Is the Selective Tuning Model of Visual Attention Still Relevant?

John K. Tsotsos
York University, Toronto, Canada
tsotsos@eecs.yorku.ca

Abstract

The Selective Tuning (ST) model of visual attention is a theory-based model, derived from first principles with little or no appeal to data-directed derivation. It seems a natural question to ask whether the model might still have relevance in the present research milieu. This presentation examines this.

To begin, ST and its theoretical roots will be briefly described. Its relevance will then be addressed on the basis of ST's ability to engender new knowledge of human visual processes and ST's contribution in terms of the quantitative performance of systems which embody its elements. To the former point, a number of recent human behavioral and neurophysiology results are offered that add to the growing body of similar past results:

a. Feature-Based Attention - Attending to a particular feature value causes local suppression in that feature dimension. New human experimental work demonstrates this for color and for motion, to supplement previous work on orientation.

b. Properties of Attentive Surround Suppression - The suppressive surround is affected by stimulus eccentricity, and the two mechanisms of spatial and feature based suppression interact.

c. Development of Attentive Surround Suppression - New experiments demonstrate that the attentive surround suppression does not develop until quite late in childhood.

On the performance side, some recent computational results will be described:

d. Fixation Control - A fixation controller has been defined that is based on the interaction of four different kinds of conspicuity representations and leads to saccade sequence prediction that matches human sequences.

e. The Functional Breadth of Recurrent Computation - There is no reason to think that all top-down connections in the brain have the same functionality. ST permits different top-down functions and examples of multi-scale spatial context modulation, attentive localization, priming, and signal interference reduction will be shown.

f. Single Neuron Comparisons - ST is based upon a particular formulation of neurons and several simulations have shown that its single neuron responses compare favorably with those of real neurons in attentive scenarios.

It is important to stress that most of the attentive behaviors of ST are observed in the time period immediately following the first feed forward pass of visual processing that dominates the concern of most research. And this brings up one additional argument supporting ST's relevance. The ultimate purpose of a vision system is to play its appropriate role within a larger intelligent agent beyond its immediate perception of a single image. We have defined such a larger agent in STAR (Selective Tuning Attentive Reference model), a cognitive architecture for visually-guided behavior. ST is embedded within STAR and this provides a framework within which it is possible to explore how an attention system might interact with working memory, task instructions, behavior, and decision-making, where the agent is an active observer in a 3D dynamic world. A short summary of STAR will be given, highlighting its key elements, underlying principles, and most interestingly, showing how a computational counterpart to emerging research on functional connectivity of brain networks may be approached by STAR. A key reason why this is possible is the fact that ST has semantically meaningful representations and provides explicit mechanistic explanations for how each function is determined.

The dimensions of attentional behavior addressed by ST go far beyond what dominates current systems, namely region of interest computations, which may in fact be a computational attention counterpart not part of human vision. Moreover, it is evident that modern computer vision systems might be seriously challenged to exhibit many of ST's behaviors. ST is definitely not relevant for tasks that may be solved without attention (e.g., categorization of image targets on non-conflicting backgrounds). But for more demanding tasks and as long as the model continues to generate successfully tested predictions leading to new knowledge, and provides for mechanistic explanations as to how vision might be accomplished in the brain and in machines, ST remains relevant.