Talk 306, Poster: 36

Characterizing Bayesian Surprise in Humans and Monkeys

(NSF 0515261 FY 05)

Laurent Itti
University of Southern California

Pierre Baldi
University of California, Irvine

Doug Munoz
Queen's University

We investigate the role of visual surprise in guiding eye movements in humans and rhesus monkeys under free viewing conditions for a variety of natural stimuli. Surprise differs from other models of bottom-up visual attention in that it quantifies how data affects an observer by measuring the difference between posterior and prior beliefs of the observer. We recorded eye movements from naïve subjects, 4 humans and 4 monkeys, while they watched 95 video clips, 6-30 seconds long, lasting approximately 27 minutes. This resulted in 6,672 saccades for humans and 12,263 for monkeys. Clips ranged in semantic content, including video of natural, non-natural, building-city, indoor, and sporting-outdoor scenes both with and without main actors. A surprise model of bottom-up visual attention then predicted in real-time how surprising every location was in the display. The distribution of surprise at the endpoint (target) locations of human or monkey saccadic eye movements was then compared to the distribution of surprise at random locations using a standard information theoretic technique, Kullback-Leibler distance. Both humans and monkeys saccade toward locations containing surprise significantly higher than chance ($p < 10^{-75}$). For humans the type of clip being observed has a significant effect on the amount of model agreement($p < 10^{-6}$), but for monkeys it does not($p > 0.5$). The presence of a main actor in the scene greatly effects the eye movements of humans, but has much less of an effect on monkeys($p < 10^{-10}$, $p < 10^{-5}$). Behavioral measures such as saccade velocity, fixation length and amplitude of the eye movements all correlate significantly with the amount of surprise at the saccade endpoint for monkeys but do not for humans. Additionally inter-observer analysis reveals that on some clips the model may predict monkey saccades 80% as well as another monkey, while the model can only predict 50% as well as another human. This indicates that monkeys are employing more bottom-up strategies than humans while freely viewing dynamic scenes and can serve as a good model for studying bottom-up attention in humans.

Work of: David Berg (USC), Susan Boehnke (Queen's), Robert Marino (Queen's), Pierre Baldi (UCI), Doug Munoz (Queen's), Laurent Itti (USC)

Project (or PI) Website

http://iLab.usc.edu

Publications

1. L. Itti, P. Baldi, Modeling what attracts human gaze over dynamic natural scenes, In: Computational Vision in Neural and Machine Systems, (L. Harris, M. Jenkin Ed.), Cambridge, MA:Cambridge University Press, 2006.

2. D. Berg, S. Boehnke, R. Marino, P. Baldi, D. Munoz, L. Itti, Characterizing Surprise in Humans and Monkeys, In: Proc. Vision Science Society Annual Meeting (VSS06), May 2006.

3. L. Itti, P. Baldi, Bayesian Surprise Attracts Human Attention, In: Advances in Neural Information Processing Systems, Vol. 19 (NIPS*2005), pp. 1-8, Cambridge, MA:MIT Press, 2006.

4. P. Baldi, L. Itti, Attention: Bits versus Wows, In: Proc. IEEE International Conference on Neural Networks and Brain, Beijing, China, Oct 2005.

5. L. Itti, P. Baldi, A Principled Approach to Detecting Surprising Events in Video, In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 631-637, Jun 2005.