13) and quicker (1:25 vs. 2:68), showing greater ability to move in the apartment. Similarly volunteer 1 was bending two time more (30 vs. 15) and in a quicker way (4 s vs. 20 s). Volunteer 1 was also using more the upper-cupboard (25 vs. 9) and in a quicker way (0:51 vs. 4:42) and was more able to use the stove (i.e. less trials).
All these measures show the greater ability of volunteer 1 to evolve in the apartment and to use the house equipment.

Conclusion

These preliminary results show that an automatic monitoring system is able to detect meaningful ADLs which enable to qualify how a person lives at home. A finer video analysis (especially by analysing fine postures and interactions with equipment) is needed to better measure people activities. More evaluation especially in real and natural environment on long periods is required to assess the robustness of the monitoring system.

Similarly, more work is needed to establish the behaviour profile of the observed person and to determine the significant changes in this profile.

Future work includes also establishing the relationships between the measured activities and the degree of frailty of the elderly.

Acknowledgments

We would like to thank the members of the "Conseil Général des Alpes Maritimes" for their financial support, and we also thank the fourteen volunteers who participated in the Gerhome experiments.

References

1. Avanzi, A., Bremond, F., Tornieri, C., and Thonnat, M.: "*Design and assessment of an intelligent activity monitoring platform*". EURASIP Journal on Applied Signal Processing, Special Issue on 'Advances in Intelligent Vision Systems: Methods and Applications'. (August. 2005).

2. Vu, V., Bremond, F., Thonnat, M.: "*Automatic video interpretation: A novel algorithm based for temporal scenario recognition*". In: The Eighteenth International Joint Conference on Artificial Intelligence. (September 9-15 2003).

3. H. Bouma. "*Gerontechnology: A framework on technology and aging*". In H. Bouma and J. Graafmans, editors, Gerontechnology, pages 1 – 6, Amsterdam, 1993. ISO Press.

4. H. Sidenbladh and M. Black. "*Learning image statistics for Bayesian tracking*". IEEE International Conference on Computer Vision (ICCV), 2:709–716, 2001.

5. B. Clarkson, N. Sawhney, and A. Pentland. "*Auditory context awareness via wearable computing*". In Proceedings of the Perceptual User Interfaces Workshop, 1998.

6. D. Moore, I. Essa, and M. Hayes. "*Exploiting human actions and object context for recognition tasks*". In Proceedings of IEEE International Conference on Computer Vision 1999 (ICCV99), 1999.

7. T. Barger et al., "*Objective Remote Assessment of Activities of Daily Living: Analysis of Meal Preparation Patterns*". Medical Automation Research Center, Univ. of Virginia Health System, 2002.

8. S. Wang, W. Pentney, A-M. Popescu, T. Choudhury, and M. Philipose, "*Common sense joint training of human activity recognizers*". Proceedings of IJCAI 2007, 2007.

9. F. Foerster and J. Fahrenberg. "*Motion pattern and posture: correctly assessed by calibrated accelerometers*". Behaviour Research Methods, Instruments, & Computers, 32(3):450–7, 2000. Find United States ISSN: 0743-3808; NLM Unique Journal Identifier: 8413015 English Index Medicus Journal Article.
10. J. Hoey, A. V. Bertoldi, P.P., Mihailidis, A.: "*Assisting persons with dementia during handwashing using a partially observable markov decision process*". In: International Conference on Computer Vision Systems (ICVS), Germany. (March 2007.).

11. E. Munguia-Tapia, S. S. Intille, and K. Larson, "*Activity Recognition in the Home Setting Using Simple and Ubiquitous Sensors*", in *Proceedings of PERVASIVE 2004*, vol. LNCS 300, B. Heidelberg, Ed.: Springer-Verlag, 2004, pp. 158-175.

12. Boulay, B., Bremond, F., Thonnat, M.: "*Applying 3d human model in a posture recognition system*". Pattern Recognition Letter. (2006.)

13. Allen, J.F.: "*Maintaining knowledge about temporal intervals*". In Communications of the ACM. (1983).