Poster abstract - Evaluation of hidden Markov models robustness in uncovering focus of visual attention from noisy eye-tracker data

Neil Cooke* Birmingham University

Martin Russell† Birmingham University

Antje Meyer‡ Birmingham University

1 Introduction

Eye position, captured via an eye tracker, can uncover the focus of visual attention by classifying eye movements into fixations, pursuit or saccades [Duchowski 2003], with the former two indicating foci of visual attention. Such classification requires all other variability in eye tracking data, from sensor error to other eye movements (such as microsaccades, nystagmus and drifts) to accounted for effectively.

The hidden Markov model provides a useful way of uncovering focus of visual attention from eye position when the user undertakes visually oriented tasks, allowing variability in eye tracking data to be modelled as a random variable.

2 Proposed Models

Two types of hidden Markov models are investigated- a standard hidden Markov model (HMM) and a hidden semi-Markov model (HSMM) [Russell and Moore 1985]. The HMM represents state duration inherently, decaying geometrically from the time of entering the state. The HSMM represents state duration explicitly, allowing more accurate representation of visual attention duration.

The visual field is represented by an ergodic HMM/HSMM with each hidden state representing a foci of visual attention. These foci are explicitly identified as regions of interest in the visual field. A state output PDF for each state describes the probable distribution of eye positions over an object while that object is the focus of visual attention.

The HMM and HSMM are implemented in two variants. The first uses a standard two-dimensional gaussian distribution as the observation PDF to represent the distribution of eye positions over an object. The second adds a second gaussian component to each dimension in the observation PDF with the same mean as the first component but with larger standard deviation (the Richter Distribution [Richter 1986]). This ensures less differentiation in the observation PDF between eye movement positions that are far away from the object compared to that of the standard gaussian PDF in the first variant.

*e-mail: cooken@eee-fs7.bham.ac.uk

†e-mail:m.j.russell@bham.ac.uk

‡e-mail:A.S.Meyer@bham.ac.uk

© 2004 ACM 1-58113-825-3/04/0003 $5.00

3 Evaluation

We evaluate performance of the HMM and HSMM in classifying focus of visual attention from eye tracking data from users undertaking a visually-oriented task. Increasing levels of random gaussian noise are added to the data. Performance degradation is measured in terms of:

Accuracy : proportion of states decoded correctly compared to a baseline non-HMM method with no added noise.

Instability : proportion of state transitions occurring compared to a baseline non-HMM method with no added noise.

4 Results

- Performance of all HMM and HSMM-based methods to added noise is better than the baseline non-HMM method.
- Best performance results from using an explicit state duration PDF (HSMM) with Richter distribution observation PDF.
- Weakened performance results from using the Richter distribution observation PDF in a HMM (as opposed to a HSMM).

5 Conclusion

Our findings show that a hidden semi-Markov model that has an explicit state duration PDF representing task-constrained visual attention is a more stable and accurate way to uncover the focus of visual attention from (simulated) noisy eye tracker data compared to using standard HMMs with inherent state duration PDF. This performance gain is only evident when the observation distribution PDF's dominance on classification is relaxed by adding an additional gaussian component to the observation distribution PDF to reduce discrimination between eye positions far away from the object.

Hidden semi-Markov models have promising uses in uncovering focus of visual attention from noisy eye tracking data. HMM-based formalisms of decoding eye movement analogous to decoding speech (and other modalities) may also facilitate development of robust decoding schemes in multi-modal systems.

References

DUCHOWSKI, A. T. 2003. Eye Tracking Methodology: Theory & Practice. Springer-Verlag, ch. 4, 43–51.

RICHTER, A. 1986. Modelling of continuous speech observations. In Advances in Speech Processing Conference, IBM Europe Institute.

RUSSELL, M. J., AND MOORE, R. K. 1985. Explicit modelling of state occupancy in hidden markov models for automatic speech recognition. In Proceedings International Conference on Acoustics Speech and Signal Processing (ICASSP 85), Tampa, 1.2.1–1.2.4.