Journal of Engineering and Applied Sciences 11 (8): 1851-1854, 2016
ISSN: 1816-949X

Simultaneous Activity Segmentation and Recognition in Smart Homes

Saed Sa'deh Juboor, Sook-Ling Chua and Lee Kien Foo
Faculty of Computing and Informatics, Multimedia University,
Persiaran Multimedia, 63100 Cyberjaya, Malaysia

Abstract: Activity recognition systems learn from sensors readings to recognise activities of the occupants in a smart home. In order to recognise activities, the sensor stream has to be segmented before any classification can be carried out. Many methods handle segmentation and recognition separately. In this study, we propose a method that can segment and recognise activity simultaneously by using a set of trained hidden Markov models and Viterbi algorithm. We evaluate our proposed method on two publicly available smart home datasets.

Key words: Sensors readings, sensor stream, segmentation, recognition, viterbi algorithm

INTRODUCTION

The number of people (aged 60 year and above) is expected to double from 841 million in 2013 to >2 billion in 2050 (UN, 2013). As people age, they are prone to develop some degree of decline in cognitive such as memory problem, hearing problem, poor vision, etc. Many older adults prefer to stay in their own home and it is difficult to rely solely on caregivers given the high cost of care. These have made the research in smart homes a topic of interest to address the challenges of the aging populations by monitoring their activities.

State-change sensors are commonly used in smart homes to collect information about the occupant. These sensors are normally attached to the household objects such as fridge, light, etc and are activated when the occupant performs their daily activities, e.g., opening the fridge, turning on the light, etc. The aim of activity recognition is to infer the activities of the occupant from a series of sensor readings. One of the challenges in activity recognition is the segmentation of the sensor readings into appropriate sequence that represent individual activity before classification can be performed.

The majority of the researcher in activity recognition assume that the sensor stream has been segmented or treat activity segmentation and recognition separately. However, both activity segmentation and recognition should not be treated separately if the recognition system is to be used in the real world. In this study, we address the issue of activity segmentation and recognition simultaneously in a smart home. We evaluate our proposed method on two distinct smart home datasets.

Literature review: The majority of the existing researcher used a fixed window length to partition the sensor stream. Tapia *et al*. (2004) used a set of feature windows each determined by the average duration that each activity takes. They used the naive Bayes classifier to recognise the activity by shifting the feature windows over the sensor stream and the class with the highest likelihood was selected as the activity. Discretized the sensor stream into fixed length of 60 sec and used the hierarchical Hidden Markov Model (HMM) for activity recognition. The research of Pham and Phuong (2013) used a fixed window length of 1.8 sec for segmenting accelerometer signals and used the HMM for recognising human physical activities. Assam and Seidl (2014) segment the sensor signals into several windows, with 1024 samples per window and used the HMM for activity recognition. Wu *et al*. (2014) used support vector machine with HMM to recognise daily human activities. The data stream is segmented into fixed frame image. However, one of the problems with fixed window length is that it may lead to inaccurate segmentation since the window may contain more than one activity.

Okeyo *et al*. (2014) proposed a dynamic varied time window algorithm for activity segmentation. A default time window is used and the size of the window is dynamically shrunk or expanded depending on whether the activity in that window can be identified by the recognition algorithm based on ontological reasoning.

Corresponding Author: Saed Sa'deh Juboor, Faculty of Computing and Informatics, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Malaysia

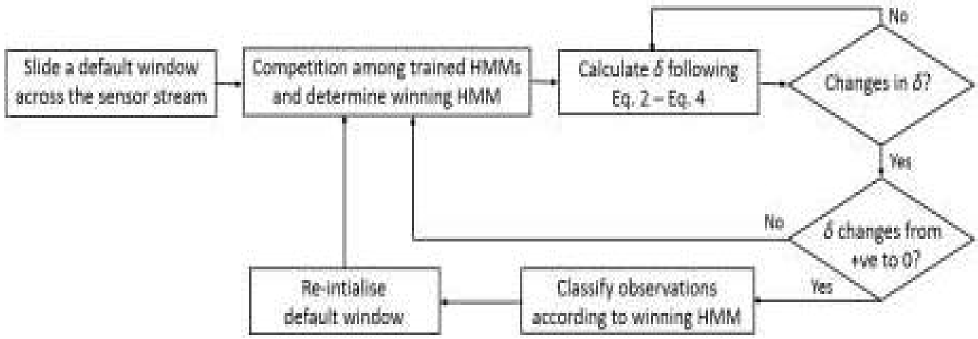

Fig. 1: Summary of our method for simultaneous activity segmentation and recognition.

Karaman *et al*. (2014) used a hierarchical two-level HMM for activity recognition on motion data. The data stream is segmented using global motion estimation. Another approach that is closely related to our method is the study of which performed a re-segmentation on the sensor sequence that has been recognised from a competition among a set of trained HMMs.

MATERIALS AND METHODS

Proposed method: In our study, we used the Hidden Markov Model (HMM) for activity recognition. We trained a set of HMMs that each recognises an activity. The observations are the sensor readings and the hidden states are the activity actions that arise from the observations. Following the method proposed, we first use a default window and slide it over the sensor stream. We then compute the likelihood among the set of trained HMMs. A winner HMM is chosen based on the one that maximizes the likelihood of the sensor sequence in that window.

Since not all the sensor observations belong to the same activity we perform a re-segmentation using the Viterbi algorithm. In contrast to the method proposed in they use the forward algorithm for re-segmentation while we use the Viterbi algorithm. The forward algorithm sums all the paths generating a particular sequence while Viterbi algorithm finds the single most probable path.

For each sensor observation in the window we use the Viterbi algorithm to find the most probable state path according to the model of the winning HMM, i.e., $P(Q,O|\lambda w)$ for HMM λw, state sequence Q $q_1, q_2, \ldots, q_T$ and observation sequence $O = o_1, o_2, \ldots, o_T$ using:

$$\delta_t(i) = \max_{q_1, q_2, \ldots, q_{t1}} P[q_1, q_2, \ldots q_t = S_i, o_1, o_2, \ldots o_t \mid \lambda_w] \qquad (1)$$

$\delta_t(j)$ can be computed recursively using:

$$\delta_t(j) = \max_{1 \leq i \leq N} d_{t1}(i) a_{ij} b_j(o_t) \qquad 1 \leq j \leq N,\ 2 \leq t \leq T \qquad (2)$$

where the initialization of:

$$d_1(i) = \pi_i b_i(o_1) \quad 1 \leq i \leq N \qquad (3)$$

and termination is:

$$p^* = \max_{1 \leq i \leq N} \delta_T(i) \qquad (4)$$

By monitoring the Viterbi variable δ for each observation we can determine how well the winning HMM explains the observed sequences. A change in the δ value indicates a change of activity from the sensor stream. If δ >0 for the first observation at the beginning of the window, then the δ values are calculated for the subsequent observations are classified according to the winning HMM until the δ value changes to 0. On the other hand, if β = 0 for the first observation at the beginning of the window, then the ä values are calculated for the subsequent observations until the δ value changes to >0. The HMMs competition will then rerun to determine the winning HMM for this sequence of observations and the process iterates. The main reason to rerun the competition on these observations where δ = 0 is due to the fact that the current winning HMM does not explain these observations. The segmentation and recognition processes are recursively computed until the end of the sensor stream. Figure 1 summarizes the process of our proposed method.

RESULTS AND DISCUSSION

We evaluate our proposed method on datasets obtained from Kasteren *et al.* (2008) and MIT PlaceLab. Kasteren *et al.* (2008) dataset used 14 state-change

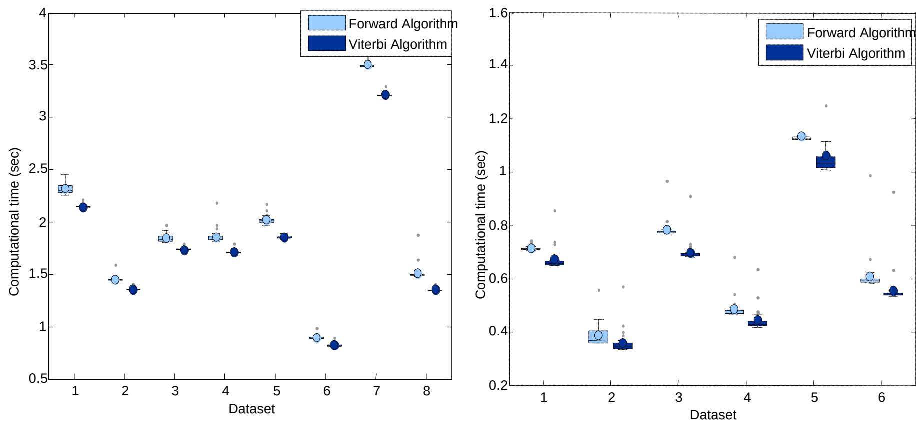

Fig. 2: Boxplot for the computational time (sec) between the proposed method and the baseline method on two datasets: a) Van Kasteren; b) MIT PlaceLab using default window size = 5

Table 1: Results on recognition accuracy (%) and computational time (sec.) of the proposed method (HMM + Viterbi) and baseline method (HMM + Forward) on (a) van Kasteren and (b) MIT PlaceLab datasets. Van Kasteren dataset

	Default window size = 5				Default window size = 10			
	Proposed method		Baseline method		Proposed method		Baseline method	
Test sets	Acc (%)	Time (sec)	Acc (%)	Time (sec)	Acc (%)	Time (sec)	Acc (%)	Time (sec)
1st	100	0.67	100	0.71	100	0.78	100	0.85
2nd	100	0.36	100	0.39	100	0.35	100	0.39
3rd	100	0.70	100	0.78	100	0.65	100	0.75
4th	100	0.44	100	0.48	100	0.55	100	0.62
5th	100	1.06	100	1.14	100	1.25	100	1.37
6th	97	0.55	97	0.61	97	0.67	97	0.73
Avg	99.5	0.63	99.5	0.69	99.5	0.71	99.5	0.79

sensors to capture the activities of a subject living in a three-room apartment over a period of 28 days. For the MIT PlaceLab. Dataset, 77 state-change sensors were installed in an apartment with a subject living in it over a period of 16 days. These datasets were annotated with activities by the subjects themselves.

We used leave-four days-out cross validation on van Kasteren dataset and leave-two days-out cross validation on MIT PlaceLab dataset. The main reason for such splits is to ensure that every activity is observed in the test sets. The HMMs were each trained using the Expectation-Maximization (EM) algorithm.

We compared our proposed method with the baseline method. We looked at the recognition performance of the algorithm to recognise activity and also the effects on the computational time performance.

Experimental results 1: recognition performance: The aim of this experiment is to compare the recognition performance of the proposed method with the baseline method. The data from the sensor stream are presented to the HMMs by using a default window that slides over the sensor stream. For this, we investigated two different default window sizes, i.e., 5 and 10 and tested on both datasets. The results are presented in Table 1.

There is no significant different in terms of the recognition accuracy for the proposed and the baseline methods on both datasets. This is expected since both methods use the HMMs for recognition. The default window size may seem to play an important role when there are more variations in the activities as can be seen from the MIT PlaceLab dataset. Thus, a smaller default window size is preferred.

Experimental results 2: computational time performance: This experiment aims to evaluate the computational time performance of the proposed method. We conducted 30 runs for each test set on both datasets using default window size = 5. The results in Fig. 2 clearly show that the proposed method has a shorter computational time compared to the baseline method across all the test sets on both datasets. This is important for real time recognition since time efficiency is vital.

We have also investigated whether the default window size has any effect on the computational performance. The results are presented in Table 1. As the table shows, the default window size = 5 has a shorter computational time than window size = 10 on both datasets. Thus a shorter window size is preferred Table 2.

Table 2: MIT PlaceLab dataset

	Default window size = 5				Default window size = 10			
	Proposed method		Baseline method		Proposed method		Baseline method	
Test sets	Acc. (%)	Time (sec.)	Acc. (%)	Time (sec.)	Acc. (%)	Time (sec.)	Acc. (%)	Time (sec.)
1st	78	1.33	78	1.47	78	2.15	78	2.32
2nd	77	0.88	77	0.97	77	1.36	77	1.45
3rd	76	1.08	76	1.18	76	1.74	76	1.85
4th	75	1.08	75	1.11	75	1.71	75	1.86
5th	78	1.24	78	1.32	78	1.85	78	2.02
6th	85	0.55	85	0.61	85	0.82	85	0.89
7th	78	2.26	78	2.44	75	3.21	75	3.50
8th	73	0.77	73	0.88	71	1.35	71	1.51
Avg	77.5	1.15	77.5	1.25	76.9	1.77	76.9	1.93

CONCLUSION

This study presents a method that can perform activity segmentation and recognition simultaneously using a set of trained HMMs and Viterbi algorithm. Experiments were conducted to evaluate the recognition performance and computational time on two publicly available datasets.

We also compared the results with the baseline method based on HMMs and Forward algorithm. Although there is no significant difference in terms of recognition performance, results have shown that the proposed method has shorter computational time performance compared to the baseline method. We have evaluated the effects of default window size and found that a shorter window size is preferred. The researcher presented in this study focuses on recognizing activities from a single inhabitant. We plan to extend our research to recognise activities from multiple inhabitants.

ACKNOWLEDGEMENTS

This research is supported by the Ministry of Education (MOE), Malaysia under the Exploratory Research Grant Scheme (No: ERGS12013ICT01 MMU0302). We would like to thank MIT PlaceLab and van Kasteren *et al*. for providing access to their datasets.

REFERENCES

Assam, R. and T. Seidl, 2014. Activity recognition from sensors using dyadic wavelets and hidden markov model. Proceedings of the 2014 IEEE 10th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), October 8-10, 2014, IEEE, New York, USA., ISBN:978-1-4799-5041-6, pp: 442-448.

Karaman, S., J.B. Pineau, V. Dovgalecs, R. Megret and J. Pinquier *et al*., 2014. Hierarchical hidden markov model in detecting activities of daily living in wearable videos for studies of dementia. Multimedia Tools Appl., 69: 743-771.

Kasteren, T.V., A. Noulas, G. Englebienne and B. Krose, 2008. Accurate activity recognition in a home setting. Proceedings of the 10th International Conference on Ubiquitous computing, September 21-24, 2008, ACM, New York, USA., ISBN:978-1-60558-136-1, pp: 1-9.

Okeyo, G., L. Chen, H. Wang and R. Sterritt, 2014. Dynamic sensor data segmentation for real-time knowledge-driven activity recognition. Pervasive Mob. Comput., 10: 155-172.

Pham, C. and T.M. Phuong, 2013. Real-Time Fall Detection and Activity Recognition Using Low-Cost Wearable Sensors. In: International Conference on Computational Science and its Applications, Murgante, B., M. Sanjay, M. Carlini, M.T. Carmelo and H.Q. Nguyen *et al*. (Eds.). Springer, Berlin, Germany, ISBN:978-3-642-39637-3, pp: 673-682.

Tapia, E.M., S.S. Intille and K. Larson, 2004. Activity Recognition in the Home Using Simple and Ubiquitous Sensors. In: International Conference on Pervasive Computing, Ferscha, A. and M. Friedemann (Eds.). Springer, Berlin, Germany, ISBN:978-3-540-24646-6, pp: 158-175.

UN, 2013. Department of economic and social affairs, population division. Current World Population, United Nations. http://www.un.org/ en/development/ desa/population/publications/.

Wu, H., W. Pan, X. Xiong and S. Xu, 2014. Human activity recognition based on the combined SVM & HMM. Proceedings of the IEEE International Conference on Information and Automation (ICIA), July 28-30, 2014, IEEE, New York, USA., ISBN:978-1-4799-4.