Spatio-Temporal and Context Reasoning in Smart Homes

Sook-Ling Chua, Stephen Marsland, and Hans W. Guesgen

School of Engineering and Advanced Technology, Massey University
Palmerston North, New Zealand
{s.l.chua,s.r.marsland,h.w.guesgen}@massey.ac.nz
http://muse.massey.ac.nz/

Abstract. Ambient intelligence has a wide range of applications, including smart homes. Smart homes can support their inhabitants in a variety of ways: monitoring for potential security risks, adapting the home to environmental conditions, reminding the inhabitant of tasks to be performed (e.g. taking medication), and many more. This often requires the ambient intelligence to recognise the behaviour of the inhabitants, which can be done more effectively if spatial, temporal, and contextual knowledge is taken into consideration. For example, cooking in the evening is pretty normal, but it is unusual if somebody has been cooking only an hour previously. However, if we know the context of the current situation, we may be able to explain this anomaly: having a party in the house may require several dishes to be cooked over a short interval. In this paper, we discuss spatio-temporal and context-based reasoning in smart homes and some methods by which it may be achieved.

Key words: Spatio-temporal reasoning, context-awareness, smart home, behaviour recognition, activity segmentation

1 Motivation

It is well-known that we are facing a demographic change towards an aging population. In Europe, for instance, it was reported that the number of people aged 65 and over is projected to increase from 10% of the total population in 1950 to more than 25% in 2050. In New Zealand, the number of people over 65 has doubled between 1970 and 2005 [2]. Furthermore, the global life expectancy at birth is projected to increase from 58 years in 1970–1975 to 75 years in 2045–2050 [3]. The group of elderly (aged 65 years and above) is the fastest growing segment of the world population.

Aging often results in some degree of physical disability, and even when the elderly are physically healthy, the aging process can be accompanied by cognitive impairment such as diminished sense and touch, slower ability to react, physical weakness, and memory problems. It is impossible to rely solely on increasing the number of caregivers, since even now it is difficult and expensive to find care. Additionally, many people are choosing to stay in their own homes

as long as possible, and hope to remain independent. In order for them to remain autonomous, they need to perform the basic self-help tasks also known as the 'Activities of Daily Living (ADLs)' that include bathing, dressing, toileting, eating, and so on [1]. This has lead to a large number of monitoring systems also known as 'smart homes', or 'ambient intelligence systems' that consist of a network of sensors connected to household appliances, with the aim of assisting in the activities of daily living, either directly through involvement with the person, or by alerting carers when a problem arises. Examples of such smart home projects include the Adaptive Home [5], iDorm [6], MavHome [7], PlaceLab [4], Georgia Tech Aware Home [8], and Gator Tech Smart House [9].

For a smart home to react intelligently to its inhabitant's needs, the system needs to recognise their behaviour and to use spatio-temporal information, such as where (in which room?) and when (at what time) did a particular event occur? Additionally, and possibly more importantly, the contextual information (how was the current situation reached? what else is happening? what is the state of the environment?) needs to be considered.

Although, to some extent, the location of the sensors is known *a priori* from the sequence of sensor observations, it is not enough to allow effective reasoning. Using the example from [12], it would be unusual for somebody to walk around in a triangle repeatedly in the living room from the sofa to the window and then to the television. But it makes sense if that person is walking around a triangle repeatedly between fridge, cabinet and stove in the kitchen. And even in the kitchen, this behaviour would be unusual if it occurs in the middle of the night. One issue with both spatial and temporal resolution is that the scale on which it is measured can change the analysis. For example, an event can happen once a year (birthday, Christmas, etc.), weekly (visit from a health worker), daily (showering), or repeatedly during the day (phone calls). If the birthday was being celebrated every day then this would be something that the smart home should recognise as unusual. However, the temporal resolution may have to be even finer to recognise many potentially dangerous events (e.g., microwave oven used for too long).

Besides spatio-temporal information, context awareness also play an important role in interpreting behaviour, and should not be treated independently. For example, it is normal for a person to take an afternoon tea in the garden (spatial) if it is summer and it occurs during the day (temporal). The context information could include details such as whether the cup is filled with tea, or whether the person is standing, squatting, or sitting in the garden, as well as what they were doing earlier in the day and who else is around. For example, it may be normal for a person to boil water in the middle of the night if the weather is cold, and the person has just finished watching a movie. If we know the context of the situation, we can reason that the person was in the living room and then goes to the kitchen (spatial), and since it is Saturday (temporal), the person stays up longer.

The above examples clearly show that space, time and context all play an important role in behaviour recognition. Representing all of this information

in the smart environment is a significant challenge. This paper discusses the importance of spatio-temporal and context awareness, and how to represent them in behaviour recognition, the first part of the smart home problem.

2 Behaviour Recognition

The smart home uses sensors to collect information about the inhabitant's activities. These could be from a wide variety of sensors, including video cameras, microphones, on-body sensors, or Radio Frequency Identification (RFID) tags. However, we are not directly interested in the types of sensors used, but rather how the sensory signals are being processed independent of the sensor type. We assume that the sensor output arrives in the form of 'tokens' in a sequence over time. The tokens could be the direct representation of the current sensor states being triggered (i.e., kitchen light is turned off, heater is switched on, bedroom door closed, etc.), but they do not have to be. Table 1 shows an example of a sequence of tokens from the sensors.

Date Activation	*Time Activation*	*Room*	*Object Type*	*Sensor State*
16/6/2008	18:05:23	living room	television	off
16/6/2008	18:08:19	living room	curtain	closed
16/6/2008	18:09:48	kitchen	light	on
16/6/2008	18:10:35	kitchen	cabinet	open
16/6/2008	18:25:06	kitchen	fridge door	open
16/6/2008	19:00:02	laundry	washing machine	on
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

Table 1. Example of a sequence of tokens from the sensors where the sensor states are in the form of on/off or open/closed. A representation of actual tokens may contain more attributes such as sensor identification, the source of sensor, types of sensor (reed switch, leak detector, motion), etc.

Since human behaviours periodically change and the exact activities are not directly observed, it is certainly hard to model using a deterministic approach, which relies on some known properties of the behaviours, which is difficult to determine beforehand. One way to deal with these is to use a stochastic approach where the variable states (activities) are determined using a probability distribution. Given that we have this sequence of tokens obtained from the sensors, the question is then how to recognise behaviours. The challenges in this task are that behaviours are rarely identical on each use; the order in which the individual components happen can change, the length of time each piece takes can change, and components can be present or absent at different times (for example, making a cup of tea may involve milk, or may not, the milk could be added before or after the water, and the length of time the teabag is left in

the cup can vary). Adding in the fact that the exact activities are not directly observed and that sensor observations are themselves intrinsically noisy, it is no surprise that Hidden Markov Models (HMMs), and variants of them, have been the most popular method of recognising behaviours [17–19].

The HMM is a probabilistic graphical model that uses a set of hidden (unknown) states to classify a sequence of observations over time. For example, the sensor observations could be that the bathroom fan is on and the water is running where the possible state could be that someone is taking a shower. By linking all these sequences of sensor observations over time into activities, it can help to recognise behaviours. Fig. 1 shows a simple representation of a HMM where the nodes represent the variables and the edges represent the conditional dependencies between the nodes. Further details on HMMs can be found in [10, 11].

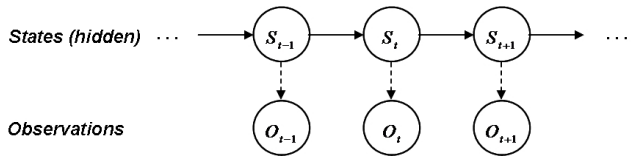

Fig. 1. The graphical representation of a Hidden Markov Model. The nodes represent the variables (top line is the hidden states, and bottom line is the observations). The edges (arrows) represent the conditional dependencies: solid lines represent the conditional dependencies between states, $P(S_t|S_{t-1})$ i.e. the probability of transition to a state S_t, which depends only on the current state S_{t-1}, and dashed lines represent the conditional dependencies of the observation on the hidden states, $P(O_t|S_t)$ i.e. the observation at O_t depends only on the hidden state S_t at that time slice.

In order to use HMMs, there are a few problems that have to be solved. One is to break the token sequence into appropriate pieces that represent individual behaviours (i.e., segmentation), and another is to classify the behaviours using the HMM. In previous work [16] we have proposed a method of automatic segmentation of the token stream, based on competition between a set of trained HMMs. In this paper, we propose methods by which this can be augmented with spatial, temporal, and contextual information in order to improve the classification accuracy. We do not cover the problem of training the individual HMMs; at present we use hand-labelled data to do this, although we plan to identify a better approach in future work.

Although many methods have been proposed to label inhabitant's activities, these approaches relied heavily on hand-labelled and manual segmentation, both which are time-consuming and error-prone. Some works even progressed towards using the sliding window technique to partition the input sensory stream but they do rely on a fixed window length, which results the segmentation being biased towards the size of window used [13, 14]. To address this, an initial study has been conducted that uses a set of Hidden Markov Models (HMMs) with

each HMM recognises different behaviours and compete among themselves to explain the current sensor observations. We will discuss how we can further incorporate the spatio-temporal and context information to the existing model. Before discussing this further, we demonstrate that such a system is capable of identifying human behaviours from a real smart home.

2.1 Experiment: Behaviour Recognition using HMMs

This experiment was conducted to perform behaviour recognition and segmentation so that behaviours can be individually segmented based on competition between trained HMMs. Given a set of HMMs trained on different behaviours, we present data from the sensory stream to all of the HMMs, which each computes the likelihood of the sequence of activities according to the model of each behaviour. We posit that a typical behaviour is a sequence of activities that occur close to one another in time, in one location. While this is not always true, for now we are focussing on these types of behaviour, which includes activities such as cooking and preparing beverages. It is interesting to note that it would not necessarily include common activities such as laundry, which may well be separated in time (while waiting for the washer to finish) and in space (for example, if clothes are hung outside rather than using a dryer).

In order to demonstrate our algorithm, we took a dataset from the MIT PlaceLab [4]. They designed a system based on a set of simply installed state-change sensors that were placed in two different apartments with real people living in them. The subjects kept a record of their activities that form a set of annotations for the data, meaning that there is a 'ground-truth' segmentation of the dataset. We trained the HMMs using this hand-segmented and labelled data. While this is a simplification of the overall aims of the project, it enables us to evaluate the method properly.

We assume for now that activities take place in one room, and that the location of the sensors is known *a priori*. For this reason, we concentrated on just one room, namely the kitchen, which contained more behaviours than any other room. The behaviours that were originally labelled in the kitchen were (i) prepare breakfast, (ii) prepare beverage, (iii) prepare lunch, and (iv) do the laundry. We split behaviour (i) into two different ones: prepare toast and prepare cereal. This made two relatively similar behaviours.

We partitioned the data into a training set consisting of the first few days, followed by a test set consisting of the remainder. The HMMs were each trained on the relevant labelled data in the training set using the standard Expectation-Maximization (EM) algorithm [10]. The data that is presented to the five trained HMMs is chosen from the sensor stream using a Parzen window that moves over the sequence. The choice of the size of this window is important, because it is unlikely that all of the activities in the sequence belong to one behaviour, and so the HMM chosen to represent it will, at best, represent only some of the activities in the sequence. Rather than using a fixed window length, we proposed a variable window length that moves over the sequence of observations. We do not discuss the details on how our algorithm self-determines the Parzen window size (see [16]

for further details), but focus on how effective our algorithm is at recognising behaviours based on the competition among HMMs in that particular room in the house. To simplify this experiment, we use a Parzen window of size 10.

The results of sliding a Parzen window of size 10 over the data consisting of 727 sensor observations is shown in Fig. 2, which displays the outputs of the algorithm, with the winning behaviour at each time being clearly visible. The winning behaviour is classified when the $\alpha = 1$. As the figure shows, we can determine that the subject is doing laundry at observation 150 since $\alpha = 1$ at that observation (see first graph in Fig. 2). The classification accuracy of this experiment was high enough (over 90% accuracy) to encourage us to look further.

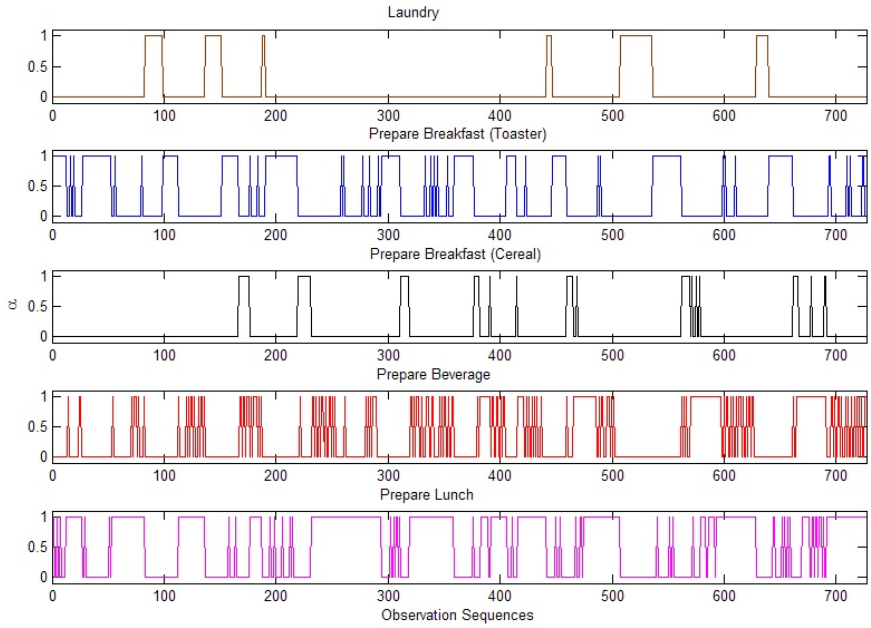

Fig. 2. Illustration of competition between HMMs based on a testing set of 727 observations

2.2 Results

The experimental results show that the method works effectively to detect changes of activities based on relatively small amount of training data. As the model is relatively simple and based on recursive computation, the computational costs are significantly lower than many other methods.

One of the limitations found in the study is that the method misclassified the behaviour in situations where the end of one behaviour contains observations

that could be at the start of the next. For example, the last activity for preparing lunch could be to put the leftover food in the fridge. After preparing lunch, the inhabitant proceeds to make a cup of coffee, and the first activity to make a cup of coffee is to take the milk from the fridge (see observation O_5 in Fig. 3). This will not pose a problem if the second behaviour happens immediately after the first. However, if the second behaviour happened two hours after the first, that would be a totally different unrelated behaviours.

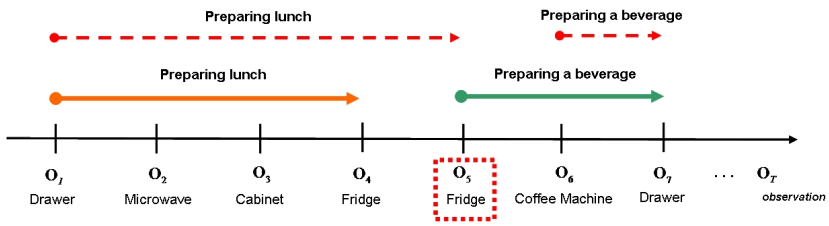

Fig. 3. Misclassification (shown in dashed arrow) occurs when the end of one behaviour contains the observations that could be in the start of the next behaviour

One way to reduce the misclassification is by adding extra information. If temporal information is included, then places where two behaviours abut one another can be reduced (see Fig. 4). If spatial information is included, then places where the two different unrelated behaviours occur can be reduced, since now the room location will change before the second behaviour occurs.

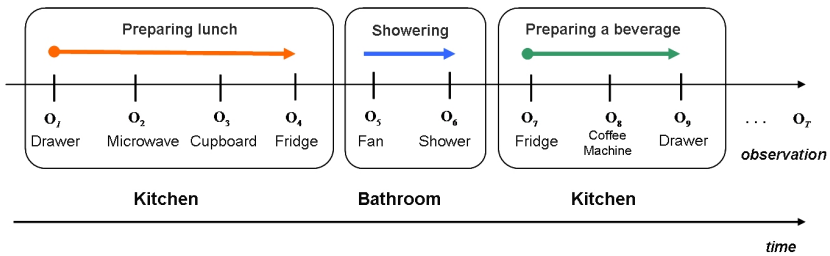

Fig. 4. Illustration of how spatio-temporal information can be included to classify behaviours.

The current study assumes that actions in a behaviour are contiguous, and that all of the separate parts of the behaviour as different instances of that behaviour. This may not be the case in the real environment, as behaviours are normally interleaved: a person may well make a beverage at the same time as preparing lunch, which could be done while the laundry was running. Where a behaviour is split into relatively short pieces (e.g., a visit to the toilet), it should

be possible to recognise this. However, care needs to be taken to ensure the fact that somebody cooks 3 times a day (breakfast, lunch, and dinner) does not get bunched into one behaviour interspersed with breaks.

3 Spatio-Temporal and Context Reasoning

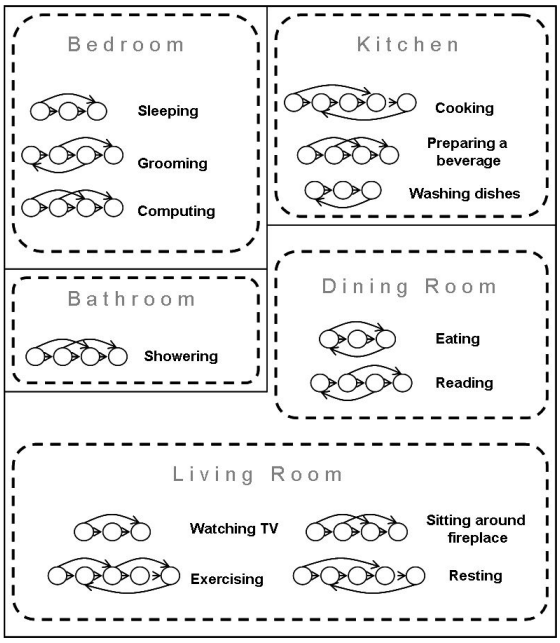

Fig. 5. Illustration of how each behaviour is defined using separate HMMs in different locations in the home (e.g. sleeping, grooming, and computing are the representation of separate behaviours in the bedroom)

We have presented a simple system that performs behaviour recognition and segmentation, and our results suggest that the method works effectively. In the experiment presented here we used very basic spatial information by concentrating on just the kitchen. To extend the method to other rooms such as bedroom, dining room, living room, and bathroom, we will need a better spatial model. We have also shown that by not using temporal information, the system makes errors that could be avoided. Fig. 5 summarises how the possible behaviours to be recognised can be identified based on the different locations (rooms) that the inhabitant is in. As the figure shows, if the behaviour takes place in the living room, the possible activities are 'watching TV', 'exercising', 'sitting around fireplace' or 'resting'. Assuming that the system is based on data of what the person does in their house (for example, during some training phase), it is reasonable

to assume that these behaviours cover the possible actions. Note that we are not interested in the exact coordinates of the activity, but rather in the room where the activity occurs; it may be that for other activities, a finer spatial resolution is needed.

Spatial information is not enough to classify whether a person's behaviour is typical or not, however. Without temporal information, the system cannot differentiate between a person showering in the bathroom at 3am and at 8am. Referring to the example shown in Fig. 6, when 'watching TV' is chosen as the winner behaviour, the system may know that this behaviour can occur throughout the day, but is mostly seen during the night, and when 'sitting around fireplace' is chosen as the winner, the system might recognise that this behaviour only occurs in the winter. Thus, time gives us another way in which we can segment the activities that should be recognised. This is shown in Fig. 6. Of course, once we have both spatial and temporal data, we need to fuse them in some way. This can be relatively simple in this case, since we can merge our two partitions.

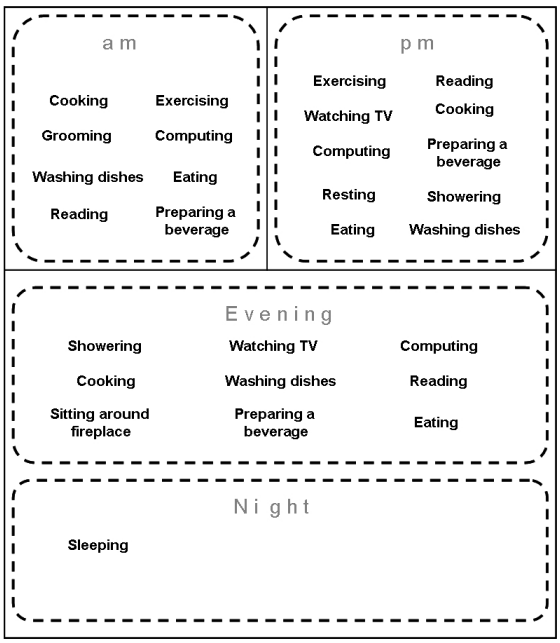

Fig. 6. Illustration of how each behaviour is defined in different time of a day

The second place where temporal information can be useful is when we form a pattern of how the smart home inhabitant spends their days. The output of the original system is a list of behaviours and the time they occur, and another system can therefore be used to learn a model of the behaviour sequence. This

could be another HMM, for example. By adding in knowledge of the length of time that people spend on particular activities, a meta-level description of behaviours can be achieved. Fig. 7 shows an interpretation of this.

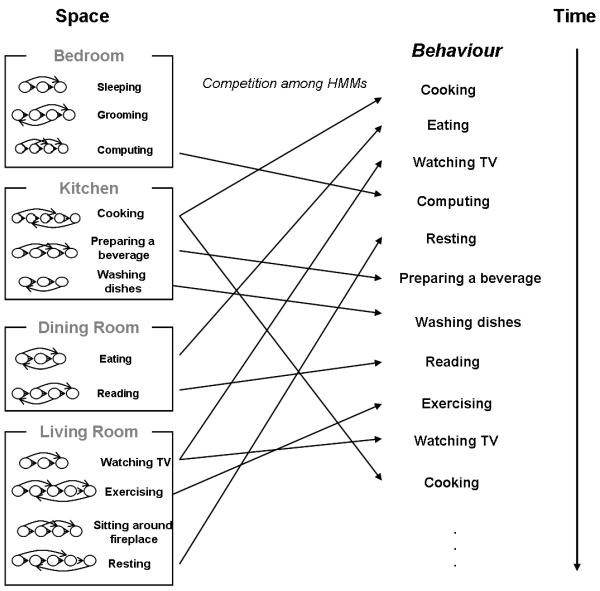

Fig. 7. Illustration on how temporal information can be useful to generate a sequence of behavioural patterns

In this case, we are starting to use information about previous activities that the person has been performing to classify their current behaviour. However, even here, we are missing relevant data. There is a host of other information that could be useful, either about the state of the environment (temperature, whether the heater is on, etc.) or how the current situation was reached. A person who reaches home from work in the middle of the night and proceeds to the kitchen to make some snack is considered normal, while a person that has been soundly asleep throughout the night and suddenly wakes up to cook may not be.

Having shown that context awareness and spatio-temporal information are useful in the recognition process, we need to decide how to include them. One option is implicit in Figs. 5 and 6, since we can use the current time and place data from the sensor stream to limit the number of HMMs that are allowed to compete. However, this may make mistakes, particularly with time, if the person is late one day and makes lunch at 3pm. Rather, some form of weighting system could be used to suggest how relevant each behaviour is based on the time and spatial data. This could be implemented in the form of a fuzzy logic system, for example.

4 Conclusion

We have shown that competition between HMMs is a possible mechanism for behaviour recognition and segmentation in a smart home. In this paper, we have discussed how context awareness and spatio-temporal information can improve the accuracy of this system, and how it can be used to better recognise when behaviours are not typical.

Algorithms for behaviour recognition generally fall into two categories: those that are based on an explicit representation of behaviours together with the events that characterise them, and those that mine them from sensor streams. The second has the advantage that we don't need to know what events constitute a behaviour, and therefore they are the preferred approach by many researchers. However, most approaches do not fully exploit the data but mainly focus on which sensors are triggered, and use these sensor sequences for learning. We are using the extra information in the stream (based on implicit information to represent behaviours where we use a set of hidden states to classify a sequence of observations over time), which falls into three categories: spatial, temporal, contextual.

For a smart home to support inhabitant's daily activities, the system should not only recognise behaviours, but also to monitor potential abnormality. A system could learn whether a novel input presented to it is a completely new behaviour, additional information to describe the current learned behaviours (which may be due to newly added sensors), or an abnormal behaviour. We are particularly interested in the use of novelty detection methods to address this, although there may be other suitable machine learning methods. The idea behind novelty detection is to train on a set of normal behaviours consisting of spatial, temporal and contextual information and then using the learned 'normal' behavioural models to identify inputs that do not fit into the pattern of the training set [15]. Applying these ideas to behaviour recognition in smart homes is the future direction of this research.

Acknowledgments. This study is conducted as part of PhD research in the School of Engineering and Advanced Technology at Massey University. The financial support from Massey University is gratefully acknowledged. We also acknowledge the support of the other members of the MUSE group (`http://muse.massey.ac.nz`).

References

1. Newcomer R., Kang T., laPlante M., Kaye S.: Living Quarters and Unmet Need for Personal Care Assistance among Adults with Disabilities. J. Gerontology: Soc. Sci. 60(B), 4, S205–S213 (2005)
2. Dunstan K., Thomson, N. : Demographic Aspects of New Zealand's Ageing Population, 2006. `http://www.stats.govt.nz/products-and-services/papers/demographic-aspects-nz-ageing-population.htm`

3. Department of Economic and Social Affairs, Population Division. : World Population Prospects: The 2006 Revision. http://www.un.org/esa/population/publications/wpp2006/WPP2006_Highlights_rev.pdf
4. Tapia E.M., Intille S.S., Larson K.: Activity Recognition in the Home Using Simple and Ubiquitous Sensors. Pervasive, pp. 158–175. (2004)
5. Mozer M.C.: Lessons from an Adaptive Home: Smart Environments: Technology, Protocols, and Applications. In: Cook, D.J. and Das, S.K. (eds.) pp. 273–298. Wiley, New York (2005)
6. Hagras H., Callaghan V., Colley M., Clarke G., Pounds-Cornish A., Duman H.: Creating an Ambient-Intelligence Environment using Embedded Agents. Intelligent Systems, IEEE. 19, 6, 12–20 (2004)
7. Youngblood G.M., Cook D.J.: Data Mining for Hierarchical Model Creation. IEEE Transactions on Systems, Man, and Cybernetics, Part C. 37, 4, 561–572 (2007)
8. Kidd C.D., Orr R., Abowd G.D., Atkeson C.G., Essa I.A., MacIntyre B., Mynatt E.D., Starner T., Newstetter W.: The Aware Home: A Living Laboratory for Ubiquitous Computing Research, Composing a Complex Biological Workflow through Web Services. In: CoBuild '99: Proc. of the Second International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture. pp. 191–198. Springer-Verlag (1999)
9. Helal S., Mann W., El-Zabadani H., King, J., Kaddoura, Y.,Jansen, E.: The Gator Tech Smart House: A Programmable Pervasive Space. Computer. 38, 3, 50–60 (2005)
10. Rabiner L.R.: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proc. of the IEEE. 77, 2, 257–286 (1989)
11. Juang B.H., Rabiner L.R.: Hidden Markov Models for Speech Recognition. Technometrics. 33, 3, 251–272 (1991)
12. Guesgen H.W., Marsland, S.: Handbook of Ambient Intelligence and Smart Environments: Spatio-Temporal Reasoning and Context Awareness. In Nakashima H., Aghajan H., Augusto J.C. (eds.) Springer (2009).
13. Niu F., Abdel-Mottaleb M.: HMM-based Segmentation and Recognition of Human Activities from Video Sequences. In: IEEE International Conference on Multimedia and Expo, pp. 804-807. (2005)
14. Kim D., Song J., Kim D.: Simultaneous Gesture Segmentation and Recognition based on Forward Spotting Accumulative HMMs. Pattern Recognition. 40, 11, 3012-3026 (2007)
15. Marsland S., Nehmzow U., and Shapiro J.: Environment-Specific Novelty Detection.From Animals to Animats, the 7th International Conference on Simulation of Adaptive Behaviour, pp. 36–45.(2002)
16. Chua S.L., Marsland S., Guesgen H.W.: Behaviour Recognition from Sensory Streams in Smart Environments. *Submitted for Review.* (2009)
17. Duong T.V., Bui, H.H., Phung D.Q., Venkatesh S.: Activity Recognition and Abnormality Detection with the Switching Hidden Semi-Markov Model. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1, 838–845. (2005)
18. Liao L.,Patterson, D.J., Fox D., Kautz, H.: Learning and Inferring Transportation Routines. Artificial Intelligence. 171, 5-6, 311–331 (2007)
19. Nguyen N.T., Phung D.Q., Venkatesh S., Bui H.: Learning and Detecting Activities from Movement Trajectories using the Hierarchical Hidden Markov Model. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2, 955–960. (2005)