Hierarchical Learning Machines and Neuroscience of Visual Cortex

tomaso poggio
McGovern Institute
CBCL,BCS, CSAIL
MIT

Learning is the gateway to understanding the brain and to making intelligent machines.

Problem of learning:
a focus for
- math
- computer algorithms
- neuroscience

ECML, Barcelona, September 21, 2010

Intelligence in minds, brains and machines

At the intersection
between neuroscience and AI+CS
learning
is key
for new science and new applications...

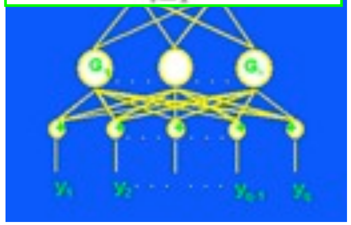
Intelligence in minds, brains and machines

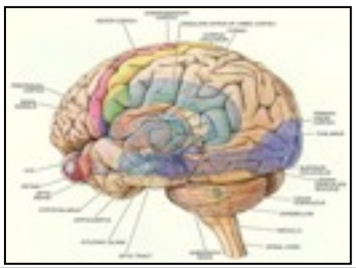
1. The past: a bit of personal history of learning theory and applications
2. The future: from neuroscience to smarter machines
   - Learning in Visual Cortex
   - Implications for Computer Vision and Machine Learning
   - Hierarchical Learning Machines
   - Beyond Classification

Learning:
Math, Engineering, Neuroscience

$$\min_{f \in H} \left[\frac{1}{\ell} \sum_{i=1}^{\ell} V(y_i, f(x_i)) + \mu \|f\|_K^2 \right]$$

$$f(x) = \sum_{i=1}^{l} c_i K(\mathbf{x}_i, \mathbf{x})$$





LEARNING THEORY + ALGORITHMS

Theorems on foundations of learning

Predictive algorithms

ENGINEERING APPLICATIONS

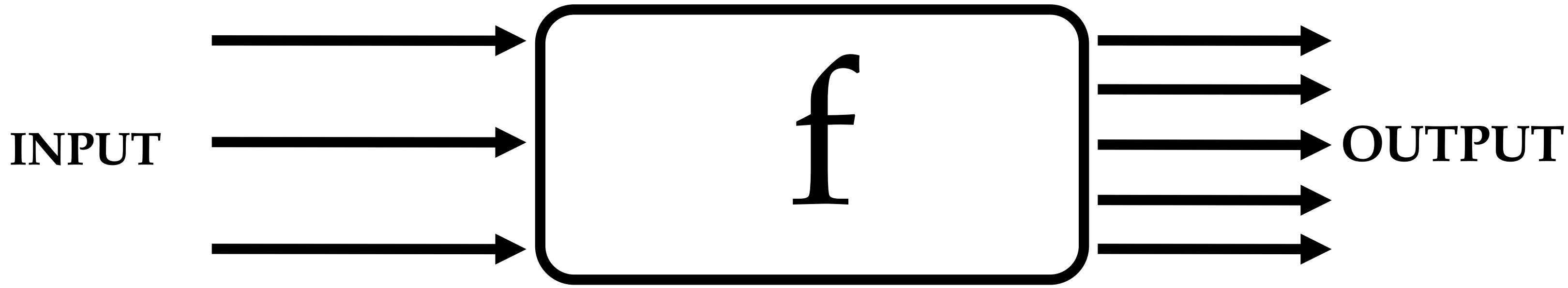
- Bioinformatics
- Computer vision
- Computer graphics, speech synthesis, creating a virtual actor

COMPUTATIONAL NEUROSCIENCE: models+experiments

How visual cortex works

Learning:
Math, Engineering, Neuroscience

Supervised learning



Given a set of l examples (data) $\{(x_1, y_1), (x_2, y_2), \ldots, (x_\ell, y_\ell)\}$

Question: find function f such th $f(x) = \hat{y}$

is a **good predictor** of y for a **future** input x (fitting the data is **not** enough!)

Learning:
Math, Engineering, Neuroscience

Classical learning algorithms: Kernel Machines (eg Regularization in RKHS)

$$\min_{f \in H} \left[\frac{1}{n} \sum_{i=1}^{n} V(f(x_i) - y_i) + \lambda \; \|f\|_K^2 \right]$$

implies

$$f(\mathbf{x}) = \sum_i^n \alpha_i K(\mathbf{x}, \mathbf{x}_i)$$

Equation includes splines, RBF, SVMs (depending on choice of V). Bayesian interpretation in terms of MAP, forget about margin and kernel trick...

For a review, see Poggio and Smale, 2003; see also Schoelkopf and Smola, 2002; Bousquet, O., S. Boucheron and G. Lugosi; Cucker and Smale; Zhou and Smale...

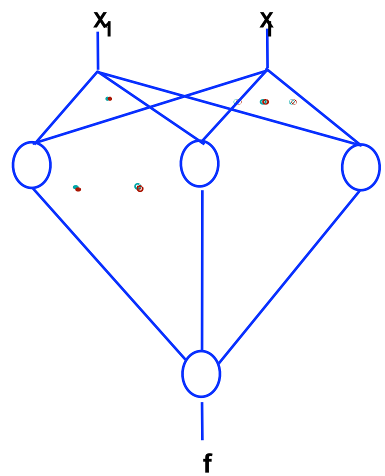
Learning:
Math, Engineering, Neuroscience

Classical learning algorithms: Kernel Machines (eg Regularization in RKHS)

$$\min_{f \in H} \left[\frac{1}{n} \sum_{i=1}^{n} V(f(x_i) - y_i) + \lambda \; \|f\|_K^2 \right]$$

implies

$$f(\mathbf{x}) = \sum_i^n \alpha_i K(\mathbf{x}, \mathbf{x}_i)$$



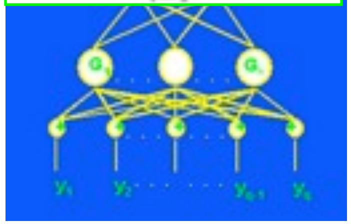
Remark (for later use):

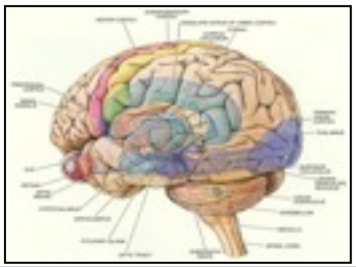
Kernel machines correspond to
shallow networks

Learning:
Math, Engineering, Neuroscience

$$\min_{f \in H} \left[\frac{1}{\ell} \sum_{i=1}^{\ell} V(y_i, f(x_i)) + \mu \|f\|_K^2 \right]$$

$$f(x) = \sum_{i=1}^{l} c_i K(\mathbf{x}_i, \mathbf{x})$$





LEARNING THEORY + ALGORITHMS

Theorems on foundations of learning

Predictive algorithms

ENGINEERING APPLICATIONS

- Bioinformatics
- Computer vision
- Computer graphics, speech synthesis, creating a virtual actor

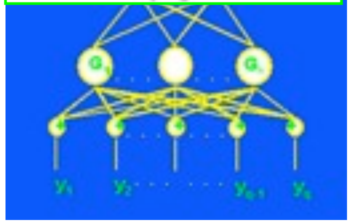
COMPUTATIONAL NEUROSCIENCE: models+experiments

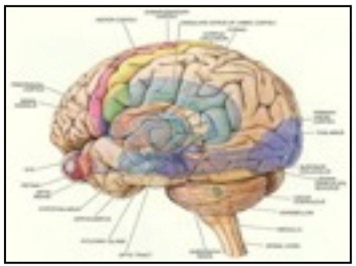
How visual cortex works

Learning:
Math, Engineering, Neuroscience

$$\min_{f \in H} \left[\frac{1}{\ell} \sum_{i=1}^{\ell} V(y_i, f(x_i)) + \mu \|f\|_K^2 \right]$$

$$f(x) = \sum_{i=1}^{l} c_i K(\mathbf{x}_i, \mathbf{x})$$





LEARNING THEORY + ALGORITHMS

Theorems on foundations of learning

Predictive algorithms

Sung & Poggio 1995, also Kanade& Baluja....

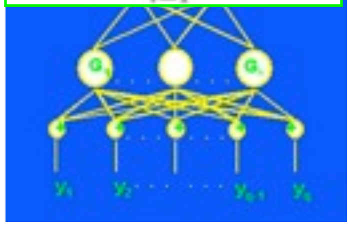
COMPUTATIONAL NEUROSCIENCE: models+experiments

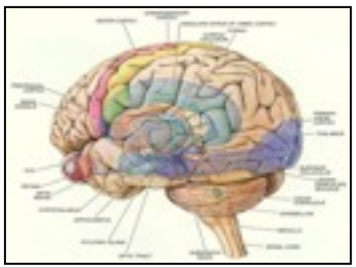
How visual cortex works

Learning:
Math, Engineering, Neuroscience

$$\min_{f\in H}\left[\frac{1}{\ell}\sum_{i=1}^{\ell}V(y_i,f(x_i))+\mu\|f\|_K^2\right]$$

$$f(x)=\sum_{i=1}^{l} c_i K(\mathbf{x}_i,\mathbf{x})$$





LEARNING THEORY
+
ALGORITHMS

Theorems on foundations of learning

Predictive algorithms

Sung & Poggio 1995

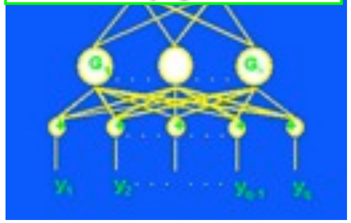
COMPUTATIONAL NEUROSCIENCE: models+experiments

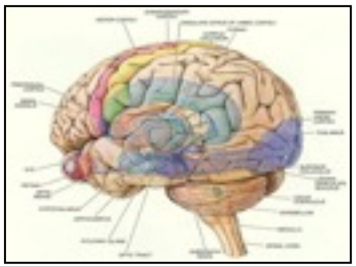
How visual cortex works

Learning:
Math, Engineering, Neuroscience

$$\min_{f \in H} \left[\frac{1}{\ell} \sum_{i=1}^{\ell} V(y_i, f(x_i)) + \mu \|f\|_K^2 \right]$$

$$f(x) = \sum_{i=1}^{l} c_i K(\mathbf{x}_i, \mathbf{x})$$





LEARNING THEORY
+
ALGORITHMS

Theorems on foundations of learning

Predictive algorithms

Face detection is now available in digital cameras (commercial systems)

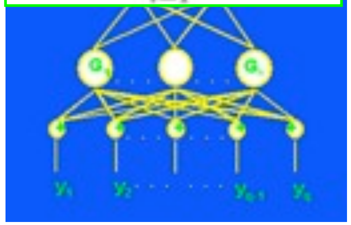
COMPUTATIONAL NEUROSCIENCE: models+experiments

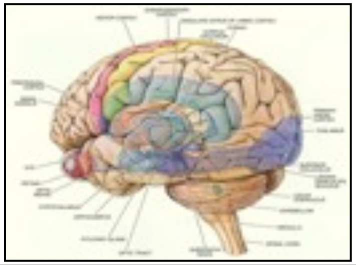
How visual cortex works

Learning:
Math, Engineering, Neuroscience

$$\min_{f \in H} \left[\frac{1}{\ell} \sum_{i=1}^{\ell} V(y_i, f(x_i)) + \mu \|f\|_K^2 \right]$$

$$f(x) = \sum_{i=1}^{l} c_i K(\mathbf{x}_i, \mathbf{x})$$

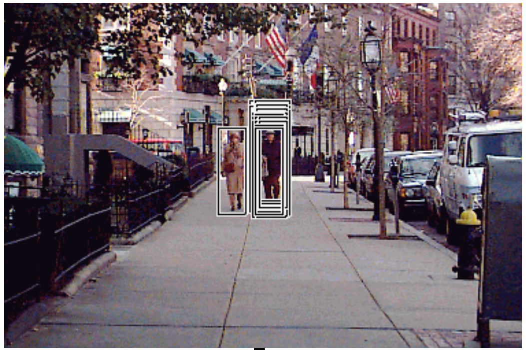




LEARNING THEORY + ALGORITHMS

Theorems on foundations of learning

Predictive algorithms



Papageorgiou&Poggio, 1997, 2000
also Kanade&Scheiderman

COMPUTATIONAL NEUROSCIENCE: models+experiments

How visual cortex works

Learning:
Math, Engineering, Neuroscience (now)

- Since the introduction of *learning/statistics* in the '90s, computer vision has made significant (and not well known) advances in a few problem areas:

- *Face identification* under controlled conditions is probably "solved" (commercial systems)

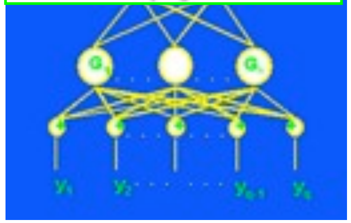
- *Face detection* is available in cheap digital cameras (commercial systems)

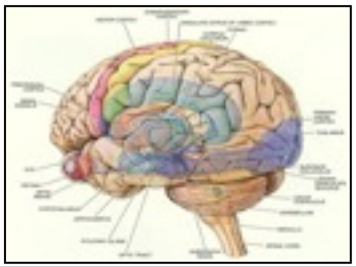
- *Pedestrian and car detection* are also "solved" (commercial systems)

Learning:
Math, Engineering, Neuroscience (now)

$$\min_{f\in H}\left[\frac{1}{\ell}\sum_{i=1}^{\ell} V(y_i, f(x_i)) + \mu \|f\|_K^2\right]$$

$$f(x) = \sum_{i=1}^{l} c_i K(\mathbf{x}_i, \mathbf{x})$$





LEARNING THEORY + ALGORITHMS

Poggio, T. and F. Girosi. Regularization Algorithms for Learning that are Equivalent to Multilayer Networks, *Science*, 1990

Poggio, T. and F. Girosi. Networks for Approximation and Learning, *Proceedings of the IEEE* , 1990

Poggio, T. and S.Smale. The Mathematics of Learning: Dealing with Data, *Notices American Mathematical Society (AMS)*, 2003

Poggio, T., R. Rifkin, S. Mukherjee and P. Niyogi. General Conditions for Predictivity in Learning Theory, *Nature*, 2004

ENGINEERING APPLICATIONS

Beymer, D. and T. Poggio. Image Representation for Visual Learning, *Science*, 272., 1905-1909, 1996

Brunelli, R. and T. Poggio. Face Recognition: Features Versus Templates, *IEEE PAMI*, 1993

Sung, K.K. and T. Poggio. Example-Based Learning for View-Based Human Face Detection, IEEE PAMI, 1998 (1995)

Ezzat, T., G. Geiger and T. Poggio. "Trainable Videorealistic Speech Animation," *ACM SIGGRAPH 2002*

COMPUTATIONAL NEUROSCIENCE: models+experiments

Freedman, D.J., M. Riesenhuber, T. Poggio and E.K. Miller. Categorical Representation of Visual Stimuli in Prefrontal Cortex, *Science*, 291, 312-316, 2001.

Riesenhuber, M. and T. Poggio. Hierarchical Models of Object Recognition in Cortex, *Nature Neuroscience*, 2, 1019-1025, 1999.

Serre, T., A. Oliva and T. Poggio. A Feedforward Architecture Accounts for Rapid Categorization, *(PNAS)*, Vol. 104, No. 15, 6424-6429, 2007.

Poggio, T. and E. Bizzi. Generalization in Vision and Motor Control, *Nature*, Vol. 431, 768-774, 2004.

Intelligence in minds, brains and machines

1. The past: learning theory and applications
2. The future: from neuroscience to smarter machines
   - Learning in Visual Cortex
   - Implications for Computer Vision and Machine Learning
   - Hierarchical Learning Machines
   - Beyond classification

Learning in visual cortex

The MIT Press

YOUR PROFILE | TO ORDER | CONTACT US

The MIT Press is the only university press in the United States whose list is based in science and technology. This does not

Vision

A Computational Investigation into the Human Representation and Processing of Visual Information

David Marr
Foreword by Shimon Ullman
Afterword by Tomaso Poggio

David Marr's posthumously published *Vision* (1982) influenced a generation of brain and cognitive scientists, inspiring many to enter the field. In *Vision*, Marr describes a general framework for understanding visual perception and touches on broader questions about how the brain and its functions can be studied and understood. Researchers from a range of brain and cognitive sciences have long valued Marr's creativity, intellectual power, and ability to integrate insights and data from neuroscience, psychology, and computation. This MIT Press edition makes Marr's influential work available to a new generation of students and scientists.

In Marr's framework, the process of vision constructs a set of representations, starting from a description of the input image and culminating with a description of three-dimensional objects in the surrounding environment. A central theme, and one that has had far-reaching influence in both neuroscience and cognitive science, is the notion of different levels of analysis—in Marr's framework, the computational level, the algorithmic level, and the hardware implementation level.

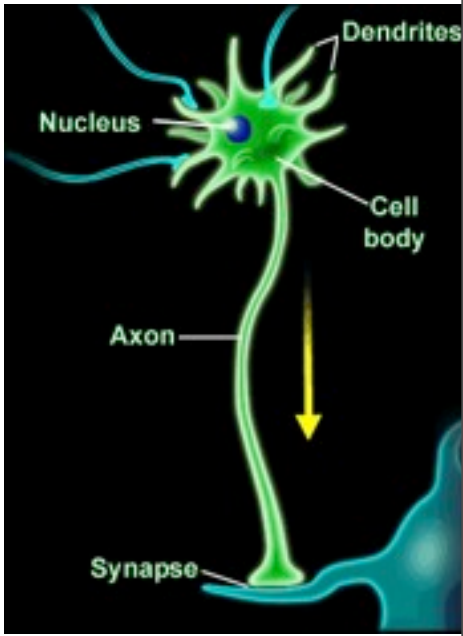
Now, thirty years later, the main problems that occupied Marr remain fundamental open problems in the study of perception. *Vision* provides inspiration for the continui

Learning in visual cortex

tomaso poggio
McGovern Institute
CBCL,BCS, CSAIL
MIT

~ 1979 , with David Marr and Francis Crick, Borego Desert

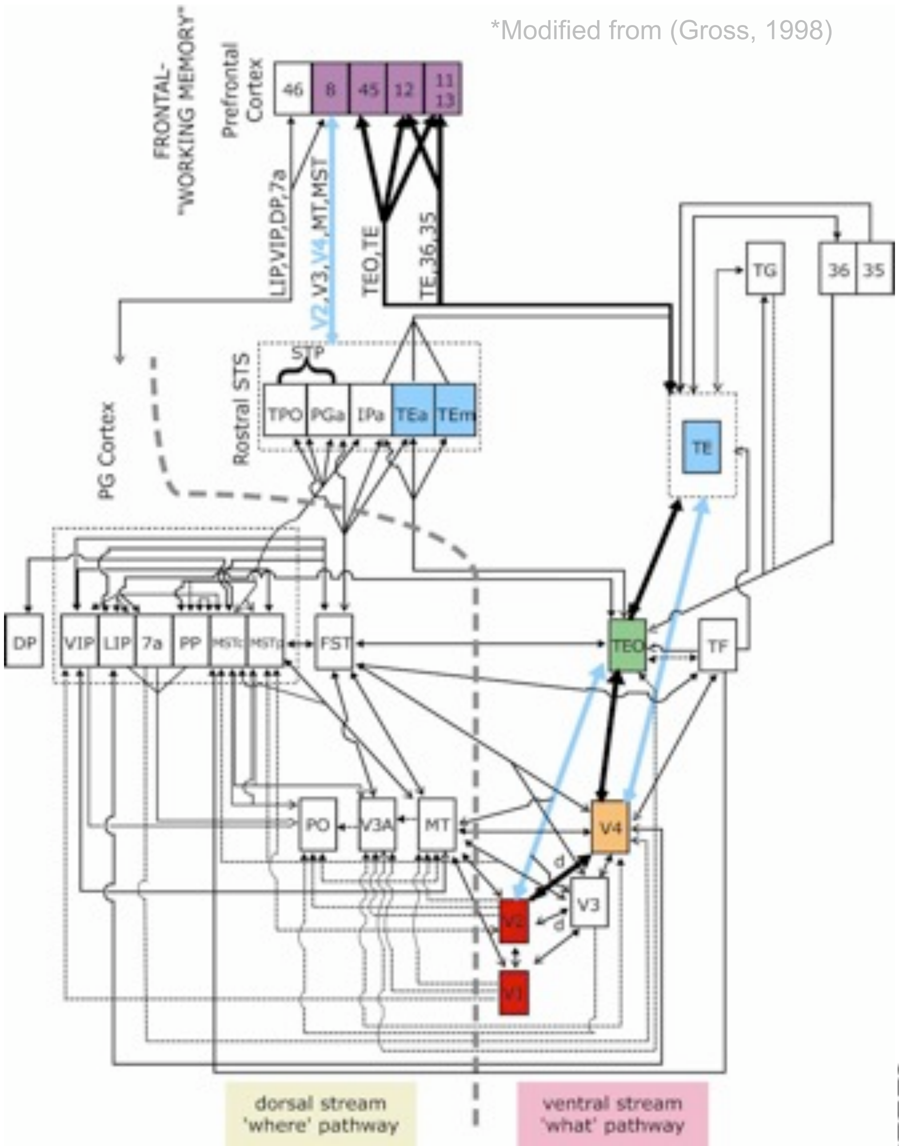
Intelligence in minds, brains and machines

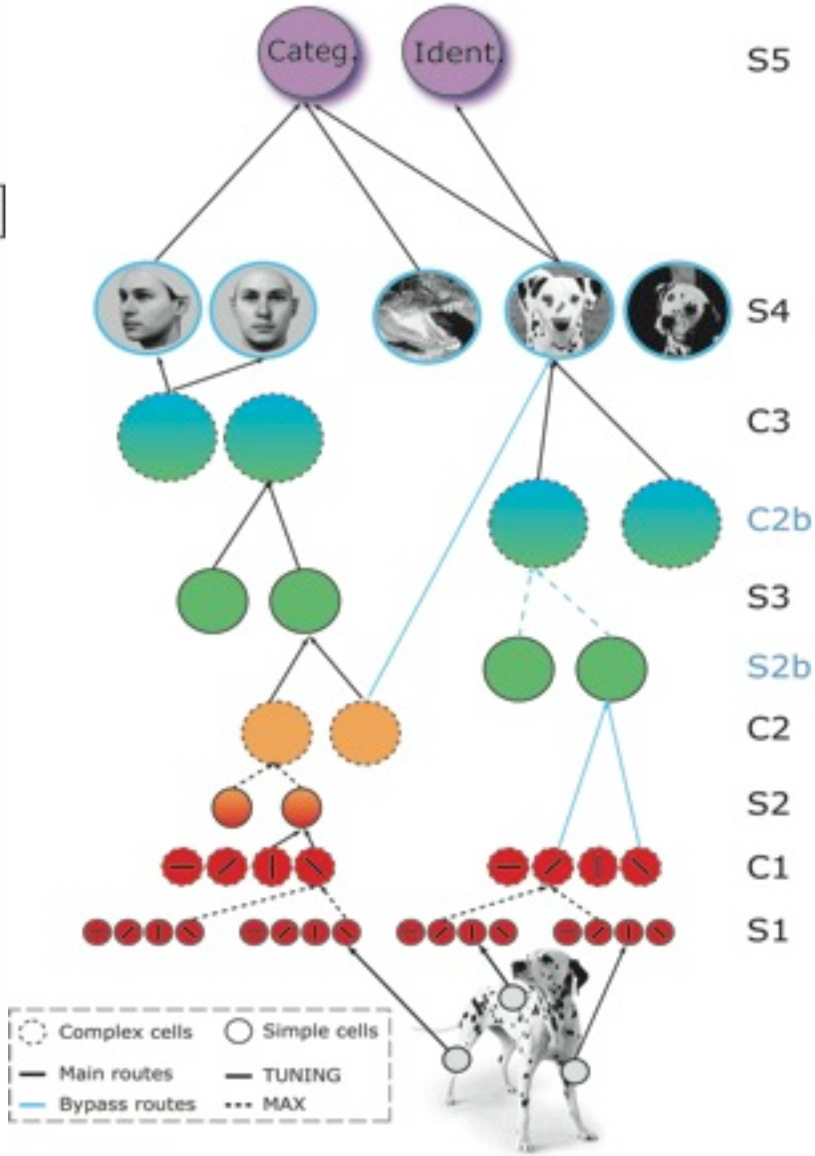


- Human Brain
  - 10^{10}-10^{11} neurons (~1 million flies ☺)
  - 10^{14}- 10^{15} synapses

- Ventral stream in rhesus monkey
  - ~10^{9} neurons in the ventral stream (350 10^{6} in each emisphere)
  - ~15 10^{6} neurons in AIT (Anterior InferoTemporal) cortex

Learning in visual cortex





[software available online
with CNS (for GPUs)]

Riesenhuber & Poggio 1999, 2000; Serre Kouh Cadieu
Knoblich Kreiman & Poggio 2005; Serre Oliva Poggio 2007

tomaso poggio
McGovern Institute
BCS, CSAIL
MIT

MIT Intelligence Initiative