Human gaze control in natural exploration

Frank Schumann¹, Wolfgang Einhäuser²,³, Stanislavs Bardins⁴, Klaus Bartl⁴, Guido Böning⁵, Erich Schneider⁶, Peter König¹.
¹Institute of Cognitive Science, University of Osnabrück, ²Institute of Computational Science, ETH Zürich, 
³Division of Biology, California Institute of Technology, ⁴Neurology, Hospital of the Ludwig-Maximilians University Munich

Abstract

Models of mammalian vision are often based on two assumptions: (1) Visual input is a sequence of quasi-stable images generated by large volitional gaze shifts. (2) There is some form of attentional guidance on the basis of salient visual features. Here, we continuously recorded gaze-contingent visual input in a natural, ecologically valid setting to test both assumptions under free exploration.

Methods

Eye-Tracking

Our recording device (EyeSeeCam) continuously aligns a head-mounted camera with the observer’s orientation of gaze (“gaze camera”). A second camera is fixed relative to the observer’s head (“head camera”). See accompanying poster of Thomas Villgrattner.

Subjects and Environments

Four volunteers freely explored three different environments in Munich: a local forest, Munich’s central train station, and an apartment. This yielded a total of 67 minutes of valid recording with an tracking precision of 0.5° to 1°.

Natural eye-head coordination

Gaze- and head velocity were estimated by the shift between subsequent frames of each camera using cross-correlation. Eye velocity was computed as the difference between gaze- and head-velocity.

We quantified the relative distribution of movement directions (and velocities) between the cameras going in compensatory (opposite) and synergistic (similar) directions. We find that:

1. Eye and head movements are not independent (C, D).
2. Eye and head show compensatory interaction to stabilize gaze (E, second diagonals).
3. Yet gaze is only slightly more stable than eye and head alone (A, B).
4. Eye and head also show synergistic interaction to adjust gaze (E, first diagonal).
5. Hence, synergistic interaction of eye- and head movements yields a rich temporal dynamics in natural retinal input.

1. Gaze can select features with high values under ecological recording conditions

During periods of stable gaze (fixations), high feature values accumulate at the centre of gaze, but unexpectedly also in the periphery. However, the peripheral effect is also present during fast movements. This suggests systematic heading differences between gaze and head also on a global scale. Hence we used frames of high-velocity as a control for peripheral differences to isolate only the local effect of features at fixations (last column Fig A. and Fig. B). The found effect is relatively small.

2. Situational context dominates local gaze control

In the forest environment, we find local effects for barnes, luminance-, texture-, and color contrasts at different spatial scales. By contrast, in the train station, we find only weak effects of color contrast (Fig B). This indicates a decisive role of context for the control of gaze.

Conclusions

Under conditions of natural exploration:

1. Gaze stabilization is not the only dominant function of eye-head coordination. A substantial number of (non-saccadic) eye-movements interact synergistically with the head to adjust gaze. Hence gaze dynamics needs to be explicitly modeled.
2. Gaze can select locations with high feature values. This quantifies laboratory results in an ecologically valid setting for a large variety of visual features.
3. Under natural conditions - context dominates the allocation of visual attention, employing spatial and feature selection in a top-down manner. Local feature effects at the centre of gaze vary with the environment.

Contact: Frank.Schumann@uni-osnabrueck.de

Acknowledgements: This work was financially supported by the Bavarian-Californian Technology 680 Center (BaCaTelC), the Swiss National Science Foundation (grant no. P400A1-111447 to WE), Bayerische Forschungsstiftung (FORBAS, Az 167/03), in part within the DFG excellence initiative research cluster Cognition for Technical Systems - CoTelSys, and the EU Project Perception on Purpose (POP).

References