Visual Recognition in Primates and Machines

Tomaso Poggio

McGovern Institute for Brain Research, Center for Biological & Computational Learning,
Department of Brain & Cognitive Sciences, Computer Science and Artificial Intelligence Laboratory,
Massachusetts Institute of Technology,
Cambridge, MA 02139, U.S.A.

Abstract

Our visual abilities are computationally amazing and we are still far from imitating them with computers. Thus, visual cortex may well be a good proxy for the rest of the cortex and indeed for intelligence itself. But despite enormous progress in the physiology and anatomy of the visual cortex, our understanding of the underlying computations remains fragmentary. In this tutorial, I will briefly review the anatomy and the physiology of primate visual cortex and then describe a class of quantitative models of the ventral stream for object recognition, which, constrained by physiology and biophysics, have been developed during the last two decades and which have been shown to be moderately successful in explaining several physiological data across different visual areas. I will discuss their performance and architecture from the point of view of computer vision system. Such models can mimic several aspects of human performance in difficult rapid image categorization tasks in which human vision is forced to operate in a feedforward mode.

I will then focus on key limitations of such hierarchical feedforward models for object recognition, discuss why they are incomplete models of vision and suggest possible alternatives focusing on the computational role of attention and its likely substrate – cortical backprojections.

The broad thesis of this tutorial is that computational neuroscience is beginning to provide novel insights into the problem of how our visual cortex is computing and is thus increasingly relevant for computer vision.

Syllabus: Visual cortex, Learning theory, Kernel machines, Object recognition.