

# **An Activity Monitoring System for Real Elderly at Home: Validation Study**

Nadia Zouba, François Bremond, Monique Thonnat

► **To cite this version:**

Nadia Zouba, François Bremond, Monique Thonnat. An Activity Monitoring System for Real Elderly at Home: Validation Study. 7th IEEE International Conference on Advanced Video and Signal-Based Surveillance AVSS10, Aug 2010, Boston, United States. inria-00503864

**HAL Id: inria-00503864**

**<https://hal.inria.fr/inria-00503864>**

Submitted on 19 Jul 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# An Activity Monitoring System for Real Elderly at Home: Validation Study

Nadia ZOUBA, Francois BREMOND, Monique THONNAT  
INRIA Sophia Antipolis - Mediterranee, PULSAR Team  
2004 route des Lucioles, BP 93, Sophia Antipolis, France

Nadia.Zouba@sophia.inria.fr, Francois.Bremond@sophia.inria.fr, Monique.Thonnat@sophia.inria.fr

## Abstract

*Since the population of the elderly grows highly, the improvement of the quality of life of elderly at home is of a great importance. This can be achieved through the development of technologies for monitoring their activities at home. In this context, we propose an activity monitoring system which aims to achieve behavior analysis of elderly people. The proposed system consists of an approach combining heterogeneous sensor data to recognize activities at home. This approach combines data provided by video cameras with data provided by environmental sensors attached to house furnishings. In this paper, we validate the proposed activity monitoring system for the recognition of a set of daily activities (e.g. using kitchen equipment, preparing meal) for 9 real elderly volunteers living in an experimental apartment. We compare the behavioral profile between the 9 elderly volunteers. This study shows that the proposed system is thoroughly accepted by the elderly and it is also well appreciated by the medical staff.*

## 1. Introduction

Demographic changes associated with the aging population and the increasing numbers of elderly people living alone are leading to a significant change in the social and economic structure of our society. Elderly population is expected to grow highly over the next 20 years. By 2030, third of Europeans will be over than 65 and 40% of them will require assistance. Without receiving sufficient care, elderly are at risk of loosing their independence.

Activities of daily living (ADLs), such as cooking, bathing and toileting, are good indicators of the cognitive and physical capabilities of elderly [6]. Therefore, a system that automatically recognizes these activities allows automatic health monitoring [2], and provides an objective measure for medical staff. An activity monitoring system is therefore a crucial step in the future health applications. Advances in technology have provided the ability to equip the home environment with a large number of different sensors.

In this context, we propose an activity monitoring system to recognize activities observed by video cameras and environmental sensors attached to house furnishings. In this paper, we validate the proposed activity monitoring system for the recognition of a set of daily activities (e.g. using kitchen equipment, preparing meal) for 9 real elderly volunteers living in an experimental apartment.

Considerable research has been devoted towards activity monitoring through the deployment of sensing technology to detect interactions with objects, from visual sensors like video cameras [7] to sensors which provide binary "on" or "off" outputs such as contact sensors that are used to detect for example a door being opened or closed [11]. One approach is to tag a large number of objects in a home with RFID tags. An RFID reader in the form of a bracelet is worn by the user to detect which objects are used [5]. Another approach is to use video. Duong et al. [4] use four video cameras to capture a scene from different angles. From the videos they extract the location of a user and use it for activity recognition. Wu et al. [13] use a single camera combined with an RFID bracelet. Ogawa, et al. [9] have used motion sensors to detect movement, use of appliances, and presence in a room. They used this information to analyze behavior patterns of two elderly ladies living alone. Nambu, et al. [8] found that analyzing TV watching patterns was effective at identifying and analyzing behavior patterns, without the need for additional customized sensors. These systems have two limitations. They are either intrusive (wearing a glove or bracelet) or too specific (specialized for a particular activity (e.g. TV watching) or based on unique cue).

The next section gives an overview of our activity monitoring system. Section 3 presents our experiments and the obtained results of the recognition of a set of daily activities and results of elderly profile. Finally, in section 4 we present our conclusion and the future work.

## 2. Activity Monitoring System

Our activity monitoring system is based on multisensor analysis and a human activity recognition. This system includes detecting people, tracking people as they move, and

recognizing their postures and activities of interest. It exploits three major sources of knowledge: the models of activities (see section 2.2), a 3D model of an observed scene (see section 2.3), and a 3D model of the mobile objects present in the observed scene (e.g. a 3D model of a person). The proposed system takes as input the data provided by the different sensors (i.e. video cameras and environmental sensors) and produces as output a set of recognized activities which are saved in XML files for evaluation purpose.

## 2.1. Multisensor Analysis

Multisensor analysis consists in collecting multisensor data of the observed person and process them. Sensors we used are: video cameras to detect and track an observed person, contact sensors to measure open-close status of doors, cupboards, drawers and fridge; pressure sensors to measure sitting in a chair, armchair or lying in bed; presence sensors to detect presence of people near sink, cooking stove and washbowl; electrical sensors to measure the use of the stove or other appliances; water flow sensors to measure water consumption in the kitchen and bathroom. In the proposed system, we use a fusion process at the decision level to address the problem of heterogeneous sensors. For this, we combine the video events with the environmental events in order to detect rich and complex events (i.e. multimodal events). The multimodal events (i.e. daily activities) can include video events and / or environmental events. Therefore, when the video and the environmental events are recognized, then the global multimodal event is also recognized.

In the next sections, we describe briefly video analysis, environmental sensor analysis, activity modeling and activity recognition, which are the main components of the proposed activity monitoring system. For more details see [14].

### 2.1.1 Video Analysis

Video analysis aims at detecting and tracking people moving in the scene (see figure 1). To achieve this task, we have used a set of vision algorithms coming from a video interpretation platform described in [1]. To detect postures of interest, we have used a human posture recognition algorithm [3] in order to recognize in real time a set of human postures once the person moving in the scene is correctly detected. This algorithm determines the posture of the detected person using the detected silhouette and its 3D position. The human posture recognition algorithm is based on the combination between a set of 3D human models with a 2D approach. These 3D models are projected in a virtual scene observed by a virtual camera which has the same characteristics (position, orientation and field of view) than the real camera. The 3D silhouettes are then ex-

tracted and compared to the detected silhouette using a 2D technique which projects the silhouette pixels on the horizontal and vertical axes. The most similar extracted 3D silhouette is considered to correspond to the current posture of the observed person. The algorithm is real time (about eight frames per second), and does not depend on camera position. The input of the video analysis part is a set of video stream. Its output is a set of XML files which contain parameters of the detected person. The most important parameters which we have used in this work are the 3D position of the person and his/her 3D posture.

For homecare applications, in collaboration with gerontologists from Nice hospital in France, we have proposed ten 3D key human postures which are useful to recognize activities of interest at home. These 3D human postures are based on a 3D geometrical human model. These 3D postures are: standing, standing with arm up, standing with hands up, bending, sitting in a chair, sitting on the floor with outstretched legs, sitting on the floor with flexed legs, slumping, lying on the side with flexed legs, and lying on the back with outstretched legs. Each of the proposed 3D human postures plays a significant role in the recognition of the targeted activities of daily living or of abnormal activities (e.g. falling down). For example, the posture "standing with hands up" is used to detect when a person is carrying an object such as plates. The posture "standing with arm up" is used to detect when a person reaches and opens kitchen cupboard and his/her ability to do it. These proposed human postures are not an exhaustive list but represent the key human postures taking part in everyday activities.

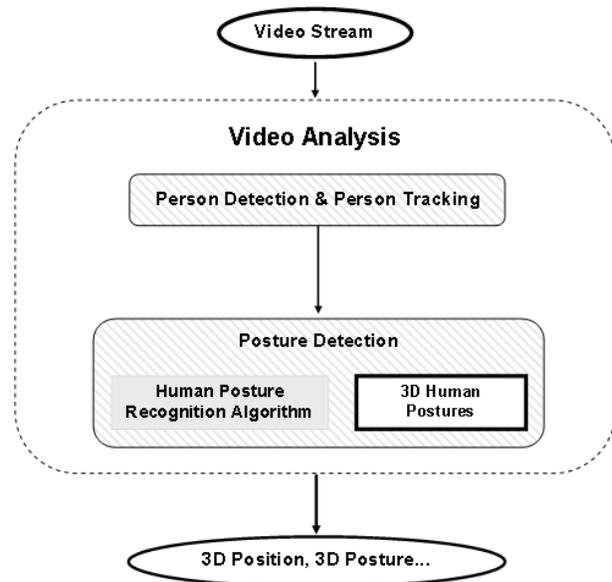


Figure 1. The video analysis architecture. Input: video stream, output: 3D position, 3D posture

### 2.1.2 Environmental Sensor Analysis

The environmental sensors provide data when an event occurs. For instance the contact sensor determines an opening and closing events for various devices (i.e. kitchen cupboards, kitchen drawers, kitchen fridge, bedroom closets).

We have defined the following form for the provided data:

- TimeStamp: represents the moment when the data was provided (YYMMDD-HHMMSS.MS);
- SensorClass: represents the class of information provided by the sensor (i.e. contact, presence, electrical, pressure and water);
- SensorLocation: represents the location of the sensor (e.g. upper cupboard);
- SensorValue: the value provided by the sensor (i.e. "On" if the sensor is activated and "Off" if the sensor is deactivated);
- SensorID: single sensor identifier which is transmitting the data.

The provided data are stored in an XML file and transmitted via a parser to the event detection process. From these data, we infer the corresponding environmental event. For example, if the provided data is "On" and the sensor class is "contact" then we infer the contact event "Open". If the provided data is "Off" and the sensor class is "contact" then we infer the contact event "Closed".

## 2.2. Activity Monitoring

### 2.2.1 Activity Modeling

To give the meaning of the activities of interest happening at home a modeling effort is needed. To model postures and activities of interest at home we have used the event description language proposed by Vu et al. [12]. This language is both declarative and intuitive (in natural terms) so the experts of the application domain can easily define and modify the event models. However, this language presents some significant drawbacks for modeling daily activities. His first drawback is that it is dedicated for data provided by only video cameras and does not take into account data provided by other types of sensors. His second drawback is that it does not allow to model complex activities by combining data from different sensors.

So, we propose 2 extensions of this language. The first extension concerns the adding of data provided by non-video sensors. The second extension allows the combination of different sensors in order to address complex activity modeling in a scene observed by video cameras and environmental sensors and over an extended period of time. The extended language uses a declarative representation of events that are defined as a set of spatio-temporal and logical constraints. Three main types of events have been designed. A **primitive state** which is a spatio-temporal property valid at a given instant or stable over a time interval that is di-

rectly inferred from the visual attributes of physical objects computed by vision routines (e.g. a person is located inside a kitchen) or by other sensors (e.g. a fridge is open). A **primitive event** which is a primitive state transition and represents the finest granularity of events (e.g. a person is staying close to a table). A **composite event** which is a combination of primitive states and events (e.g. a person is preparing a lunch). To model an event E, we distinguish the set of physical objects (e.g. persons, tables) involved in E, a set of components (i.e. sub-events) composing E and a set of constraints on these physical objects and/or these components.

For homecare application, we have done a strong effort in event modeling. The result is 100 models which is our knowledge base of events: 58 customized video events for household activities (among them 26 are posture-based events), 26 environmental event models, and 16 daily activities models. These daily activities are: using (i) fridge, (ii) cupboards, (iii) drawers, (iv) microwave, (v) stove, (vi) telephone, (vii) watching TV, (viii) dish washing, (ix) slumping in armchair, (x) taking a meal, and (xi) 6 variations of preparing a meal: breakfast, lunch, dinner, warming a meal, cold meal and hot meal. Each activity is modeled with sub-activities relating to objects involved in that activity. For example, in the definition of the model of preparing lunch, the person should be located close to the countertop in the kitchen and staying at this location for a while, the person opens cupboards to take ingredients and dishes (e.g. plates, fork, knife), opens the fridge to take foods, uses the stove to cook the meal. Figure 2 shows a model of preparing lunch. This model involves 5 physical objects (the person, and 4 equipments), five components: person close to countertop (video camera), person stays at countertop (video camera), three multisensor composite events related to the using of the kitchen equipment (Cupboards, Fridge and Stove) and three temporal constraints.

### 2.2.2 Activity Recognition

To address complex activity recognition involving several physical objects of different types (e.g. person, chairs) in a scene observed by video cameras and environmental sensors over an extended period of time, the recognition process takes as input video and environmental events and the a priori knowledge of multimodal events to be recognized. These events are first processed to synchronize them. Then, the event recognition process takes as input the synchronized events and tries to understand which events (i.e. video-environmental events or activities) are occurring. The algorithm operates in 2 stages: (i) at each incoming frame, it computes all possible primitive states related to all mobile objects present in the scene, and (ii) it computes all possible events (i.e. primitive events, and then composite

```

CompositeEvent ( PreparingLunch,
PhysicalObjects ( (p : Person), (Countertop : Equipment), (UpperCupboard : Equipment),
(Fridge : Equipment), (Stove : Equipment)
Components ( (c1: PrimitiveState CloseToCountertop(p, Countertop))
(c2: PrimitiveEvent stays_at(p, Countertop))
(c3: CompositeEvent UsingCupboards(p, UpperCupboard))
(c4: CompositeEvent UsingFridge(p, Fridge))
(c5: CompositeEvent UsingStove(p, Stove))
Constraints ( (c1 During c2)
(c2 Duration >= threshold1)
(c5 Duration >= threshold2))
Alert ( AText("Person prepares lunch")
AType("NOTURGENT")) )
```

Figure 2. Example of "preparing lunch" model. The threshold1 and the threshold2 are calculated using 5 annotated "PreparingLunch" event.

events) that may end with the previously recognized primitive states.

### 2.3. Scene Modeling

The defined 3D model of an apartment contains both geometric and semantic description of the specific zones, walls and the equipment located in the observed apartment and contains also geometric information of the installed sensors. In this 3D model, we have defined:

- \* A 3D referential which contains the calibration matrices and the position of the video cameras;
- \* A list of environmental sensor positions. To define these positions, we have provided for each installed sensor the associated equipment with its location in the scene.
- \* A list of geometric areas corresponding to the different rooms (i.e. entrance, kitchen, livingroom, bedroom and bathroom) in the observed environment (i.e. an apartment);
- \* A list of geometric zones corresponding to the different zones of interest in the observed environment (i.e. entering zone, exiting zone, cooking zone, eating zone, sleeping zone and bathing zone);
- \* A list of walls to describe for instance home walls (e.g. kitchen north wall, bedroom west wall);
- \* A list of the different equipment (e.g. table, fridge, microwave) present in the observed scene with its characteristics (e.g. static objects (e.g. stove), objects with displacements (e.g. chairs)).

The geometric description of areas contains a polygon defined in a plane (i.e. a ground). The geometric description of equipment is defined by its size (i.e. height, width, length) and its coordinates in a plane. The semantic description of an area, of a zone, of a wall, and of an equipment contains two attributes: its type (area, zone, wall



(a) External view of the experimental apartment (b) The kitchen in the experimental apartment

Figure 3. Views of the experimental apartment

or equipment) and its name (e.g. cooking zone, kitchen, table). The proposed 3D model of an apartment can be used in another environment, by redefining the geometric information of the new observed environment.

A 3D model of a mobile object is composed by a name of a model, and by a set of Gaussian functions which describe the 3D width, 3D height, and 3D depth of the mobile object. The availability of a 3D model of mobile objects allows us to have a more precise description of the mobile objects present in the scene (e.g. person, pets).

## 3. Results and Validation of the proposed system

In order to evaluate the whole proposed activity monitoring framework, an experiment has been performed. The main objectives of this experiment are to validate the different phases of the activity monitoring framework, to highlight interesting characteristics of the approach, and to evaluate the potential of the framework for real world applications.

### 3.1. Experimental Site

Developing and testing the impact of the activity monitoring solutions requires a realistic environment in which training and evaluation can be performed. To attain this goal we have set up an experimental apartment (see figure 3). This laboratory is located in the CSTB (Scientific Center of Technical Building) at Sophia Antipolis in France. This experimental site looks like a typical apartment of an elderly person: 41m<sup>2</sup> with an entrance, a livingroom, a bedroom, a bathroom, and a kitchen. The kitchen includes an electric stove, a microwave, a fridge, cupboards, and drawers. This apartment is equipped with different sensors (see figure 4). Commercially available sensing devices were used for data gathering including video cameras, and environmental sensors embedded in the home infrastructure.

### 3.2. Performed Experiment

We have evaluated the proposed activity monitoring framework in collaboration with gerontologists from Nice

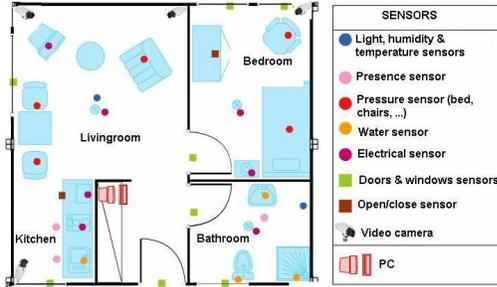


Figure 4. Position of the sensors in the experimental apartment

hospital in France. These gerontologists have defined ADLs scenarios of human activities. These scenarios have been tested in the experimental apartment with real elderly volunteers they have selected [15]. While living in the experimental site, nine real elderly people, aged from 64 to 85 years old, have been observed, each one during 4 hours, and 36 video sequences have been acquired by 4 video cameras (at ten frames per second), each video sequence contains about 144 000 frames. The collected data include the 36 video streams, and also data provided by 24 environmental sensors. These data are available on [www.sop.inria.fr/members/Francois.Bremond/topicsText/gerhomeProject.html](http://www.sop.inria.fr/members/Francois.Bremond/topicsText/gerhomeProject.html). The volunteers were encouraged to behave freely and to maintain as normal as possible their behaviors and were asked to perform a set of household activities such as preparing meal, taking meal, washing dishes, cleaning the kitchen, watching TV and taking a nap while staying in the apartment. Each volunteer was alone in the apartment during the observation period and was observed during 4 hours (i.e. between 10h and 14h). All the volunteers were interviewed separately, after the study, about the experience of living in the experimental apartment. They were also asked about the proposed scenario, the acceptance of the sensor technologies and the proposed system. This study shows that the proposed system is thoroughly accepted by the elderly and is also well appreciated by the medical staff.

### 3.3. Results of Recognition

To estimate threshold values (i.e. the different thresholds introduced in the definition of activity models in section 2.2 in figure 2), we use 5 ground truth video sequences for 5 observed elderly people among the all experimental data. Results of the recognition of 6 daily activities in the 5 ground truth video sequences are shown in table 1. The duration of these video sequences is 20 hours. Table 1 summarizes the ground truth (GT), the true positive (TP), the false negative (FN), the false positive (FP), the precision ( $P=TP/(TP+FP)$ ) and the sensitivity ( $S=TP/(TP+FN)$ ) of the recognition of six daily activities. These daily activities are well recognized by the proposed monitoring system. The errors in the

Multimodal activity	GT	TP	FN	FP	P	S
<i>Use fridge</i>	65	54	11	9	86%	83%
<i>Use stove</i>	177	165	11	15	92%	94%
<i>Sitting on chair</i>	66	54	12	15	78%	82%
<i>Sitting on armchair</i>	56	49	8	12	80%	86%
<i>Prepare lunch</i>	5	4	1	3	57%	80%
<i>Wash dishes</i>	16	13	3	7	65%	81%

Table 1. Results of the recognition of 6 daily activities for 5 video sequences

recognition are sometimes due to the sensor measurement errors and sometimes due to the predefined event durations used in the definition of the activity model (i.e. the predefined thresholds).

- The false positives in the recognition of "sitting in a chair" and "sitting in an armchair" activities are detected when a person drops his bag on a chair or in an armchair which activates the pressure sensors installed under the chair and/or under an armchair.

- The missing event in the recognition of "prepare lunch" activity is due to the fact that in one video sequence, a person who prepared lunch has used the microwave instead of the stove. Thus, in this case the activity "prepare lunch" should be modeled differently.

- The false positives in the recognition of "wash dishes" activity are detected when someone opens the tap water to wash hands. The false negatives are caused when a person washes a cup which takes less time than washing dishes.

#### 3.3.1 Results of the Recognition without and with using Dempster Shafer Uncertainty

In this work we used environmental sensors which provide two binary values "On" if the sensor is activated and "Off" if the sensor is deactivated. The challenges posed with the use of binary sensor technology and the determination if a sensor provides a value of "On" or "Off" how sure can we be about this measurement are huge. By applying Dempster-Shafer (DS) theory of evidence for the representation and management of sensor uncertainty will provide a possible solution to this problem. For this, we use Dempster-Shafer [10] theory through the fusion of contextual information inferred from uncertain sensor data. Using the Dempster-Shafer (DS) theory, we calculate the uncertainty in sensor measurements of 4 activities for one video sequence. Comparison between the results obtained without using uncertainty and the results obtained with using uncertainty (see table 2) shows some improvements in the recognition of activities. The new results (using uncertainty in sensor measurements) show a good recognition, com-

Multimodal activity	GT	Without uncertainty			With uncertainty		
		TP	FN	FP	TP	FN	FP
Use Fridge	24	18	6	3	21	3	1
Use Stove	18	15	3	4	16	2	1
Sitting on a Chair	29	25	4	7	27	2	2
Sitting in an Armchair	8	6	2	5	7	1	2

Table 2. Comparison between obtained results without and with using DS uncertainty of recognition of 4 daily activities for 1 video sequence

pared to the results obtained without using uncertainty in sensor measurements, of the "use fridge" (Precision= 95% vs. 85% and Sensitivity= 87% vs. 75%), of the "use stove" (Precision= 94% vs. 78% and Sensitivity= 88% vs. 83%), of the "sitting in a chair" (Precision= 93% vs. 78% and Sensitivity= 93% vs. 86%), and of the "sitting in an armchair" (Precision= 77% vs. 54% and Sensitivity= 87% vs. 75%) activities.

### 3.4. Behavioral Profile

The basic goal of defining behavioral profile is to measure variables from individuals during their daily activities in order to capture deviations of activity and posture to facilitate timely intervention or provide automatic alert in emergency cases. In this work, to define a behavioral profile for a person  $P_j$ , we measure two main variables: the duration and the number of occurrence of each daily activity for a person  $P_j$ .

#### 3.4.1 Event Duration

In this work we propose to estimate the duration of each event  $E_i$  by calculating the mean duration value of  $E_i$ . We calculate the mean duration  $\mu_{E_i, P_k}$  of  $E_i$  for each volunteer  $P_k$  by summing all the durations  $d_{E_i}$  of  $E_i$  and dividing by their number of occurrence  $N_{E_i}$  (see equation 1).

$$\mu_{E_i, P_k} = \frac{\sum_{i=1}^n d_{E_i}}{N_{E_i}} \quad (1)$$

#### 3.4.2 Leave-One-Out Cross Validation

For each observed elderly person among the 9 observed elderly people, we have done the leave-one-out cross validation on the activity duration. This technique involves using a single observation as the validation data, and the remaining observations as the training data. This is repeated such that each observation is used once as the validation data.

We calculate firstly the mean duration of each event for  $R$

persons ( $R$  is the number of training set, the number of people minus 1) among the all observed persons by using the following equation:

$$MD_{E_i, P_j} = \frac{\sum_{P_k \in P, P_k \neq P_j} \mu_{E_i, P_k}}{R}, \quad \forall P_j \in P \quad (2)$$

Where:  $MD_{E_i, P_j}$  represents the mean duration for a given event  $E_i$  for each person but a person  $P_j$ ;  $\mu_{E_i, P_k}$  represents the mean duration for each event  $E_i$  for each person  $P_k$  (see equation 1);  $P = \{P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9\}$ ; and  $R$  represents the number of the training set of data (i.e.  $R = P - 1 = 8$ ).

Secondly, we calculate the standard deviation  $\sigma_{E_i, P_j}$  for each event  $E_i$  for each person but a person  $P_j$  by using the following equation:

$$\begin{aligned} \sigma_{E_i, P_j} &= \sqrt{\frac{1}{R} \sum_{k=1}^R (\mu_{E_i, P_k} - MD_{E_i, P_j})^2} \quad (3) \\ &= \sqrt{\frac{1}{R} \sum_{k=1}^R \mu_{E_i, P_k}^2 - MD_{E_i, P_j}^2} \end{aligned}$$

Where:  $\sigma_{E_i, P_j}$  represents the standard deviation for each event  $E_i$  for each person but a person  $P_j$ .

Finally, we calculate for each person but a person  $P_j$  and for each event  $E_i$  the interval  $I_{E_i, P_j} = [MD_{E_i, P_j} - \sigma_{E_i, P_j}; MD_{E_i, P_j} + \sigma_{E_i, P_j}]$ .

After that we compare the mean duration  $\mu_{E_i, P_k}$  for each event  $E_i$  with the interval  $I_{E_i, P_k}$  (results are shown in section 3.4.3).

#### 3.4.3 Comparison between results for 9 elderly people

In this study we compare the behavioral profile of 9 elderly people observed during 4 hours by using results of the recognition of daily activities. We compare the number of occurrence  $N_{E_i}$  of 5 daily activities for the 9 volunteers during the four hours of observation. The obtained results are shown using graphics. Figure 5 shows the results comparing the number of occurrence of 5 activities for 9 elderly people. In this figure, there are some difference in the occurrence of the 5 daily activities between the 9 elderly people. Results of the number of occurrence of the 5 daily activities show:

- The volunteer P9 (woman of 85 years) was using more the "stove" than the remaining volunteers.
- The volunteer P1 (man of 64 years) was using more the "fridge" than the remaining volunteers.
- The volunteers P1 (man of 64 years) and P7 (woman of 66 years) were more often seen "sitting on chair" than the remaining volunteers, and the volunteers P4 (man of 66

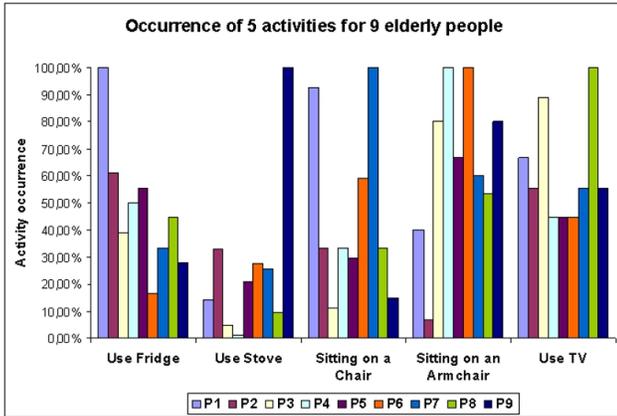


Figure 5. Results of normalized activity occurrence comparing 9 elderly people (100% is given to the person who made the maximum occurrence of an activity)

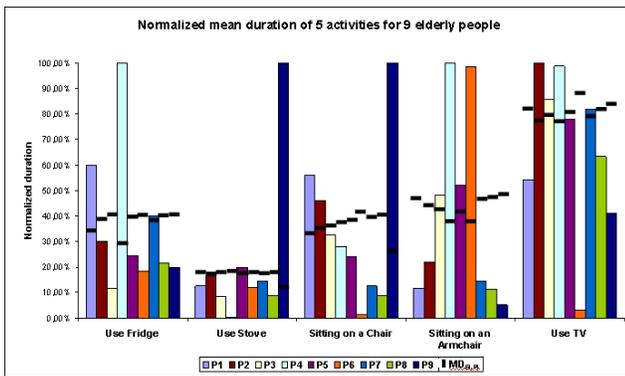


Figure 6. Normalized mean duration of 5 activities comparing 9 elderly people. The horizontal rectangles represent the normalized mean durations  $MD_{E_i, P_j}$

years) and P6 (woman of 70 years) were more often seen "sitting in an armchair" than the remaining volunteers.

Using the cross-validation technique described in section 3.4.2, we have obtained 54 intervals (i.e. 6 events x 9 persons). We validate each activity by comparing the mean duration  $\mu_{E_i, P_k}$  of each event  $E_i$  for each person  $P_k$  with the corresponding interval. The obtained results are shown in figure 6. For example, for the person  $P_1$ , to validate an event  $E_1$  which represents the event "Use Fridge", we compare his mean duration value  $\mu_{E_1, P_1}$  to the interval  $I_{E_1, P_1} = [MD_{E_1, P_1} - \sigma_{E_1, P_1}; MD_{E_1, P_1} + \sigma_{E_1, P_1}]$ .

Figure 6 shows:

- The volunteer P4 has used a fridge during a long time compared to the others, and the volunteer P3 has used a fridge for a shorter time compared to the others.
- The volunteer P9 was "sitting in a chair" during a longer



Figure 7. The recognition of "preparing a meal" activity of a real elderly man (person P2)

duration than the others, and person P6 was "sitting in a chair" for a shorter time. Using these results, we can deduce that a person P6 is more able to the person P9 to move in the apartment.

- The volunteers P4 and P6 were "sitting in an armchair" for a longer duration than the others, and person P9 was "sitting in an armchair" for a shorter time.
- The volunteer P6 has used a TV for a shorter time than the others, and person P2 has used a TV during a longer duration.

The main deductions of all the obtained results show that:

- The person P9 (woman of 85 years) has a fairly different profile from the others. This person shows some inabilities in using kitchen equipment (e.g. on using stove) and also shows some difficulties to move in the laboratory (e.g. sitting in a chair for a long duration), which may be the first sign of the frailty of this person.
- The person P6 (woman of 70 years) shows different profile in using TV, in sitting in a chair and in sitting in an armchair. After viewing the videos, we found that this person had difficulties in turning on the TV. This may be due to the fact that this person does not have TV in her own home or has difficulties using the remote control. About sitting in a chair and sitting in an armchair it is due to the sensor failures, the pressure sensors are very sensitive to weight (e.g. they are activated for very light weight like bags).
- The person P4 (man of 66 years) shows different profile in using TV, using fridge and sitting in an armchair.

Figure 7 shows the recognition of "preparing a meal" activity for the volunteer P2, and figure 8 shows the recognition of "taking a meal" activity for the volunteer P9.

The obtained results demonstrate that the proposed activity monitoring system allows to detect and recognize a set of



Figure 8. The recognition of "taking a meal" activity of a real elderly woman (person P9)

activities of a person by using the data provided by the combination of the selected sensors. The study shows that the proposed system is thoroughly accepted by the elderly and it is also well appreciated by the medical staff.

#### 4. Conclusion and Future Work

In this paper we have presented an activity monitoring system to analyze the behavior of elderly people. The proposed system has been successfully tested for a set of daily activities of 9 real elderly volunteers observed in an experimental apartment. We have obtained good results with few false alarms. In the current work, the proposed activity monitoring approach was evaluated in the experimental laboratory with 9 elderly people. The next step of this work requires to test this approach for a long time in nursing homes and in hospital environment involving more people with different wellness and different health status (e.g. Alzheimer patients).

#### ACKNOWLEDGEMENTS

We would like to thank the elderly people who have participated in the GERHOME experiments and we would like to thank them for allowing us to use their data (i.e. pictures) in our scientific publications. We would like to thank also medical staff from Nice hospital for their help.

#### References

- [1] A. Avanzi, F. Bremond, C. Tornieri, and M. Thonnat. Design and assesment 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
- [2] T. Barger, D. Brown, and M. Alwan. Health status monitoring through analysis of behavioral patterns. In *8th congress of the Italian Association for Artificial Intelligence (AIIA) on ambient intelligence*, 2003. 1
- [3] B. Boulay, F. Bremond, and M. Thonnat. Applying 3d human model in a posture recognition system. *Pattern Recognition Letter*, vol.27, no.15:pp.1785–1796, 2006. 2
- [4] T. Duong, H. Bui, D. Phung, and S. Venkatesh. Activity recognition and abnormality detection with the switching hidden semi-markov model. In *Proceedings of the IEEE computer society conference on computer vision and pattern recognition (CVPR)*, 2005. 1
- [5] K. Fishkin, M. Philipose, and A. Rea. Hands-on rfid: wireless wearables for detecting use of objects. In *9th IEEE international symposium on wearable computers (ISWC)*, IEEE Computer Society, 2005. 1
- [6] S. Katz. Assessing self-maintenance: Activities of daily living, mobility, and instrumental activities of daily living. *Journal of the American Geriatrics Society*, vol.31, no.12:pp.721–727, 1983. 1
- [7] T. Moeslund, A. Hilton, and V. Kruger. A survey of advances in vision based human motion capture and analysis. *Computer Vision Image Understanding (CVIU)*, vol.104, no.2:pp.90–126, 2006. 1
- [8] M. Nambu, K. Nakajima, M. Noshira, and T. Tamura. An algorithm for the automatic detection of health conditions. *IEEE Engineering Medicine Biology Magazine*, vol.24, no.4:pp.38–42, 2005. 1
- [9] M. Ogawa, R. Suzuki, S. Otake, T. Izutsu, T. Iwaya, and T. Togawa. Long term remote behavioral monitoring of elderly by using sensors installed in ordinary houses. In *IEEE Engineering in Medicine and Biology Society (EMBS) special topic conference on microtechnologies in medicine and biology*, October 2002. 1
- [10] G. Provan. The validity of dempster-shafer belief functions. *International Journal of Approximate Reasoning*, vol.6, no.3:pp.389–399, 1992. 5
- [11] D. Tran, D. Phung, H. Bui, and S.Venkatesh. A probabilistic model with parsimonious representation for sensor fusion in recognizing activity in pervasive environment. In *Proceedings of the 18th International Conference on Pattern Recognition*, 2004. 1
- [12] V. Vu, F. Bremond, and M. Thonnat. Automatic video interpretation: A novel algorithm based for temporal scenario recognition. In *The Eighteenth International Joint Conference on Artificial Intelligence (IJCAI)*, September 9-15 2003. 3
- [13] J. Wu, A. Osuntogun, T. Choudhury, M. Philipose, and J. Rehg. A scalable approach to activity recognition based on object use. In *Eleventh IEEE International Conference on Computer Vision (ICCV)*, October 2007. 1
- [14] N. Zouba. Multisensor fusion for monitoring elderly activities at home. In *PhD Thesis at Nice University, France*, 2010. 2
- [15] N. Zouba, F. Bremond, M. Thonnat, A. Anfosso, E. Pascual, P. Mallea, V. Mailland, and O. Guerrin. A computer system to monitor older adults at home: Preliminary results. *Gerontechnology Journal*, vol.8, no.3:pp.129–139, July 2009. 5