Bayesian Theory of Visual Search

Wei Ji Ma¹, Vidhya Navalpakkam², and Jeffrey M. Beck¹
¹University of Rochester, ²California Institute of Technology

Human performance in searching for a tilted target among vertical distractors is well described by signal detection theory (SDT) models which combine the outputs of noisy detectors using a maximum or summation operation [1]. These theories assume that each detector is a neuron best tuned to the target, and that a detector’s response to the target or distractor follows a Gaussian distribution. However, recent work has questioned the validity of SDT measures in search tasks with more complex stimulus distributions, for instance ones in which distractors flank the target in orientation space [2]. Instead, a saliency-based signal-to-noise ratio was proposed to measure task difficulty. A drawback of this theory is that it cannot predict receiver operating characteristics (ROCs).

We reconcile and generalize both approaches by developing a fully Bayesian model of visual search. We study the detection of a target among distractors, where target and distractor features are drawn from arbitrary distributions. At each location in the display, the feature (such as orientation) is coded in a population of neurons with Poisson-like variability [3]. Decisions are based on the log odds of target presence given the responses of all neurons at each location (unlike SDT, where decisions are based on the neuron at each location that is best tuned to the target). These global log odds are obtained through a nonlinear operation on the location-specific log odds, which incorporate the feature distributions. Search difficulty is quantified by a number of metrics including mutual information, Kullback-Leibler divergence, area under the ROC, and proportion correct. This model 1) reproduces behavioral effects and ROCs in simple visual search, such as the increase of search difficulty with set size, target-distractor similarity, and distractor heterogeneity; 2) explains behavior in complex search conditions, such as when distractors flank the target; 3) can in special cases be approximated by the maximum or sum rule. In simple search, the stimulus-conditioned distributions of the local log odds are approximately equivariant Gaussian distributions, indicating that a decision variable from Bayes-optimal computation can be linked to the SDT models. Moreover, for any distribution of target and distractor features, the Bayesian model predicts which neurons are most informative. For instance (see figure), when distractors are similar to the target (either flanking the target at ±10° or on one side at +10°), this model predicts that neurons tuned to orientations slightly away from target or distractor are most crucial to search performance.

It is important to note that Bayes-optimal performance requires full knowledge of the generative model, including location and feature distributions of target and distractors. In many cases, this information is not available or known with certainty, leading to extra behavioral variability and suboptimal performance. We suggest that the effect of attention is to provide prior knowledge that improves the quality of the approximation to the generative model. This notion is compatible with human experiments.

References