Modeling the Transition from Bottom-up to Top-down Gaze Control Strategies in the Context of Gaze Following

Jochen Triesch[1,2], Hector Jasso[3], and Gedeon Deák[2]

[1] Frankfurt Institute for Advanced Studies, J.W. Goethe University Frankfurt, Germany
[2] Department of Cognitive Science, University of California, San Diego {triesch@cogsci.ucsd.edu}
[3] Department of Computer Science and Engineering, University of California, San Diego {hjasso@cs.ucsd.edu}

Eye movements can be controlled by bottom-up, stimulus-driven mechanisms [1], as well as by top-down, goal-driven mechanisms [2]. The relative importance of the two is a topic of current debate. Arguably, the extent to which bottom-up or top-down systems are dominant may depend on the individual's age and experience, with infants' gaze likely being driven largely by bottom-up mechanisms, and adults controlling their gaze using substantial amounts of top-down knowledge. We are interested in better understanding the developmental transition from largely bottom-up to progressively more top-down control of looking behavior. Here we present a model of this transition in the context of gaze following, which is the ability to redirect one's visual attention towards an object that someone else is looking at.

Infants progressively acquire competence in gaze following during the first two years of life. Our model of this development, which is an extension of [3], is implemented using a virtual reality platform designed for studying cognitive development [4] (see Figure 1). The virtual environment consists of a graphical environment, the *Virtual Living-room*, built using preexisting computer graphics and computer vision libraries. The infant model includes a bottom-up attention model based on color, motion, and contrast information, as well as a face detection cue. These cues are subject to habituation and serve as inputs to a reinforcement learning model with linear function approximation. In the model, bottom-up saliency cues and information about the caregiver's direction of gaze are combined in a body-centered coordinate frame. We find that initially the infant model's looking behavior is dominated by the bottom-up saliency mechanism along with random exploration of the environment. Later it learns to better and better utilize knowledge about his caregiver's looking behavior to localize rewarding objects in the environment. In particular, the infant model learns to "purposefully" look at the caregiver to determine where he is looking, and then to shift gaze in the same direction. Over time, these inferences about the caregiver's locus of attention supplement and sometimes override the bottom-up saliency mechanism, i.e. the infant model may learn to ignore salient stimuli in order to follow the cargiver's gaze.

In conclusion, we present a parsimonious, reinforcement-based model of the emergence of gaze following skills in infancy, which captures the transition from utilizing purely bottom-up mechanisms for gaze control to incorporating top-down strategies. Our virtual modeling platform is a useful complement to very abstract, simplistic modeling platforms on the one hand and full-fledged robotic models on the other hand.

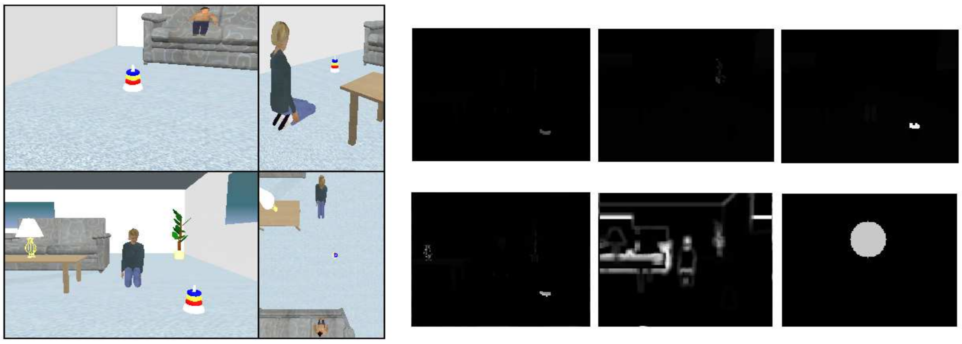

Figure 1: Left: four views of the virtual living-room environment. Right: result of simple saliency processing. The saliency maps from six saliency channels are depicted: four color channels, one contrast channel, and one "face detection" channel. The saliency maps correspond to the view of the virtual living-room shown in the lower left.

References

[1] Itti, L. and Koch, C. A saliency-based search mechanism for overt and covert shifts of visual attention, *Vis. Res.,* **40(**10-12), 2000.
[2] Hayhoe, M. and Ballard, D. Eye Movements in natural behavior. TICS **9**(4), 2005.
[3] H. Jasso, J. Triesch and C. Teuscher. A Reinforcement Learning Model Explains the Stage-wise Development of Gaze Following. Joint Symposium on Neural Computation (JSNC 2005) Los Angeles, CA, USA, May 14, 2005.
[4] H. Jasso and J. Triesch. A virtual reality platform for studying cognitive development. Proceedings of the Third International Conference on Development and Learning (ICDL04), La Jolla, California, October 20-22, 2004.

Acknowledgments

This work was supported in part by the National Alliance for Autism Research (NAAR). We thank Christof Teuscher and all members of the MESA project at UC San Diego (http://mesa.ucsd.edu) for their continuing collaboration.