Focusing on Selection for Fixation

John K. Tsotsos, Calden Wloka, Yulia Kotseruba

Dept. of Electrical Engineering and Computer Science
York University
Toronto, Canada

Building on our presentation at MODVIS 2015, we continue in our quest to discover a functional, computational, explanation of the relationship among visual attention, interpretation of visual stimuli, and eye movements, and how these produce visual behavior. Here, we focus on one component, how selection is accomplished for the next fixation. The popularity of saliency map models drives the inference that this is solved; we suggested otherwise at MODVIS 2015. Here, we provide additional empirical and theoretical arguments. We then develop arguments that a cluster of complementary, conspicuity representations drive selection, modulated by task goals and history, leading to a blended process that encompasses early, mid-level and late attentional selection and reflects the differences between central and peripheral processes. This design is also constrained by the architectural characteristics of the visual processing pathways, specifically, the boundary problem, as well as retinal photoreceptor distribution. These elements combine into a new strategy for computing fixation targets and a first simulation of its performance is presented.

The main elements of this strategy are captured in Figure 1 and will be detailed in the presentation.

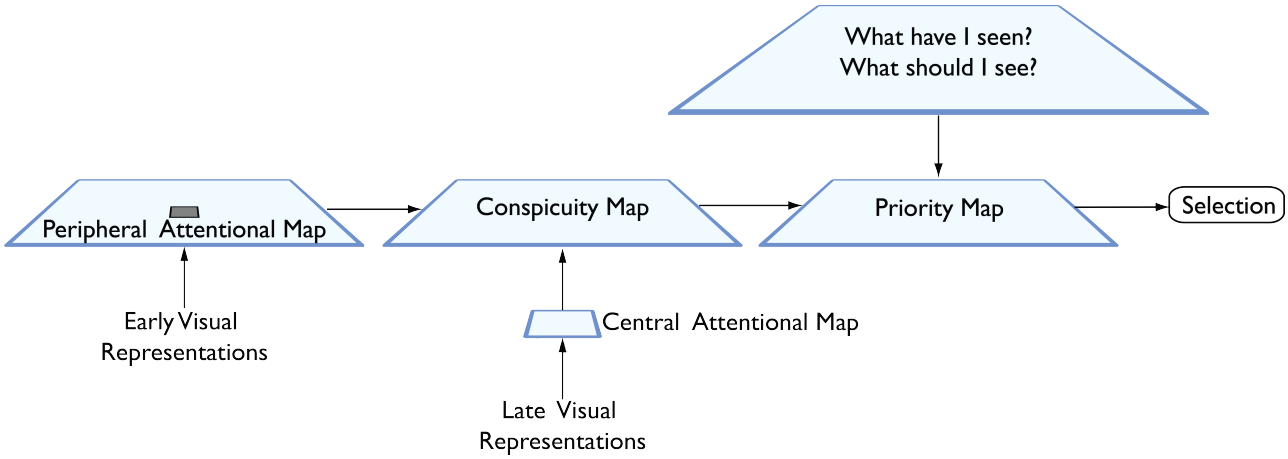

Figure 1. The proposed solution to the boundary problem requires the integration of three representations: early visual conspicuity based on early visual representations, interpretation based on late visual representations, next fixation history (what have I seen?) and priority (what should I see?). Selection of the next target to fixate is made competitively from this combined representation. The dark rectangle of the Peripheral Attentional Map is the portion of the visual field not represented. The Central Attentional Map is smaller because it represents only the veridical components of late visual representations.