Rapid, global image processing: Powerful, but capacity-limited

Karla Evans and Jeremy Wolfe

Abstract

It has been well documented that people are able to get some semantic and/or statistical information out of a briefly presented image. They are able to report on the gist of a scene in 100 ms (Potter, 1975), detect the presence of an animal in an image presented for 20 ms (Thorpe et al., 1996) and determine the mean size of sets of objects (Ariely, 2001; Chong & Treisman, 2003). These findings suggest that some advanced scene processing is possible without attentional selection of specific objects. What are the limits on this non-selective processing? During natural viewing, do we have simultaneous access to multiple non-selective possibilities: mean orientation, scene gist, animal detection, etc? Alternatively, perhaps we only have access to currently task-relevant capability. If you are looking for animals, you might not automatically compute mean orientation. In a series of psychophysical experiments we tested whether there is a cost to non-selective processing and the nature of that cost.

We compared conditions where observers know the relevant global image property before seeing the stimulus to conditions where the relevant global property is specified after the stimulus has been presented. The results show that observers' post-cued performance was well beyond what could be achieved if that processing ability had to be set to one attribute at a time. However, performance was below what would be predicted if observers computed all non-selective properties without cost. In a second set of experiments, observers monitored RSVP streams for one target category (e.g. beach) in a block of trials where two others (e.g. person, vehicle) could be targets. Images containing a trial-relevant and a block-relevant item (car on beach) produced more errors as the presence of the car seemed to block encoding of the beach. Non-selective processing of semantic information cannot encode all available signals.