Journal of Mobile Multimedia, Vol. 6, No.1 (2010) 015-025

ACTIVITY INFERENCE FOR RFID-BASED ASSISTED LIVING APPLICATIONS

JUDITH SYMONDS, BOON-CHONG SEET, JINGWEN XIONG

Faculty of Design and Creative Technologies, Auckland University of Technology
City Campus, Auckland 1010, New Zealand
{ jsymonds, bseet, vfj0504}@aut.ac.nz

Received August 6, 2009
Revised September 2, 2009

Technology assisted living is a practical solution to the increasing demands for access to healthcare services in an era of aging populations and dwindling supply of professional healthcare workers. Radio Frequency Identification (RFID) technology with complementary sensors is widely considered as a very promising approach to realizing the vision of technology assisted living. At the core of any assisted living systems is the important function of human activity inference, which is what enables such systems to be intelligently perceptive and responsive to the humans under their care. In this paper, we review the current state-of-the-art in activity inference for RFID-based assisted living applications, and present our ongoing work on an assisted living prototype for 'goal training' or brain rehabilitation of patients with cognitive impairment in their home environments, with a discussion on the potential design issues involved.

Key words: Activity Inference, Radio Frequency Identification, Assisted Living, Goal Training

1 Introduction

Human activity inference is an important field that is associated with ubiquitous computing, context-sensitive computing, and artificial intelligence. For computers to become useful and capable of independently assisting humans, activity inference through sensing and machine learning is a key aspect for computers to understand our behavior, situation, and needs.

In general, the process of activity inference may involve the use of any of the following six questions of context: When, Where, Who, What, How, and Why [1]. A number of challenges exist in human activity inference. Firstly, is the sheer number of activities that humans can possibly engage in. For instance, just for household activities, there are hundreds of different activities, and each one of them can be accomplished in great number of ways. Secondly, human activities are notoriously variable and unpredictable. This makes modeling human activities extremely complex and difficult given the very large size of graphical activity models with potentially tens of thousands of nodes.

In the past, researchers have used various sensors to achieve human activity inference, such as motion-detection using accelerometers, computer-vision using cameras, or other sensing technologies including acoustic, pressure, optical, and thermal sensors. However, the limitations of using sensors such as motion and video sensors are that they can be difficult, inflexible and costly to deploy. On the other hand, RFID technology has gained increasing attention as a low cost, flexible, and relatively fast solution for wireless identification. RFID has been widely used for product tracking in industry supply-chain management, and more recently has been experimented for detecting human activities through user-object interactions.

With the worldwide growing needs for healthcare services due to the continued aging of our societies and the increasing shortage of professional healthcare workers, there has been a serious contemplation for alternative healthcare arrangements where technology is being used to assist human living. RFID complemented with sensor technology has been identified as a key enabler of systems for technology assisted living [2], and activity inference, which infers the activities of the users through cues from RFID tags and sensors, is a crucial functional component of any assisted living systems.

The remaining of the paper is organized as follows. In Section 2, we review and discuss the different sensing modalities that have been used in RFID-based applications for assisted living. In Section 3, we overview the current two major classes of inference approaches, namely the rule-based and probabilistic based inference. In Section 4, we present our ongoing work on an assisted living prototype that we have developed for goal management training of cognitively impaired users and provide a discussion on some of the potential design issues. Finally, in Section 5, we conclude the paper with some remarks on future work.

2 Sensing Modalities

There is a myriad of applications for assisted living that adopt different sensing approach towards activity inference. In this paper, we are interested in applications that use RFID or RFID in conjunction with other sensing modalities such as inertial and vision-based motion tracking for inference. The following reviews selected works based on their adopted sensing approach with a discussion of their respective merits and limitations.

2.1 RFID

RFID has been recently used by many researchers in what we call identity-based activity inference. The identity here could refer to the identity of the user, identity of the object that has interacted with the user (such as being touched, picked up, or used), or identity of the location of the user. For any given application, a combination of such information with common-sense or specialized domain knowledge could be used to efficiently identify or infer the high-level activities of the user, such as meeting, studying, dining, etc. [3-5]. However, the inference of a RFID-only based system could be limited by several factors.

First and foremost, is the possible ambiguity in the inference when multiple activities could be identified by similar sets of RFID identifiers. For instance, it is not easy to differentiate between meeting and group studying, simply based on detected RFID objects and landmarks. A system may erroneously infer that an object has been picked up, when in fact the RFID reader just happened to be positioned near the object. Inference could also be affected by inherent problems of current RFID technology, in particular missing tag detection due to tag collisions, tag detuning, tag misalignment, and the presence of metal and water in the tag vicinity [6]. These have led the researchers to propose augmenting RFID with other sensing technologies as explained in the following.

2.2 Accelerometers

Accelerometers are a type of inertial sensors that could be attached to parts of human body for sensing their rate and orientation of movements in order to infer primitive actions such as sitting, standing, walking, or arm swinging [7]. Such sensor data from the accelerometers could complement object-use data from RFID tags to enable an understanding of how an object is being used, e.g. a detected object 'hammer' with sensed action 'arm swing' could imply a 'hammering' activity. Both the object-use data and acceleration data could be jointly used in different ways to infer an activity. For instance, the authors in [8] formulated a joint probabilistic model of object-interaction (RFID) and physical action (accelerometers) for activity inference. On the other hand, the authors in [9] proposed a two-stage classification in which RFID-based classification is used as the baseline method, and only if it fails to infer the activity, then the acceleration based classification is used. Both authors reported better inference is attained than when either accelerometer or RFID is used alone.

2.3 Video Sensors

Vision-based motion tracking is another effective approach to augment RFID, in particular in its ability to eliminate ambiguities in RFID-based inference, such as through using video for tracking the motion of objects detected (or missed) by the RFID reader to validate that an object has indeed been manipulated, e.g. being picked up, by the user. In [10], the authors proposed an unsupervised method to learn object models automatically from the video of household activities and employ these models for activity inference using a combination of video sequences and RFID data. In [11], the authors adopted a different approach in which cameras are used for tracking humans instead of objects, such as for monitoring the motion of their heads and hands. On the other hand, RFID readers were modified to capture received signal strength from the tags in order to allow estimation of their positions and orientation. The resulting fusion of visual human motion and RFID object modalities enables the authors to make inference about an user's activities such as picking up an object, holding at eye-level to examine it, and then taking it away or putting back, in the context of a retail application. Despite the possible high cost and complexity involved in vision-based inference, integrating RFID with video sensors appears to be a very interesting approach.

3 Inference Approaches

In this section, we overview and discuss two major classes of approaches for activity inference, namely the rule-based and probabilistic-based approaches.

3.1 Rule-based Inference

In the rule-based approaches, logic rules are formulated and applied to correlate patterns of sensor events with the activities. As an example, the authors in [12] presented a system that uses knowledge-engineered rules for in-home detection of daily living activities such as meal preparation and possible emergency conditions of elderly people. The rules are expressed in the form IF A (AND/OR) B THEN C, where A and B are sensor events, and C relates to the activity that is being inferred. Such inference rules are computationally efficient to execute given their simplicity, but their accuracy depends on the specificity of the rules, and improving the specificity normally implies more instrumentation using sensors on aspects of environment that are key to the activity of interest in order to reduce the rate of false positive detection.

Related to rule specificity is the information granularity of the input (sensor events) and output (inferred activity) of the system. Depending on the type of sensors, different granularity or 'intensity' of sensor events may be captured, and depending on the type of applications, different level of details about an inferred activity could be desired. To support such information granularity in a rule-based system, fuzzy-type rules that use multi-value logic to represent notions of granularity, such as <high, medium, low> or <good, fair, poor> will be necessary. The iDorm (Intelligent Dormitory) [13] is one such system that uses a fuzzy rule-based inference engine to map sensor states to actuator commands representing user's actions for automating a living environment.

3.2 Probabilistic-based Inference

More prevalent are approaches based on probabilistic frameworks such as Hidden Markov Models (HMM), Dynamic Bayesian Networks (DBN), and more recently Conditional Random Fields (CRF). The reason is because probabilistic approaches are often more capable than their rule-based counterparts in handling uncertainties, such as noisy sensor readings due to imperfect hardware and operating environments, variable action sequence due to the fact that the same activity could be performed in various ways, or ambiguous situations when different activities could generate similar sensor patterns. Thus, activities are more often represented as probabilistic distributions over sequences of object use/sensor events, and the one associated with a sequence found having the highest probability (and above a threshold) is chosen as the inferred activity.

HMM [14, 15] is a temporal model that considers an activity as being composed of a sequence of sub-activities with corresponding object use/sensor events occurring in consecutive time slices. As shown in Fig. 1, it consists of a hidden state yt representing a sub-activity for each time slice t, and an observation or outcome xt representing a vector of object use/sensor events occurring during that time slice. The states are 'hidden' as only the outcomes and not the states are visible to the external observer. The model is specified with a set of transition probabilities $p(y_t \mid y_{t-1})$ that governs the transition of each state to its next state, and a set of observation probabilities $p(x_t \mid y_t)$ that governs the generation of observation x_t in each state. A HMM can be built for each activity and its probabilistic parameters learned by maximizing the joint probability, $p(X,Y) = \prod_{t=1}^{T} p(y_t \mid y_{t-1}) p(x_t \mid y_t)$ in the training with actual activities. Inference is then performed by finding the model that best matches the new observed outcome (sequence of object use/sensor events). Both parameter learning and inference can be performed using algorithms commonly in use with HMM, e.g. forward-backward, Baum-Welch, and Viterbi algorithms [16].

DBN [17, 18] is a more general temporal model where HMM and their variants can be seen as special cases of DBN. Key difference between HMM and DBN lies in their model structure and dependency relationship: In HMM, only one hidden state and one observation variable is allowed in each time slice. Moreover, a hidden state at time t depends only on the previous hidden state at time t-1, while the observation variable at time t depends only on the hidden state in that time slice. DBN relaxes both of these requirements by allowing arbitrary number of hidden states and observation variables in each time slice, and dependencies among the states and variables that were assumed independent in HMM (Fig. 2). While this introduces greater computational complexity, it does offer more flexibility to the designer of DBN. Furthermore, it is possible to have one observation variable per object use/sensor event, instead of packing them all into a vector represented by one observation variable as in HMM. The one hidden state in HMM can also be represented by several constituting hidden states. From such decomposition, DBN can take advantage of the conditional independencies that may exist among them to simplify the computation of joint probability distribution, e.g. by factorization, when a complex system with many variables is involved. In contrast, compressing all hidden states and observation variables into one 'super' hidden state and observation variable in HMM causes the loss of dependency and interdependency relationships, which in turn will require more training data for HMM to converge to the true distribution. In terms of inference accuracy, DBN is found to surpass HMM in a comparative study in [19] as it uses learned dependencies between variables that are not available in HMM. The presence of missing observations is also found to have less impact on DBN than on HMM.

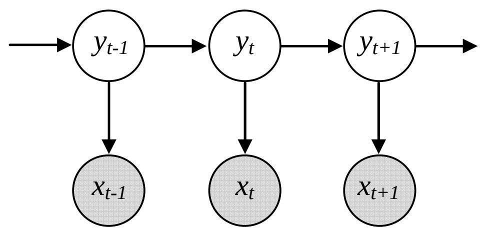


Figure 1. HMM. The clear and shaded nodes represent hidden states and observation variables, respectively.

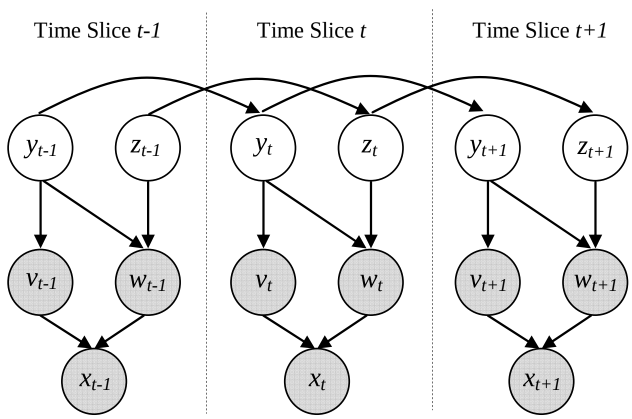
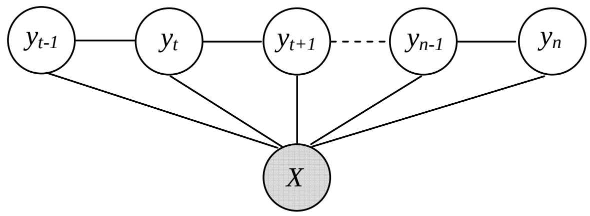


Figure 2. DBN. Allows arbitrary number of states and variables and arbitrary dependencies within a time slice. However, Markov assumption still holds for relationship between states across time-slices

Finally, CRF [20, 21] is another probabilistic model that can have many forms, but the form of most interest is the linear-chain CRF as shown in Fig. 3, due to the sequential nature of activities to be inferred. Unlike HMM and DBN, there are no directed edges in the model, and the hidden states are all linked to a single observation variable representing the entire sequence of observations over time. This means that CRF is an 'undirected' model, and it represents the conditional probability $p(Y \mid X)$ of the state sequence given the observation sequence X, rather than the joint probability $p(X, Y)$ of both the states and observations as in HMM. Since the model conditions on the entire observation sequence X, it avoids the need for independence assumption between observations as required by HMM. Thus, a hidden state y_t in CRF is not constrained to look solely at observation x_t, but has the flexibility to incorporate observations from any time slice. This makes CRF an attractive model to use for inference of activities with complex time-overlapped observations. However, training a CRF requires significantly more computation than training a HMM. For more details of this model, readers may refer to [20].



$X = x_{t-1}, x_t, x_{t+1}, .., x_{n-1}, x_n$

Figure 3. Linear-chained CRF. The model conditions on the entire sequence of observations X and therefore does not need to assume that the observations across time are independent.

4 GMT-PD

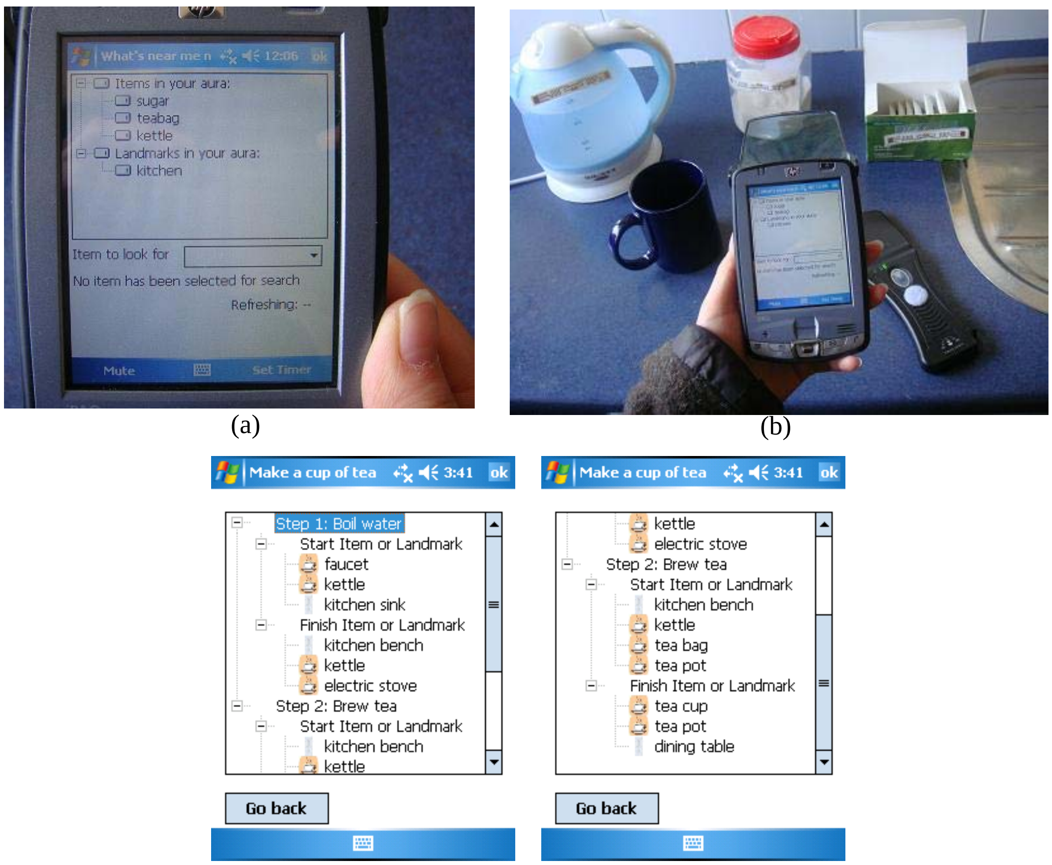
Having discussed the existing approaches for activity inference, we present an assisted living prototype, which we call a Goal Management Training Personal Device (GMT-PD) that we have developed (and still developing) for cognitively impaired patients, and discuss some of the potential design issues for inference in such applications.

Goal setting is an approach aimed to assist people to adapt and recover after injury and illness. For those experiencing cognitive decline due to brain injury or natural aging, goal attainment can be facilitated by Goal Management Training (GMT), which is a process of steps that guide the patient to complete a set of defined tasks in order to achieve everyday goals such as making a cup of tea [22]. However, GMT traditionally relies on one-on-one interaction with a rehabilitation therapist over an extended period and is thus costly and limited to weekly sessions. Thus, we want to explore whether electronic personal digital assistant (PDA) coupled with a RFID reader and tags could help patients to apply GMT principles on their own in their own time and own environment, in addition to the one hour professional training they receive to achieve a GMT outcome such as improved memory, attention and organization ability.

4.1. System Description

The GMT prototype is built on an existing RFID home support system that our group has developed for tracking and recovering losable personal objects such as eye glasses [23]. The system was developed on a HP iPAQ Pocket PC paired with a portable Tracient RFID reader (UHF) using Bluetooth. The application software was developed using C#.NET and Microsoft SQL Server CE as database engine. The system maintains a database of RFID tag identifiers corresponding to static landmarks such as kitchen sinks and cabinets, and user movable objects such as tea cups and pots. A generic customizable framework for GMT [24] was implemented with the following features:

Goal setting: For the rehabilitation therapist to set up specific activity tasks for the patient's goal training. This is done by customizing a generic goal-setting template that defines the steps or milestones required for completion of each task. Each step is further defined by a start and finish point, each identified by a set of known objects and landmarks involved. Fig. 4 illustrates the goal setting for an example two-step task: 'Making a cup of tea'. It begins with the initial registration of objects and landmarks associated with the task (Fig. 4a-b), then defines 'Boil water' and 'Brew tea' as the two steps required to complete the task, with each step being further specified with a set of objects (items) and landmark for both its start and finish items of landmarks (Fig. 4c-d). The start finish approach was adopted to allow the system to be able to prompt the user during the process and not to prompt the user when the task has been completed.



(a) (b)

(c) (d)

Figure 4. Goal setting

Goal recognition: The system regularly monitored the user's surrounding (or 'aura') to identify potential beginnings of goal tasks, i.e. by monitoring the set of objects and landmark associated with the start point of each task. When a goal task is recognized, the system started the dialogue by asking the user if he/she needed help in 'relearning' how to perform the task at hand. Goals were registered separately and may reuse objects and landmarks that were tagged in the environment. However, due to the nature of this task, participants do not complete two goals at the same time.

Goal training: Once a goal task was recognized, the system was used as a virtual goal training guide to monitor the progress of the user, and respond with sets of escalating prompts when the system detected that the user was not carrying through the task correctly. The training process goes through three phases of support: i) Pause; ii) Hint; and iii) Tell. The first phase encourages the user to pause and think of the step that he/she should be doing next. If the user did not react with the correct next step, the system prompted the user again and executed second phase of support by providing him/her with some hints. If user still fails to react as expected, the system carries out the last phase of support by telling what objects he/she needs, or where he/she should go to perform the next step. To achieve the therapeutic effect of goal training, it is desirable that the user works through the current task to completion and not abandon the task due to frustration.

4.2. Design Issues for GMT

The most difficult challenge for the GMT system was dealing with false tag non-detection. The instance of a false non-detection of an RFID tag already in range of the system was an issue. Because the system reacts directly to the presence or absence of tags to sense progress with the goal activity, the system was very vulnerable to false tag non-detection. That is, the tag is in range of the reader but is not detected on that read cycle. Where false non-detection of tags occurs, there are two implications. Firstly, the display will incorrectly show the environment around the person. That is, an item of interest is present, but because the tag was not detected, it will not show on the display. This was very confusing for our test users as they will be confronted with an object not showing on the screen that is clearly within the users environment. Secondly, if the tag that was falsely non-detected is a finishing tag, the system will recognize that the step in the goal being practiced is completed and will move on to monitor the next step when the user of the system has not actually completed the step.

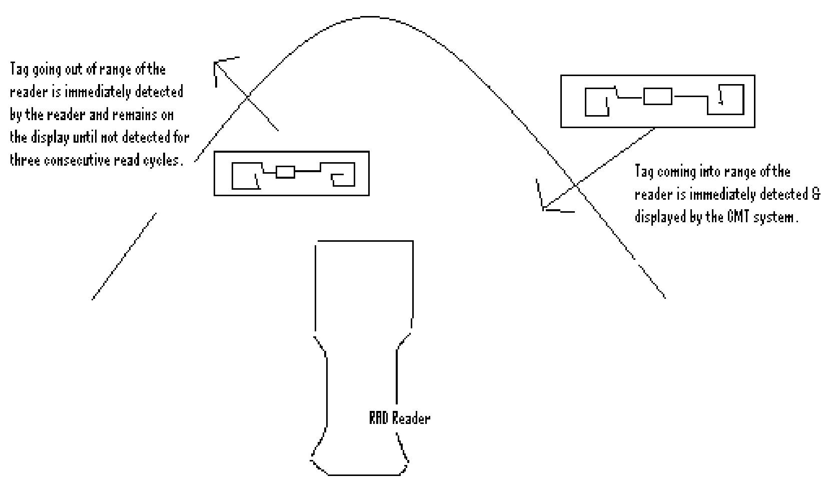


Figure 5. Approach with dealing with false tag non-detection

Our approach to overcome the false tag non-detection problem was to delay the removal of tags from the display until they were not detected in three consecutive reader cycles (refer to Figure 5). The implications of this approach for the user was that there was an approximately three second delay on completing each activity (depending on the read interval set on the RFID reader device) before the display recognized that the particular step in the activity was complete.

5 Conclusion

In this paper, we have reviewed and discussed the different sensing modalities and inference approaches used in human activity inference for RFID-based applications. We have also presented our first GMT-PD prototype, which is currently undergoing performance and usability evaluation, as well as further development. The current prototype is solely based on RFID and rule-based inference, which maps user's activities to sets of landmarks and objects. As future work, we plan to adopt a more robust probabilistic-based approach for inference, taking into account the environmental needs and unique characteristics of the users as highlighted in this paper. We also plan to revise the user interface particularly where interaction will be required during the goal training (as opposed to the setup sections of the software). We will also investigate minimizing the delay in the system recognizing task completion through the use of probability sampling. That is, using historical data from the system to calculate the probability that tag non-detection is legitimate. The contribution of this paper has been to investigate development of an Assisted Living device applied to a health context and to explore how human activity inference models can be applied to improve the reliability of a prototype utilising RFID to provide human activity inference data.

Acknowledgements

The authors gratefully acknowledge the financial support from the AUT Faculty of Design and Creative Technologies under Grant No. 08-08-JZ01, the collaborative support and input to the project on the whole from Professor Kathryn McPherson and Nicola Kayes and their team at the Health and Rehabilitation Research Centre in the Faculty of Health and Environmental Sciences, AUT and the RFID equipment sponsorship and technical support by Tracient Technologies (www.tracient.com) for this project.

References

1. K. P. Fishkin, "Ubiquitous Computing Challenges in Recognizing and Predicting Human Activity", *Proceedings of the IEEE International Conference on Fuzzy Systems*, Budapest, Hungary, 25-29 July 2004, pp. 3-7.
2. R. Magjarevic, "Home Care Technologies for Ambient Assisted Living", *Proceedings of the 11th Mediterranean Conference on Medical and Biomedical Engineering and Computing*, Ljublijana, Solvenia, 26-30 June 2007, pp. 379-400.
3. L. Ravindranath, V. N. Padmanabhan, and P. Agrawal, "SixthSense: RFID-based Enterprise Intelligence", *Proceedings of the International Conference on Mobile Systems, Applications, and Services*, Breckenridge, CO, USA, June 2008, pp. 253-266.
4. N. Yamada, K. Sakamoto, G. Kunito, K. Yamazaki, and S. Tanaka, "Human Activity Recognition Based on Surrounding Things", *Embedded and Ubiquitous Computing*, November 2005, pp.1-10.

5. D. Surie, T. Pederson, F. Lagriffoul, L.-E. Janlert, and D. Sjolie, "Activity Recognition Using an Egocentric Perspective of Everyday Objects", *Ubiquitous and Intelligence Computing*, July 2007, pp. 246-257.
6. C. Floerkemeier, and M. Lampe, "Issues with RFID Usage in Ubiquitous Computing Applications", *Pervasive Computing*, April 2004, pp. 188-193.
7. L. Bao, and S. S. Intille, "Activity Recognition from User-Annotated Acceleration Data", *Pervasive Computing*, April 2004, pp. 1-17.
8. S. Wang, W. Pentney, and T. Choudhury, "Common Sense Based Joint Training of Human Activity Recognizers", *Proceedings of the International Conference on Artificial Intelligence*, India, January 2007, pp. 2237-2242.
9. M. Stikic, T. Huynh, K. V. Laerhoven, and B. Schiele, "ADL Recognition based on the Combination of RFID and Accelerometer Sensing", *Proceeding of the Second International Conference on Pervasive Computing Technologies for Healthcare*, Tampere, Finland, February 2008.
10. J. Wu, A. Osuntogun, T. Choudhury, M. Philipose, and J. M. Rehg, "A Scalable Approach to Activity Recognition based on Object Use", *Proceedings of the IEEE 11th International Conference on Computer Vision*, Brazil, October 2007, pp. 1-8.
11. N. Krahnstoever, J. Rittscher, P. Tu, K. Chean, and T. Tomlinson, "Activity Recognition using Visual Tracking and RFID", *Proceedings of the 7th IEEE Workshop on Application of Computer Vision*, Breckenridge, CO, USA, January 2005, pp. 494-500.
12. S. Dalal, M. Alwan, S. Seifrafi, S. Kell, and D. Brown, "A Rule-Based Approach to the Analysis of Elder's Activity Data: Detection of Health and Possible Emergency Conditions", *Proceedings of the AAAI Symposium on Caring Machines*, Arlington, VA, USA, November 2005.
13. F. Doctor, H. Hagras, and V. Callaghan, "A Fuzzy Embedded Agent-Based Approach for Realizing Ambient Intelligence in Intelligent Inhabited Environments", IEEE *Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, vol. 35, no. 1, 2005, pp. 55-65.
14. R. C. Ugarte, "Human Activity Recognition from Video Sequences Using Hidden Markov Models", *Proceedings of the First International Seminar on New Issues of Artificial Intelligence*, Madrid, Spain, February 2008.
15. J. He, S. Hu, and J. Tan., "Layered Hidden Markov Models for Real-time Daily Activity Monitoring using Body Sensor Networks", *Proceedings of the 5th International Workshop on Wearable and Implantable Body Sensor Networks*, Hong Kong, June 2008, pp. 326-329.
16. L. R., Rabiner "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", *Proceedings of the IEEE*, vol. 77, no. 2, 1989, pp. 257-286.
17. T., Kasteren, and B. Krose, "Bayesian Activity Recognition in Residence for Elders", *Proceedings of the 3rd International Conference on Intelligent Environments*, Ulm, Germany, September 2007, pp. 209-212.
18. S. Park, and H. Kautz, "Hierarchical Recognition of Activities of Daily Living Using Multi-Scale, Multi-Perspective Vision and RFID", *Proceedings of the International Conference on Intelligent Environments*, Seattle, USA, July 2008, pp. 1-4.
19. N. Oliver, and E. Horvitz, "A Comparison of HMMs and Dynamic Bayesian Networks for Recognizing Office Activities", *Proceedings of the 10th International Conference on User Modeling*, Edinburgh, UK, July 2005, pp. 199-209.
20. D. L. Vail, M. M. Veloso, and J. D. Lafferty, "Conditional Random Fields for Activity Recognition", *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*, Honolulu, Hawaii, USA, May 2007, pp. 1331-1338.
21. T. L. M. Kasteren, A. K. Noulas, B. J. A. Krose, "Conditional Random Fields versus Hidden Markov Models for Activity Recognition in Temporal Sensor Data", *Proceedings of the Annual Conference of the Advanced School for Computing and Imaging*, Netherlands, June 2008.

22. B. Levine, Stuss, D.T, Winocur, G., Binns M.A., Fahy, L., Mandic M., Bridges, K. and Robertson, I.H "Cognitive Rehabilitation in the Elderly: Effects on Strategic Behaviour in Relation to Goal Management", *Journal of International Neuropsychological Society*, vol. 13, 2007, pp. 143-152.
23. D. Parry, and J. Symonds, "RFID and Assisted Living for the Elderly", *RFID and Smart Technologies for Information Convergence*, IGI Global, USA, 2009, pp. 119-135.
24. K.M. McPherson, N. Kayes and M. Weatherall, "A pilot study of self-regulation informed goal setting in people with traumatic brain injury", *Clinical Rehabilitation*, 2009; 23, pp. 296-309.