



HAL
open science

Introducing Knowledge in the Process of Supervised Classification of Activities of Daily Living in Health Smart Homes

Anthony Fleury, Norbert Noury, Michel Vacher

► **To cite this version:**

Anthony Fleury, Norbert Noury, Michel Vacher. Introducing Knowledge in the Process of Supervised Classification of Activities of Daily Living in Health Smart Homes. IEEE HealthCom 2010 - 12th International Conference on E-health Networking, Application & Services, Jul 2010, Lyon, France. pp.322-329. hal-00503242

HAL Id: hal-00503242

<https://hal.science/hal-00503242>

Submitted on 18 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.

Introducing Knowledge in the Process of Supervised Classification of Activities of Daily Living in Health Smart Homes

Anthony Fleury

Univ Lille Nord de France,
F-59000 Lille, France
EMDouai, IA
F-59500 Douai, France

Email: Anthony.Fleury@mines-douai.fr

Norbert Noury

INL-INSA Lyon Lab. Team MMB
UMR CNRS/ECL/INSA/UCBL 5270
Av. Einstein, F-69621 Villeurbanne, France
& TIMC-IMAG Lab. Team AFIRM
Faculté de Médecine de Grenoble
Email: Norbert.Noury@insa-lyon.fr

Michel Vacher

LIG Laboratory
Team GETALP
UMR CNRS/UJF/INPG/UPMF 5217
385, avenue de la Bibliothèque - B.P. 53
F-38041 Grenoble Cedex 9, France
E-mail: Michel.Vacher@imag.fr

Abstract—Telemedicine and Telemonitoring of elderly people is an actual challenge that is explored to prevent some problems linked to the constant growing of the mean age of the population. It requires to recognize the behavior and the actions of a person inside his own home with non-intrusive sensors and to process data to check the evolution of the person. Activities of Daily Living can be learned and automatically recognized using supervised classification on sensor data. This paper presents the results of the study of prior introduction, in Support Vector Machine, to improve this automatic recognition of Activities of Daily Living. We started from a set of data acquired in daily life during an experimentation in the Health Smart Home of the TIMC-IMAG Lab. From this restricted set of data, we obtained models for seven activities of Daily Living and test, with leave-one-out method, the performance of this classification. This first step gave baseline results that this paper tends to improve using consistent priors to compute more specific and accurate models of the different activities that are learned and obtain better results on the leave-one-out method on the sensors data.

Index Terms—ADL ; SVM ; Health Smart Home ; Knowledge Introduction

I. INTRODUCTION

Miniaturization and low prices of sensors make Ubiquitous Environments and sensor networks easiest to deploy. However, this leads to new challenges for the process of the data acquired on these networks and the information extraction among them.

One of the field that uses these possibilities of acquiring data is telemedicine. Indeed with the evolution of the world population, we face a lack of institution and one of the solution could be to remotely monitor elderly people to detect, as early as possible, a dangerous evolution of the state of the person. This can be achieved in multiple ways, including Health Smart Homes [1] that uses sensors integrated in the environment of the person to analyze his activity during long-term measurements. This leads to different applications, but the one that we would like to favor with this work is an assistance to the geriatricians to complete the autonomy evaluation of the person using scales like Activities of Daily

Living (ADL) introduced by Katz [2]. Such index evaluates the autonomy by considering the different activities and for each one, considering if it can be performed without assistance or if the person needs partial or complete assistance. These are filled during interrogation of the patient by the geriatricians and consider only the subjective view of the elderly people and his family.

To fill in such grid automatically and help the diagnostic of geriatricians, we should first consider to automatically recognize the ADL that is currently performed by the person in his home. For this purpose, different automatic recognition techniques and sensor configurations can be used. For instance, Philipose *et al.* [3] explored the thematics using RFID tags on a large number of objects (more than a hundred) and tried to infer the activity that was performed considering the objects touched by the person using a glove with a RFID receiver and also a Gaussian curve describing the mean time of execution of each activity. They considered 14 activities and made experimentations on 14 persons. Dynamic Bayesian Networks gave them 88% of global detection accuracy. Hong *et al.* [4] also used RFID tags on foods and objects to create models for the activities “preparing a drink (cold or hot)” and “Hygiene”. Using Dempster-Shafer Theory, they presented the values of the belief and mass functions that show the possibility to distinguish between both activities with this configuration. For this work, Nugent *et al.* [5] also tested the impact of sensor failures on recognition using the evidential theory. Kröse *et al.* [6], in the CARE project, also tried to differentiate between two activities (“going to the toilets” and “exit from the flat”) using the data from a lot of sensors (switch, environmental, etc.) and considering Hidden Markov Models for classification. They achieved promising results and presented them for two elderly people (contrary to other studies made on young individuals). Finally, Tsukamoto *et al.* [7] and Berenguer *et al.* [8] tested the use of electrical signatures to detect different various activities of daily living. Indeed, by using pattern recognition on the electrical network

it is possible to infer what materials are being used, and when they are turned on and off. In the last paper, the authors presented the detection of the activity “take a meal” on 18 aged people whose flat were monitored.

In our previous works [9], we used Support Vector Machines (SVM) with a minimum set of sensors to recognize seven Activities of Daily Living. The AILISA project [10], that used presence infra-red sensors to monitor the activity of the person, was the starting point for our research. Our experimentations took place in a modified version of the flat set-up. Using microphones, temperature and hygrometry, a wearable sensor, presence infra-red sensors and finally on/off switches for the use of three commodities, we achieved **13.8%** of Global Error Rate (GER) for the seven activities performed by 13 subjects.

In this paper, we consider the dataset acquired for this previous work and we try to introduce prior knowledges in our training dataset, to learn more accurate models. In the following, section II describes the Health Smart Home used for this experimentation, the flat set-up and also the dataset and its contents. Section III then introduces the use of Support Vector Machines for ADL classification, the different kind of knowledge that we considered and for each one the results of the classification on our dataset. Finally, section IV discusses these results and concludes the paper.

II. HEALTH SMART HOME FOR ACTIVITIES OF DAILY LIVING RECOGNITION

A. The Grenoble Health Smart Home - TIMC-IMAG Lab

In 1999, researchers of the TIMC-IMAG laboratory created the AFIRM Team, devoted to the application of Technologies of Information and Communication to Medical applications and especially to Telemedicine. During the AILISA project, several flats have been equipped with Infra-Red Presence sensors and a computer in charge to transmit data to our lab. This allowed to obtain years of activity data in the flats of elderly people (real flats and hospital suits). One of these flats have been installed in the Faculty of Medicine of Grenoble. This flat was the basis of this work.

Even if there are real flats set-up with elderly people living inside, we used, for our experiments, the flat that is in the Faculty of Medicine of Grenoble. Indeed, as it is very intrusive, only this flat includes webcams that allow us to record the activities and test supervised learning algorithms.

The sensors included in the flat for this work are described in the following section and shown on Figure 1.

1) *Infra-Red Presence Sensors*: These sensors detect all the movements that occur in a defined cone. A message is then sent using their wireless connection at each event. They are placed in the flat to monitor specific regions that are of interest, for instance the bed, the table of the kitchen, etc. Between two detections, we can consider that the person did not change of room. With this assumption, we are able to complete this signal. These detectors are all connected to the CAN bus of the flat and the AILISA project software is in charge to store the received data.

2) *Microphones*: Seven microphones are placed all around the flat and connected to a computer equipped with a National Instrument card in a computer. Each channel of the card is analyzed continuously to detect the beginning and the end of an occurring sound. Then, when an event is detected, it is classified either as sound or as speech and sent to a classifier that determines which kind of sound it is (one of eight classes, e.g. step, dishes, door clap, door lock) in the first case or translates the sound into the five most probable sentences in case of speech. This complete system, named AuditHIS, is described more precisely and tested individually in the flat in [11].

3) *Environmental Sensor*: This sensor is placed in the bathroom and gives temperature and hygrometry measured every 5 minutes.

4) *Actimetry*: An home-made sensor containing accelerometers and magnetometers is placed under the left armpit of the person, in a pocket specifically designed on the shirt, and is used to determine, at each time, the posture of the person (sit, stand, lie down) and if he is in a walking period [12].

5) *Door contacts*: Three contacts are placed on the fridge, the cupboard, and chest of drawers to detect the beginning and end of their respective use.

6) *Webcameras*: In addition to that, five wide-angle webcams recorded the actions in the flat. These cameras are used only for indexation. They are not processed as another sensor.

B. Activities of Daily Living

Based on the both ADL International Scale and AGGIR french one used by geriatricians, we defined seven activities of daily living that we would like to monitor because of their interest considering the autonomy of the person. To test this, a protocol, in which the subject entered alone the flat and live his life as desired, was used. The only requirement was to perform at least once each activity. No notion of minimum time of execution was given so that the activity was as natural as possible. Moreover, the flat was deeply presented to the person before the beginning of the session so that he felt like home.

The seven chosen activities are described in the following sections.

1) *Sleeping*: A bed was available in the bedroom for the individual to sleep as long as necessary.

2) *Preparing and having a breakfast*: The fully equipped kitchen also contained material and foods necessary for breakfast. Everything was available for the individual to choose and to prepare in his own way; he would then clean up the kitchen and do the dishes.

3) *Dressing and undressing*: Clothes were available for this activity.

4) *Resting*: This activity was the broadest one. The individual could do whatever he wants and enjoys doing during his leisure time at home. He could read a book or a magazine, listen to the radio, or watch the TV, etc.

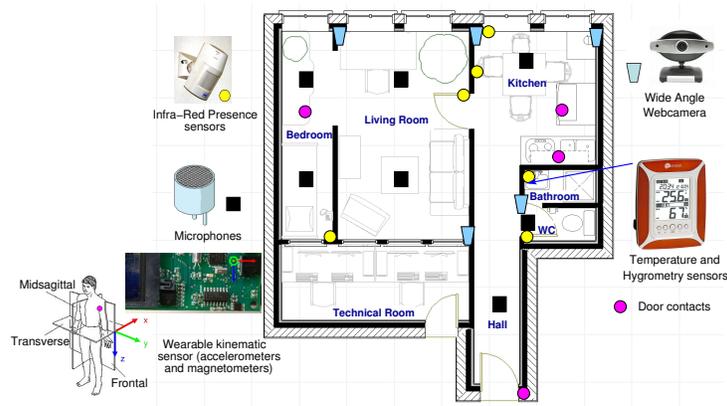


Fig. 1. Map of the demonstration flat of the TIMC-IMAG Laboratory with localization of the different sensors

5) *Hygiene activities*: During this activity the individual was inside the bathroom and performed normal hygiene activities. It was difficult to ask the individuals to take a shower for a first experiment, thus we only asked them to wash their hands and teeth. To respect privacy, neither the bathroom nor the toilets were recorded on video. We asked the individuals to close the door completely or partially when they were respectively in the toilets or in the bathroom, so as to differentiate the activities.

6) *Bowel movement*: For this activity, the subject was in the toilets.

7) *Communication*: This last activity consisted in answering a phone call and having conversations. In our protocol, the subject was called five times on the phone and has to answer with given previously created phone conversations which were randomly selected.

C. Data Acquired

This study brought validation in real conditions, in the health smart home of the TIMC-IMAG previously presented. For this validation, thirteen participants (including six women), young and healthy (mean age 30.4 years \pm 5.9, 24-43 min-max), were enrolled. While the experimentation runs, they have no contact with the supervisors.

The data from this experimentation are used in supervised classification algorithms. It is then mandatory to construct the training and testing datasets. To do so, the experimentations were all recorded using the webcams of the laboratory and manually labeled afterward using these video records. For each experimentation, an XML file describing the complete execution were created. It contains the information on all the sensors (e.g. synchronization between video and actimetry, paths for the files of each sensor) and on the activities performed (time of beginning and end, label). From these data, we worked with fixed size time frames without overlapping. These windows are of length 3 minutes (that corresponds to the mean time to correctly execute the shortest activity). For instance, an activity that lasts 19 minutes will give us 6 windows of 3 minutes and the last minute will be dropped.

Table I gives the organization of the training dataset from this experimental protocol.

The validation method that has been chosen is the leave-one-out (due to the weak number of samples). Data are normalized (centered and reduced) at each execution of the algorithm. The coefficients of normalization are determined using the training database (without the tested sample) and the last sample is normalized using the previously computed coefficients.

TABLE I
REPARTITION OF THE DIFFERENT CLASSES IN THE LEARNING DATASET

Class	Training size	Repartition
<i>Sleeping</i>	49	21.2%
<i>Resting</i>	73	31.5%
<i>Dressing</i>	15	6.4%
<i>Eating</i>	45	19.4%
<i>Elimination</i>	16	6.9%
<i>Hygiene</i>	14	6%
<i>Communication</i>	20	8.6%
Total	232	100%

D. Features

In order to automatically classify activities, we must find indications, computed from the different sensors and named features, whose value will be discriminative between two different activities and equivalent between two realizations of the same activity, even from two different participants.

These features are summed-up in the table II. They have been selected from previous experimentations that allowed to select the most interesting ones. These features compound a vector that will represent the 3 minutes window of the learning or testing samples.

III. SVM CLASSIFICATION OF ACTIVITIES OF DAILY LIVING WITH AND WITHOUT PRIOR KNOWLEDGES

The goal of our project is, in this condition, to classify activities of daily living automatically in an health smart home. For this classification, we studied different possibilities of algorithms and finally selected Support Vector Machines. As in

TABLE II
SUM-UP OF THE DIFFERENT MODALITIES WITH SELECTED FEATURES AND THE ACTIVITY FOR WHICH IT IS INFORMATIVE

Modality	Features selected
Actimeter	Percentage of time spent in the different posture (stand, sit, lie) and walking
Microphones	Number of events per class and number of events per microphones
PIR	Percentage of time in each rooms and number of events for each detectors
Door contacts	Percentage of time in "open" position and predominant state (open/close) in the considered time slot
Environmental	Differential measure for the last 15 minutes for temperature and hygrometry

all classification problems, we first extracted the most relevant features for our problem, and then we learned the models before testing them on new data.

Our first choice to achieve this goal was to construct generic models of activities [9] in order to be non-specific and being able to adapt to every situations. However, generic models give, in general, lower performances. One of the question raised by this study was to know whether it could be interesting to adapt the models to the knowledge that we have on ways to perform activities or to keep these models.

In this section, we will present the results of the classification of activities of daily living in four different conditions. The first one is the generic models that are remembered here. All these conditions use support vector machines for classification and the features are those presented in the previous section. For all these algorithms, the results presented are from the leave-one-out validation method on the dataset previously introduced.

A. Classification of Activities with SVM

Numerous methods are available to classify samples from training data. In our case, the weak number of available samples makes the learning process difficult for a large number of them (Bayesian, Neural Networks...). For this reason, we decided to test the SVM (Support Vector Machines) method that seem to fit more to our problem and that can be used for training with small sets of data. Considering two classes of points, labeled -1 and 1 ; and a set of N vectors $\mathbf{x}_i \in X \subset \mathbb{R}^d, i \in [1; N]$ (d is the dimension of our input space) with their associated class $y_i \in \{-1; 1\}$. Supervised learning is the problem of inferring a function f such as:

$$f : X \subset \mathbb{R}^d \rightarrow \{-1; 1\}$$

from a set of observations and that will correctly classify the maximum number of vectors x_i and more important, that will correctly describe the phenomenon responsible of the separation between the two classes so that a new and unknown point will be classified into the correct class (capacity of generalization of the classifier).

Vapnik et al. designed a classification algorithm, based on both ideas of linear separation and maximization of the margin between the separation and the nearest points of the training database [13]. This margin will give the maximum of "safety" for the generalization of the algorithm and its application to new points. SVM are widely used and showed a good capacity of generalization in various applications. Their construction is performed determining an hyperplane whose equation is $\langle \mathbf{w}, \mathbf{x} \rangle + w_0 = 0$ (where \mathbf{w} and w_0 are the parameters of the equation of the hyperplane to determine). From this hyperplane, we construct the function f given by:

$$\begin{cases} \langle \mathbf{w}, \mathbf{x}_i \rangle + w_0 > 0 \Rightarrow f = 1 \\ \langle \mathbf{w}, \mathbf{x}_i \rangle + w_0 < 0 \Rightarrow f = -1 \end{cases} \quad (1)$$

f is the output of the algorithm for a new point x_i . The hyperplane is constructed by solving the following equation that maximizes the margin:

$$\arg \max_{w, w_0} \min_{i=1..N} \{ \|x - x_i\| : x \in \mathbb{R}^d, \langle \mathbf{w}, \mathbf{x} \rangle + w_0 = 0 \}$$

by solving the linear problem (using the lagrangian):

$$\text{Min} \quad \frac{1}{2} \|\mathbf{w}\|^2 \quad \text{s.t.} \quad f(\langle \mathbf{w}, \mathbf{x}_i \rangle + w_0) \geq 1, i = 1..N \quad (2)$$

The previous equation stands for the linearly separable case but a large part of real problems are not included. In all this formulation, every equation in which the points of the training dataset appear relies only on inner product. Some functions K can act as an inner product ($K(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$). Such function will map the input space into a high-dimensional space (even infinite), named feature space in which the separation could be linear. The problem of determining the best kernel for a given application is an open address. The resolution is obtained by replacing the dot products $\langle \cdot, \cdot \rangle$ in the lagrangian by the kernel $K(\cdot, \cdot)$. The construction of such a function follows the Mercer conditions. In this application, we obtained, in every case, better results with the Gaussian Kernel than with other ones. This kernel is defined as follow:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp \left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2} \right)$$

In the following, we will have to use SVM for binary classification and for multiclass classification. For this second method, we used the one-versus-one algorithm, because it minimizes the indetermination zone (compared to the one-versus-all) and it does not necessitate larger training datasets to solve more complex problems. This scheme consists in constructing $\frac{N \cdot (N-1)}{2}$ classifiers, using all the pairwise combinations of the N classes. In this case, we will construct binary classifiers to differentiate the classes C_i and C_j , $0 < i \leq N$ and $0 < j < i$. To determine the class of a new point, a majority voting is applied. Mathematically speaking, the decision is given by: $C = \max_{k=1..N} \text{Card}(\{y_{i,j}\} \cap \{k\})$ where $y_{i,j}$ is the decision given, for this new point, by the SVM trained to distinguish the classes i and j . In case of equality, the class chosen is the one with the maximal margin from the final subset.

TABLE III
 CONFUSION MATRIX FOR GENERIC MODELS, GAUSSIAN KERNEL WITH OPTIMAL VALUE OF σ . GLOBAL ERROR RATE (GER): 13.79%

		Classification Results						
		<i>Sleeping</i>	<i>Resting</i>	<i>Dress/undress</i>	<i>Eating</i>	<i>Toilet use</i>	<i>Hygiene</i>	<i>Communication</i>
Activity	<i>Sleeping</i>	97.96%	2.04%	0%	0%	0%	0%	0%
	<i>Resting</i>	16.44%	78.08%	0%	1.37%	4.11%	0%	0%
	<i>Dressing/undressing</i>	13.3%	6.7%	80%	0%	0%	0%	0%
	<i>Eating</i>	0%	0%	2.2%	97.8%	0%	0%	0%
	<i>Toilet use</i>	0%	6.25%	0%	6.25%	81.25%	6.25%	0%
	<i>Hygiene</i>	7.14%	0%	0%	7.14%	14.29%	71.43%	0%
	<i>Communication</i>	5%	10%	5%	0%	0%	0%	80%

B. Generic model with no prior knowledge

1) *Construction*: For those first models of activities, we only consider the Support Vector Machine classification of the frames using the previously described features. We construct $\frac{N \cdot (N-1)}{2} = 21$ classifiers constructed to distinguish class i from class j , i and j being two different activities. A majority voting is used to determine the class of the frame after the classification of the new point (the one that has been left out) by the 21 classifiers.

2) *Results*: The results of this first classification are given by the confusion matrix in table III. These results are for the optimal value of the hyper-parameter of the kernel. The Global Error Rate (GER) is then **13.79%**.

This matrix shows us the correct classification rate for each class but also the most common mistakes that are made by the classifier. For instance, we can see that *Resting* is difficult to define with such a low number of samples. Indeed, this activity is not perfectly bounded and can be whatever the participant wants. The consequence of it is that it is very diversified and difficult to classify (because for each kind of *Rest* we have very few samples, and we remove one for the test). Still for *Rest* and *Sleep*, we can see that the posture should have an important place considering the incorrect classification of this matrix. We will investigate on the future on the misclassification problems of posture in daily living.

We can also notice that some classes are close one to another, *Rest* and *Sleep* are two examples but *Toilets use* and *Hygiene* also, because *Toilets use* include *Hygiene* activity after.

Finally, one of the problem that is due to the few number of data that we have is that the training database for the different classes have different numbers of samples. For this classification and for all the following, this non-balanced dataset creates a distortion in the classifiers that are created, for instance between an important class as *Sleeping* and a small one like *Communication*.

C. Introduction of spatial knowledge

1) *Construction*: The seven activities chosen are typical ones that can be assigned to a specific room of the flat. Those activities do not need changes of room. When we aim at determining the activity performed by a participant, the first thing to look at could be the room in which the frame occurs

mostly. This first information could be of high interest. Table IV gives the rooms that are associated to the activities. In this table, we can notice that we have associated the Kitchen and the Hall (depending on where we are located in the Kitchen, we are detected by one of the two sensors) and also we have associated the Bathrooms and Toilets because there are a part of the two rooms (without wall) in which we can be detected by both sensors (aliasing of the areas). This is not due to a misconception in the flat but this correspond to kind of flats that we can have in real conditions. Indeed, lots of flats and especially the small ones (and elderly people living alone are in this kind of flats), have not got Toilets and Bathroom separated and can have the Kitchen associated to a Hall (or a Living-room) without wall.

TABLE IV
 ROOM OF THE FLAT ASSIGNED TO THE ACTIVITIES

Class	Dominant Room
<i>Sleeping</i>	Bedroom
<i>Resting</i>	Living Room
<i>Dressing</i>	Bedroom
<i>Eating</i>	Kitchen and Hall
<i>Elimination</i>	Bathroom and Toilets
<i>Hygiene</i>	Bathroom and Toilets
<i>Communication</i>	Living Room

To perform this classification, we first determine the percentage of time spent in the four different "rooms": Bedroom, Living room, Bathroom/Toilets and Kitchen/Hall. If this percentage of time is greater than 80% for one of the room, then we associate the room to the frame. If none of them is over 80% then no room is associated.

From this room, we then deduce the possible activities and try to classify the frame in one of them. If the frame is associated to the Kitchen/Hall, as the only possible activity is *Eating*, then the decision is taken. If it is for instance Living Room, then we use the classifier that allow to differentiate *Communication* from *Resting* to classify the activity. It is the same for Bedroom, we use the classifier that distinguish *Sleeping* from *Dressing/Undressing* and finally for Bathroom/Toilets we use the classifier for *Hygiene* and *Elimination*. In the case of a frame that has not been associated to a room, we keep the same classifiers (the 21) as in the previous section.

TABLE V
 CONFUSION MATRIX FOR MODELS INCLUDING SPATIAL KNOWLEDGES, GAUSSIAN KERNEL WITH OPTIMAL VALUE OF σ . GER: 21.12%

		Classification Results						
		<i>Sleeping</i>	<i>Resting</i>	<i>Dress/undress</i>	<i>Eating</i>	<i>Toilets use</i>	<i>Hygiene</i>	<i>Communication</i>
Activity	<i>Sleeping</i>	69.39%	2.04%	0%	28.57%	0%	0%	0%
	<i>Resting</i>	10.96%	78.08%	0%	6.85%	4.11%	0%	0%
	<i>Dressing/undressing</i>	13.3%	6.7%	73.3%	6.7%	0%	0%	6.25%
	<i>Eating</i>	0%	0%	2.2%	97.8%	0%	0%	0%
	<i>Toilets use</i>	0%	0%	0%	6.25%	87.5%	6.25%	0%
	<i>Hygiene</i>	0%	0%	0%	14.29%	14.29%	71.42%	0%
	<i>Communication</i>	5%	10%	10%	10%	0%	0%	65%

As a consequence, the processing of these data is extremely faster. Instead of 21 classifications we make 0 or 1 classification for each frame in most of the case. On our dataset, only 8 frames (over the 232) were not associated to a specific room.

2) *Results*: The results of this second classification are presented by the Table V. Unexpectedly, the GER is greater than for the generic models and is **21.12%**. If we examine it more carefully, it could be explained by one simple fact: the missing detections of the presence infra-red sensors. Indeed, we can for instance focus on the first line of the table, *Sleeping* activity, and we examine the results. The frames that are classified as *Resting* are still classified as *Resting* and correspond to ones that are not assigned to a room. But we can see appearing more than 28% of classification as *Eating*. If our classifier that distinguishes the possible activities in a room goes wrong, the classification should be *Dressing/Undressing*. The only possibility to obtain *Eating* is to be in the Kitchen. That corresponds to non-detections on the presence infrared sensors. We examined for instance this particular case of *Sleeping*, and we noticed these non-detections. The participant number 12 entered the flat and directly went to the bed to perform the first activity that she chose, *Sleeping*. She stayed in bed for 25 minutes. If we look at the file that contains the location sensors data, we have a first detection at the entrance and then in the hall. After, when she entered the Living room we have no detection, and the same when she entered the Bedroom. The first detection appears at 22 minutes of *Sleeping* activities. For the 22 first minutes, for the algorithm, as there are no other detection, the person is still in the hall, not moving. As a consequence, all these frames are 100% in the Kitchen/Hall and the decision taken is *Eating*. We have other non-detection like this that bring a high number of misclassification in the *Eating* activity for all the classes.

This knowledge based on the spatial occupation is interesting and allow to disambiguate some other situations. For instance, we can see that *Toilets uses* are better classified with this method (the frames classified previously as *Resting* are now correctly classified), and some *Resting* frames classified as *Sleeping* could be corrected if there were no problem with the presence sensors.

With a faultless presence sensor, this knowledge allow to have a lower computing time and to disambiguate situations (as it considers less classifiers). However, with faulty sensors,

some frames for which the incorrect data were compensated by the other sensors are now incorrectly interpreted because the correct class is not even considered as possible and is, as a consequence, not tested.

D. Introduction of temporal knowledge

1) *Construction*: When we think about the most relevant variables to determine an activity, after the location we get the time of the day. Indeed, for instance, to differentiate two activities that occurs in the Bedroom, if it occurs at 3:00 PM, we could consider that it is more likely to be *Sleeping* than *Dressing*.

For these models, we first determine a coding of this hour of the day. The day could be cut in seven pieces that are the following: T1 representing the breakfast (from 7 AM to 9 AM), T2 the morning (from 9 AM to 12 AM), T3 the lunch (from 12 AM to 2 PM), T4 the afternoon (from 2 PM to 7 PM), T5 the diner (from 7 PM to 9 PM), T6 the evening (from 9PM to 11 PM) and finally T7 for the night (from 11 PM to 7 AM).

From these time slots, we can now attribute them to the different activities, as it has been done for the locations:

TABLE VII
 POSSIBLE TIME SLOTS FOR THE DIFFERENT ACTIVITIES

Class	Slots
<i>Sleeping</i>	T1, T4, T6, T7
<i>Resting</i>	T1, T2, T4, T5, T6
<i>Dressing</i>	T1, T2, T6, T7
<i>Eating</i>	T1, T3, T5
<i>Elimination</i>	T1, T2, T3, T4, T5, T6, T7
<i>Hygiene</i>	T1, T2, T3, T5
<i>Communication</i>	T1, T2, T3, T4, T5, T6, T7

As we can observe, the restrictions on the time are different for the activities. Indeed, we can say that taking a meal is on breakfast, lunch or diner time, but we can not precise the time at which the person will be called or will call someone or will go to the toilets (so no restrictions for the *Communication* and *Toilets uses*). The models try to be general and to correspond to most of the people. This can be adapted for different populations if the targets are known and their uses well identified.

TABLE VI
 CONFUSION MATRIX FOR MODELS INCLUDING TEMPORAL KNOWLEDGES, GAUSSIAN KERNEL WITH OPTIMAL VALUE OF σ . GER: 9.91%

		Classification Results						
		<i>Sleeping</i>	<i>Resting</i>	<i>Dress/Undress</i>	<i>Eating</i>	<i>Toilets use</i>	<i>Hygiene</i>	<i>Communication</i>
Activity	<i>Sleeping</i>	100%	0%	0%	0%	0%	0%	0%
	<i>Resting</i>	10.96%	83.56%	0%	0%	2.74%	1.37%	1.37%
	<i>Dressing/Undressing</i>	0%	13.3%	86.7%	0%	0%	0%	0%
	<i>Eating</i>	0%	0%	2.2%	97.8%	0%	0%	0%
	<i>Toilets use</i>	0%	12.5%	0%	0%	87.5%	0%	0%
	<i>Hygiene</i>	0%	0%	0%	7.1%	14.3%	78.6%	0%
	<i>Communication</i>	0%	5%	5%	5%	0%	0%	85%

To test it, we can not use the timestamps of our dataset. Indeed, the participants was not here all the day and made the different activities in a duration of for instance 1 hour and 30 minutes. As a consequence, we simulated a perfect timing by assigning randomly a coherent time to each frame. Indeed, as we know the possible timing for each activity, we picked-up randomly one of these timing for each of the frames of the labeled dataset.

Then, when we want to classify a frame whose time is T3, we have four choices: *Eating*, *Hygiene*, *Elimination* and *Communication*. For these four activities, we have $\frac{N \cdot (N-1)}{2} = 12$ classifiers and a majority voting gives us the result.

2) *Results*: The results of this third algorithm are given in Table VI. As the time parameter is given randomly to the different frames, the results could change with these 232 assignments. To avoid this, the results given are the mean of five executions of the same algorithm, so five different assignments for the 232 frames. Globally, they are better than the generic ones with the GER is only **9.91%**. Indeed, this time introduction allows to disambiguate between some classes for instance *Sleeping* and *Resting* and reduce the misclassification rate in both cases.

These results are logical because what we have done here is nothing else than reducing the number of possible classes for each frame, as it was done previously with the algorithm including spatial knowledges, but here it was performed using a faultless sensor contrary to the case of the location.

Moreover, as previously, the classification is faster than for the generic models because we reduce the number of estimations and so the computing time.

E. Hybridization of knowledges into the classifier

1) *Construction*: Finally, the last test that we have performed is to consider the three previous classifier and to aggregate them to take the decision.

Classifiers that include spatial and temporal knowledges are first estimated and the sample is tested on them. If their results are the same, the decision is taken with only them. Else, we estimate the generic one (the most time-consuming classifier) and we take the decision with a majority voting between the three classifiers. If the three answers are different, the generic one is kept (because of the more general models used).

2) *Results*: The results of this last algorithm are presented in the Table VIII. The GER is then **13.2%**, so almost the same than the generic models. These results are easily explainable.

Indeed, aggregation of classifiers should enhance the results. It is for instance the basis of boosting methods. If we eliminate the cases in which the classifier with spatial knowledges is wrong because of faulty measurements, we have the expected behavior. With a bigger training dataset or with more predictable faults in the measurement, we could have trained an artificial neural network with three entries (the three decisions). But here this would not have any sense.

Here, the information brought by the two other classifiers, that is supposed to correct the errors of the generic one, are not consistent between them. As a consequence, the decision is taken on the most general one. We improved a little (for few frames) the results of the general classifier by adding the two others but as we can see, the problems of the classifier based on the room makes these improvements negligible. The only result that has been taken into account for the final classification is for the *Toilets use*. The items classified as *Sleeping* has been corrected and are in the correct class in that case. On most of the cases, the decision is taken considering the generic classifier (because it is the one that does not bring any prior knowledge and that uses only the models determined from the samples).

IV. DISCUSSION AND CONCLUSION

This paper presents the results of the supervised classification of seven activities of daily living, using support vector machines, in four conditions. The first is the creation of generic models for each activity using only the training dataset. The second is the use of knowledge on the possible activities for each room to reduce the number of possible activities for a given frame. The third is the introduction of knowledge on the possible activities depending on the time of the day to reduce again the possibilities. Finally, the last aggregates the results of the three first models.

The best results are achieved by the third one, using the time of the day. For the second one (localization), the results are lower than the one of the generic models because of the non-detections that occurred. In the general models, the problems of detection are compensated with the other sensors. For the models with priors, classification stage cannot correct them.

TABLE VIII
 CONFUSION MATRIX FOR MODELS WITH HYBRID KNOWLEDGES, GAUSSIAN KERNEL WITH OPTIMAL VALUE OF σ . GER: 13.2%

		Classification Results						
		<i>Sleeping</i>	<i>Resting</i>	<i>Dress/undress</i>	<i>Eating</i>	<i>Toilets use</i>	<i>Hygiene</i>	<i>Communication</i>
Activity	<i>Sleeping</i>	97.96%	2.04%	0%	0%	0%	0%	0%
	<i>Resting</i>	16.44%	78.08%	0%	1.37%	4.11%	0%	0%
	<i>Dressing/undressing</i>	13.3%	6.7%	80%	0%	0%	0%	0%
	<i>Eating</i>	0%	0%	2.2%	97.8%	0%	0%	0%
	<i>Toilets use</i>	0%	0%	0%	6.25%	87.5%	6.25%	0%
	<i>Hygiene</i>	7.14%	0%	0%	7.14%	14.29%	71.43%	0%
	<i>Communication</i>	5%	10%	5%	0%	0%	0%	80%

Another deduction from these results is that the improvement using these knowledges, on the room or on the location (if they are perfects) is only few percent (4 or 5 %) and does not allow to create perfect models with our dataset. We can have doubts on the interest to use such knowledges for a low improvement. Indeed, fixing it makes the models less general, that is obvious, and it does this by fixing bounds on the way to perform activities. Indeed, for instance, with the time knowledge, if the person starts to eat at 1:30 PM and ends at 2:30 PM, the frames between 2 PM and 2:30 PM will not be considered as eating because it will be out of the time slots of the eating. It will be the same for every other activities unless we totally remove all the conditions.

This work was a test on the interest of using knowledges and it appears that it is not a path of interest to obtain the best results. Our future work will keep the generic models and we think that it would be interested to look at other solutions to improve it. One of them is to consider the frames one with the others. Indeed, for the moment, we only consider the frames as independent ones that will be classified. A path of interest is to look at the data after this classification stage and try to “understand” the day of the person and the succession of activities, by considering the series of classification as a text and analyzed it with grammar-related methods. With this, we could correct some frames that has been incorrectly classified and obtain better results for our classification.

To test such a method, we will have to acquire a larger and more representative dataset. Such acquisition is hard to schedule and must be prepared. To be representative, it has to be done on young and elderly people, and to have exhaustive results, we have to obtain an indexation of the activities. We can not use webcams in the home of the person. We can imagine a system using a PDA to annotate the activities of the day. With this system, the person indicates the activity performed before and allow us to check the concordance with the results. As people can not annotate each activity during a long period, they will do this partially and we can use the rest of the classified activities to check the coherency of the results (three meals a day, consistent use of the toilets, etc.).

ACKNOWLEDGMENT

The data have been collected during a collaboration between TIMC-IMAG laboratory, Team AFIRM and LIG Laboratory,

Team GETALP. All the measurements have been performed in the TIMC-IMAG lab., in their Habitat Intelligent pour la Santé (HIS). The authors would like to thank all the subjects for their time and also Christophe Villemazet and RBI company for their work on the AILISA project.

REFERENCES

- [1] M. Chan, D. Estève, C. Escriba, and E. Campo, “A review of smart homes- present state and future challenges.” *Computer Methods and Programs in Biomedicine*, vol. 91, no. 1, pp. 55–81, Jul 2008.
- [2] S. Katz and C. Akpom, “A measure of primary sociobiological functions,” *Intl. Journal of Health Services*, vol. 6, no. 3, pp. 493–508, 1976.
- [3] M. Philipose, K. P. Fishkin, M. Perkowitz, D. J. Patterson, D. Fox, H. Kautz, and D. Hahnel, “Inferring activities from interactions with objects,” *IEEE Pervasive Computing*, vol. 3, no. 4, pp. 50 – 57, Oct. 2004.
- [4] X. Hong, C. Nugent, M. Mulvenna, S. McClean, and B. Scotney, “Evidential fusion of sensor data for activity recognition in smart homes,” *Pervasive and Mobile Computing*, pp. 1 – 17, 2008.
- [5] C. Nugent, X. Hong, J. Hallberg, D. Finlay, and K. Synnes, “Assessing the impact of individual sensor reliability within smart living environments,” in *Proc. IEEE International Conference on Automation Science and Engineering CASE 2008*, 2008, pp. 685 – 690.
- [6] B. Kröse, T. van Kasteren, C. Gibson, and T. van den Dool, “Care: Context awareness in residences for elderly,” in *International Conference of the International Society for Gerontechnology*, Pisa, Tuscany, Italy, June 4–7 2008.
- [7] S. Tsukamoto, H. Hoshino, and T. Tamura, “Study on indoor activity monitoring by using electric field sensor,” in *International Conference of the International Society for Gerontechnology*, Pisa, Tuscany, Italy, June 4–7 2008.
- [8] M. Berenguer, M. Giordani, F. Giraud-By, and N. Noury, “Automatic detection of activities of daily living from detecting and classifying electrical events on the residential power line,” in *HealthCom’08 - 10th IEEE Intl. Conf. on e-Health Networking, Applications and Service*, 2008.
- [9] A. Fleury, M. Vacher, and N. Noury, “SVM-based multi-modal classification of activities of daily living in health smart homes: Sensors, algorithms and first experimental results,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 274 –283, March 2010.
- [10] G. Le Bellego, N. Noury, G. Virone, M. Mousseau, and J. Demongot, “A model for the measurement of patient activity in a hospital suite,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 92 – 99, Jan. 2006.
- [11] M. Vacher, A. Fleury, F. Portet, J.-F. Serignat, and N. Noury, *New Developments in Biomedical Engineering*. Intech Book, Feb. 2010, ch. Complete Sound and Speech Recognition System for Health Smart Homes: Application to the Recognition of Activities of Daily Living, pp. 645 – 673, ISBN: 978-953-7619-57-2.
- [12] A. Fleury, N. Noury, and M. Vacher, “A wavelet-based pattern recognition algorithm to classify postural transition in humans,” in *17th European Signal Processing Conference (EUSIPCO 2009)*, Glasgow, Scotland, Aug. 24–28 2009, pp. 2047 – 2051.
- [13] V. N. Vapnik, *The Nature of Statistical Learning Theory (2nd Edition)*. Springer-Verlag, 2000, ISBN: 978-0-387-98780-4.