Multisensor Fusion for Monitoring Elderly Activities at Home

Presented by: Nadia Zouba Valentin
PULSAR team

Supervisor: Monique Thonnat
Co-supervisor: François Bremond
Outline

1. Introduction
2. State of the art
3. Approach overview
4. Video analysis
5. Sensor analysis
6. Event recognition
7. Multisensor event fusion
8. Evaluation and results
9. Conclusion and future work
Introduction

- In 2035: 1/3 of Europeans more than 65 years
- 40% of people 65 and older have a disability
- Over 20% require continuous monitoring

Increasing number of high risk population (elderly, people with illness, …) living alone at home

Activities of Daily Living (ADLs)

Daily activities at home are activities that people tend to do everyday

Examples of daily activities:

- Preparing meal (in kitchen) (11h– 12h)
- Eating (in dinning room) (12h -12h30)
- Resting, watching TV (in living room) (13h– 16h)
Introduction: motivations

Increase independence and quality of life

- Enable elderly to live longer in their own home
- Maintain autonomy of elderly
- Reduce costs for public health systems
- Relieve family members and caregivers

ADLs monitoring at home

- Detect alarming situations (e.g. falling down)
- Calculate the degree of frailty of elderly people
- Detect changes in behavior
Introduction: objectives

Propose a new cognitive vision approach using ambient sensor technologies to recognize interesting activities at home;

Based on multisensor analysis and human activity recognition, it consists in:

• Detecting people
• Tracking people as they move
• Recognizing postures and activities of interest

Two main hypotheses have been considered

1- Fixed video cameras:
   • Video cameras are fixed on a wall without pan, tilt and zoom

2- Monitoring one individual:
   • One elderly person living alone in his/her own home is monitored
State of the art: sensing modalities

Fixed On

The Environment
- Motion sensors
- Contact sensors
- Water sensors
- Presence sensors
- Light sensors
- Electrical sensors
- Video sensors
- Audio sensors

The Person
- Heart Rate
- Pulse Oxymeter
- ECG
- EEG
- Accelerometers
- Pressure sensors

To Identify

Location
Interaction with Household Objects
Amount of Movement
Biomedical Function

Activities Performed
Health and Wellbeing’s
State of the art: industrial systems for monitoring ADLs

QuietCare system [Glascock and Kutzik, 2006]
- Wireless motion sensors
  - Learns the daily living activities of elderly
- Detects only motion

“i-pot” system [ipot, 2005]
- Electrical kettle
  - Detects a sudden change in an elderly person’s tea habits
- Detects only one activity

Social alarms “Vivago system” [Sarela et al., 2003]
- Long-term monitoring of circadian rhythm
  - 27-40% of users do not wear the alarm device on daily basis
State of the art: research projects for monitoring ADLs

Wearable and environmental sensors

**RFID glove** [Philipose et al., 2004]
- RFID: Radio Frequency Identification

**RFID bracelet** [Tapia et al., 2006]
- Good recognition of several human activities
  - Need to wear a glove or a bracelet
  - All objects have to be tagged (lot of sensors are needed: > 100)
State of the art: research projects for monitoring activities

Video cameras

Probabilistic and stochastic approaches: (e.g. NNs, HMM, …)

+ Easy to implement using HMM
- Difficult to modify or to add a priori knowledge

Assist person with dementia using video camera [Hoey, et al. ICVS 2007]

+ Good recognition of hand washing activity
- Limited variety of activities

Constraints resolution approaches: [Vu, et al. IJCAI 2003]

- For video surveillance applications (e.g. metro station)
+ Based on video event models and spatio-temporal constraints
- Support only video cameras (recognize only video events)
Approach overview

The proposed approach

• Based on an extended constraint resolution method:
  + Recognize multisensor activities with spatio-temporal constraints (support video and non video sensors)
  + Recognize ADLs (healthcare applications)

Main characteristics of the approach:

  + No wearable sensors: the elderly do not wear the device
  + Few sensors: 28 sensors vs. 100 sensors for [Philipose et al., 2004] and [Tapia et al., 2006]
  + Only house furniture were equipped with sensors
  + More than one activity at home: 100 vs. 1 for [Hoey et al. 2007]
Approach overview (2)

A PRIORI KNOWLEDGE BASE

3D models of person and models of events

3D model of the scene

VIDEO ANALYSIS

Person Detection → Person Tracking → Posture Detection

SENSOR ANALYSIS

Sensor Processing and Modeling → Sensor Processing and Modeling

EVENT RECOGNITION

Video Event Recognition

Environmenta l Event Recognition

ENVIRONMENTAL EVENT FUSION

Multisensor Event Fusion

Activity Recognition

Log-files (XML files)

Alarms

3D Visualization
Video analysis

Input: video stream
Output: tracked objects with their 3D postures

Person detection → Person tracking → Posture detection

- Motion detection
- Person classification

[Avanzi et al. 2005] & [Zuniga et al. 2006] → [Boulay et al. 2006]
Video analysis: posture detection

1. Input
2. Detected silhouette
3. Camera parameters
4. Virtual camera
5. 2D silhouettes comparison
6. Detected posture

- 3D posture models
- 3D position
- 3D silhouette generator
- Generated silhouettes
Video analysis: ten 3D key human postures

- Standing
- Standing with arm up
- Standing with hands up
- Bending
- Sitting on a chair
- Sitting with flexed legs
- Sitting with outstretched legs
- Slumping
- Lying with flexed legs
- Lying with outstretched legs
A PRIORI KNOWLEDGE BASE

3D models of person and models of events

3D model of the scene

VIDEO ANALYSIS

INPUT DATA

Video Sensors
Camera 1
Camera n

Person Detection → Person Tracking → Posture Detection

EVENT RECOGNITION

Video Event Recognition

MULTISENSOR EVENT FUSION

Multisensor Event Fusion

Activity Recognition

Log-files (XML files)
Alarms
3D Visualization

SENSOR ANALYSIS

INPUT DATA

Environmental Sensors
Sensor 1
Sensor m

Sensor Processing and Modeling

Environmental Event Recognition

OUTPUT DATA

40 ans - la révolution de l'information

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE
SOPHIA ANTIPOLIS - MÉDITERRANÉE

INRIA
Sensor analysis: method

Collecting information about interactions between people and the contextual objects.

Environmental sensors: binary sensors
  - Give the status $S$ for $N$ physical objects
  - Each sensor has two binary values: $s = 0$ and $s = 1$, representing "sensor Off" and "sensor On" respectively
  - Not always reliable sensors (e.g. wrong value)

Two probabilities are proposed to verify an occurred observation $O$ for $k$ sensors:
  - $P_d$ (probability of detection)
    \[ P(O_k = 1|S_s = 1) = P_d \]
  - $P_f$ (probability of false alarm)
    \[ P(O_k = 1|S_s = 0) = P_f \]
Each sensor data is represented by

$$O = \langle Id; c; x; t; m; y; \Delta y \rangle$$

- **Id**: sensor identifier
- **c**: sensor class (e.g. contact, pressure)
- **x**: sensor location (3D position)
- **t**: time when the physical property is measured
- **m**: sensor mode (e.g. continuous, by event, on request)
- **y**: value of the physical property as measured by the sensor
- **Δy**: potential error in sensor observation ($P_d$, $P_f$)
Event recognition: event description language

Four types of event are defined:

- Primitive states: perceptual property of one or several physical objects
- Composite states: combination of primitive states
- Primitive events: change of primitive state values
- Composite events: combination of primitive/composite states and/or primitive/composite events

An event is mainly constituted of four parts:

- **Physical objects**: all objects involved in an event E: i.e. mobile objects, contextual objects, zones of interest
- **Components**: list of states and sub-events involved in an event E
- **Constraints**: symbolic, logical, spatial and temporal relations between the physical objects and/or the components to be verified: e.g. before, after, duration, during, in, close to, ...
- **Alert**: a set of actions to be performed when an event E is recognized (Not-Urgent, Urgent and Very-Urgent).
Event recognition: event model

Syntax used to define an event model is:

```plaintext
EventType ( event name,
    PhysicalObjects ( (name of physical object: type of physical object) )
    Components ( (component 1)
        (component n) )
    Constraints ( (condition 1)
        (condition n) )
    Alert ( AText ("text alarm to display")
        AType ("type of the alarm") )
)
```

The event type can be a primitive state, a primitive event, a composite state, a composite event
A PRIORI KNOWLEDGE BASE

3D models of person and models of events

3D model of the scene

VIDEO ANALYSIS

Person Detection → Person Tracking → Posture Detection → Video Event Recognition

EVENT RECOGNITION

Video Event Recognition

MULTISENSOR EVENT FUSION

Multisensor Event Fusion

Activity Recognition

Log-files (XML files) Alarms 3D Visualization

INPUT DATA

Video Sensors

Camera 1

Camera n

Environmental Sensors

Sensor 1

Sensor m

SENSOR ANALYSIS

Sensor Processing and Modeling

Sensor Processing and Modeling

OUTPUT DATA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN INFORMATIQUE

INRIA

40 ans la révolution de l'information

SOPHIA ANTIPOLIS - MÉDITERRANÉE
Event recognition: new knowledge base of event models

For healthcare applications, I propose 100 event models for daily activities:

✔ **58 customized video event models**, among them 26 posture-based events, related to the human postures (e.g. standing) and transitions in human postures (e.g. standing up, sitting down, fainting, falling down)

✔ **26 new environmental event models** related to the data provided by the environmental sensors and to the status of each equipment in the scene

✔ **16 new multimodal event models** related to the complex activities at home, e.g. preparing meal, eating meal, using kitchen equipments

The proposed knowledge base is reusable in other homecare applications.
Event recognition: a priori knowledge

For homecare applications, I propose a 3D model of the scene:

- A 3D referential (calibration matrix and positions of video cameras)
- A 3D positions of environmental sensors
- Geometric areas to different rooms in the scene (e.g. kitchen, bedroom)
- Geometric zones of different zone of interest (e.g. cooking-zone, sleeping-zone)
- List of walls (e.g. kitchen right wall)
- List of different equipment (e.g. fridge, stove)

3D model of the person:

- 3D width
- 3D height
- 3D depth
Event recognition: video event models

Fainting event model

Two types of fainting situation:
• fainting without loss of balance
• fainting with loss of balance

PrimitivEvent ( PersonFainting_without_Loss_Balance,
PhysicalObjects ( (p : Person) )
Components ( (c1: PrimitiveState Standing(p))
(c2: PrimitiveState Bending(p))
(c3: PrimitiveState SittingFlexedLegs(p))
(c4: PrimitiveState SittingOutstretchedLegs(p)) )
Constraints ( (c1; c2; c3; c4)
(c4 Duration >= threshold) )
Alert ( AText ("Person is Fainting ")
AType ("URGENT") )

➢ 1 physical object (person)
➢ 4 components (human postures)
➢ 2 constraints:
  • Sequential order
  • Duration of an event
➢ 1 urgent alert

The threshold is calculated using 15 annotated fainting events
Event recognition: video event models

Falling down event model

Three models of falling down event:

- 1: Standing, sitting on the floor and lying on the floor with outstretched legs
- 2: Standing, bending and lying on the floor with outstretched legs
- 3: Standing, sitting on the floor and lying on the floor with flexed legs

Falling down 1
Event recognition: multimodal event models

• 16 multimodal event models:
  Using fridge, cupboards, drawers, microwave, stove, telephone, watching TV, washing dishes, slumping in an armchair, eating meal, preparing breakfast, preparing lunch, preparing dinner, warming up meal, preparing cold meal, and preparing hot meal

Each event E is modeled with sub-events related to objects involved in the event E.

Example: “preparing lunch”

*IF* a person is located close to the countertop in the kitchen for a long period AND (a person accesses to meal ingredients (e.g. fridge, cupboards) AND a person accesses plates or utensil cupboards) AND a person uses an appliance (e.g. microwave, stove) for a long period *THEN* a meal is prepared.
Multisensor event fusion

**Heterogeneous sensors:** fusion at the event level

- **Input:** video and environmental events + a priori knowledge
- Event processed to synchronize them
- Synchronized events
- Understand which events are occurring

An event $E$ is recognized at an instant $t$, if:

- all its components have been recognized,
- its last (using the temporal order) component being recognized at the given instant $t$
Multisensor event fusion: event synchronization

To cope with the different sensor measurement frequencies:

- **synchronization**

Necessary for different sensors to have:

- **The same temporal referential:** timestamp synchronization
  - We define the measurement frequency $f_{sensor}$: the number of times the sensor provides data to the multisensor fusion system per second
  - We define the measurement latency $T_{L,sensor}$: the measurement acquisition time and the measurement processing time (delay between sent data and the received data)

- **The same spatial and semantic referential:**
  - We propose a 3D model of the scene (geometric + sensors + semantic)
Multisensor event fusion: Dempster-Shafer

Dempster-Shafer Theory: generalization of traditional probability allows to better quantify uncertainty

Basic concepts for DST:

• Frame of discernment of sensors
  — e.g. two states of environmental sensors: Off and On
  \[ \Theta = \{0,1\} \]

• Mass function: distribution of belief
  \[ P(\Theta) = 2^\Theta \rightarrow [0,1] \]

• Belief and plausibility: lower and upper bounds of probability
  \[ \text{Bel}(A) = \sum_{B \subseteq A} m(B) \quad \text{Pls}(A) = \sum_{B \supseteq A} m(B) \]

• Uncertainty:
  \[ \delta (A) = \text{Pls}(A) - \text{Bel}(A) \]
Multisensor event fusion: sensor uncertainty

Dempster-Shafer theory for evidential reasoning

- Handle sensor measurement errors

Example:

The person which drops his/her bag on the chair, may activate the chair sensor and give a false result

- How to detect if the person or the bag is on the chair (by using pressure sensors and video sensors)

Statistics of 20 video sequences of one human actor, show that pressure sensors work at 70%, and video sensor works at 75%

1- mass function
2- belief and plausibility
Multisensor event fusion: uncertainty of “sitting on chair”

Mass function for each sensor:

\[ m_{VSCair}(\{\text{Person}\}) = 0.75 \quad \text{and} \quad m_{VSCair}(\{\neg \text{Person}\}) = 0.25 \]

\[ m_{PSChair}(\{\text{Person}\}) = 0.70 \quad \text{and} \quad m_{PSChair}(\{\neg \text{Person}\}) = 0.30 \]

Summing up mass functions:

\[ m_{PSChair,VSCair}(\text{Person}) = \frac{1}{2} \left( m_{PSChair}(\text{Person}) + m_{VSCair}(\text{Person}) \right) \]

\[ = \frac{1}{2} (0.70 + 0.75) = 0.72 \]

\[ m_{PSChair,VSCair}(\neg \text{Person}) = \frac{1}{2} \left( m_{PSChair}(\neg \text{Person}) + m_{VSCair}(\neg \text{Person}) \right) \]

\[ = \frac{1}{2} (0.30 + 0.25) = 0.27 \]
Multisensor event fusion: uncertainty of “sitting on chair”

Belief ($\text{Bel}$) and Plausibility ($\text{Pls}$):

$$\text{Bel}(\text{Person \_ Sit \_ Chair}) = m_{\text{PSChair}, \text{VSChair}}(\text{Person}) = 0.72$$

$$\text{Pls}(\text{Person \_ Sit \_ Chair}) = m_{\text{PSChair}, \text{VSChair}}(\text{Person}) + m_{\text{PSChair}, \text{VSChair}}(\neg \text{Person})$$

$$= 0.72 + 0.27 = 0.99$$

Calculate uncertainty:

$$\delta (\text{Person \_ Sit \_ Chair}) = \text{Pls}(\text{Person \_ Sit \_ Chair}) - \text{Bel}(\text{Person \_ Sit \_ Chair})$$

$$= 0.99 - 0.72 = 0.27$$

• A person is sitting on the chair with belief = 0.72
Evaluation and results

Experimental site: Gerhome laboratory

• GERHOME (Gerontology at Home) : homecare laboratory
  http://www-sop.inria.fr/orion/personnel/Francois.Bremond/topicsText/gerhomeProject.html

• Experimental site in CSTB (Centre Scientifique et Technique du Bâtiment) at Sophia Antipolis http://gerhome.cstb.fr

• Partners: INRIA, CSTB, Nice Hospital, CG06…
Evaluation and results: Gerhome laboratory

- Technical solutions to help the elderly people to stay at home
- Typical apartment of an elderly person living alone
- 41 m² with entrance, bedroom, bathroom, living room, and kitchen

<table>
<thead>
<tr>
<th>SENSORS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light, humidity &amp; temperature sensors</td>
</tr>
<tr>
<td>Presence sensor</td>
</tr>
<tr>
<td>Pressure sensor (bed, chairs, ...)</td>
</tr>
<tr>
<td>Water sensor</td>
</tr>
<tr>
<td>Electrical sensor</td>
</tr>
<tr>
<td>Doors &amp; windows sensors</td>
</tr>
<tr>
<td>Open/close sensor</td>
</tr>
<tr>
<td>Video camera</td>
</tr>
<tr>
<td>PC</td>
</tr>
</tbody>
</table>
Evaluation and results: sensors in Gerhome laboratory

- **4 fixed video cameras** (1 in kitchen, 2 in livingroom, 1 in bedroom)

- **24 environmental sensors** (contact sensors in cupboard doors, fridge doors and drawers; water flow sensors in water pipes in kitchen and bathroom; electrical sensors in outlets (stove, microwave, TV), pressure sensors under chairs, armchair and bed; & presence sensors near sink, stove and washbowl)
Evaluation and results: evaluation metrics

Classical metrics:

• **Precision P and sensitivity S**
  
  With TP: true positive, FP: false positive, FN: false negative
  
  \[
P = \frac{TP}{TP + FP} \quad S = \frac{TP}{TP + FN}\]

• **F-score**
  
  \[
  F = \frac{2P \times S}{P + S}
  \]

New specific metrics: comparison between 2 elderly people

• **Normalized Difference of Mean durations of Activity (NDA)**
  
  With m1 and m2 represent respectively mean durations of a certain activity of 2 people
  
  \[
  NDA = \frac{|m1 - m2|}{m1 + m2}
  \]

• **Normalized Difference of Instance number (NDI)**
  
  With n1 and n2 represent respectively instance number of a certain activity of 2 people
  
  \[
  NDI = \frac{|n1 - n2|}{n1 + n2}
  \]
Evaluation and results: 2 experiments

Experiment 1: with one human actor

- 15 video sequences at 10 fps
- Duration of each video sequence is about 20 minutes (about 9600 frames)
- 10 normal activities (e.g. using house equipment) and 2 abnormal activities (e.g. fainting and falling down) have been tested in Gerhome laboratory

Goal: test the **functionality** of the sensors and detect **abnormal activities** ( alarming situations )

Experiment 2: with 14 elderly volunteers

In collaboration with Nice hospital and the CSTB, 14 elderly volunteers (i.e. 6 women and 8 men aged from 60 years to 85 years) were recruited and were asked to perform a set of household activities

- 56 video sequences at 10 fps (4 cameras x 14 volunteers)
- Duration of each video is about 4 hours and each video contains about 144 000 frames
- A set of daily activities (e.g. using kitchen equipment, preparing meal) has been tested in Gerhome laboratory
- Dataset of 9 elderly among the 14 elderly people have been analyzed (5 datasets with ground truth)

Goal: analyze behavioral profile and calculate the **degree of frailty** for each volunteer
Evaluation and results: video-based evaluation (experiment 1)

- Good recognition of a set of activities and human postures (video cameras)

<table>
<thead>
<tr>
<th>States / Events</th>
<th>Ground truth</th>
<th>TP</th>
<th>FN</th>
<th>FP</th>
<th>Precision (%)</th>
<th>Sensitivity (%)</th>
<th>F-score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the kitchen</td>
<td>45</td>
<td>40</td>
<td>5</td>
<td>3</td>
<td>93</td>
<td>88</td>
<td>90</td>
</tr>
<tr>
<td>In the livingroom</td>
<td>35</td>
<td>32</td>
<td>3</td>
<td>5</td>
<td>86</td>
<td>91</td>
<td>88</td>
</tr>
<tr>
<td>Standing</td>
<td>120</td>
<td>95</td>
<td>25</td>
<td>20</td>
<td>82</td>
<td>79</td>
<td>80</td>
</tr>
<tr>
<td>Bending</td>
<td>92</td>
<td>66</td>
<td>26</td>
<td>30</td>
<td>68</td>
<td>70</td>
<td>69</td>
</tr>
<tr>
<td>Sitting</td>
<td>80</td>
<td>58</td>
<td>22</td>
<td>18</td>
<td>76</td>
<td>72</td>
<td>74</td>
</tr>
<tr>
<td>Slumping</td>
<td>35</td>
<td>25</td>
<td>10</td>
<td>15</td>
<td>62</td>
<td>71</td>
<td>66</td>
</tr>
<tr>
<td>Lying</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>66</td>
<td>66</td>
<td>66</td>
</tr>
<tr>
<td>Standing up</td>
<td>57</td>
<td>36</td>
<td>21</td>
<td>6</td>
<td>85</td>
<td>63</td>
<td>72</td>
</tr>
<tr>
<td>Sitting down</td>
<td>65</td>
<td>41</td>
<td>24</td>
<td>8</td>
<td>83</td>
<td>63</td>
<td>71</td>
</tr>
<tr>
<td>Sitting up</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>80</td>
<td>66</td>
<td>71</td>
</tr>
</tbody>
</table>

- Errors occur at the border between livingroom and kitchen
- Mixed postures such as bending and sitting due to segmentation errors
Evaluation and results: recognition of the “fainting” event

The person (human actor) is recognized with the postures "standing", "bending", "sitting on the floor with flexed legs", "sitting on the floor with outstretched legs".
Evaluation and results: recognition of the “falling-down” alarming situation

The person (human actor) is recognized with the postures "standing", "sitting on the floor with flexed legs", and "lying on the floor with outstretched legs".
Evaluation and results: comparison between two elderly volunteers

Volunteer 1: man of 64 years and volunteer 2: woman of 85 years

Big values = big difference in profile
Evaluation and results: comparison between two elderly volunteers

Volunteer 1: man of 64 years and volunteer 2: woman of 85 years

- Different profiles between the 2 volunteers
  - Greater ability for the volunteer 1 to live in Gerhome laboratory
Evaluation and results: recognition of the “eating meal” activity

The elderly person is recognized “sitting on a chair in the livingroom” for a long time after preparing meal
Evaluation and results: Dempster-Shafer (DS), (experiment 2)

Using **DS uncertainty** shows some **improvements** in the recognition of activities (with about 20%)

![Bar chart showing F-score comparison with and without DS uncertainty]

- **Sitting on an armchair**: F-score without DS uncertainty = 63%, F-score with DS uncertainty = 82%
- **Sitting on a chair**: F-score without DS uncertainty = 82%, F-score with DS uncertainty = 93%
- **Using stove**: F-score without DS uncertainty = 80%, F-score with DS uncertainty = 91%
- **Using fridge**: F-score without DS uncertainty = 80%, F-score with DS uncertainty = 91%
Evaluation and results: leave-one-out cross validation

To bring out the possible differences in the behaviors of the 9 volunteers: compare behavioral profile of 9 elderly volunteers by using the leave-one-out cross validation algorithm on the activity duration

\[ MD_{Ei,Pj} = \frac{\sum_{Pk \in P, Pk \neq Pj} d_{Ei,Pk}}{R}, \forall Pj \in P \]

\[ \sigma_{Ei,Pj} = \sqrt{\frac{1}{R} \sum_{K=1}^{R} d_{Ei,Pk}^2 - MD_{Ei,Pj}^2} \]

- \( MD_{Ei,Pj} \): mean duration for an event \( Ei \) for each person without a person \( Pj \);
- \( d_{Ei,Pk} \): duration for each event \( Ei \) for each person \( Pk \);
- \( P = \{P1; P2; P3; P4; P4; P5; P6; P7; P8; P9\} \)
- \( R \): represents the number of the training set of data (R=8)
- \( \sigma_{Ei,Pj} \): represents the standard deviation for each event \( Ei \) of each person without a person \( Pj \);
Volunteer 9 (women of 85 years) has a fairly different profile from the others (e.g. some inabilities in using stove)

Evaluation and results: behavioral profile (experiment 2)

Comparison between 9 elderly people (using stove event)
Evaluation and results: behavioral profile (experiment 2)

Comparison of 5 event durations between the volunteer 9 with the mean durations for the other 8 volunteers

![Bar chart showing event durations](chart.png)
Conclusion and future work

A new approach based on **multisensor analysis** by combining video events with environmental events to **recognize interesting activities at home**

✔ Adapted to healthcare applications

The approach:

➢ Good results in the recognition of a set of daily activities at home

➢ Good results in detecting **alarming situations** (e.g. falling down)

➢ Allows us to calculate the **degree of frailty** of an elderly person

➢ About **changes in behaviors** for an elderly person, we need a lot of data during 6 months for an elderly person
Conclusion and future work (2)

Contributions

- A knowledge base of human activities is proposed: 100 events have been modeled
  - 58 video events, 26 environmental events and 16 multimodal events

- An extended constraint-based approach for activity recognition
  - Non-video sensors
  - Multisensor fusion at the event level
  - Uncertainty in sensor measurements

- A performance evaluation of the approach has been done in a real world environment with real elderly people

- A new dataset (available on the web) of 14 elderly people performing a set of household activities has been proposed: 224 hours of video stream (14 people x 4 hours x 4 cameras) and 14 log-files of environmental data
Conclusion and future work (3)

Limitations

- **Sensors**: used environmental sensors give only **coarse information**
  - not possible to **infer which food item is removed** from the fridge by simply considering the current state of the fridge door

- **Environment**:
  - Gerhome laboratory was not the volunteer real home, volunteer behavior was not completely natural
  - Missing some activities (e.g. activities taking place in the bathroom)

Future work

- **Learning event models**:
  - Learn automatically normal behavior models of everyday data

- **Extending the approach for several people living together**:
  - Manage when several people trigger the same set of sensors
  - Extend the proposed knowledge base of event models

- **Other environments**:
  - In nursing homes with healthy and frail elderly
  - In hospital environment with different patients with different diseases (e.g. Alzheimer)
List of the publications

International Journal:


World Congress:


International Conferences:


THANK YOU FOR YOUR ATTENTION