

"A New Hierarchical Model of Object Recognition in Cortex"

Maximilian Riesenhuber

Dept. of Brain and Cognitive Science, and Center for Biological and Computational Learning, Massachusetts Institute of Technology

ABSTRACT

In the ventral visual stream of the macaque, which is believed to be crucial for object recognition, neurons' receptive field sizes and complexity of preferred features grow from simple cells tuned to bars to cells in infero-temporal cortex (IT) that are tuned to views of complex objects (such as faces) and show great tolerance to changes in stimulus size or position.

I will describe a simple hierarchical model (developed together with Tommy Poggio) that, in the spirit of the original simple-to-complex cell model by Hubel and Wiesel, attempts to explain in a biologically plausible way how view-tuned neurons can arise from localized simple cell-like inputs. Using the same stimuli as in the experiment, we find that model IT neurons exhibit tuning properties which closely parallel those of the neurons found by Logothetis et al. (1995) and are also compatible with reports on feature tuning of IT cells (Tanaka, 1996). The use of different computational mechanisms to achieve the two goals of increasing invariance and increasing feature specificity is key: While the latter is best served using a template match, the former requires the ability to select the best-matching among several afferent, in our model through a "MAX" operation postulated to be a key module in cortex which selects (an approximation to) the maximum activity over all the afferents.

Interestingly, the model also helps to interpret recent data from several other experiments regarding recognition in clutter (Missal et al., 1997), or the presence of multiple objects in an IT cell's receptive field (Sato, 1989; Rolls & Tovee, 1995), with implications with respect to the "Binding Problem": As the "MAX" operation ignores non-relevant afferents, recognition in these cases is possible in a purely bottom-up fashion, without any need for special segmentation processes.