Objects and saliency: Reply to Borji et al.

Wolfgang Einhäuser

Neurophysics, Philipps-University, Marburg, Germany
Center for Interdisciplinary Research (ZiF),
Bielefeld, Germany

We thank Borji, Sihite, and Itti (2013) for the attention they give to our 2008 paper (Einhäuser, Spain, & Perona, 2008) and for reanalyzing our data with a variety of recent salience models. The question of what drives attention in natural scenes remains both important and fascinating. With respect to their analysis, we would like to make the following comments:

(a) The authors' reanalysis of our data with respect to the ITTI* model yields very similar results numerically, so there is no disagreement on the results per se.

(b) We acknowledge that more recent "early saliency" models outperform our object-based approach. It is, however, no surprise that our naïve object model cannot live up to this comparison. For example, it is well established that fixations do not target objects uniformly (e.g., Nuthmann & Henderson, 2010), but have a tendency towards the object's center. Taking this bias into account is therefore likely to improve object-based models. In our paper we deliberately chose the simplest possible object model (uniform distribution of fixations on the object) to show that even such a simple assumption outperformed the then widely used ITTI* model. How object-based models that use more realistic fixation distributions within an object fare against the best models used in Borji et al.'s (2013) comment is an interesting question for future research.

(c) The improvement of the ITTI* model under the smoothing condition is consistent with the aforementioned distribution of fixations within objects, as smoothing tends to move high values of early saliency models from object edges towards their centers.

(d) When comparing ITTI* to our object-based approach (e.g., with the "sAUC"), the used Bonferroni correction could in principle greatly increase the probability of a type II error (erroneously failing to reject the null hypothesis stated in the comment's title). Although this does not seem to be the case here (personal communication with the authors and table in their supplemental material), absence of significance of course does not imply no effect, especially given the comparably small number of participants in our original study.

Hence, we consider the analysis by Borji et al. (2013) interesting, as it brings a new aspect into the object-based versus feature-based debate. We assert, however, that a lot of experimental work and modeling on the object side is still needed to resolve this question. In the meantime, we will of course be happy to provide any interested reader with our data.

Acknowledgments

References

Borji, A., Sihite, D. N., & Itti, L. (2013). Objects do not predict fixations better than early saliency; a re-analysis of Einhäuser et al.'s data. *Journal of Vision*, *13*(10):19, 1–4, http://www.journalofvision.org/content/13/10/19, doi:10.1167/13.10.19. [Article]

Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision*, *8*(14):18, 1–26, http://www.journalofvision.org/content/8/14/18, doi:10.1167/8.14.18. [PubMed] [Article]

Nuthmann, A., & Henderson, J. M. (2010). Object-based attentional selection in scene viewing. *Journal of Vision*, *10*(8):20, 1–19, http://www.journalofvision.org/content/10/8/20, doi:10.1167/10.8.20. [PubMed] [Article]