Explicit State Duration HMM for Abnormality Detection in Sequences of Human Activity

Sebastian Lühr[1], Svetha Venkatesh[1], Geoff West[1], and Hung H. Bui[2]

[1] Department of Computing, Curtin University of Technology,
GPO Box U1987, Perth, Western Australia
{luhrs, svetha, geoff}@cs.curtin.edu.au
[2] Artificial Intelligence Center, SRI International,
333 Ravenswood Ave, Menlo Park, CA
bui@ai.sri.com

1 Introduction

Much of the current work in human behaviour modelling concentrates on activity recognition, recognising actions and events through pose, movement, and gesture analysis. Our work focuses on learning and detecting abnormality in higher level behavioural patterns. The hidden Markov model (HMM) is one approach for learning such behaviours given a vision tracker recording observations about a person's activity. Duration of human activity is an important consideration if we are to accurately model a person's behavioural patterns. We show how the implicit state duration in the HMM can create a situation in which highly abnormal deviation as either less than or more than the usually observed activity duration can fail to be detected and how the explicit state duration HMM (ESD-HMM) helps alleviate the problem.

2 Experimentation Methodology

We recorded 150 video sequences of normal behaviour in a kitchen scenario using a single camera, each recording belonging to one of five normal classes of activity sequences one might observe in a kitchen. Motion in the room was segmented using a robust tracker and a Kalman filter was employed to track moving objects between frames. A subject's proximity to one of six areas of interest was calculated, mapped to discrete observations and recorded approximately every 1.5 seconds.

The normal classes of behaviour were designed to highlight the importance of modelling duration given the limitations of the tracking system. That is, using an impoverished observation set, the classes would have the same sequence of observations but would differ in the duration spent in a location. The first two classes, preparing cereal and making toast for breakfast, are identical in the order that the areas of interest in the room are visited and hence it is only possible to distinguish between the two classes by considering the time spent at the kitchen bench, the act of making toast taking considerably longer than the preparation of a bowl of cereal. Similarly, the classes representing dinner preparation and reheating differed only in the duration spent standing by the stove. The fifth class differs to the other classes in both the activity duration and the order in which the activities are performed. A further 24 sequences of abnormal behaviour, differing to the norm only in terms of shorter or longer times spent at any of the six locations, were recorded.

Each normal class was modelled using a standard fully connected HMM, a left-right HMM, an explicit state duration HMM (ESD-HMM) and a left-right ESD-HMM. The optimal number of states were empirically found to be 12, 2, 3 and 2 respectively. The models were trained on a random sample of 60% of the normal activity sequences and tested on the remainder. A single Gaussian distribution was used to estimate the duration probabilities in the ESD-HMM case, the model otherwise requiring an unrealistic amount of training data to accurately estimate the state duration probabilities.

3 Results

The HMM was the weakest model for classification of the unseen normal sequences with 81% accuracy, its dynamic time warping property rendering it unsuitable for use as a classifier given the type of observation sequences used in this experiment. The left-right HMM was an improvement with 97% accuracy. Although the model performed well empirically, the limited number of parameters in a two state HMM is inadequate to properly encode the sequences and hence properly discriminate between classes. Confusion between the similar activity classes was the prime source of error in both cases. The ESD-HMM models performed equally well with 100% accuracy. Explicit duration allows the model to clearly differentiate between all classes.

We classified unseen observation sequences as either normal or abnormal by thresholding on the highest log likelihood, normalised by the total length of a sequence. ROC curves were used to investigate the suitability of each of the models as a detector of abnormality. Neither the HMM nor left-right HMM models are able to reliably differentiate between our normal and abnormal sequences using the thresholding approach. The ESD-HMM increased our ability to reliably detect duration abnormality, its main cause of error a result of the model treating the observations as a cyclic activity. The left-right ESD-HMM did not exhibit this behaviour and produced good results.

Long term abnormality was investigated by artificially varying the duration of a primary activity in a randomly selected test sequence. The normalised likelihood returned by the HMM and the left-right HMM increased in proportion to the duration of the activity. The HMM and left-right HMM are therefore not suitable for the detection of highly abnormal activity duration. The ESD-HMM exhibited a similar trend, the lack of transition constraints allowing the model to briefly enter a state with a sub-optimal emission probability in order to maximise the overall sequence likelihood. The left-right ESD-HMM behaved correctly, the sequence likelihood rapidly decreasing as the activity duration was increased.

4 Conclusion

This work has highlighted the importance of explicit duration modelling for classification of sequences of human activity and the reliable and timely detection of duration abnormality. The incorporation of duration in models of human behaviour is an important consideration for systems seeking to provide cognitive support and to detect deviation in the behavioural patterns of the elderly.

References

1. Lühr, S., Venkatesh, S., Bui, H.H.: Duration abnormality detection in sequences of human activity. Technical Report TR-2004/02, Department of Computing, Curtin University (2004)