Optimal Eye Movement Strategies in Foveated Visual Systems

Wilson S. Geisler

Center for Perceptual Systems and Department of Psychology

University of Texas at Austin, USA

The human visual system encodes a large field of view with a variable resolution retina, and then uses high-speed eye movements (and slower head and body movements) to direct the highest resolution region, the fovea, at specific locations in the environment. To better understand the human visual system and foveated vision systems in general, we have been working on the theory of optimal eye movements for two different classes of task. The first task is a search task where a known target is embedded at an unknown location within a random background that has the spectral characteristics of natural scenes. Comparison with human performance shows that humans perform near optimum and employ search strategies more sophisticated than moving the eyes to the most likely (or salient) target location. The second task is a scene encoding task, where the goal is to extract as much feature information as possible, given a small number of fixations (i.e., a limited amount of time). Based on the statistics of local contrast in natural images, we derive a "contrast entropy minimization" algorithm. We show that the algorithm performs optimally at reducing total contrast uncertainty, and that it also works well at reducing the mean squared error between the original image and the image reconstructed from the multiple fixations, suggesting that measurements of local contrast alone could efficiently drive the scan paths of the eye, when the goal is to gain as much information about the spatial structure of a scene as possible.