Modeling of Attentional Modulation Effects in Object Recognition

Dirk Walther[1], Max Riesenhuber[2], Tomaso Poggio[2], Laurent Itti[3], Christof Koch[1]

([1]California Institute of Technology, [2]Massachusetts Institute of Technology, [3]University of Southern California)

We present an integrated model for the dorsal (where) and the ventral (what) pathway in the primate's visual processing system and the interaction between these two pathways. To reach our goal of modeling visual search behavior in primates, we integrate and extend the saliency-based model for bottom-up attention by Itti and Koch (Nature Review Neuroscience 2001; 2(3):194-203) and the HMAX hierarchical model for object recognition by Riesenhuber and Poggio (Nature Neuroscience 1999; 2:1019-1025).

In the combined model we use saliency-based attention to modulate object recognition at the V4 level. Interesting regions in the visual scene are successively selected by a rapidly shiftable focus of attention (FOA). Neural activity of a particular neuron in V4 is inhibited based on its distance from the current FOA.

Recognition rates for stimuli composed of two paper clip objects typically increase twofold compared to previous experiments without attention (Neuron 1999;24(1):87-93). To achieve this improvement a depression of the V4 activity outside the focus of attention by as little as 20% proves to be sufficient. With 10% activity modulation recognition still improves by 70%. We find that the twofold increase in recognition rate is robust over a large range of modulation strengths of the V4 activity. We conclude that a rather weak attentional modulation of the neural activity at the V4 level suffices to recognize multiple objects in the same display.

Our model will be extended to search for specific objects in cluttered natural scenes and include biasing of the attention system in a top-down manner.