Gaze allocation during natural behavior in the real world

Bernard Marius 't Hart[1], Johannes Vockeroth[2], Frank Schumann[3], Klaus Bartl[2], Erich Schneider[2], Peter König[3], Wolfgang Einhäuser[1]

1 AG Neurophysik, Philipps-Universität Marburg 2 Neurologische Forschung, Klinikum Großhadern, Ludwig-Maximilians-Universität München 3 Institut für Kognitionswissenschaft, Universität Osnabrück

1. Introduction

"Natural" gaze allocation is typically measured while observers view photographs of complex scenes. Such laboratory experiments do not adequately capture the rich dynamics of natural input. In part these are generated by head-in-world movements, which provide a coarse centering on potentially interesting ("salient") regions of the visual field, subsequently refined by eye-in-head movements. We address the question as to what extent laboratory experiments are informative for gaze allocation under truly real-world conditions.

2. Methods

Using a novel, fully mobile setup ("EyeSeeCam") we measure gaze position during free exploration of various environments and record head-centered videos. These are replayed in a laboratory setup, either as recorded ("video replay") or as sequences of temporally unrelated 1-second frames ("static replay").

Figure 1a. EyeLink 2000. Laboratory setup.

Figure 1b. EyeSeeCam. Mobile setup.

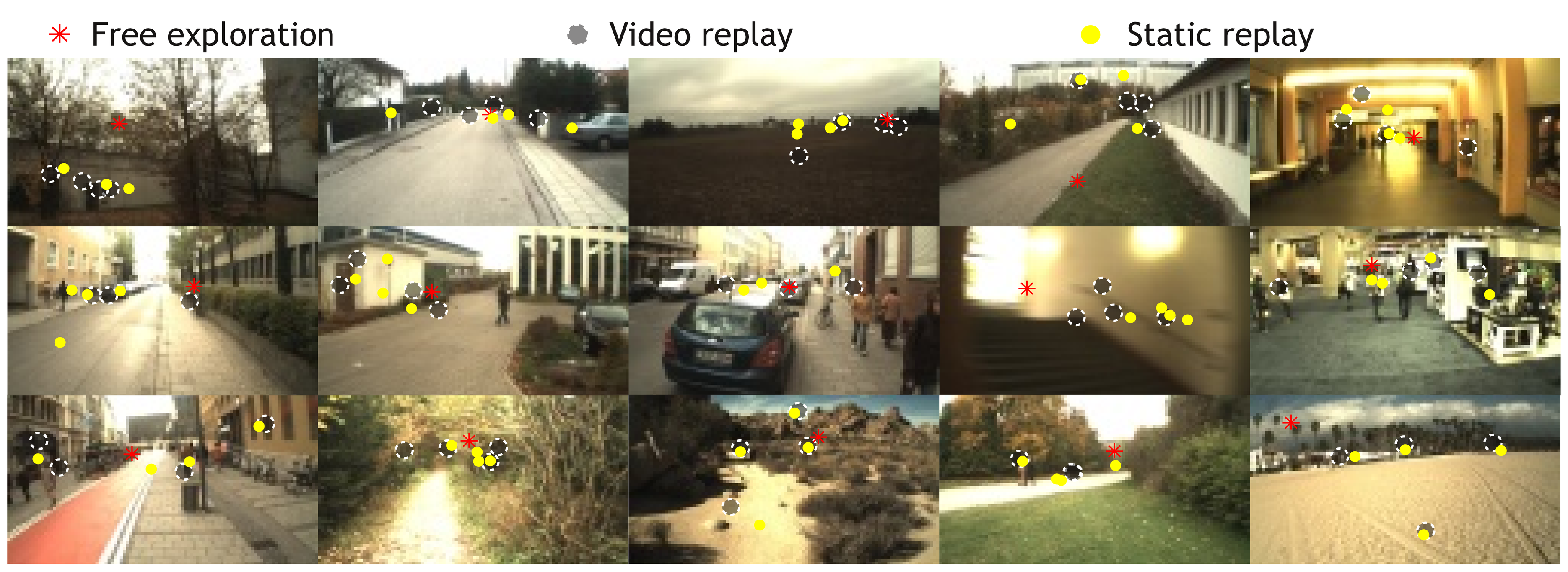

Figure 2. Environments. Example frames with fixation locations from all three conditions.

3. Spatial Bias

We find that the statistics of natural gaze allocation and thus natural retina-centered input exhibit rich dynamics, which are not adequately modeled by a series of static images interleaved by saccades. Laboratory conditions show a pronounced bias of eye position to the center of the stimulus, which is stronger for static than for dynamic presentations.

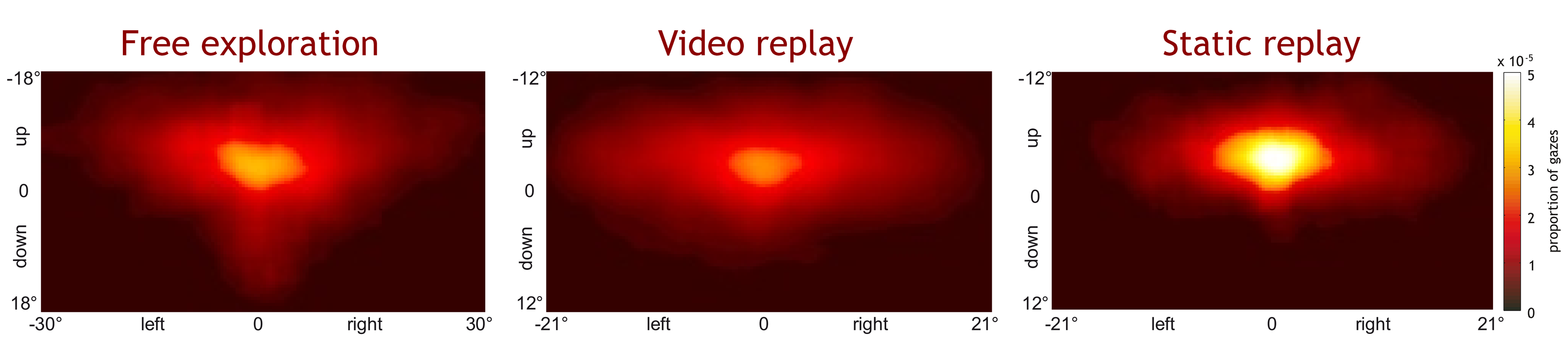

Figure 3. Gaze distributions. The static replay condition elicits a more centered gaze, while the lobe at the bottom of free exploration seems to indicate gaze to the path being actively navigated.

In static replay it appears that observers 'reorient' themselves in each frame by making saccades to the center of the screen (Fig. 4; horizontal distance). This increases inter-subject consistency between subjects in static replay (Fig. 5). Consistency of free exploration with video replay (82.6% ± 1.6%) is only slightly higher than with static replay (81.9% ± 1.0%).

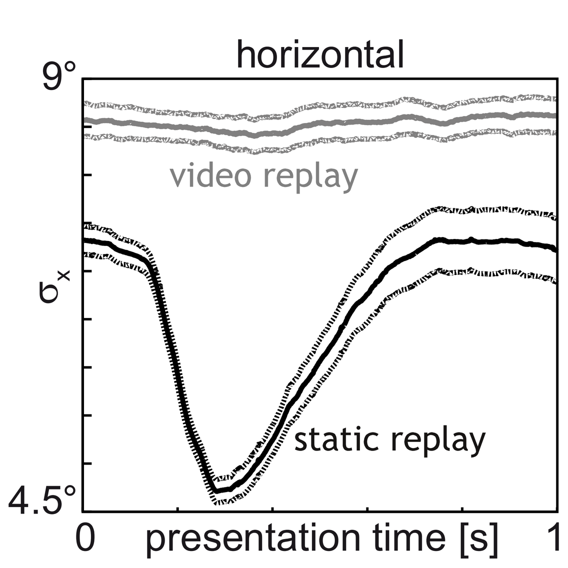

Figure 4. Horizontal distance to center.

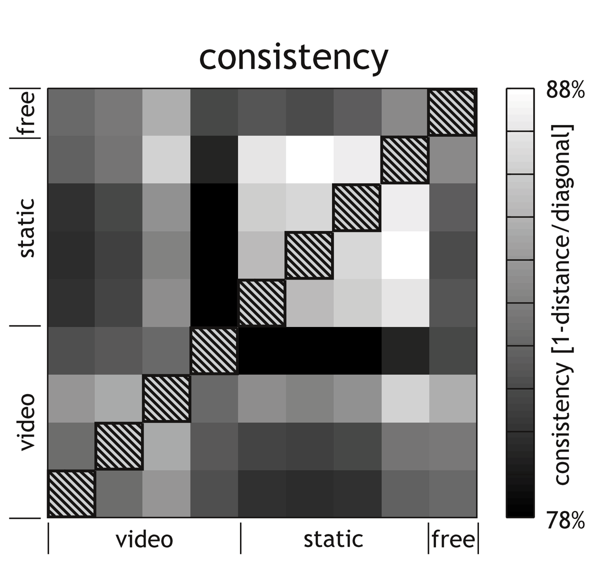

Figure 5. Inter-subject consistency.

4. Video replay predicts real-world gaze

Real world gaze is better predicted by the video replay condition than by the static replay condition in 14/15 environments (Fig. 6; $p<0.001$, sign-test). However, since prediction of real-world gaze allocation is far from optimal in either laboratory condition (maximally 69% area under ROC curve), real-world recordings are inevitable to probe the validity and restrictions of laboratory results for the real world.

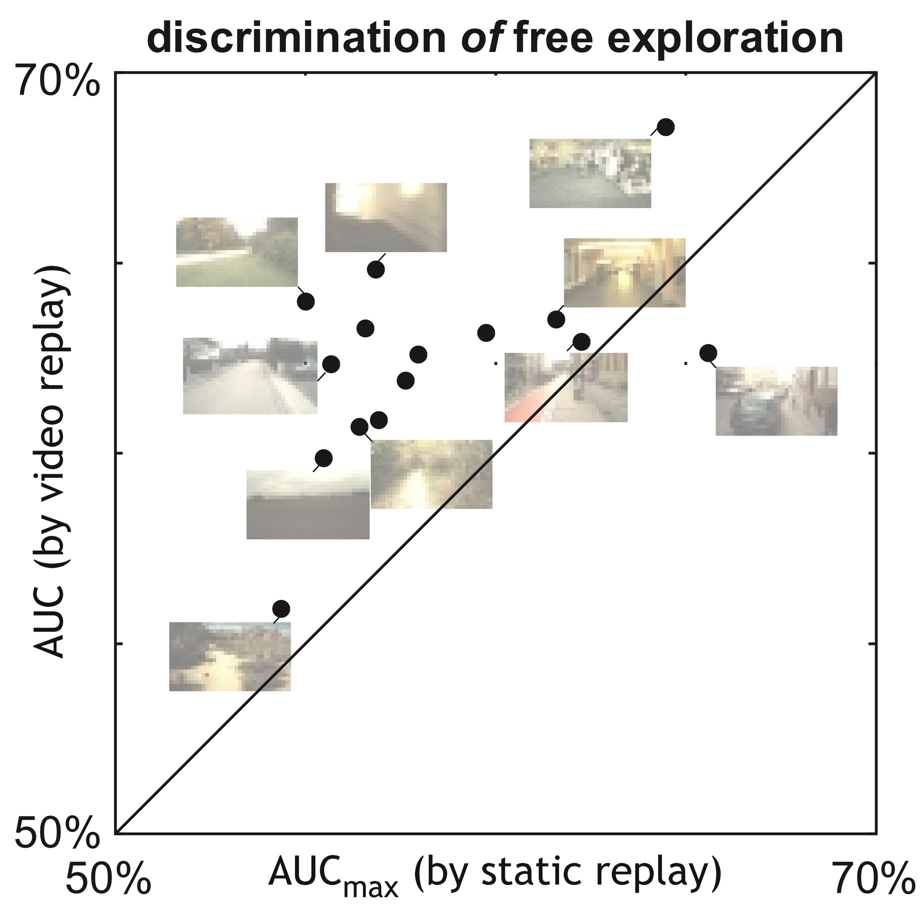

Figure 6. Discrimination of free exploration. Each dot represents the average over one movie.

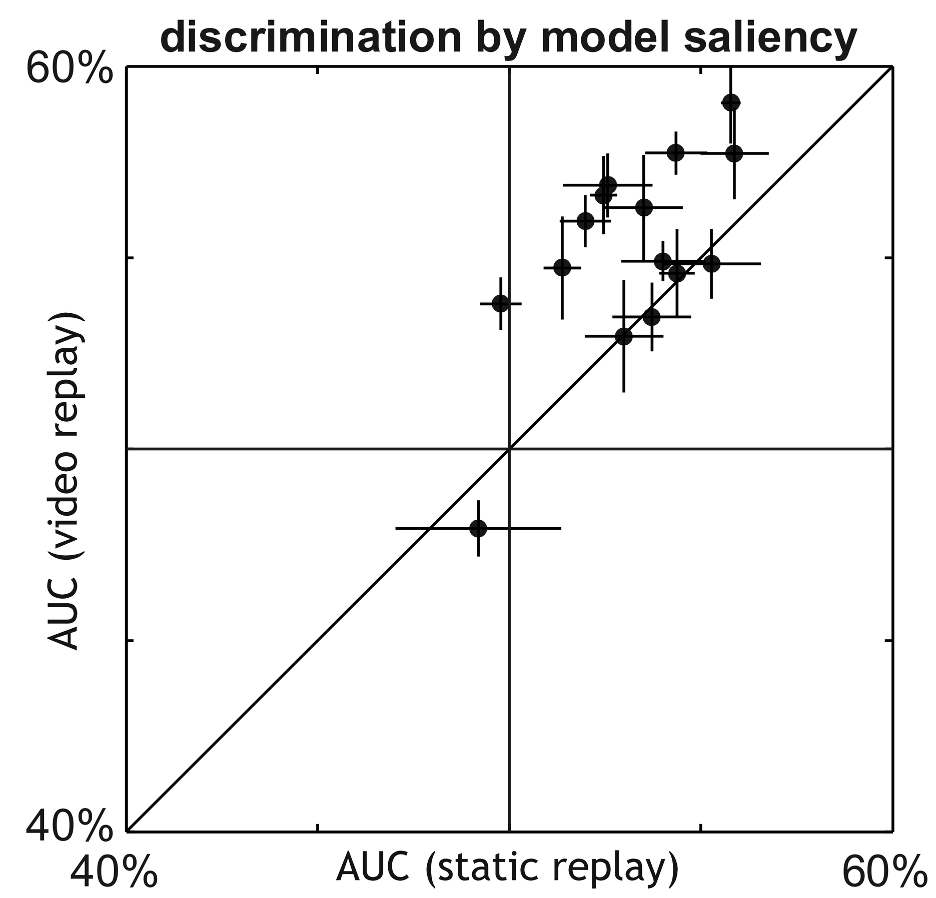

Figure 7. Discrimination by model saliency. Each dot represents the average over one movie.

A standard bottom-up model of attention, Itti & Koch's saliency map (VisRes 40:1489-1506,2000) predicts eye position better than chance in all 3 conditions, and best for continuous replay (Fig. 7).

5. Gaze centered saliency maps

Gaze centered saliency maps show a weak relationship between gaze position and saliency. In static replay this changes over the 1 second interval.

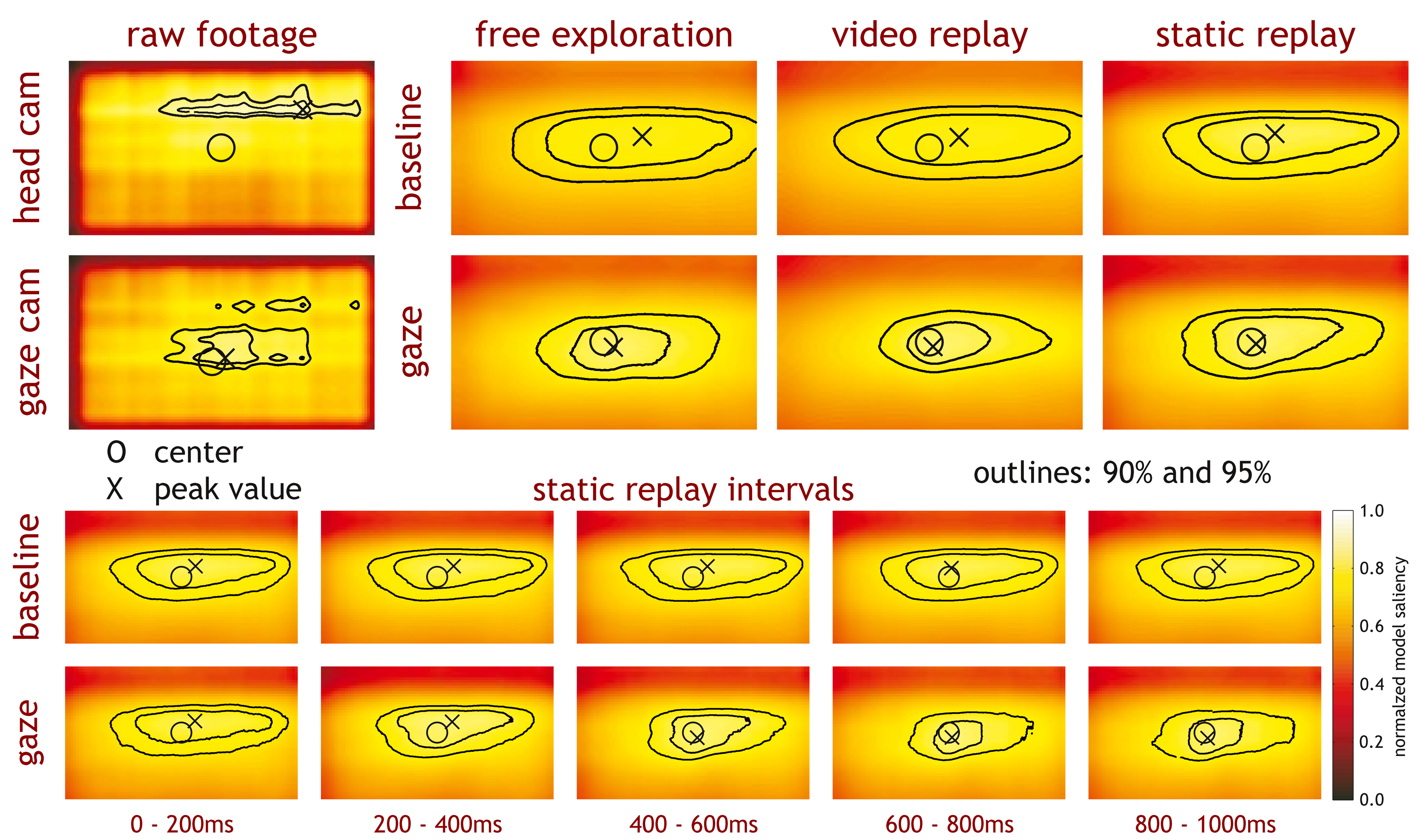

Figure 8. Gaze centered saliency. Model saliency maps (see above) centered relative to gaze and averaged across movies. Actual data show a more pronounced peak than a shuffled baseline that preserves all eye-movement characteristics except the relationship to the visual input.

6. Conclusion

These results show that experiments and models benefit from preserving the spatial statistics and temporal continuity of natural stimuli to improve their validity for real-world gaze behavior.

Contact: thart@staff.uni-marburg.de

Details: 't Hart, B.M., Vockeroth, J., Schumann, F., Bartl, K., Schneider, E., König, P. and Einhäuser, W. (in press). Gaze allocation in natural stimuli: comparing free exploration to head-fixed viewing conditions. Visual Cognition.