# ADL Recognition Based on the Combination of RFID and Accelerometer Sensing

Maja Stikic*, Tâm Huỳnh†, Kristof Van Laerhoven† and Bernt Schiele†
*Fraunhofer IGD, Darmstadt, Germany
†Computer Science Department, TU Darmstadt, Germany
{stikic, huynh, kristof, schiele}@mis.tu-darmstadt.de

***Abstract*—The manual assessment of Activities of Daily Living (ADLs) is a fundamental problem in elderly care. The use of miniature sensors placed in the environment or worn by a person has great potential in effective and unobtrusive long term monitoring and recognition of ADLs. This paper presents an effective and unobtrusive activity recognition system based on the combination of the data from two different types of sensors: RFID tag readers and accelerometers. We evaluate our algorithms on non-scripted datasets of 10 housekeeping activities performed by 12 subjects. The experimental results show that recognition accuracy can be significantly improved by fusing the two different types of sensors. We analyze different acceleration features and algorithms, and based on tag detections we suggest the best tags' placements and the key objects to be tagged for each activity.**

## I. Introduction

In order to fulfill the special needs of an increasing elderly population, elderly care is becoming a rapidly growing problem, especially in western societies. The most common way of detecting the first changes in behaviour of an elderly person is monitoring of everyday activities that are usually performed on a daily basis. For that purpose, two specific sets of activities that describe the functional status of a person (*Activities of Daily Living (ADL)* - bathing, dressing, toileting, transferring, continence, feeding), as well as interaction with the physical and social environment (*Instrumental Activities of Daily Living (IADL)* - using telephone, shopping, food preparation, housekeeping, doing laundry, transportation, taking medications, handling finances) have been defined [1]. The assessment of ADLs/IADLs is mostly done manually through interviews and questionnaires. As this is a time consuming and error prone process [2], it could highly benefit from automatic assessment technology.

Various approaches to the ADL recognition problem can be found in literature. The two most common approaches are based on complementary assumptions. The first approach is based on the assumption that the objects people use during the execution of an activity robustly categorize the activity. In this approach, sensors are typically placed in the environment to detect user's interactions with objects [3]. Radio Frequency Identification (RFID) tags and readers are usually used in these activity recognition systems, because of their durability, small size, and low costs. The second approach adopts the assumption that the activity is defined by motion of the body during its execution. Research in the wearable computing community has shown that characteristic movement patterns for activities such as running, walking or lying can effectively be inferred from body-worn accelerometers (e.g. [4], [5]). More specialized activities that have been recognized with accelerometers include household activities [4], physical exercises [6], wood workshop activities [7], and American Sign Language [8].

The goal of the research presented in this paper is to improve the recognition results by integrating these two approaches, while also aiming to compensate for the shortcomings of both. In order to be able to accurately recognize different activities, the RFID approach requires a large number of objects to be tagged. However, we argue that it is not feasible to tag all objects, because of several reasons. First, the deployment of the large number of tags is still time consuming and error prone. Second, it is not practical to tag some objects because of their material (e.g. metal) or specific usage (e.g. objects used in microwave). We propose to use only the key objects for a specific set of activities by augmenting the object usage with a complementary sensing technique (i.e. accelerometers). On the other hand, the accelerometers approach requires multiple sensors to be placed on strategic body locations, such as wrist, hip, and thigh [4] for accurate recognition. We propose to use only a single 3D accelerometer at the dominant wrist of the user. Since the target user group of elderly people might not be familiar with modern information technologies, limiting the hardware to a single wrist-mounted device containing both the RFID tag reader and the accelerometer could increase user acceptance of our ADL monitoring system.

There have been attempts to combine accelerometers with other sensor modalities, such as microphones (e.g. [7], [9]), wearable cameras [8] and recently, RFID tag readers [10]. We extend this promising approach [10] in the following directions: 1) Since we want to tag only a few important objects per activity, we perform the experiments with different numbers of tagged objects. Results of the experiments show that satisfactory recognition results can be achieved with fewer tagged objects than are usually used. 2) We evaluate our algorithms on a challenging multi-person dataset. The dataset is released and publicly available [11]. 3) We analyze different features and window lengths, as well as different ways of combining the activity recognition results from the two sensor modalities. Unlike in [10], our primary source of information is the RFID data, because it provides accurate high-level information for the activity inference. We rely on accelerometers in cases when the RFID data is not sufficient, such as in the periods when the antenna can not detect person-object interactions and in case of tag ambiguities in terms of objects shared among the activities. 4) We do not incorporate in our approach decomposition of activities into phases (i.e. simple actions performed on objects), because we want to avoid scripted activity stages, as well as their temporal modeling, training and labeling.

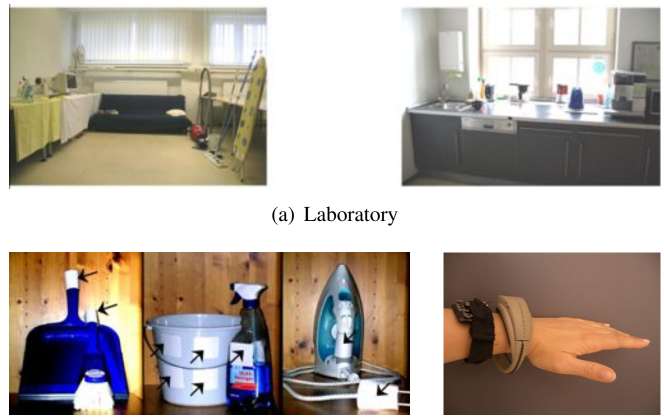
The rest of the paper is organized as follows. In Section II we introduce the hardware used in our experiments and the recorded dataset. Section III describes the three approaches used for ADL recognition. In Section IV we report on the results of all the approaches used. Finally, in Section V we summarize our results and give an outlook on future work.

## II. Experiment Setup

We have focused in our experiment on recognition of a specific class of IADLs, i.e. housekeeping activities. So far, there have been efforts on recognizing various selected instances of different ADL/IADL classes such as hand washing [12] or eating activities (e.g [13], [14]). Housekeeping activities are an important and often occurring IADL class, for which assessment can highly help in the early detection of symptoms of different age-related diseases. We use a dataset [15] of four standard IADL housekeeping activities (*vacuuming*, *ironing*, *dusting*, and *brooming*) for the evaluation of our approach that we augment with six additional activities (*mopping*, *cleaning windows*, *making bed*, *watering plants*, *washing dishes*, and *setting the table*).

The overall length of the dataset is 240 minutes. The duration strongly varies among some of the activities, reflecting the natural distribution of activities in daily life. The data was recorded by 12 subjects (3 females and 9 males), in a controlled lab environment that was converted in a living space with typical objects found and used in a domestic setting, to make it resemble a common home environment (Figure 1(a)). We wanted to avoid bias in the dataset by the subject's online annotations, so we recorded the whole process by video camera and did the annotation offline.

Although a recent study [16] showed that the movements of elderly subjects are less pronounced, it was still impractical to ask elderly people to participate in the experiment at this early stage. However, we plan to do that in the future when all the necessary requirements are fulfilled. Since the scenario presented to the subjects was kept as vague as possible, there was no significant bias in the dataset. The subjects were told to choose a certain set of activities to perform based on the list of 10 targeted activities. We did not require them to follow any description of the tasks which resulted in a wide variety of ways different people performed the same activity. In order to examine the feasibility of person-independent activity recognition, we perform a 12-fold leave-one-person-out cross validation on the data. This is a highly important requirement for our system, considering the target user group of elderly people.



(a) Laboratory

(b) Tagged Objects

(c) Sensors

Fig. 1. Experiment setup.

We deployed 191 tags on 55 objects (Figure 1(b)). The number of tags per object varies between 18 tags for a pillow and 1 tag for a dusting cloth. We aimed to tag as many objects as possible with multiple tags for the following reasons: 1) to find the key objects for the targeted set of housekeeping activities, 2) to evaluate the influence of the number of deployed tags on the recognition results by using different number of tags, and 3) to optimize tag detection for objects that are difficult to detect because of their size, shape or material.

Figure 1(c) shows the sensor platform that was used for recording of the dataset. The users wore the sensors on their dominant wrist. Three subjects that participated in our experiment were left-handed, which later influenced the recognition of the targeted set of activities. We used the Porcupine [17], a wearable multi-sensor platform that includes a 3D accelerometer to infer relevant arm movements. For detection of person-object interactions, the iBracelet [18], a wireless RFID reader, was used.

## III. Approach

The main goal of our experiment is to study the combination of RFID and accelerometer sensing technology for ADL recognition. In order to do that, we first use the accelerometer and RFID tags separately, and afterwards we apply an integrated approach to overcome the shortcomings of both approaches. In the following, we describe all three approaches used.

*a) Recognition based on Acceleration Data:* The 3D-acceleration data as recorded from the sensor was downsampled from 250Hz to 100Hz for our experiments. We computed the following features from the raw signal: *mean*, *variance*, *energy*, *spectral entropy*, *pairwise correlation between the three axes*, the first ten *FFT coefficients* and *exponential FFT bands* [19]. Each feature is computed over a sliding window shifted in increments of 0.5 seconds. We evaluated the performance of the features both individually and in combination, and over different window lengths (0.5sec-128sec).

For classification of activities we evaluate three different approaches, namely Naive Bayes, Hidden Markov Models (HMMs) and Joint Boosting. Previous work has shown that Naive Bayes and HMM classifiers are well-suited for low-level activities such as sitting, standing and walking [20] or wood workshop activities [7]. Boosting methods have also been successfully applied to data from wearable sensors [19]. The third method we employ, Joint Boosting [21], is a multi-class variant of traditional boosting approaches in which multiple weak classifiers are combined into a single strong classifier.

*b) Recognition based on RFID Data:* We associate all the tagged objects with the activities in which they are usually involved. This process is done manually, but as we aim at tagging fewer objects in our final ADL recognition system, that should not be a major constraint for the implementation of our approach. With this object-activity mapping, each detected tag clearly indicates a candidate set of possible activities.

The used dataset contains very few accidentally detected interactions with objects. Also, many interactions with objects are not detected, mostly because of the short range of the RFID reader's antenna. Generally, the number of tag detections highly varies among the activities. In the entire dataset, 0.764 tag detections happened per second on average. To overcome the problem of sparse detections, we use a sliding window over the detected RFID tags and classify each window based on the majority voting scheme of the tag readings, i.e. mapped activity labels in that window. We shift the window in increments of 1 second and we evaluate the recognition performance for different window lengths (1sec-120sec). Additionally, the tags' votes are weighted proportionally to their relative position in the window, in order to avoid bias from the previous activity at activity transitions for longer window lengths.

For the evaluation of the influence of the number of used tags on the recognition results, we use the following procedure (Figure 2). In the first run, we use all the deployed tags. Since we aim to tag as few objects as possible, we decrease the number of used tags in each run by half by using the following procedure. We rank the tags for each activity based on the number of detections and then use the best 50% of these tags until in the last fifth run we only have one tag per activity. Some of the activities include fewer objects than others, which is reflected in the dataset. E.g. activities such as *watering plants* and *brooming* require fewer interactions with objects. In some other activities, such as *washing dishes*, tagged objects are not detected, probably due to the absorption of the radio waves by water and metal.

*c) Combining RFID and Accelerometer Sensing:* For the combination of RFID and accelerometer sensing, we use the RFID recognition as a baseline method for the recognition of activities. In cases when we fail to recognize the activity based on RFID tags, we rely on the accelerometers' recognition. In principle, there are two different cases when the RFID approach fails.

In the first case, the majority of detected tags within a window is shared among several activities. Based on the RFID approach, the window is classified as one of the activities that share the tags in that window, each of those activities having the same probability. To decrease the classification errors, we resolve this ambiguity by using the acceleration classification. We classify the window as the activity which has the highest likelihood among the activities that share the detected tags in that window.

In the second case, the RFID reader fails to detect any tags within a window. We classify the window as an unknown activity, because we do not have any information about the current activity, based on the RFID data. We resolve the issue of gaps in RFID data by assigning the activity with the cumulative highest likelihood to windows without tags, based on the acceleration data. In this combined classification process, we accept the acceleration-based classifications only if their likelihood is above a certain threshold.

| Activity | 100% | 50% | 25% | 12.5% | Key object | Additional key object |
|---|---|---|---|---|---|---|
| Dusting | 18 | 9 | 5 | 3 | Dusting cloth | Dusters' box |
| Ironing | 7 | 4 | 2 | 1 | Iron | - |
| Vacuuming | 5 | 3 | 2 | 1 | Vacuum cleaner | - |
| Brooming | 3 | 2 | 1 | 1 | Small plastic broom | - |
| Mopping | 8 | 4 | 2 | 1 | Mop | - |
| Cleaning windows | 14 | 7 | 4 | 2 | Gloves | Window cleaning liquid |
| Making bed | 42 | 21 | 11 | 6 | Pillow | Pillow case |
| Watering plants | 2 | 1 | 1 | 1 | Water spray | - |
| Washing dishes | 3 | 2 | 1 | 1 | Gloves | - |
| Setting table | 18 | 9 | 5 | 3 | Glass | Cupboard |

Fig. 2. Key objects for activities and number of tags in different runs.

## IV. RESULTS

In this section we present the experimental results for the three approaches described in Section III. We evaluate our algorithms on the dataset presented in Section II by using standard metrics, namely precision (i.e. the number of true positives in the test set divided by the sum of true positives and false positives in the test set), recall (i.e. the number of true positives in the test set divided by the sum of true positives and false negatives in the test set), and accuracy (i.e. the number of true positives in the test set divided by all samples in the test set). The ground truth for the sliding window used in all three approaches is the label of the last sample in the window. All results are averaged over 12 cross validation runs. In each run we train our algorithms on the data recorded by 11 subjects and test them on the left out subject's data.

*a) Acceleration Results:* In the following we report on our recognition results based on features computed from the wrist-mounted accelerometer alone. For the HMMs, in addition to the window length, we vary the number of states (1-4), the number of Gaussians per state (1-4), and the observation length (1-32). For Joint Boosting we vary the number of weak classifiers (100-200). In order to filter out occasional misclassifications, the output of all classifiers is smoothed with a majority filter. The overall best result of 68% accuracy is achieved with the Joint Boosting approach when using all features and 200 weak classifiers. For Naive Bayes and HMMs, we found that using mean and variance of the signal

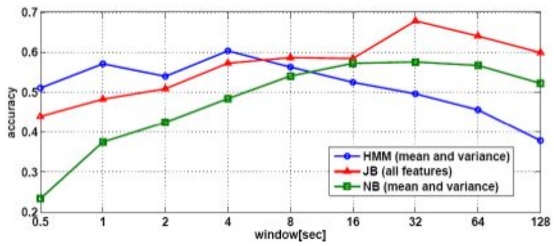



Fig. 3. Classification based on data from the accelerometer. The plot shows the accuracy across different algorithms and window lengths.

as features works best for our set of activities. From Figure 3 one can observe that both Joint Boosting and Naive Bayes work best at relatively large window sizes of 32 seconds, while HMMs perform better at smaller window sizes of up to 4 seconds. One reason for this might be that the smaller windows preserve more of the temporal structure inherent in the data, which the HMMs are able to exploit.

The seemingly low accuracy of slightly below 70% should be seen in the light that there were several factors making this recognition task more challenging than others reported in the literature: First, the use of only a single 3D accelerometer, and second the fact that we trained and tested the system on different users, some of which performed the same activity in distinctly different ways and sometimes with different hands. Third, we did not edit the recordings e.g. by cutting out only the part during which the user actually ironed, but we included the entire activity from setup (e.g. assembling the ironing board) to teardown (e.g. stowing away the ironing board).

*b) RFID Results:* In the following we report on the recognition results based on the RFID tags only. Figure 4 (left column) shows how overall precision and recall change with different window lengths. One can observe that our approach performs best in terms of precision for very short windows. On the other hand, recall is best for longer windows. That is due to the fact that with longer windows we propagate the labels to the regions where tags were not detected, so we have fewer false negatives. At the same time, longer windows increase the number of false positives, because of the tags' bias from the previous activity at the transitions between activities. When using all deployed tags, the best results for precision lie slightly above 92% when using windows of 7 seconds. Recall reaches its maximum of 72% for windows of 82 seconds. Since recall dramatically increases when we increase the window length from 1 to 40 seconds, and afterwards only a slight improvement is achieved, we propose to use 40 second windows as an optimal window length in this setting. Precision in that case still remains high (89%/70% precision/recall).

Figure 4 (left column) also shows the effect of different numbers of tags. As can be seen from the plots, by decreasing the number of tags, recall decreases as well, but surprisingly precision does not change much. On the contrary, in some cases precision even increases with fewer tags, because some of the tags shared among the activities are discarded from the dataset in that way. When we use only one tag per activity,

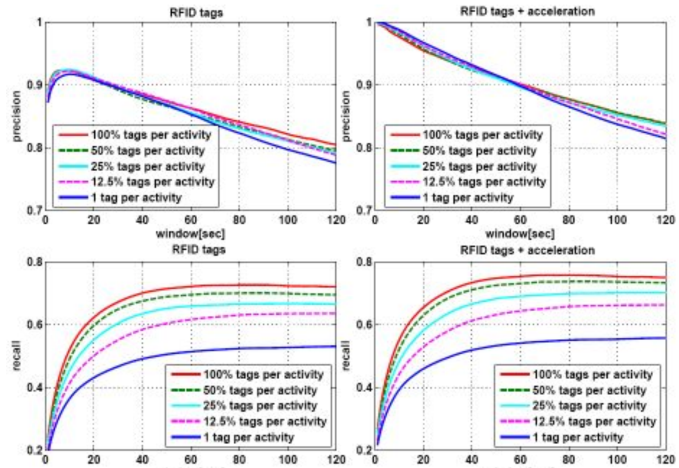



Fig. 4. Overall precision and recall for different window lengths and different number of used tags in case of the recognition based on the RFID tags only (left) and based on the RFID tags and acceleration for shared tags (right).

we choose the tag that was detected most often during the execution of each activity. That way, we define the key objects per activity (Figure 2). Since the run with 12.5% of the most detected tags performs overall better than the run when we use only one tag per activity, we add three more objects to the set of key objects. That way, we also avoid the gloves as a single key object for *window cleaning*, since they are shared among that activity and *washing dishes*.

We tagged most of the objects with multiple tags to find the best placement of the tags. For many objects (e.g. vacuum cleaner, mop, broom, iron) the tags placed on the handle of the object were detected more often than the other tags attached to the same object. That is because of the very short distance between the object handle and the RFID reader during the performed activities. For other objects the best placement is at the place where users usually grab the object (e.g. corner of the pillow) or at the place where users spend considerable time during the execution of the activity (e.g. buttons on the pillow case). For some objects (e.g. cupboard, window cleaning liquid, and dusters' box) the best placement depends on whether the subject is left or right-handed.

The results show that a satisfactory trade-off between precision and recall can be achieved with appropriate window lengths. A decreased number of tags does not influence the recognition results significantly and the key objects for activities are defined. Finally, the best placement for the tags highly depends on the person, as well as on the activity.

*c) Combining RFID and Accelerometer Sensing Results:* In the following we report on the recognition results based on the combination of RFID tags and acceleration. As the experiments show, we can improve overall precision and recall by augmenting the RFID classification with acceleration recognition scores in two cases: 1) when detected tags are shared among the activities and 2) when interactions with objects are not detected. For the combination of RFID and acceleration classification, we use the parameters that yielded the best results for the classification of acceleration data (Joint Boosting, all features over windows of 32 seconds).

We present the results of resolving tag ambiguities by means of acceleration classification in Figure 4 (right column). We have only four types of objects (tagged with 16 tags) that are shared among five activities in the dataset. Still, when we compare the results when we use all the tags to the classification based on RFID tags only, there is a clear tendency of about 3% improvement in recall. The precision increases, especially for shorter windows (from 10% increase for windows of 1 second to 6% increase for windows of 7 seconds, when the classification based on the RFID tags reaches its maximum). For larger windows, the gain in precision is smaller but still noticeable (for windows of 40 seconds, the increase is 4%, and for the largest windows of 120 seconds, there is still increase of 3%). This decrease of improvement for larger windows is due to the fact that in larger windows, we usually have not only the shared tags, but also additional tags that resolve the tag ambiguities already on the level of RFID classification.

The results of additional filling in of gaps where no RFID tags are sensed by using the acceleration classification are shown in Figure 5. Here, we present the results for the run when we use only one tag per activity. We vary the threshold between 0 (when all acceleration-based classifications are accepted) and 1 (when all acceleration-based classifications are rejected, which brings us to the previous case of using the acceleration for shared tags only). From the plot one can observe that recall increases with the number of accepted acceleration-based classifications. However, at the same time, the more accepted acceleration-based classifications we have, the more precision decreases. This is due to the fact that the recognition of higher level activities such as housekeeping is difficult using only one accelerometer placed at the dominant wrist of a user. This trade-off between precision and recall has to be taken into account based on the specific application requirements.

In the extreme case, when the threshold is 0, there is no unknown sample in the test data, which means that overall precision and recall become the same. For shorter windows the increase of recall is between 40% for windows of 1 second and 33% for windows of 7 seconds. At the same time, we observe a significant decrease of precision (from 100% to 63% for windows of 1 second, and from 99% to 69% for windows of 7 seconds). For larger windows, the trade-off between precision and recall is better, since we have higher increase of recall comparing to the decrease of precision. For example, for window length of 40 seconds, the recall increases by 24% and precision decreases by 17%. For the largest window of 120 seconds, the decrease of precision is almost three times lower than the increase of recall, i.e. precision decreases by 7% and recall increases by 19%. This is probably due to the fact that the probability that there is no detected tag is lower for larger windows than for shorter windows. Therefore, the shorter windows need to rely more often on acceleration-based classification which decreases the precision.

Classification results are slightly better in the run when we use 100% of tags, but the overall improvement in precision/recall is higher in the run when we use only one tag per activity. Thus, we can achieve good recognition results with only a few RFID tags when combining them with accelerometer sensing.

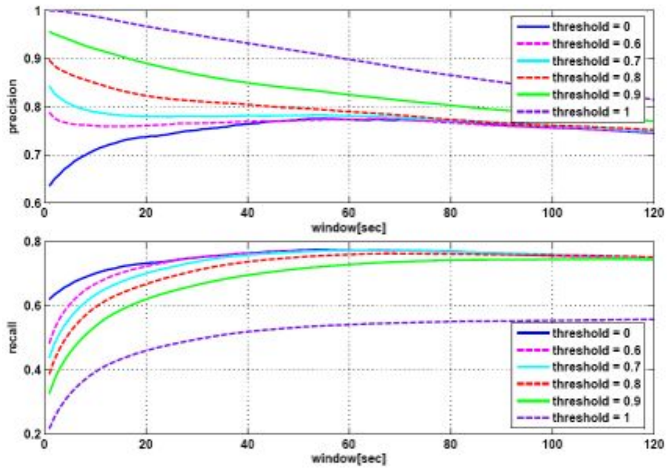



Fig. 5. Overall precision and recall for different window lengths and different likelihood thresholds in case of 1 used tag per activity.

*d) Discussion:* The activity recognition scheme based on the combination of RFID and accelerometer sensing yields better recognition scores than either sensing technology alone. Still, in order to augment the manual assessment of ADLs our system needs to overcome a few limitations.

The main issue in the RFID part of our system is a significant number of false negatives, i.e. tags were not detected even though the subjects were interacting with the tagged objects, because of the short range of the RFID reader but also because of the usage of the non-dominant hand in some activities. For example, during the *ironing* activity two subjects occasionally used their non-dominant hand for ironing some parts of the clothes that were easily reachable in that way. Also, in four cases of *cleaning windows*, the dominant hand of the subjects was occupied with cleaning utensils and subjects had to open the window with the non-dominant hand. Therefore, an additional bracelet on the non-dominant hand might improve the number of detected tags, with a risk of a lower user acceptance of the system. In the future we will aim for an RFID reader with a more appropriate antenna range. Another issue encountered during the experiment is tag ambiguities, i.e. an object is used in more than one activity. We aim to overcome these problems by relying on acceleration classification.

For the activity classification based on the accelerometers, we used state of the art algorithms, but the recognition scores for the acceleration part of our system still encountered issues most likely because of the following reasons.

First, in our experiment, we aimed at person-independent training with 12 subjects who performed activities in very different ways. For example, two subjects vacuumed not only the floor but also the sofa. Also, the subjects had different strategies for *washing dishes*. Three subjects did the washing by repetitive scrubbing and rinsing of each dish and the other subjects first scrubbed and then rinsed all the dishes.

Second, we had three left-handed subjects and even by visual inspection of the acceleration data it can be clearly seen that the range of their data is not the same as for the right-handed subjects. One possible solution to this problem would be to train and test the algorithms only on right-handed or left-handed people.

Third, we used only a single accelerometer worn on the wrist of the user which caused lower recognition scores than usually presented in the literature. Some of the specific movements during the execution of the activities could not be inferred. For example, two subjects were turning the vacuum cleaner on and off with their foot. Additional accelerometers would probably increase the accuracy of the system, but again with a risk of lower user acceptance.

Fourth, we did not divide the activities into phases because we wanted to avoid the tedious labeling and manual modeling of all the sub-activities. After comparing the recognition results with the ground truth and the video recordings, we found that the acceleration classifiers often fail to recognize parts of the activities that do not include discriminative movements typical for that specific activity. For example, during the *ironing* activity, parts when users were really ironing were correctly recognized, but parts when users were finishing ironing of one piece of clothing and preparing the next piece of clothing were usually misclassified. Also, different users performed beginning and ending of the activities differently, which introduced additional misclassifications.

## V. Conclusions and Future Work

The main goal of this paper was to demonstrate the feasibility of combining RFID and accelerometer sensing for ADL/IADL recognition. We conducted an evaluation of our algorithms' performance on 10 housekeeping activities, executed by 12 subjects. Detailed analysis of the algorithms' parameters indicates the optimal window lengths and features, which are 40 seconds window for RFID-based recognition and 32 seconds window and combination of all the acceleration features for Joint Boosting. The results show that combined recognition helps in cases when tagged objects are being shared among the activities, as well as in periods when the RFID reader can not detect interactions with objects due to its short range.

We aim to decrease the number of tagged objects and accelerometers worn by users, while keeping satisfactory recognition results when combining the two sensor modalities. By using different numbers of tags in the dataset, we explored how the number of tags influence the recognition. The results indicate that a decreased number of tags does not significantly change the precision of our system. In some cases, by decreasing the number of tags, tag ambiguities disappear from the dataset, which increases precision. This supports the scenario of Ambient Assisted Living environments with the tags placed strategically on the key objects.

In the future, we plan to validate our approach on larger datasets consisting of activities performed by elderly people. We want to investigate to which extent our results can be generalized in that case, as well as the user's acceptance of our system. In order to make the deployment of our system in home environments feasible, we aim to improve our approach in terms of hardware setup and algorithms performance. We will explore other learning methods that could enable more accurate activity recognition with less supervision.

## VI. Acknowledgments

The authors would like to express their gratitude to all the subjects that participated in the experiments, as well as Intel Research for providing the iBracelet. This research has been supported in part by Fraunhofer's Doktorandinnenprogramm.

## References

[1] 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. 721–726, 1983.

[2] D. Wilson, S. Consolvo, K. Fishkin, and M. Philipose, "In-Home Assessment of the Activities of Daily Living of the Elderly," in *CHI 2005: Workshops - HCI Challenges in Health Assessment*, April 2005.

[3] M. Philipose, K. Fishkin, M. Perkowitz, D. Patterson, D. Fox, and H. Kautz, "Inferring Activities from Interactions with Objects," *IEEE Pervasive Computing*, vol. 3, no. 4, pp. 50–57, October 2004.

[4] L. Bao and S. Intille, "Activity Recognition from User-Annotated Acceleration Data," in *Proc. Pervasive*, Vienna, Austria, April 2004.

[5] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, "Activity Classification Using Realistic Data From Wearable Sensors," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 1, pp. 119–128, January 2006.

[6] D. Minnen, T. Starner, I. Essa, and C. Isbell, "Discovering Characteristic Actions from On-Body Sensor Data," in *Proc. ISWC*, October 2006.

[7] J. Ward, P. Lukowicz, and G. Tröster, "Gesture Spotting Using Wrist Worn Microphone and 3-Axis Accelerometer," in *sOc-EUSAI*, 2005.

[8] H. Brashear, T. Starner, P. Lukowicz, and H. Junker, "Using Multiple Sensors for Mobile Sign Language Recognition." in *ISWC*, 2003.

[9] J. Lester, T. Choudhury, and G. Borriello, "A Practical Approach to Recognizing Physical Activities," in *Proc. Pervasive*, May 2006.

[10] S. Wang, W. Pentney, A. Popescu, T. Choudhury, and M. Philipose, "Common Sense Based Joint Training of Human Activity Recognizers," in *Proc. IJCAI*, 2007.

[11] Dataset. Online http://www.mis.informatik.tu-darmstadt.de/People/maja/datasetpervasivehealth2008.zip.

[12] J. Hoey, A. von Bertoldi, P. Poupart, and A. Mihailidis, "Assisting Persons with Dementia during Handwashing Using a Partially Observable Markov Decision Process," in *ICVS '07*, 2007.

[13] J. Gao, A. G. Hauptmann, A. Bharucha, and H. D. Wactlar, "Dining Activity Analysis Using a Hidden Markov Model," in *ICPR '04*, 2004.

[14] O. Amft, M. Kusserow, and G. Tröster, "Probabilistic parsing of dietary activity events," in *Proc. BSN 2007*, 2007, pp. 242–247.

[15] M. Stikic and K. Van Laerhoven, "Recording Housekeeping Activities with Situated Tags and Wrist-Worn Sensors: Experiment Setup and Issues Encountered," in *WSNHC 2007*, June 2007.

[16] R. Jafari, W. Li, R. Bajcsy, S. Glaser, and S. Sastry, "Physical Activity Monitoring for Assisted Living at Home," in *Proc. BSN 2007*, 2007.

[17] K. V. Laerhoven and A. Aronsen, "Memorizing What You Did Last Week: Towards Detailed Actigraphy With A Wearable Sensor," in *IWSAWC*, 2007.

[18] K. Fishkin, M. Philipose, and A. Rea, "Hands-on RFID: Wireless Wearables for Detecting Use of Objects," in *Proc. ISWC*, October 2005.

[19] J. Lester, T. Choudhury, N. Kern, G. Borriello, and B. Hannford, "A Hybrid Discriminative/Generative Approach for Modeling Human Activities," in *Proc. IJCAI*, August 2005.

[20] N. Ravi, N. Dandekar, P. Mysore, and M. Littman, "Activity Recognition from Accelerometer Data," *Proc. IAAI*, 2005.

[21] A. Torralba, K. P. Murphy, and W. T. Freeman, "Sharing features: efficient boosting procedures for multiclass object detection," in *CVPR'04*.