Developing C

Developing Context Sensitive HMM Gesture Recognition

Kingsley Sage, A. Jonathan Howell, and Hilary Buxton

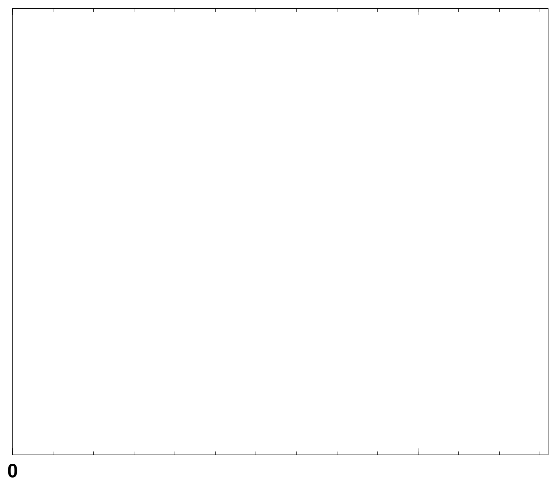
School of Cognitive and Computing Sciences,
University of Sussex, Brighton BN1 9QH, United Kingdom.

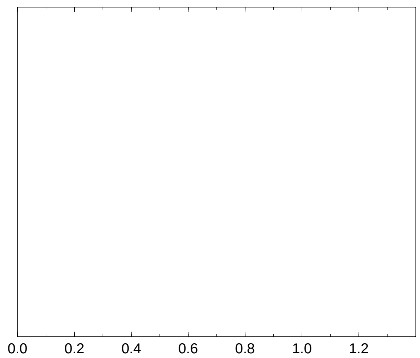
Abstract. We are interested in methods for building cognitive vision

problems of variability, is popular as it offers more sophisticated matching for many kinds of time-varying signals [15]. HMMs consist of a series of states, which nally, in sections 5 and 6, the implications of the work for task control and system integration are then discussed with conclusions and suggestions for further work.

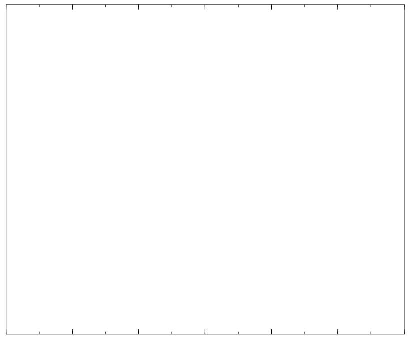
2 Hidden Markov Model for gesture recognition

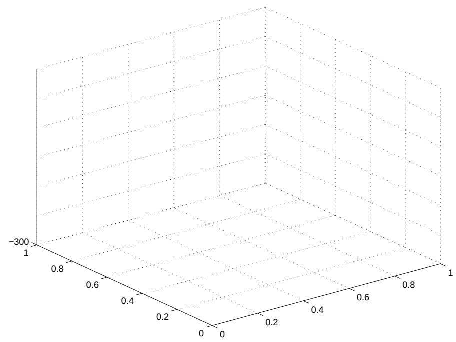
A Hidden Markov Model (HMM) is a doubly stochastic process, i.e. there is an underlying stochastic process that is not observable (hidden) but can only be observed through another set of stochastic processes that produce the sequence of observed symbols [15]. The HMM is characterised by a triple $\lambda = (\pi, A, B)$ where A is a square $N * N$ matrix of probabilities for transitions between N discrete hidden states, π is a vector of probabilities describing the initial state of the model (at time $t = 0$) and B is a $N * M$ matrix accounting for the mapusio de gnta is a


0


0.0
0.2
0.4
0.6
0.8
1.0
1.2

any given level of test set noise, generalisation is demonstrated as a diagonal classification rate matrix. Off





−300
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1

as confidence) could be used to determine model performance when context data is either unavailable of limited quality (e.g. a very early estimate from a predictive cueing process). However there is a trade-off between improvements in performance through context control and the size of the training set.

6 Conclusion

The HMM based learning can discover the temporal structure of the 3D hand gesture trajectories here from data clustering alone. The association of different interpretations with different contexts is also learnt and can allow more effective discrimination boundaries in the online system. Performance on the learning and generalisation tasks were robust to noise and scale well with task complexity, however the training with the Baum-Welch algorithm does not scale so well. The HMM developed here was coded in Matlab, thus it is premature to give computational costs but these will be established in future work. As in the discussion above, there is great potential for contextual processing using the HMM for attentive processing in the ActIPret system.

We have also proposed an approach to task control within the ActIPret system using a Dynamic Decision Network (DDN), e.g.. [9], in the Activity Reasoning Engine. However, we also want distributed control in the lower levels and one way of imposing this is by conditional probability matrices to activate the processing within each lower component, using priority metrics. Initially, it is proposed to hand code utility/task relevance nodes (e.g. watch/ignore) that determine the priority metric. In the longer term, in the context of a complete system, we hope to learn these dynamic dependencies. It may be that we can determine task-relevance automatically in this way, using a uniform Bayesian approach.

Acknowledgements

The authors gratefully acknowledge the invaluable help provided by the Laboratory for Human Motion Simulation (HUMOSIM) at the University of Michigan, USA in allowing us access to their 'Terminal Hand Orientation and Effort Reach Study, 2000' hand trajectory database; also framework concepts and funding from the EU ActIPret project.

References

1. B. Bauer, H. Heinz, and K. Kraiss. Video-based continuous sign language recognition using statistical methods. In *IEEE International Conference on Automatic Face and Gesture Recognition*, pages 440–445, Grenoble, France, 2000.
2. M.J. Black and A. Jepson. A probabilistic framework for matching temporal. In *European Conference on Computer Vision*

3. D. M. Blei and P. J. Moreno. Topic segmentation with an aspect hidden Markov model. In *Pro*