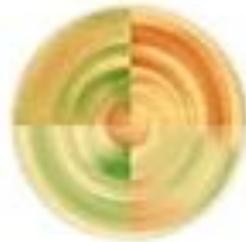


What's Missing from Deep Learning?

Bruno Olshausen

Helen Wills Neuroscience Institute, School of Optometry
Redwood Center for Theoretical Neuroscience
UC Berkeley



REDWOOD CENTER
for Theoretical Neuroscience



"We're at the beginning of a new day...
This is the beginning of the AI revolution."
— Jensen Huang, GTC Taiwan 2017



兩股力量驅動電腦的未來

深度學習點亮人工智慧紀元。

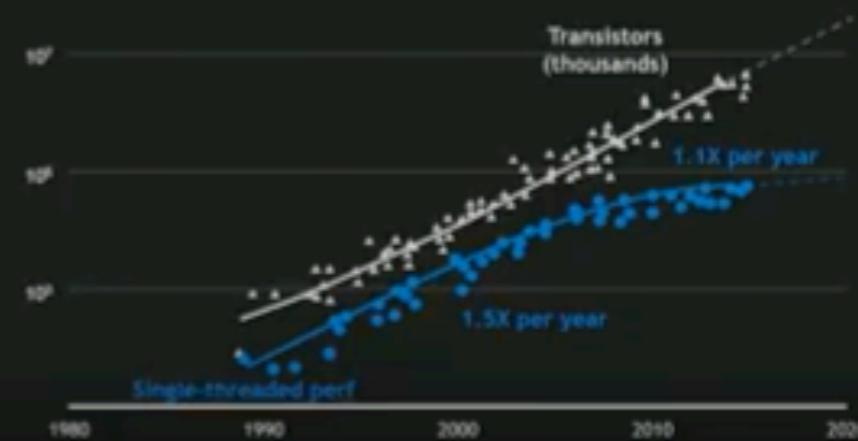
受到人腦的啟發，深度神經網路具備上億的類神經連結，藉由巨量資料來學習，這仰賴極大量的運算。

同時，摩爾定律已到了尾聲 - CPU已不可能再擴張成長。

程式設計人員無法創造出可以更有效率發現更多指令級並行性的CPU架構。

電晶體持續每年增長50%，但是CPU效能僅能成長10%。

TWO FORCES DRIVING THE FUTURE OF COMPUTING



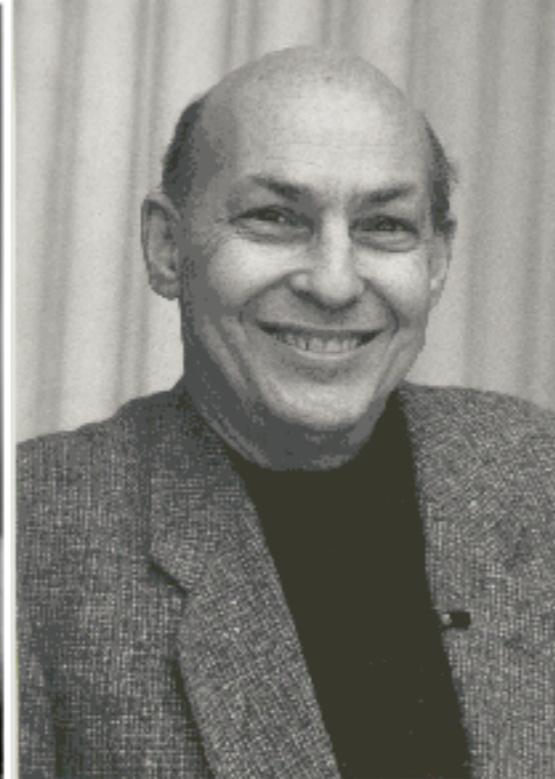
Artificial Intelligence



Alan Turing



John von Neumann



Marvin Minsky



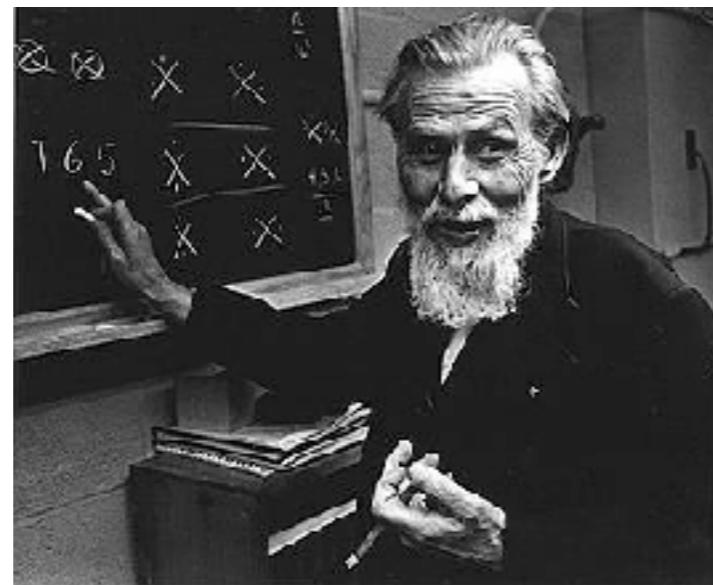
John McCarthy

Among the most challenging scientific questions of our time are the corresponding analytic and synthetic problems: How does the brain function? Can we design a machine which will simulate a brain?
-- *Automata Studies*, 1956

Cybernetics/neural networks



Norbert Wiener



Warren McCulloch & Walter Pitts



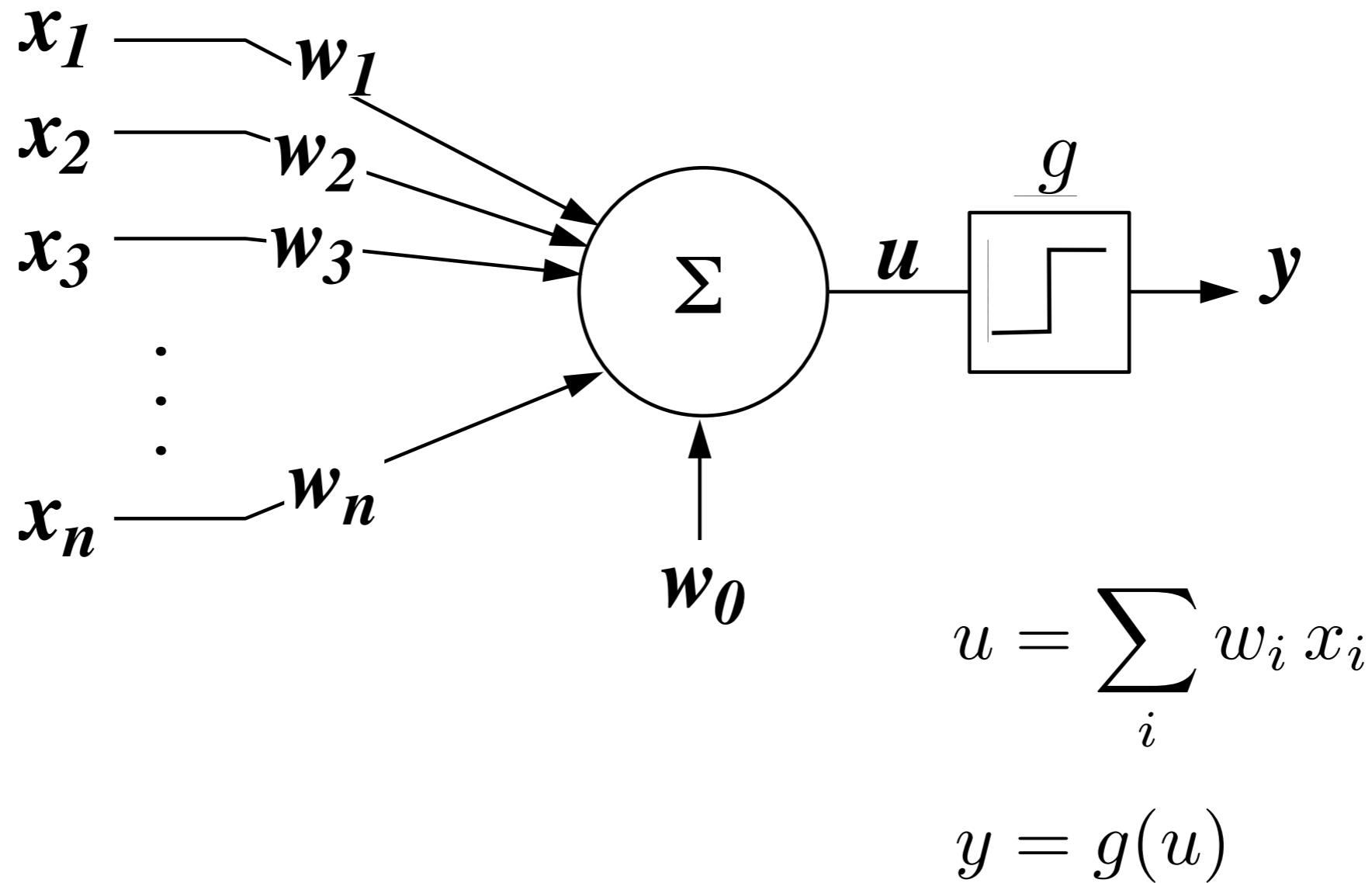
Frank Rosenblatt

“The theory reported here clearly demonstrates the feasibility and fruitfulness of a quantitative statistical approach to the organization of cognitive systems. By the study of systems such as the perceptron, it is hoped that those fundamental laws of organization which are common to all information handling systems, machines and men included, may eventually be understood.” -- Frank Rosenblatt

The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. In, *Psychological Review*, Vol. 65, No. 6, pp. 386-408, November, 1958.

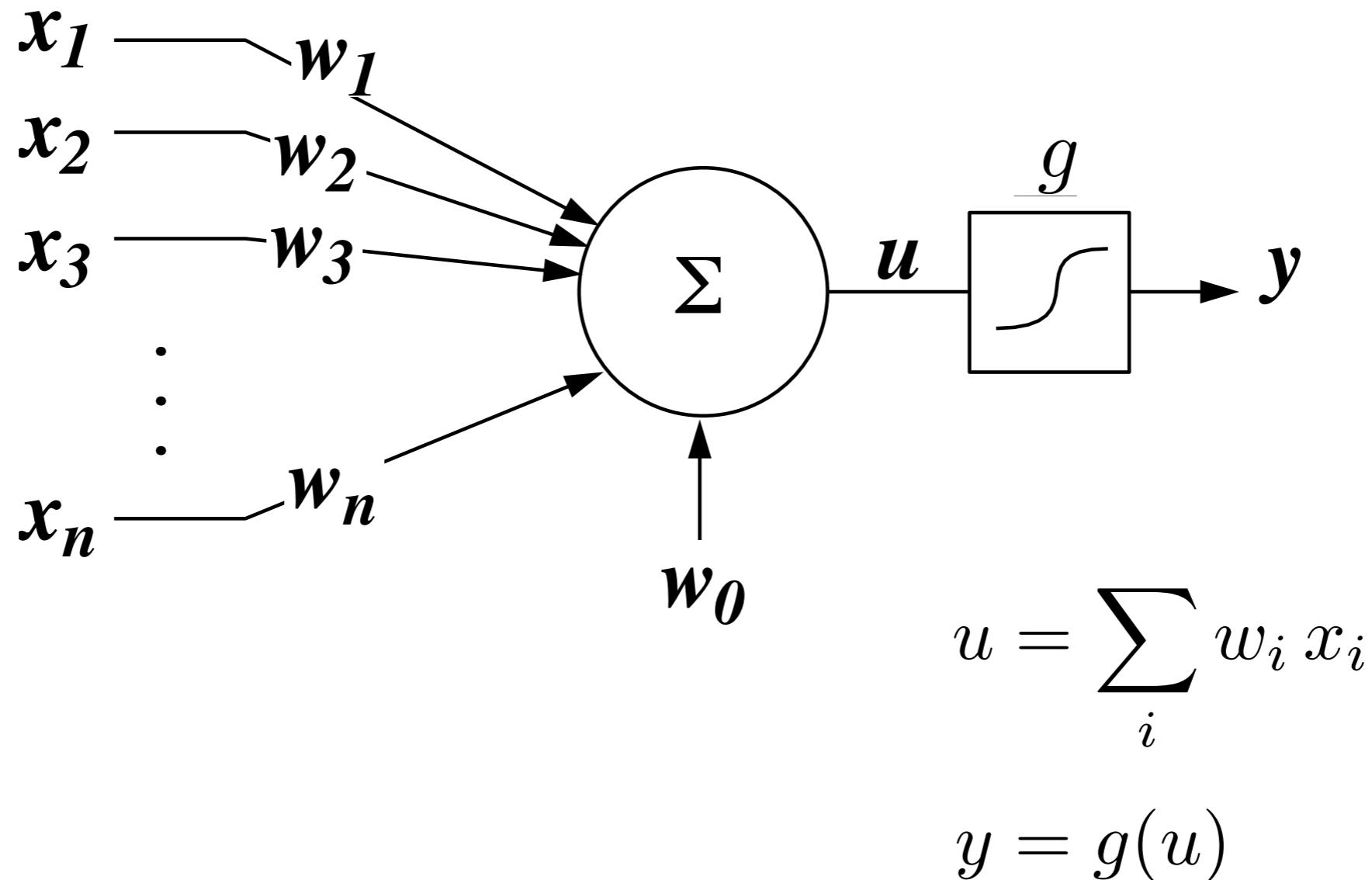
A brief history of neural networks

1960's



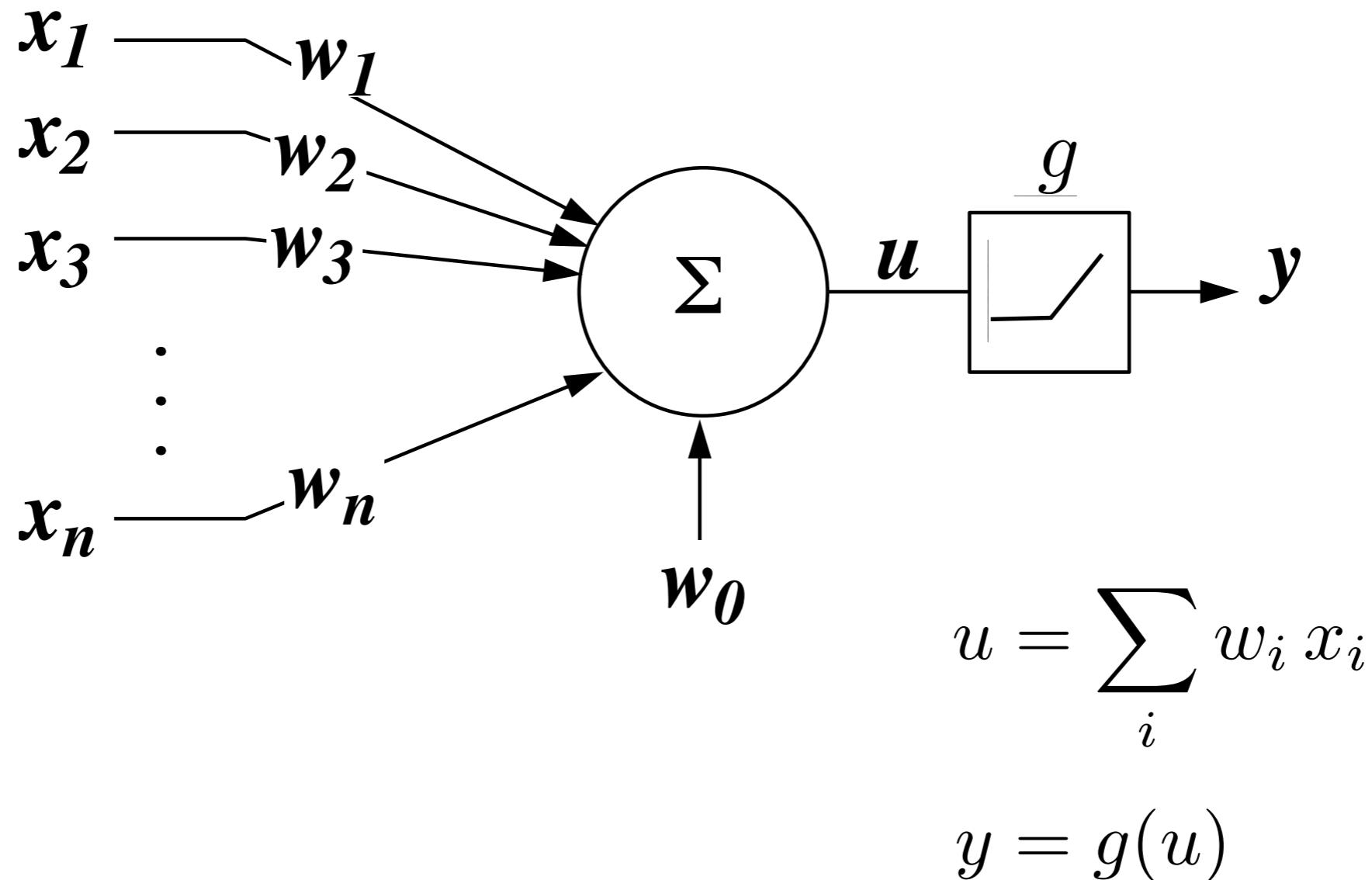
A brief history of neural networks

1980's



A brief history of neural networks

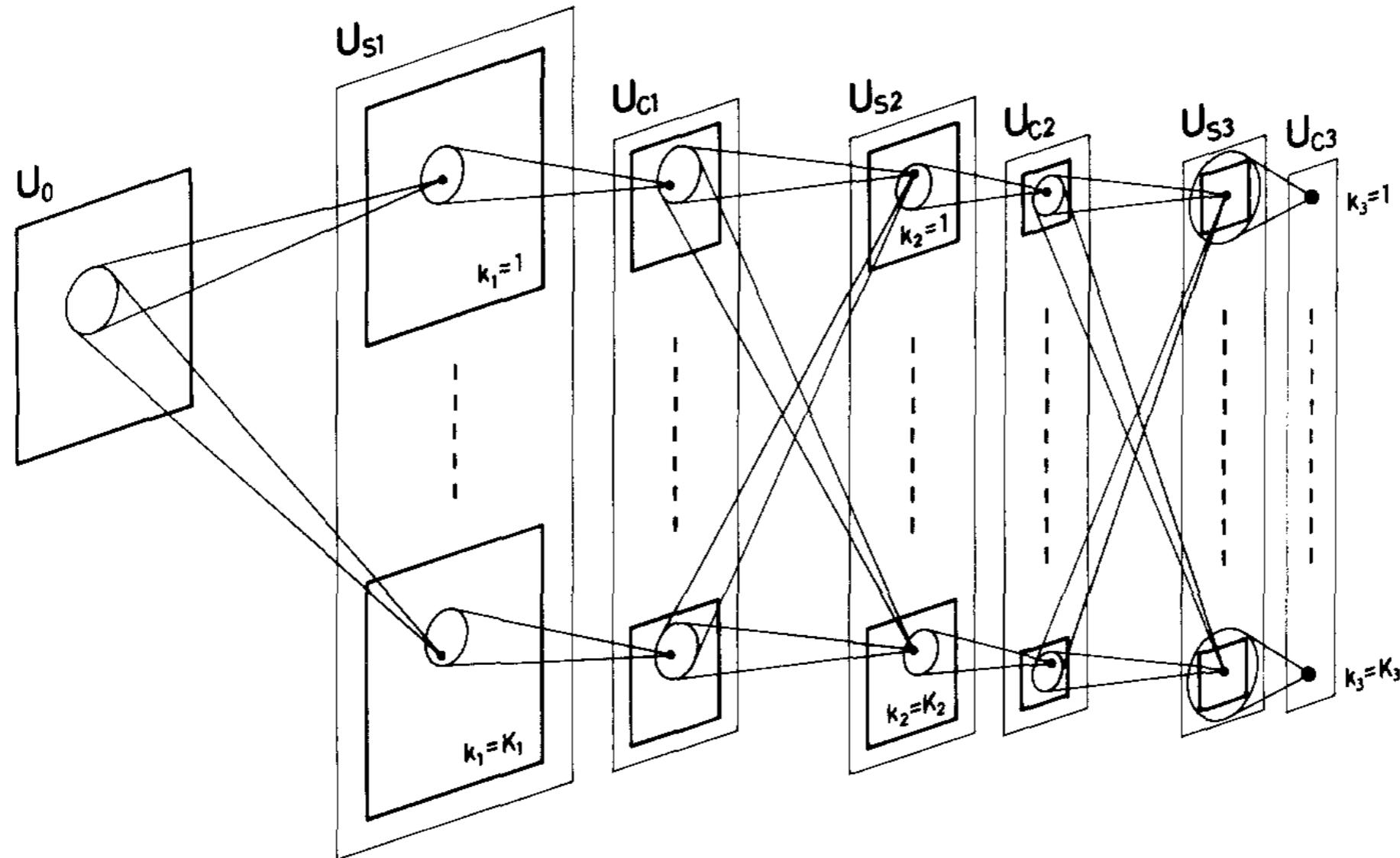
2000's



Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position

Kunihiro Fukushima

NHK Broadcasting Science Research Laboratories, Kinuta, Setagaya, Tokyo, Japan



Neocognitron: rationale

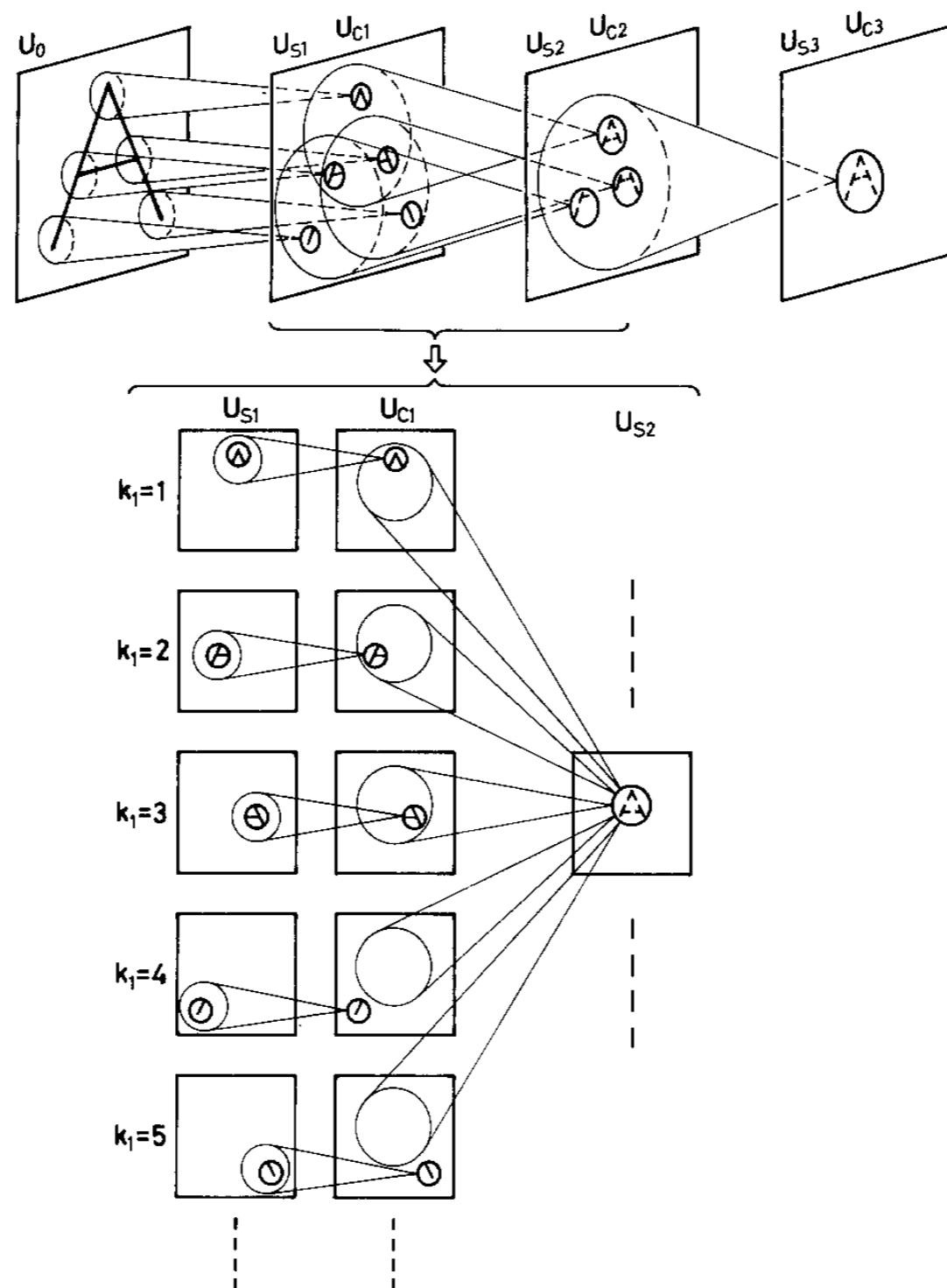


Fig. 5. An example of the interconnections between cells and the response of the cells after completion of self-organization

Neocognitron: activation rule

$$u_{Sl}(k_l, \mathbf{n}) = r_l \cdot \varphi \left[\frac{1 + \sum_{k_{l-1}=1}^{K_{l-1}} \sum_{\mathbf{v} \in S_l} a_l(k_{l-1}, \mathbf{v}, k_l) \cdot u_{Cl-1}(k_{l-1}, \mathbf{n} + \mathbf{v})}{1 + \frac{2r_l}{1+r_l} \cdot b_l(k_l) \cdot v_{Cl-1}(\mathbf{n})} - 1 \right]$$

where

$$\varphi[x] = \begin{cases} x & x \geq 0 \\ 0 & x < 0. \end{cases}$$

convolution

divisive normalization

$v_{Cl-1}(\mathbf{n}) = \sqrt{\sum_{k_{l-1}=1}^{K_{l-1}} \sum_{\mathbf{v} \in S_l} c_{l-1}(\mathbf{v}) \cdot u_{Cl-1}^2(k_{l-1}, \mathbf{n} + \mathbf{v})},$

Relu

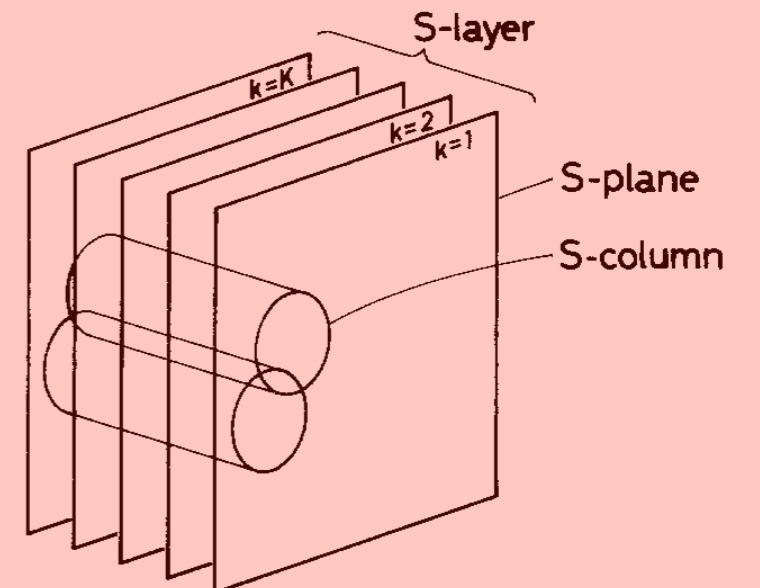
Neocognitron: learning rule

Let cell $u_{sl}(\hat{k}_l, \hat{\mathbf{n}})$ be selected as a representative.

$$\Delta a_l(k_{l-1}, \mathbf{v}, \hat{k}_l) = q_l \cdot c_{l-1}(\mathbf{v}) \cdot u_{Cl-1}(k_{l-1}, \hat{\mathbf{n}} + \mathbf{v}),$$

Hebbian learning

From each S-column, every time when a stimulus pattern is presented, the S-cell which is yielding the largest output is chosen as a candidate for the representatives. Hence, there is a possibility that a number of candidates appear in a single S-plane. If two or more candidates appear in a single S-plane, only the one which is yielding the largest output among them is selected as the representative from that S-plane. In



Local WTA

Neocognitron: performance

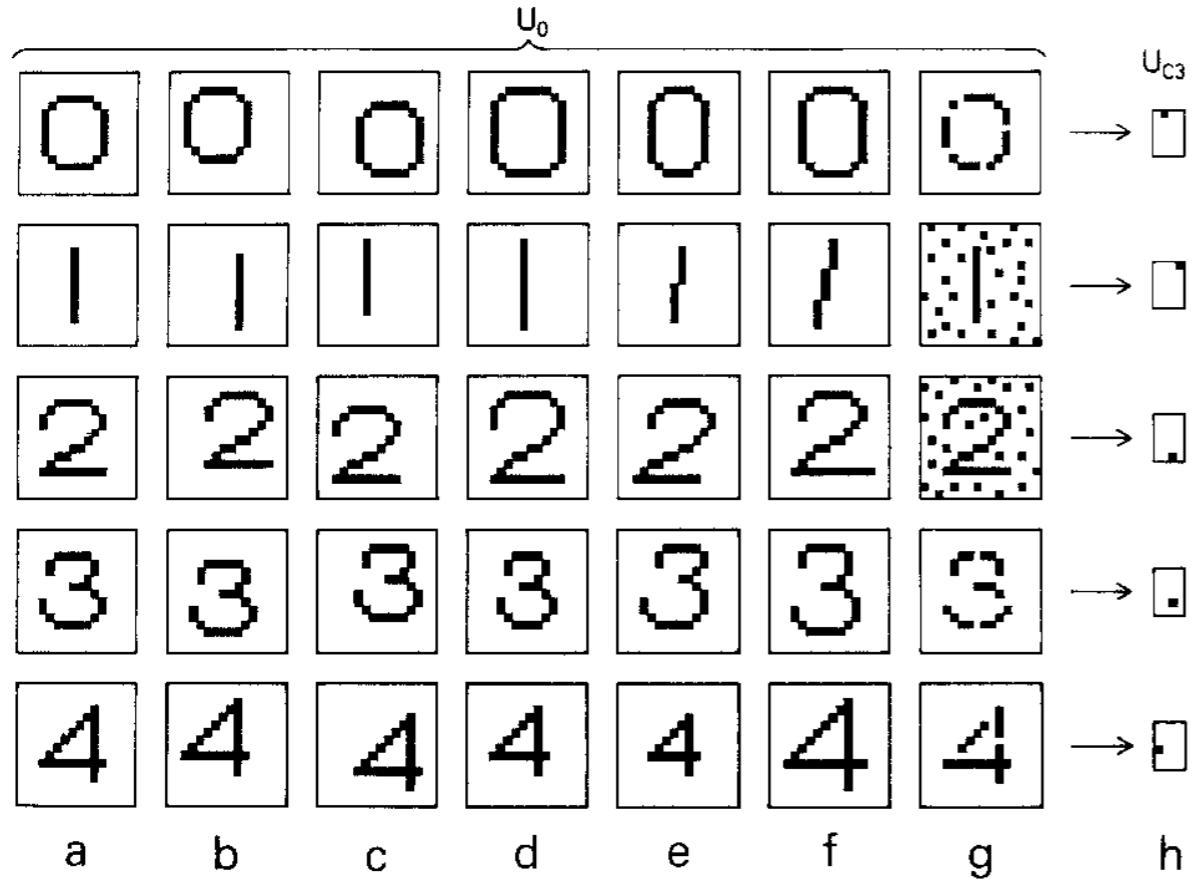


Fig. 6. Some examples of distorted stimulus patterns which the neocognitron has correctly recognized, and the response of the final layer of the network

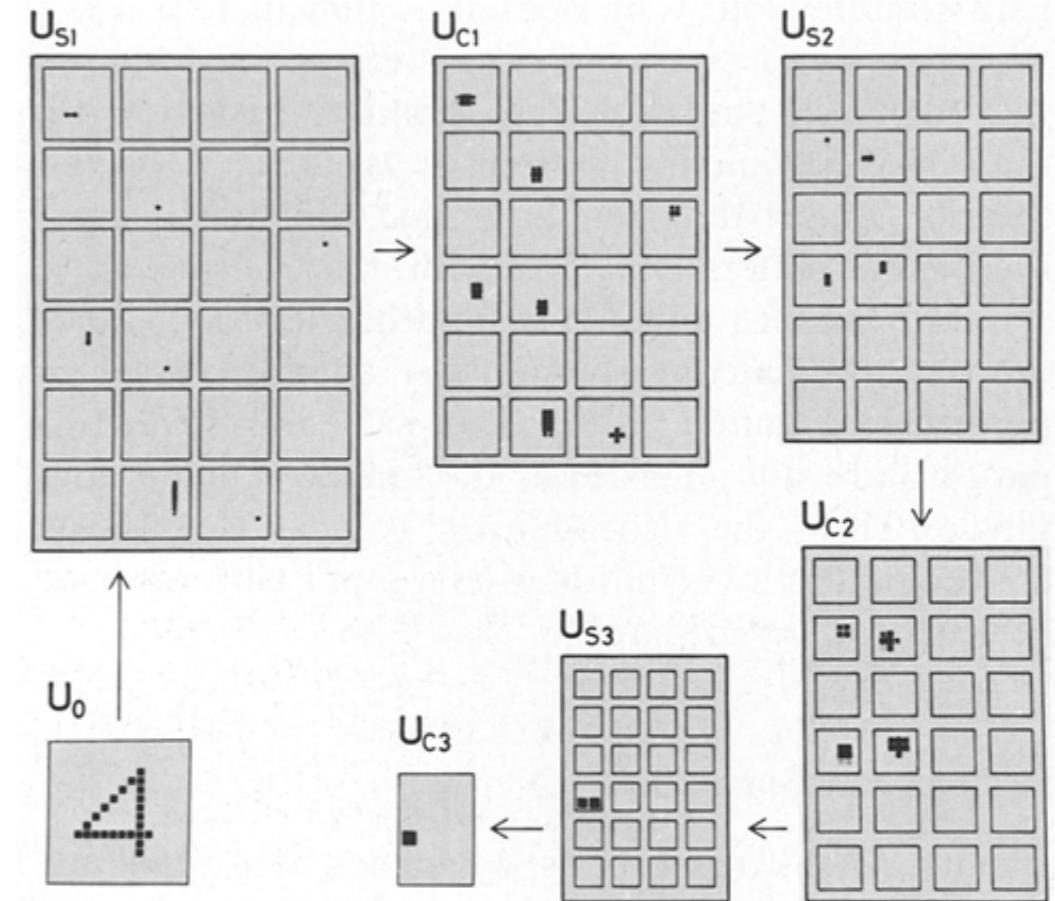
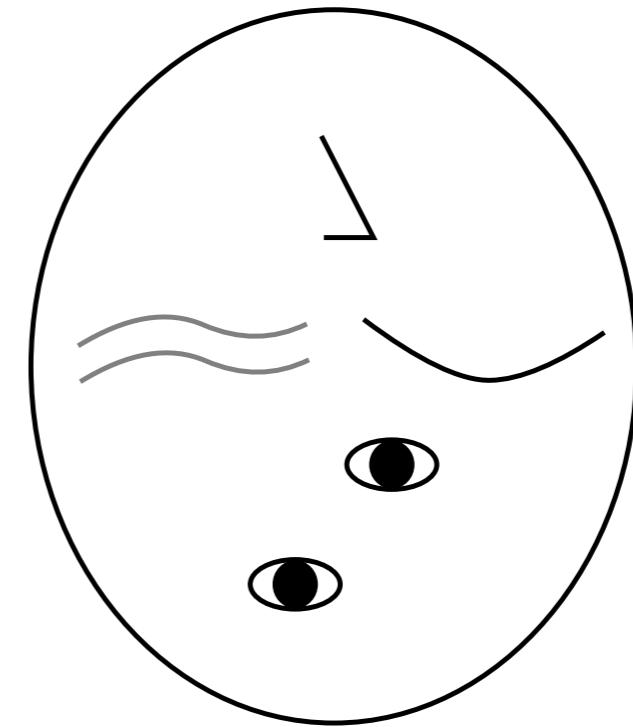
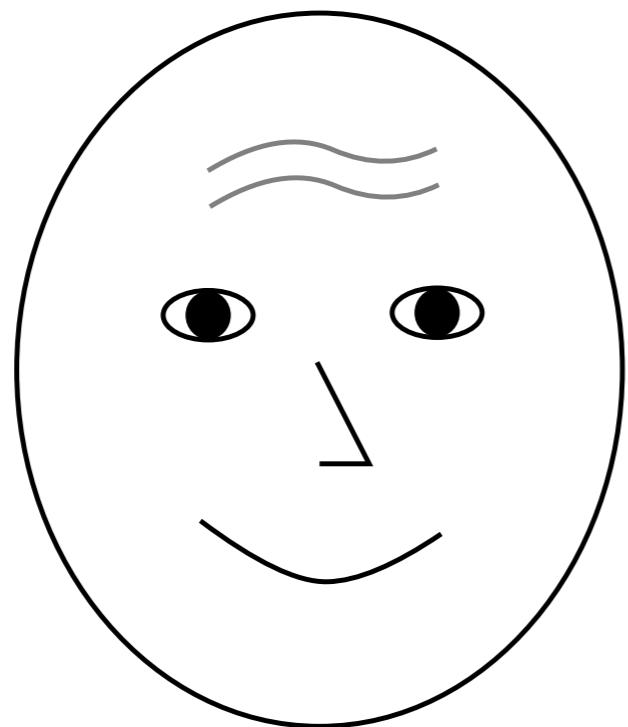


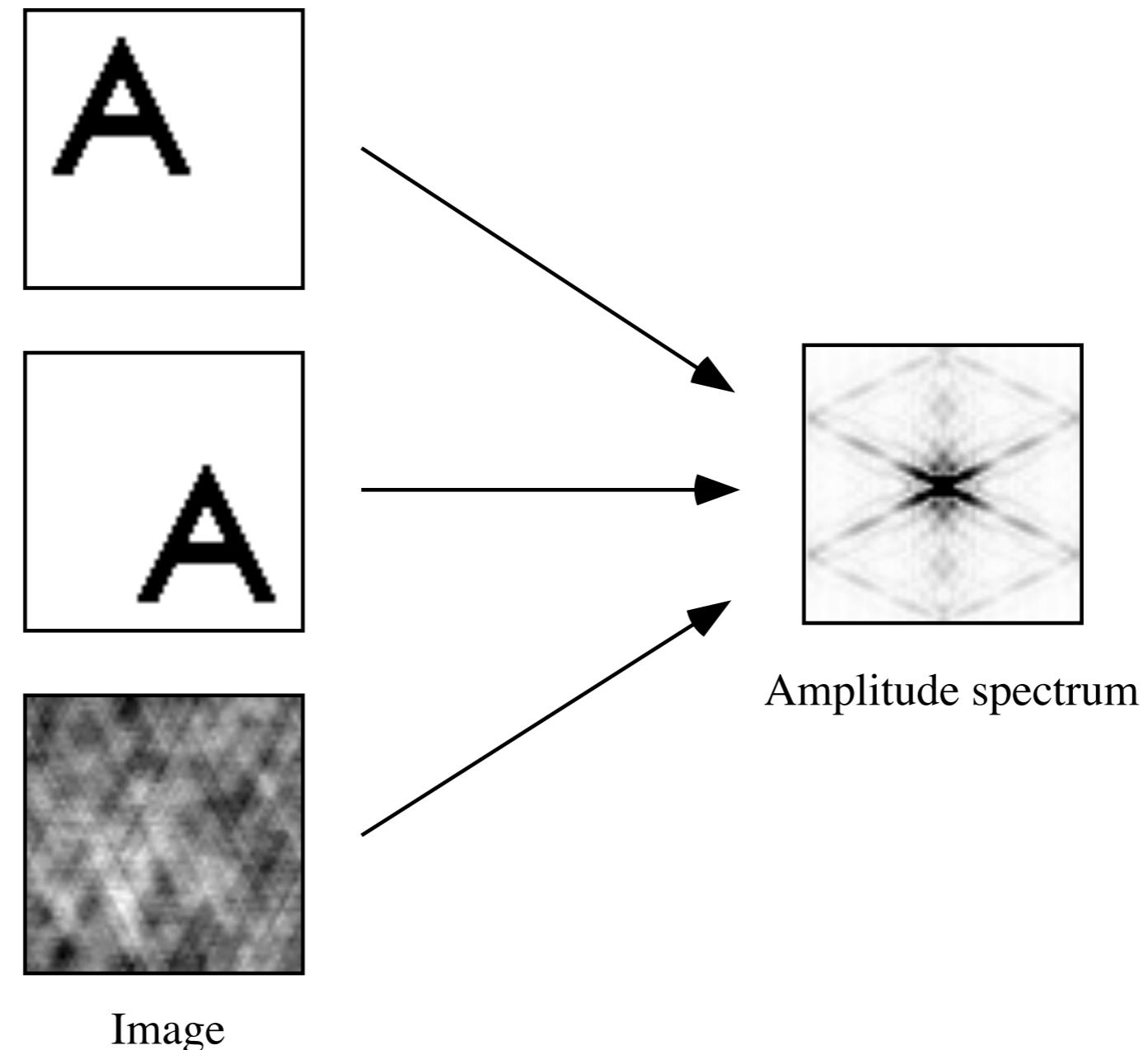
Fig. 7. A display of an example of the response of all the individual cells in the neocognitron

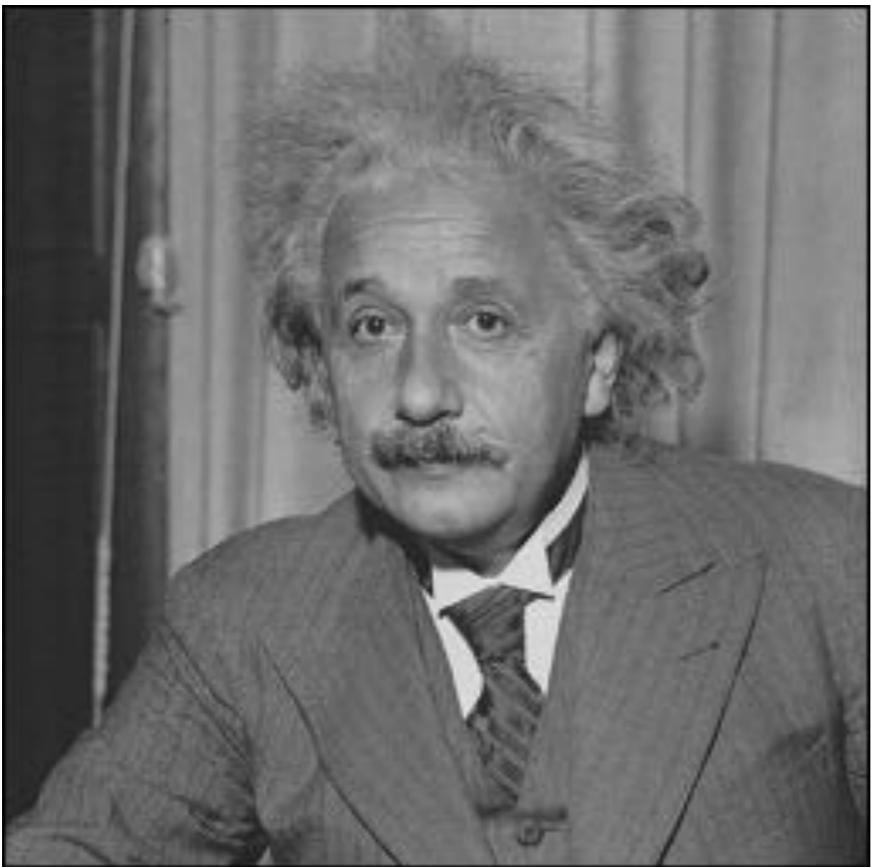
This isn't a good model of perception

Relative spatial relationships are important

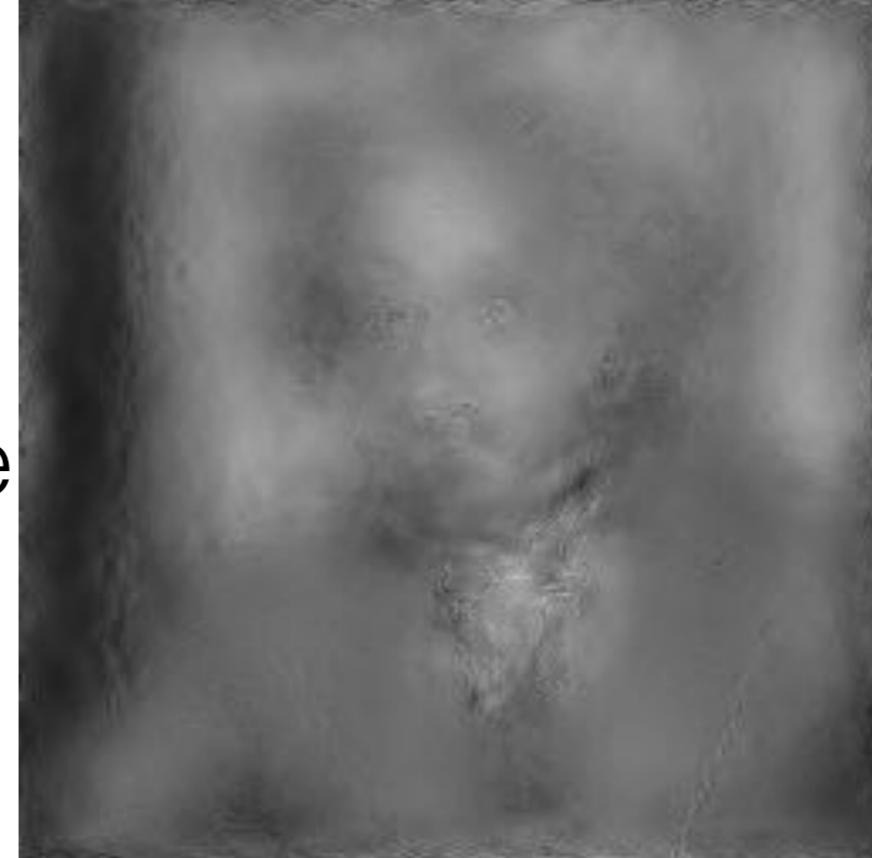
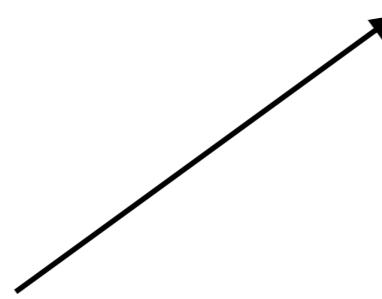


Spatial phase, not amplitude, determines shape

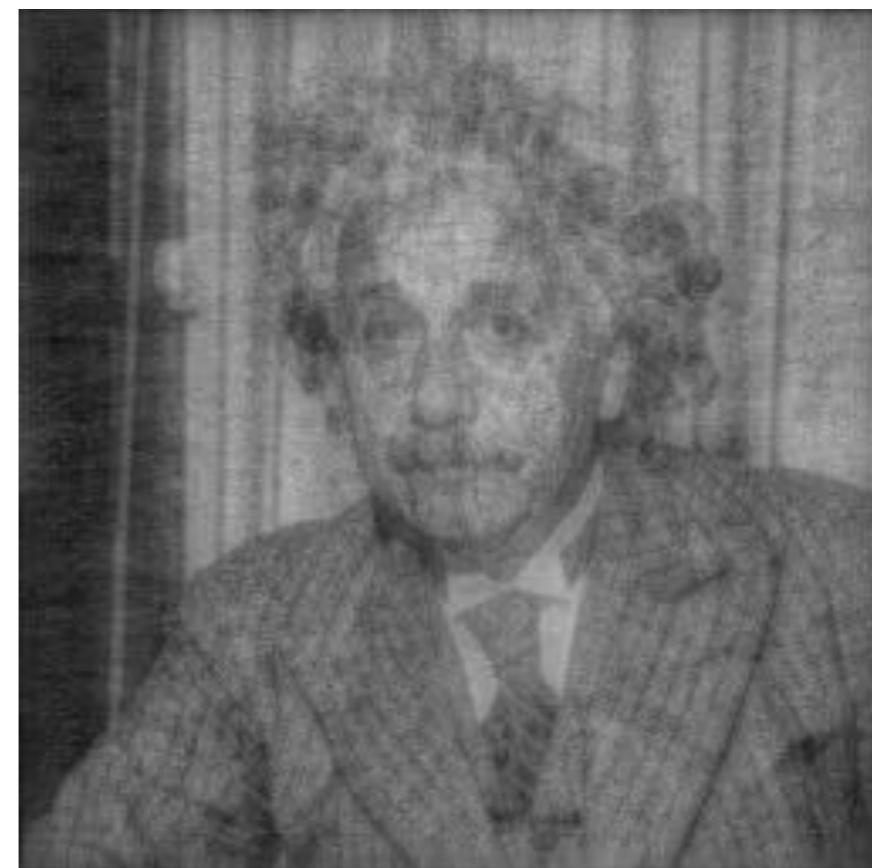
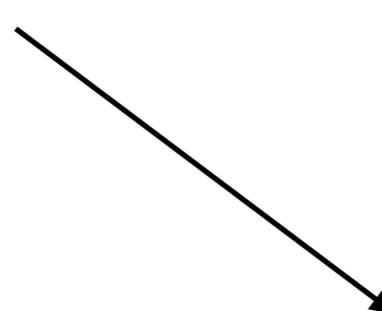




randomize
local phase

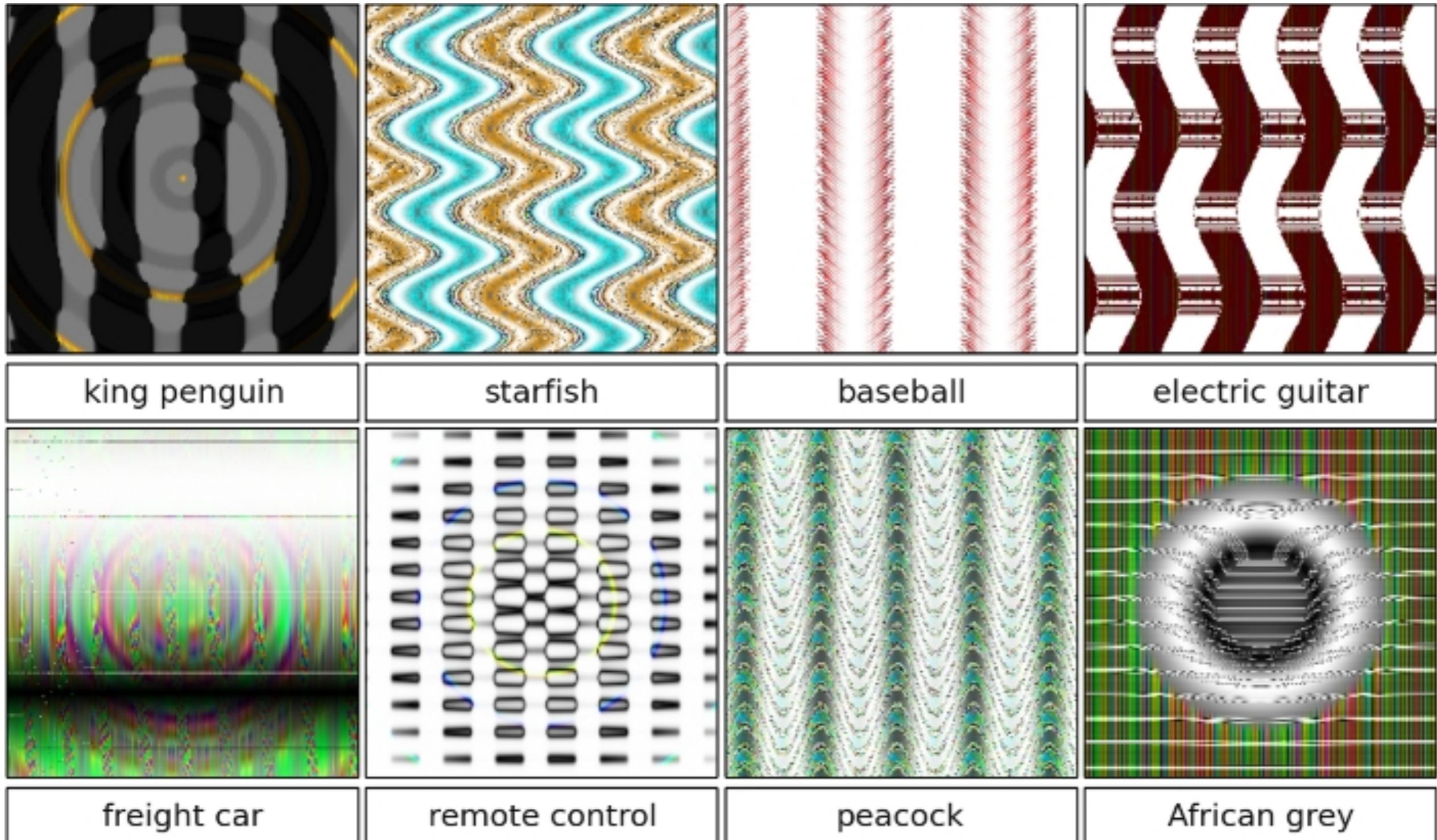


randomize
local amp.



Deep neural networks are easily fooled

(Nguyen, Yosinski & Clune 2014)





Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope*

AUDE OLIVA

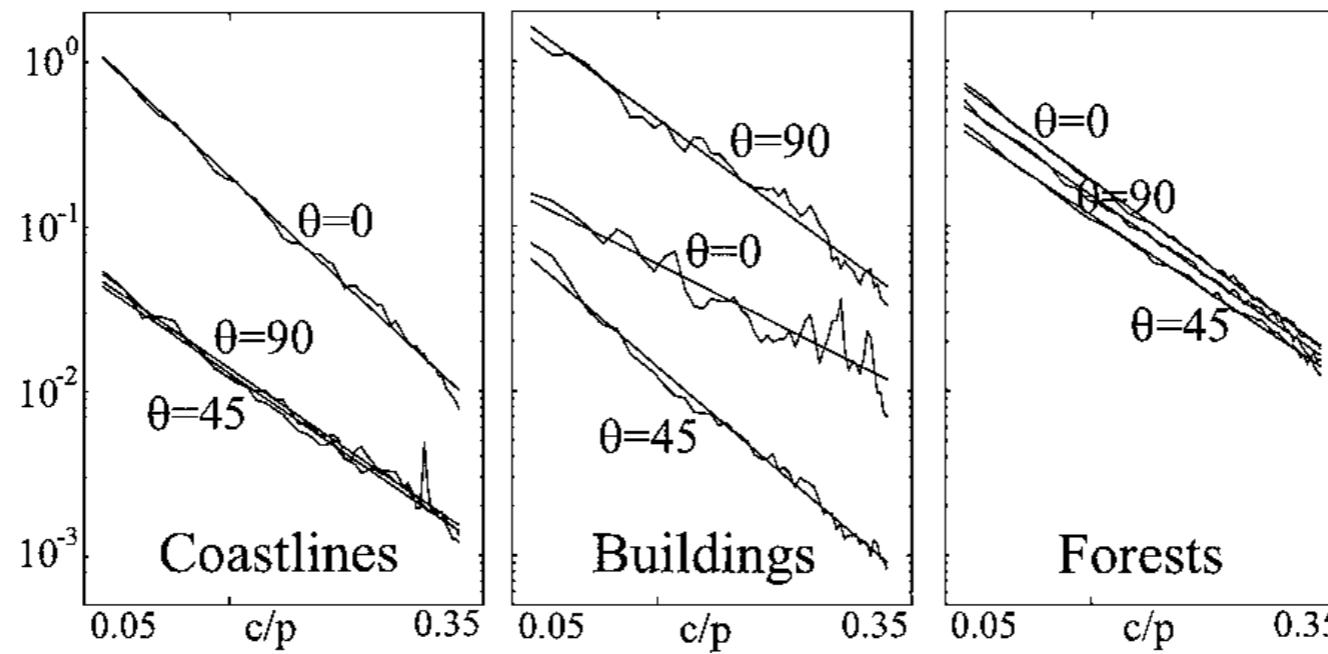
Harvard Medical School and the Brigham and Women's Hospital, 221 Longwood Ave., Boston, MA 02115

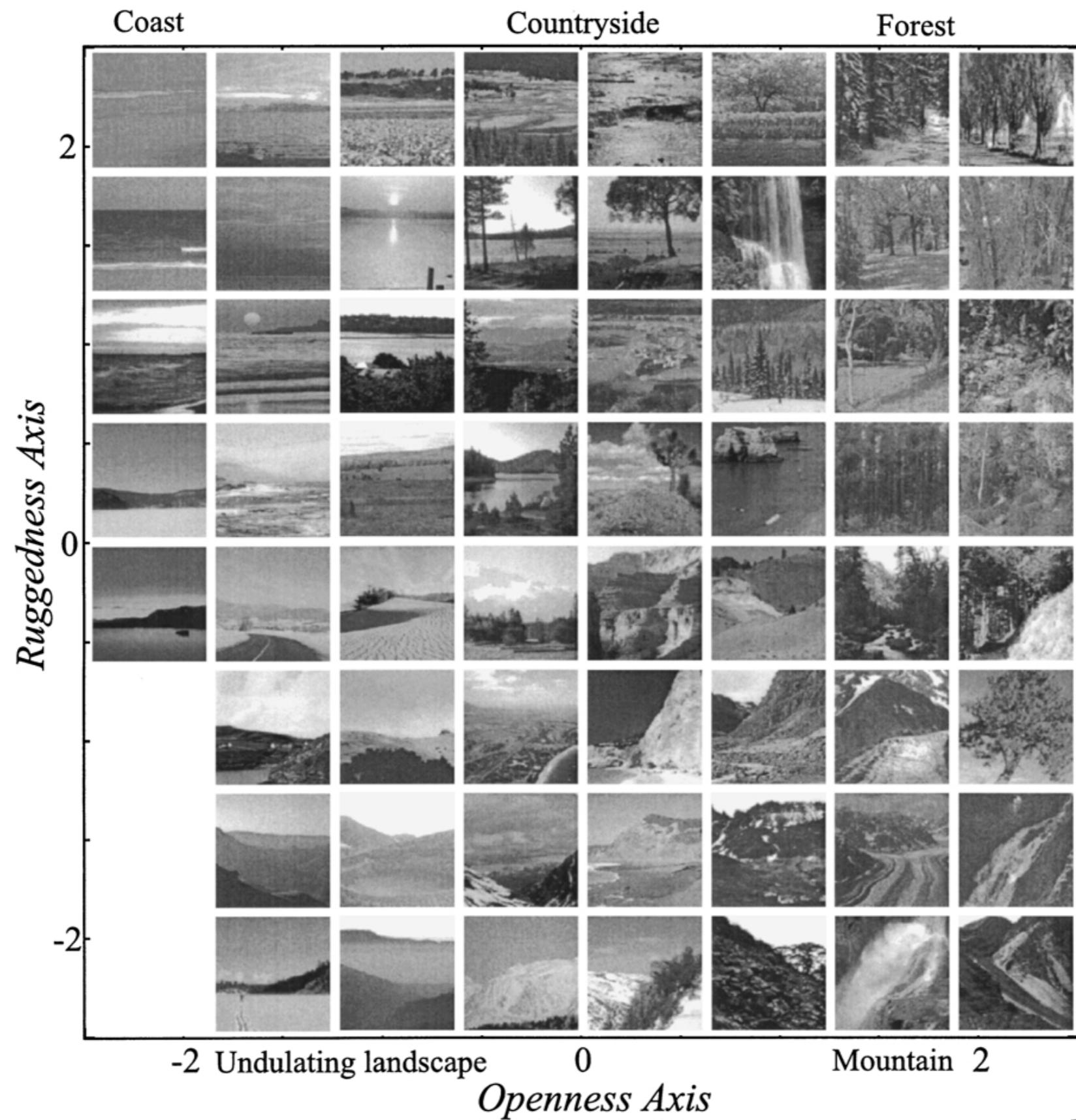
oliva@search.bwh.harvard.edu

ANTONIO TORRALBA

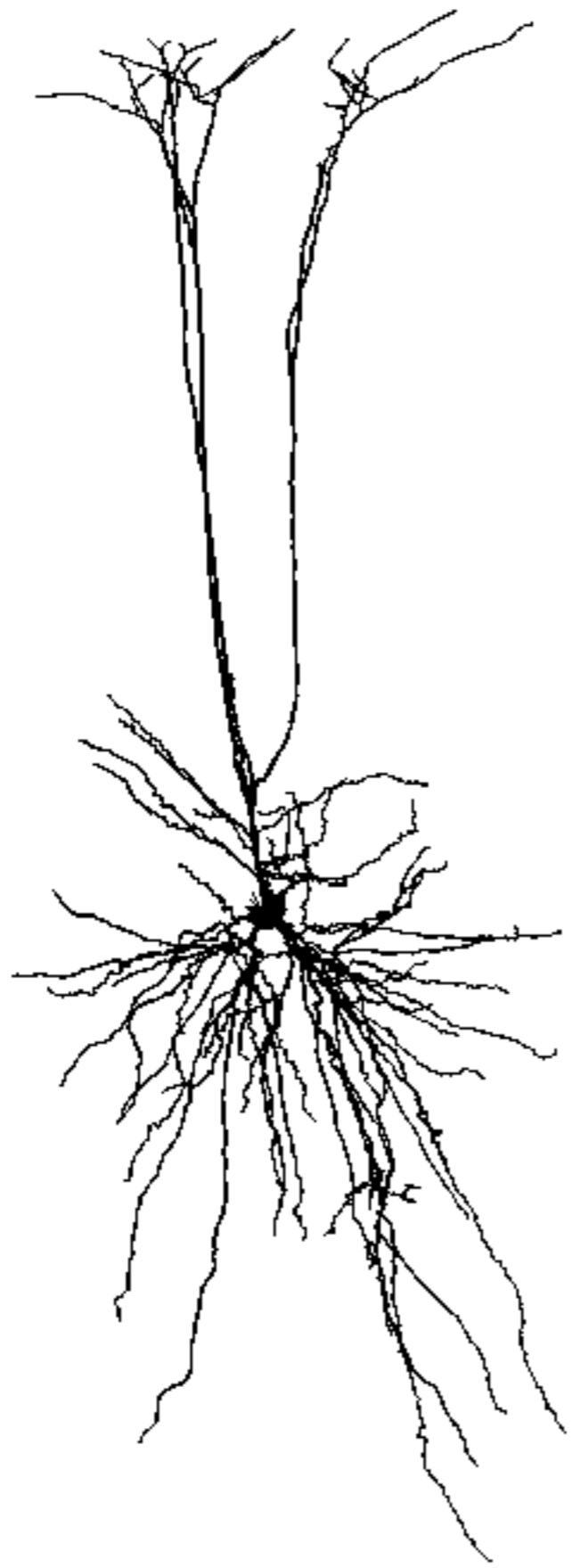
Department of Brain and Cognitive Sciences, MIT, 45 Carleton Street, Cambridge, MA 02139

torralba@ai.mit.edu

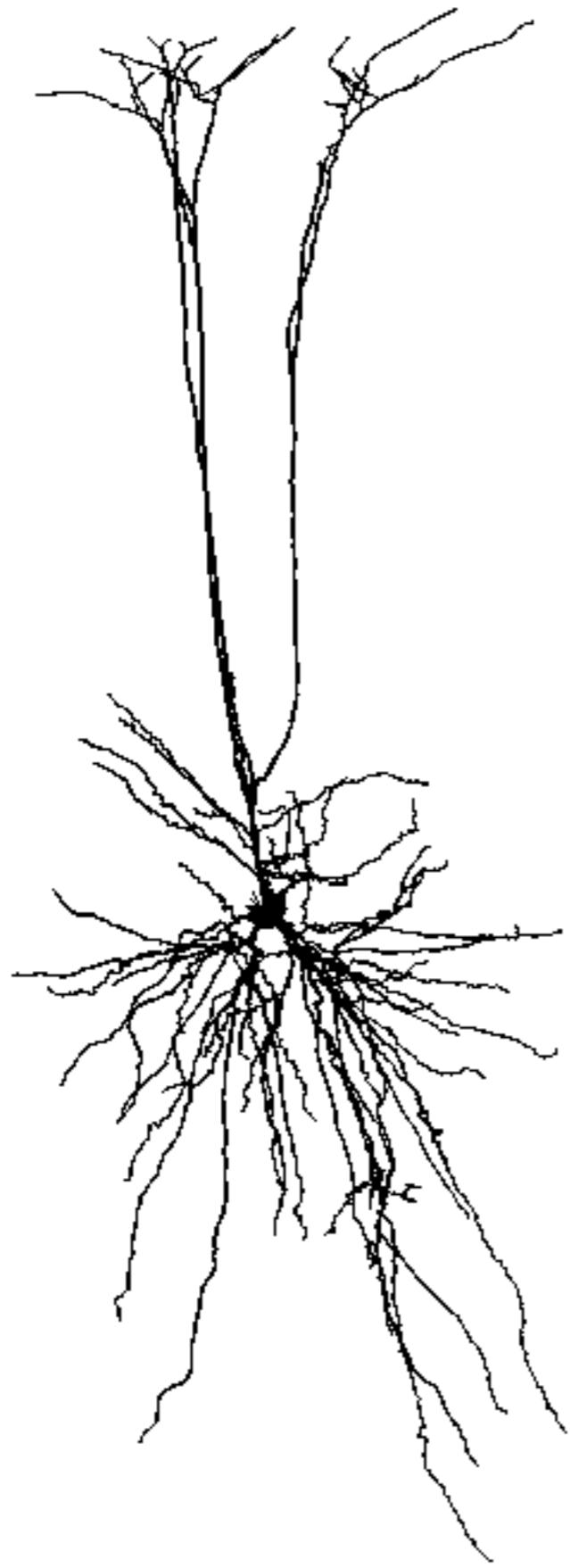




What's missing?



$$\neq \quad g\Bigl(\sum_i w_i \, x_i\Bigr)$$



~

$$g\left(\sum_i w_i \prod_{j \in G_i} x_j\right)$$

Single neuron recording ⇒ Single neuron thinking



1940

PROCEEDINGS OF THE IRE

November

What the Frog's Eye Tells the Frog's Brain*

J. Y. LETTVIN†, H. R. MATORANA‡, W. S. McCULLOCH||, SENIOR MEMBER, IRE,
AND W. H. PITTS||

Summary—In this paper, we analyze the activity of single fibers in the optic nerve of a frog. Our method is to find what sort of stimulus causes the largest activity in one nerve fiber and then what is the exciting aspect of that stimulus such that variations in everything else cause little change in the response. It has been known for the past 20 years that each fiber is connected not to a few rods and cones in the retina but to very many over a fair area. Our results show that for

factor. There are four types of fibers, each type concerned with a different sort of pattern. Each type is uniformly distributed over the

visual image in terms of local pattern independent of average illumination. We describe the patterns and show the functional and anatomical separation of the channels. This work has been done on the frog, and our interpretation applies only to the frog.

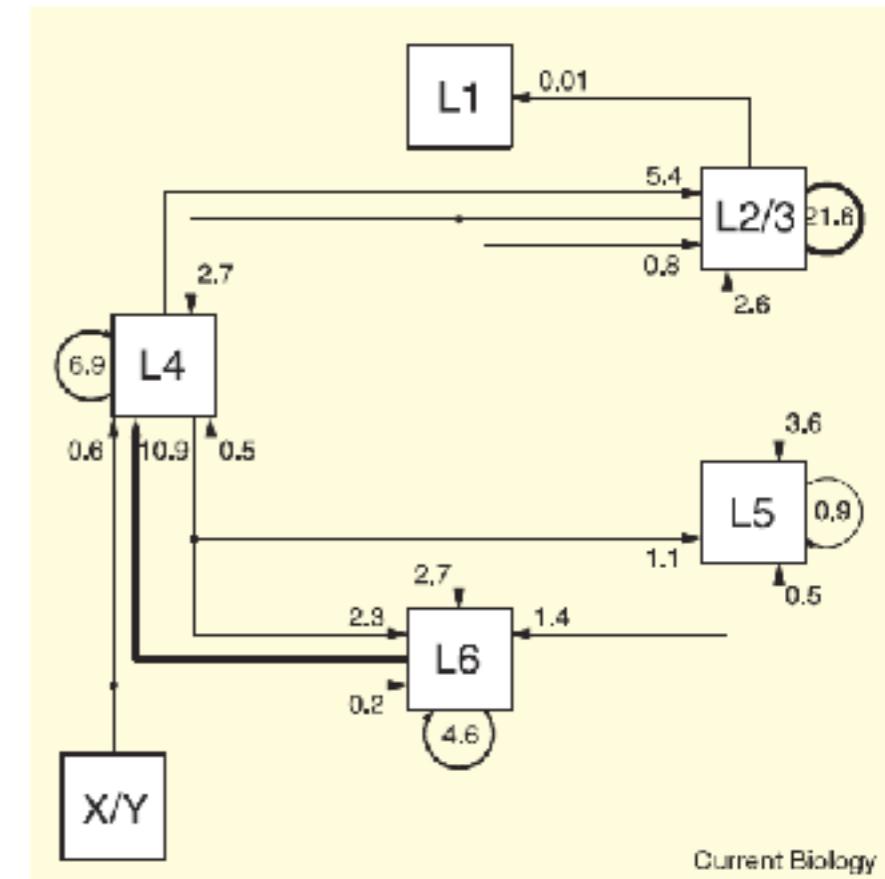
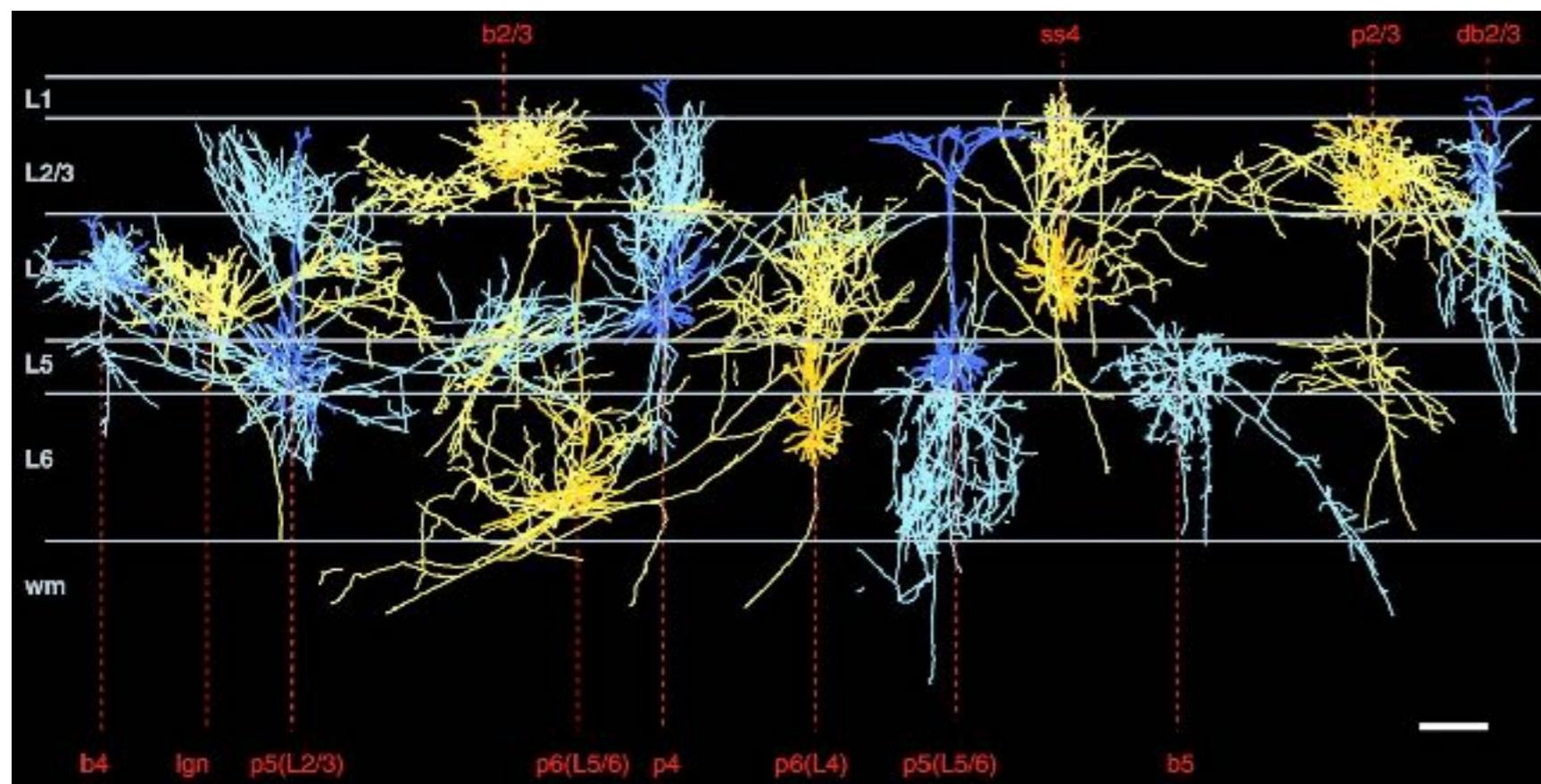
it moves like one. He can be fooled easily not only by a bit of dangled meat but by any moving small object. His sex life is conducted by sound and touch. His choice of paths in escaping enemies does not seem to be governed by anything more devious than leaping to where it is darker. Since he is equally at home in water and on

Anatomy of Frog Visual Apparatus

The retina of a frog is shown in Fig. 1(a). Between the rods and cones of the retina and the ganglion cells, whose axons form the optic nerve, lies a layer of con-

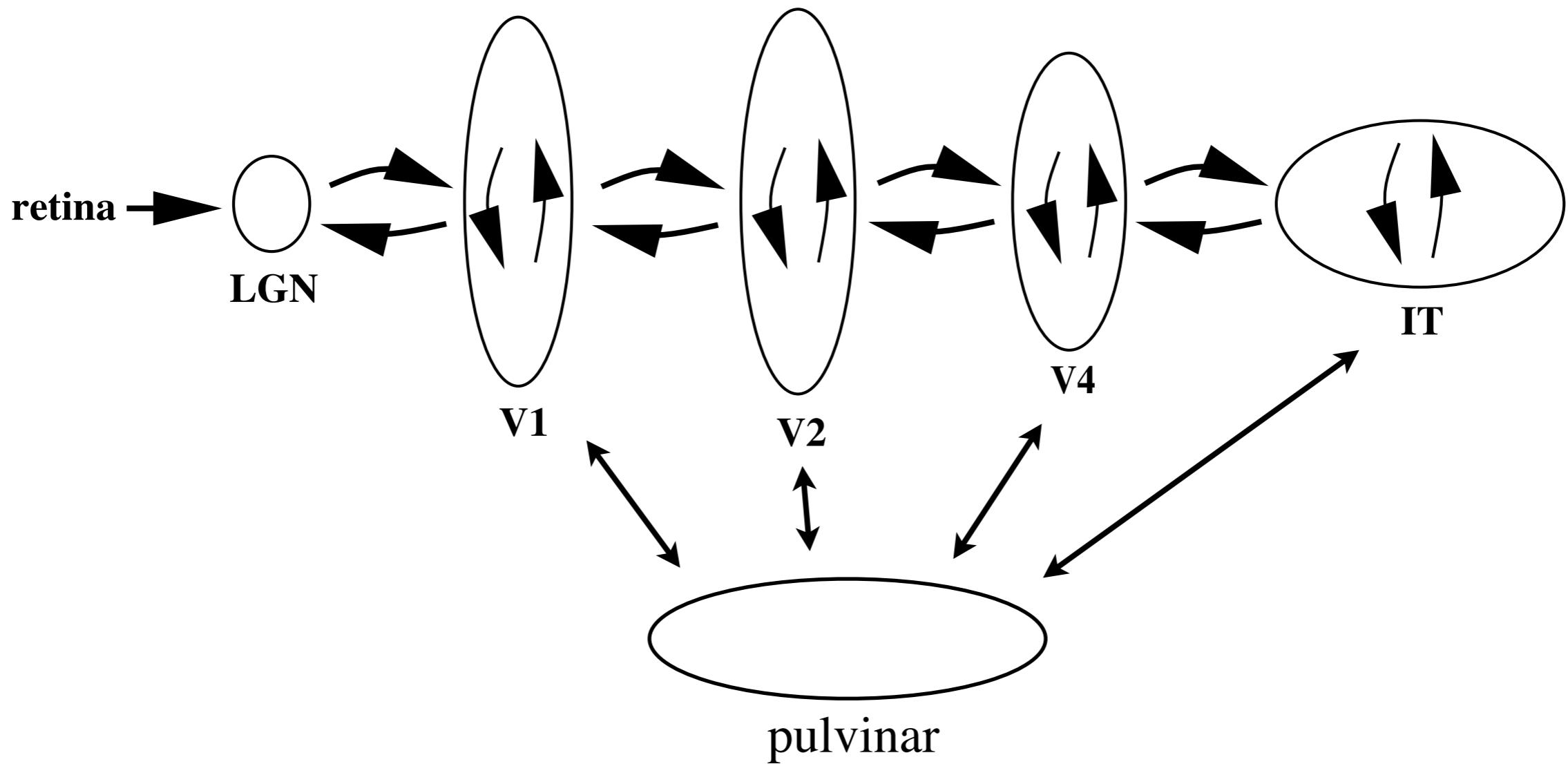
Cortical circuits

- highly organized by layer
 - layers are interconnected in a ‘canonical microcircuit’
 - signals are strongly intermixed within layers 2/3



(Douglas and Martin, 2007)

Feedback is pervasive throughout the thalamo-cortical system

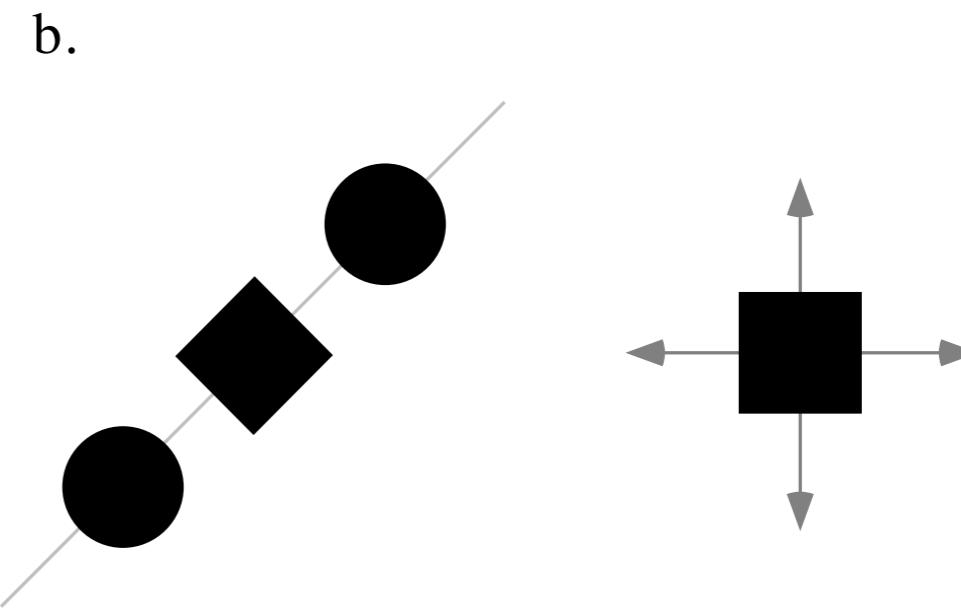
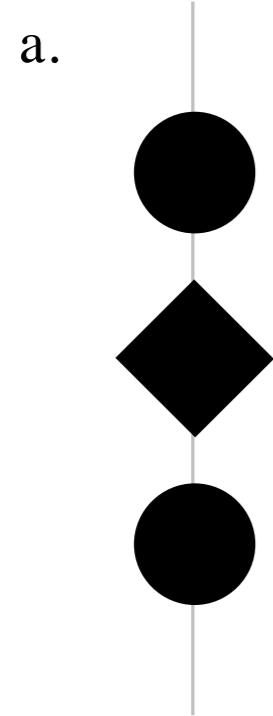


Two specific proposals

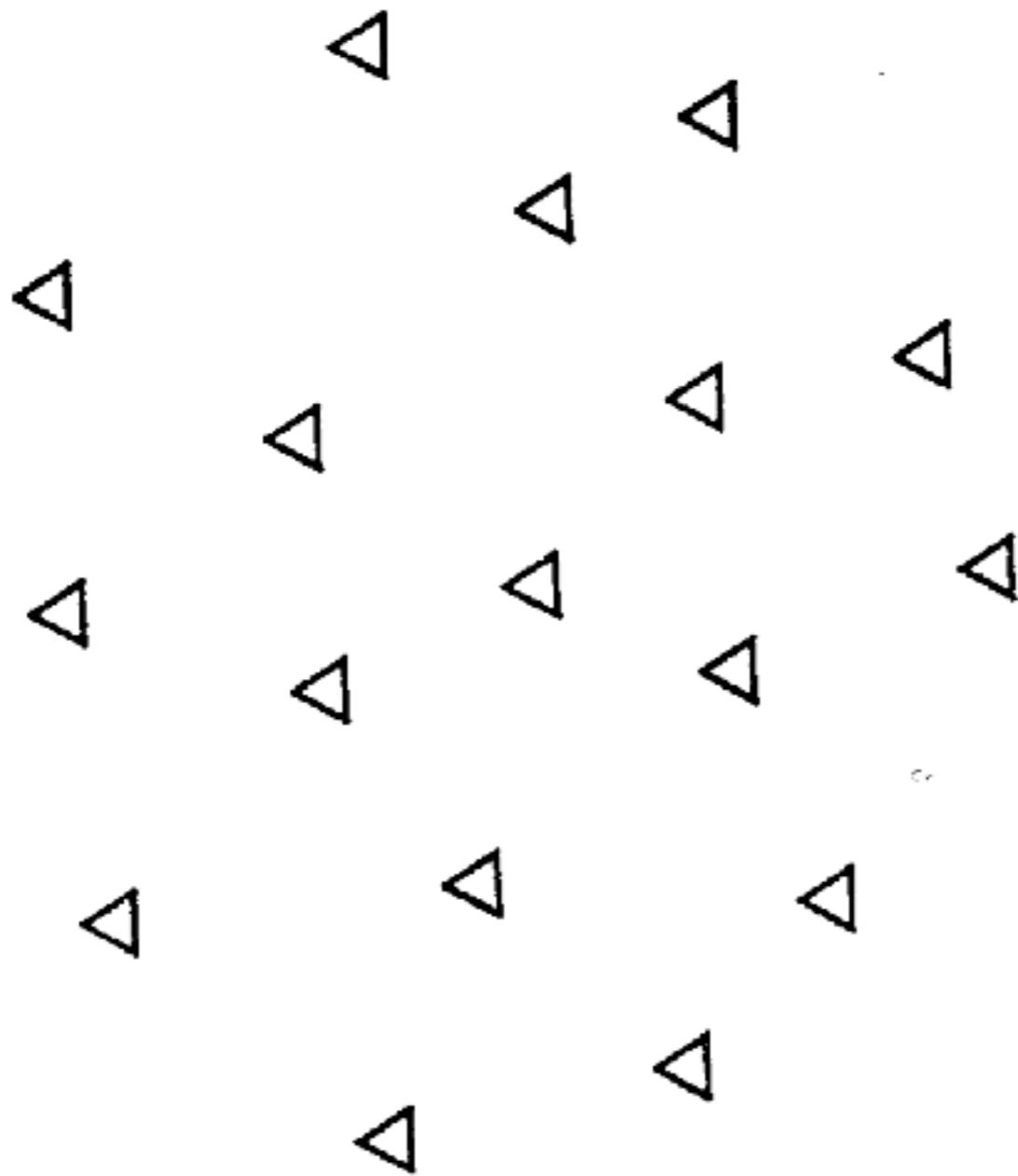
1. Dynamic routing
2. Hierarchical Bayesian inference

Reference frame effects in perception

Diamond or square?

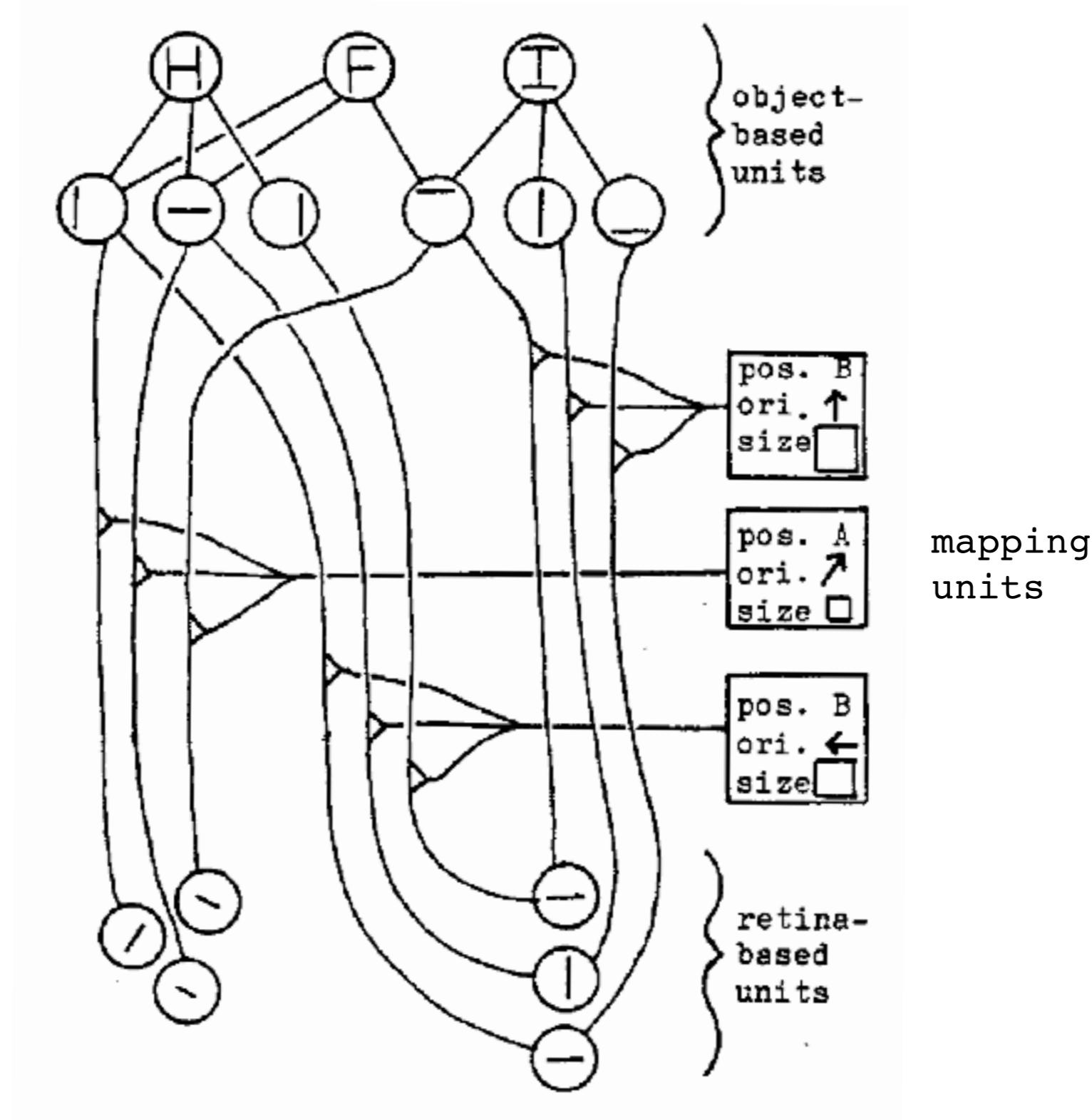


Which way are the triangles pointing?



From Attneave

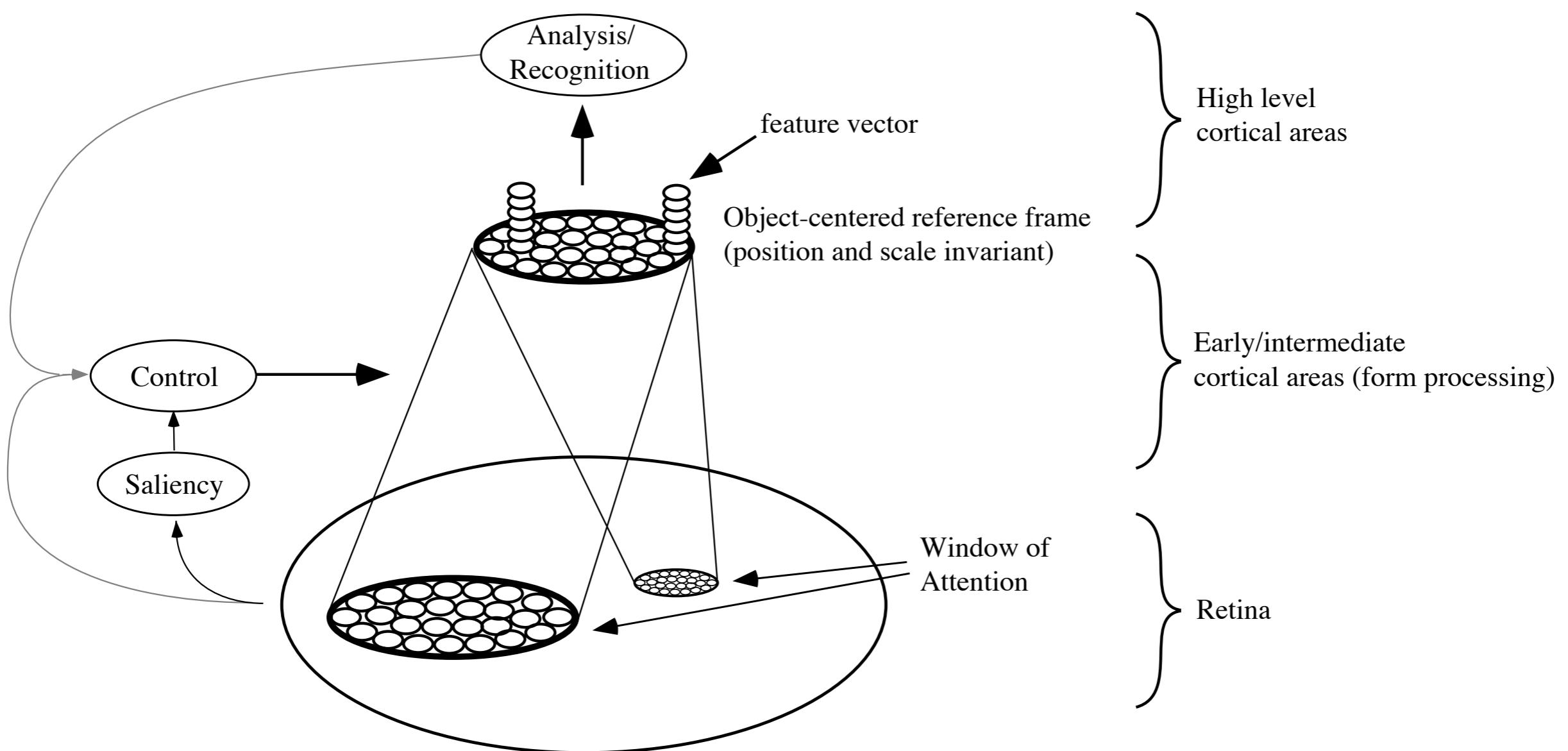
Reference frames require structured representations



Hinton (1981)

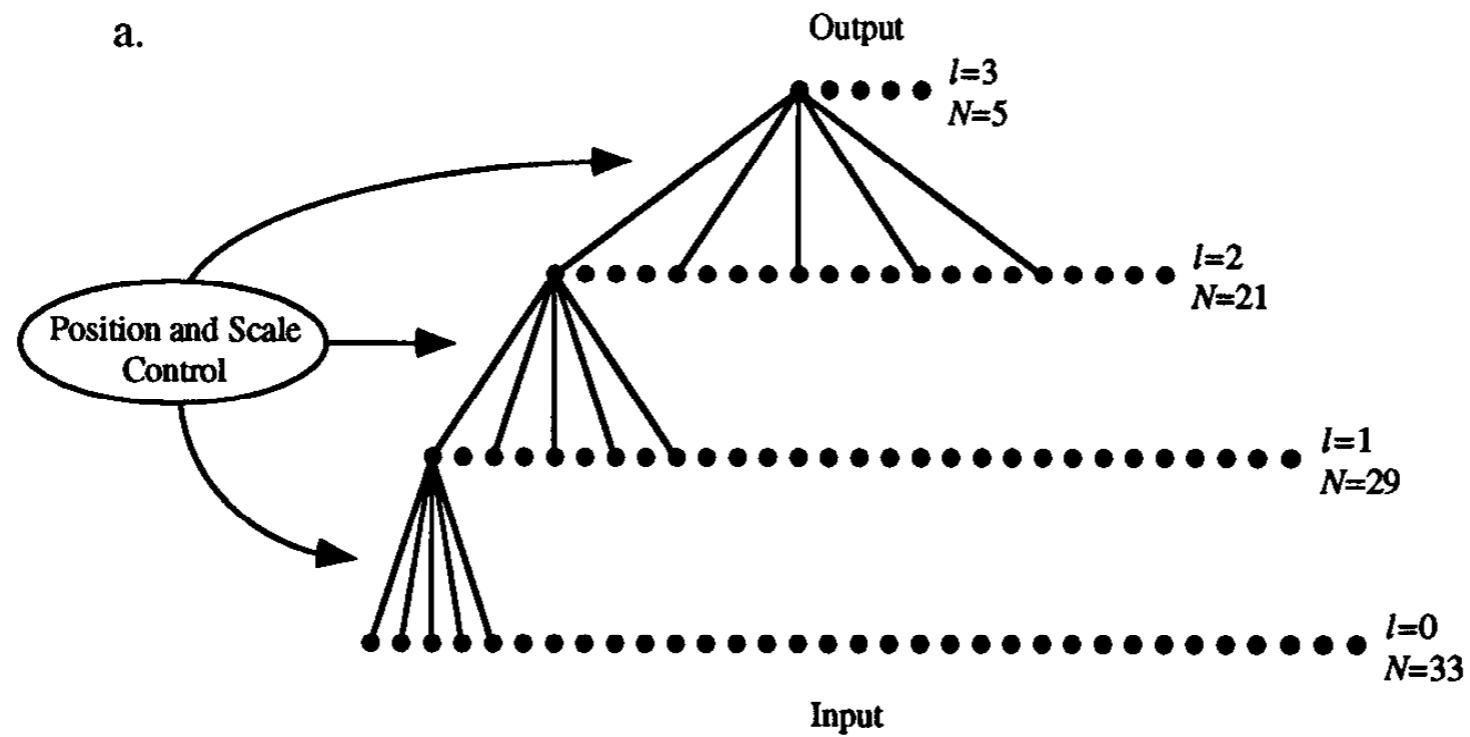
Dynamic routing

(Olshausen, Anderson, Van Essen 1993)

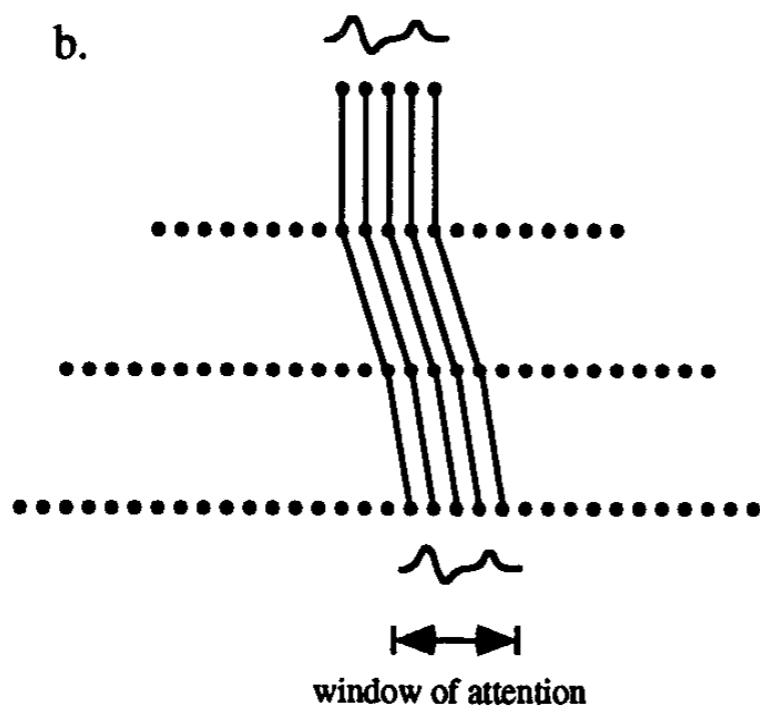


Dynamic routing circuit

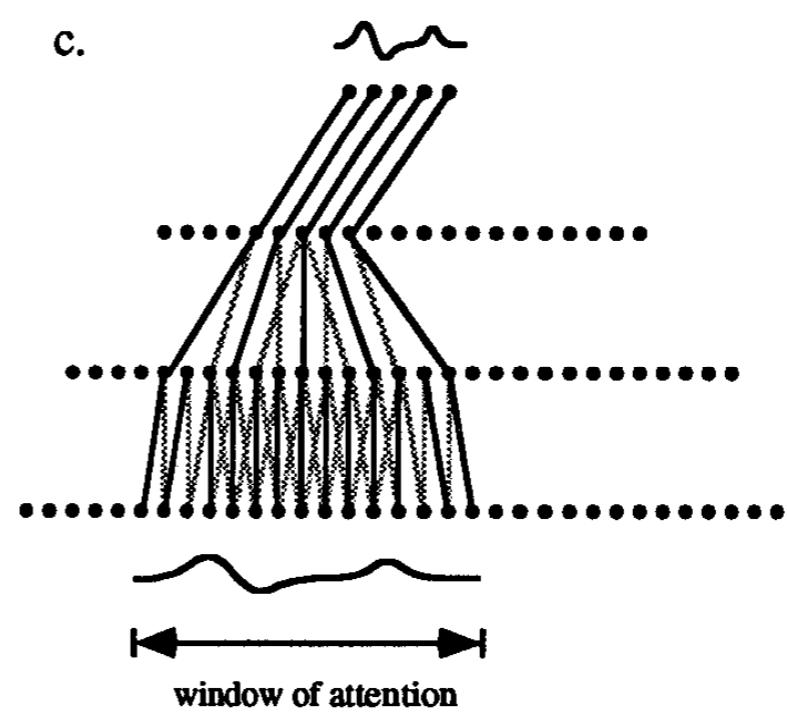
a.



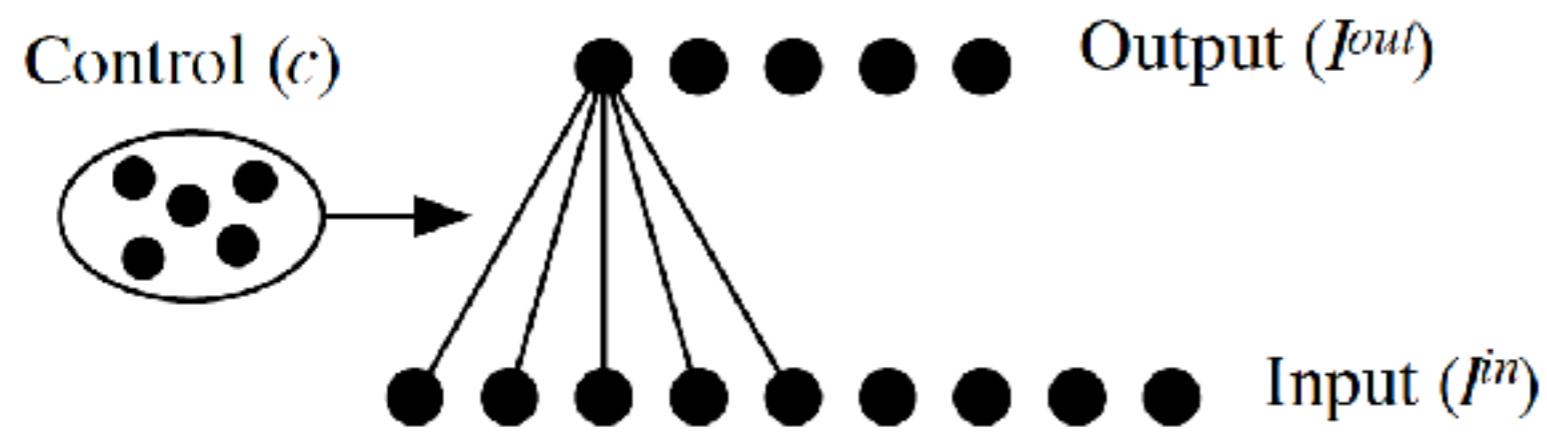
b.



c.



Dynamic routing: control

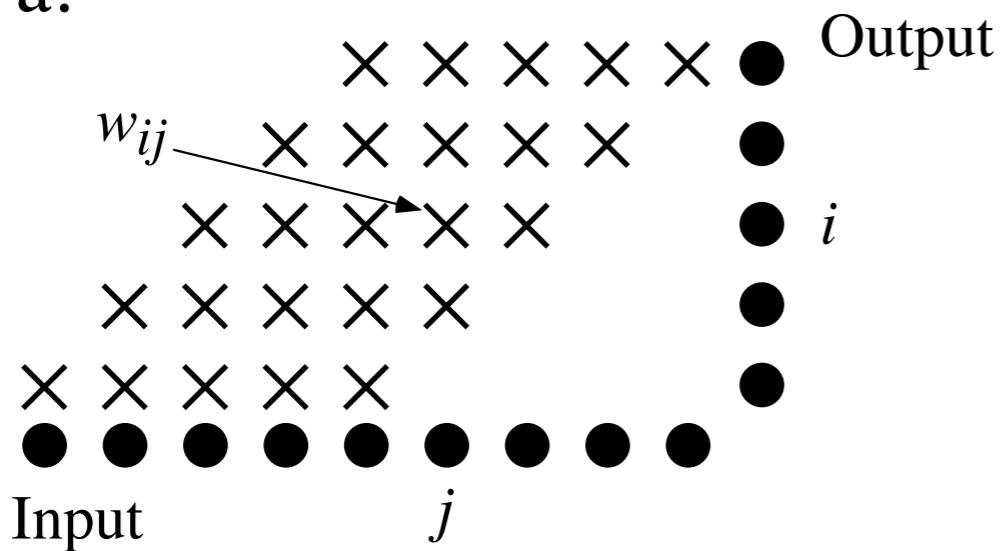


$$I_i^{out} = \sum_j w_{ij} I_j^{in}$$

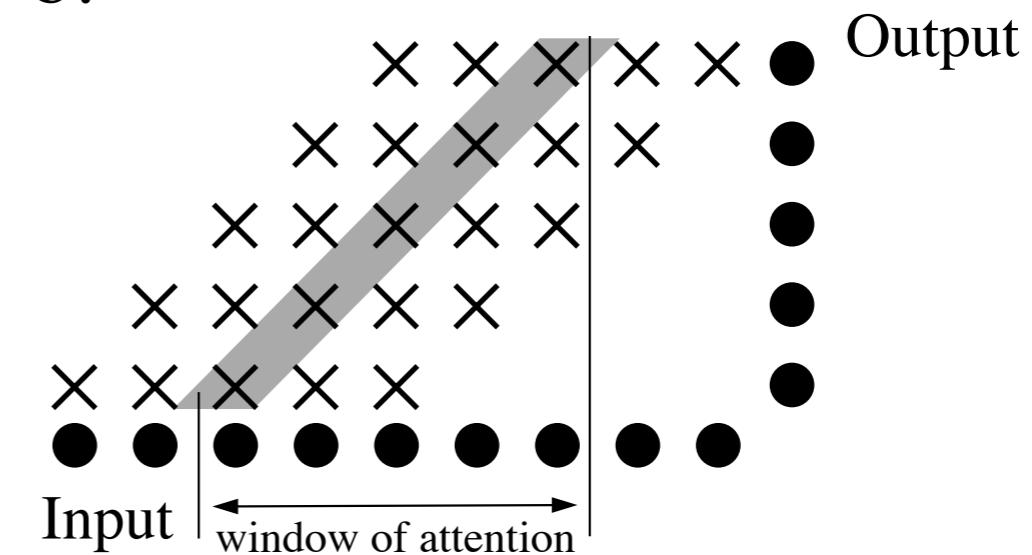
$$w_{ij} = \sum_k c_k \Gamma_{ijk}$$

Dynamic routing: control

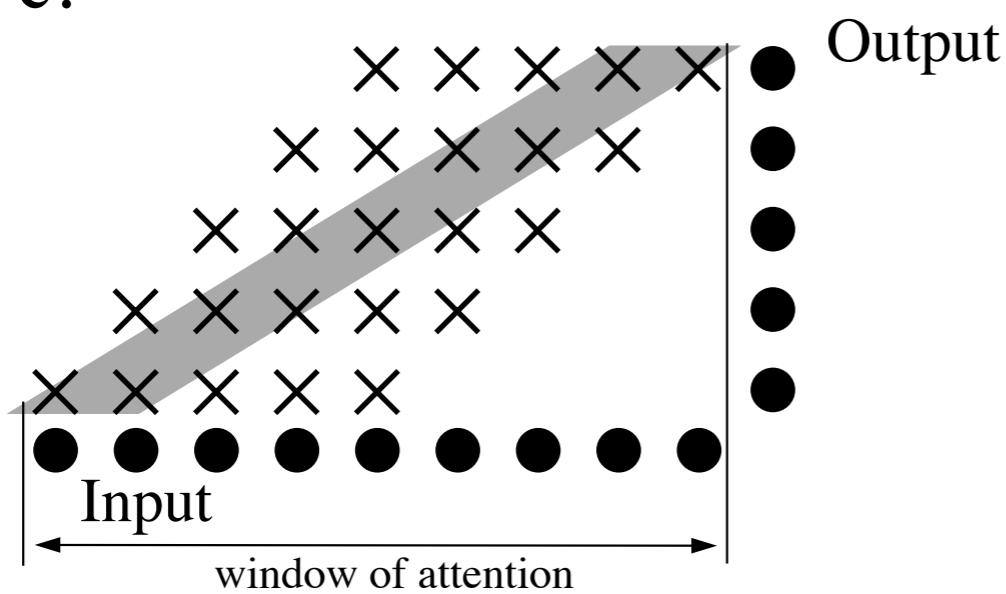
a.



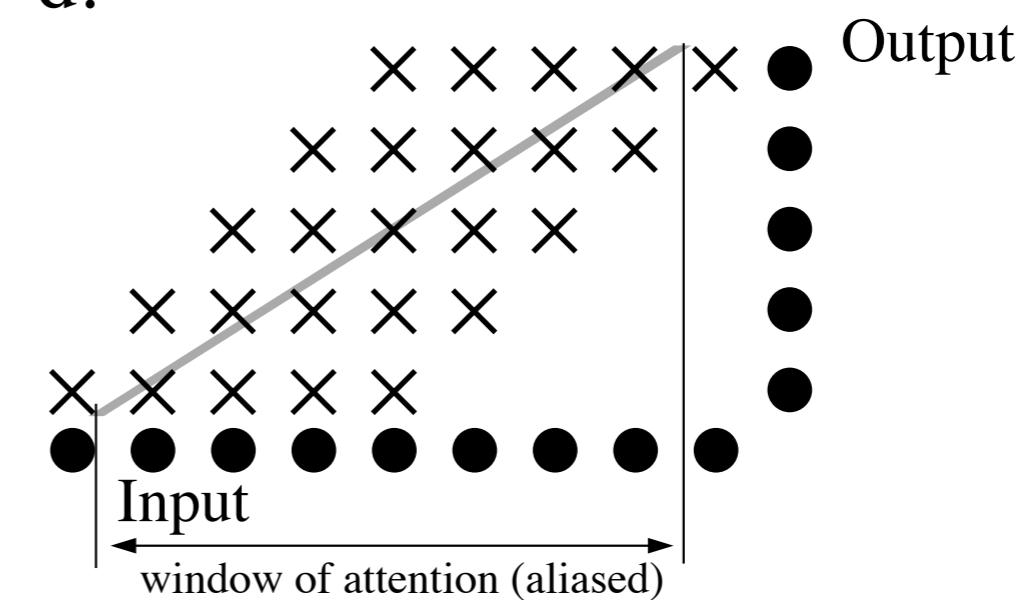
b.

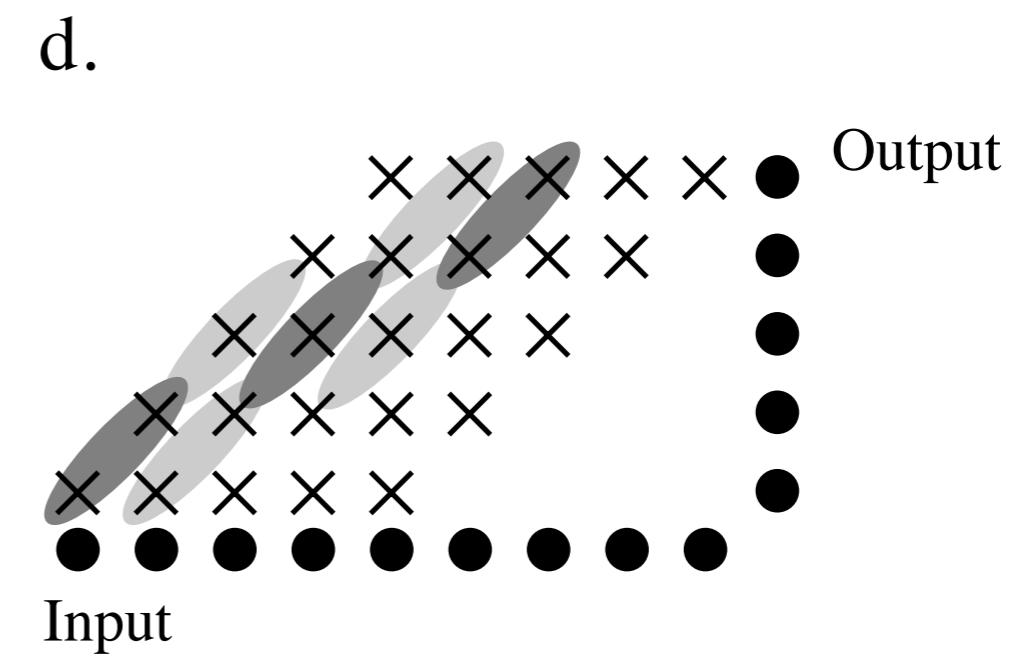
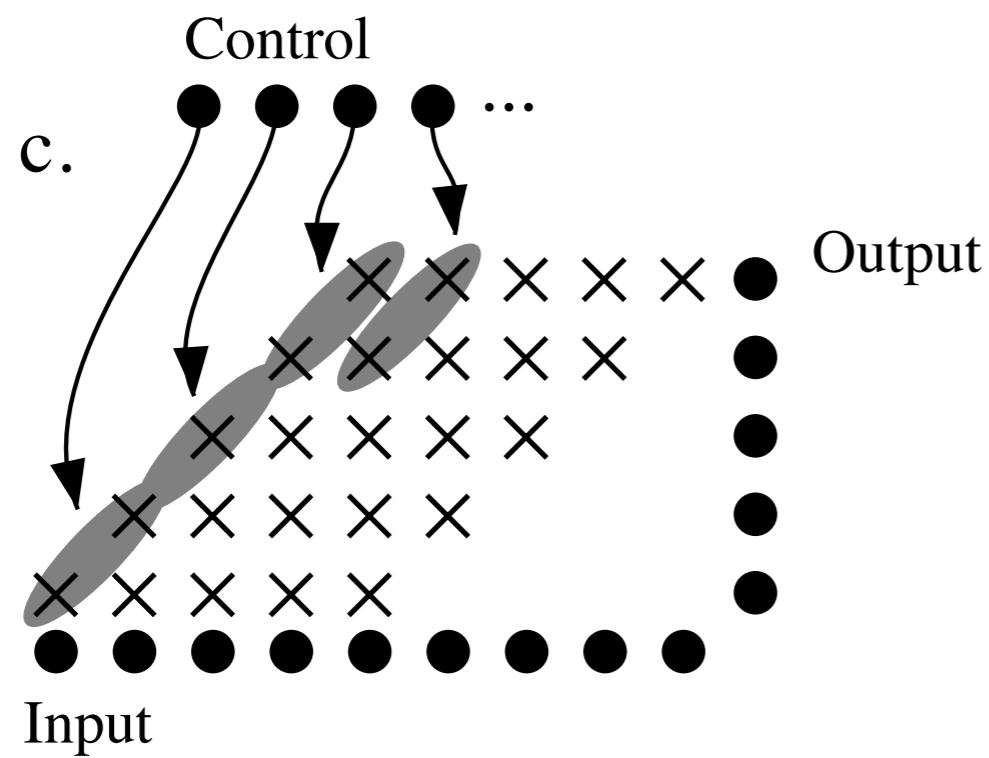
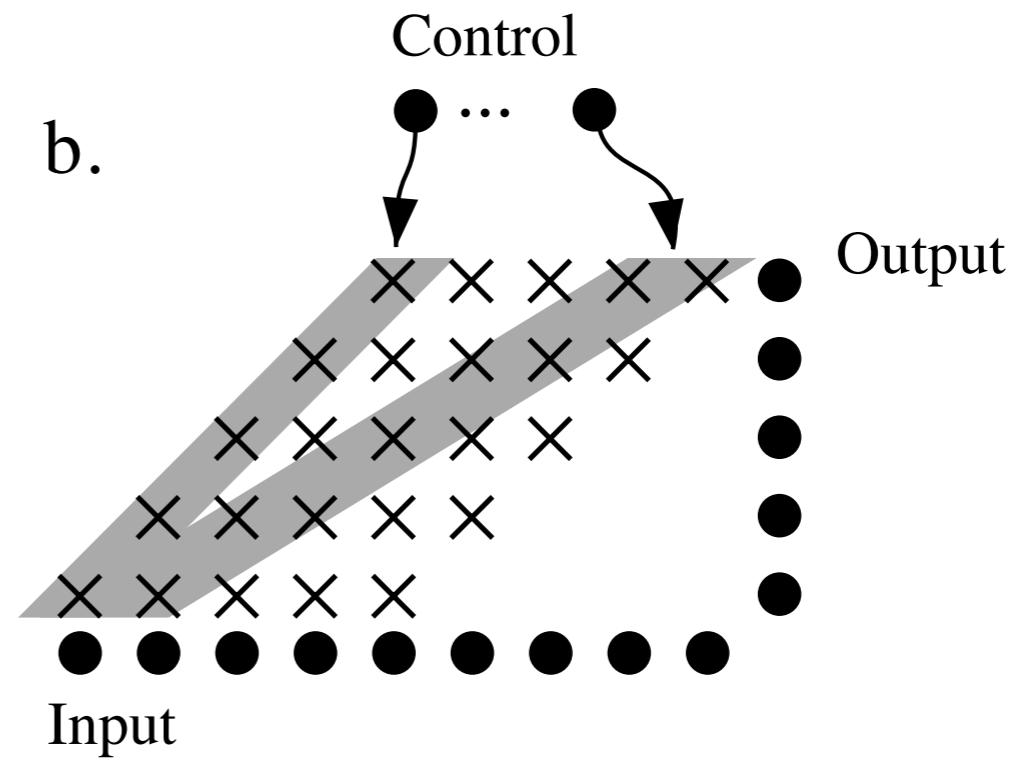
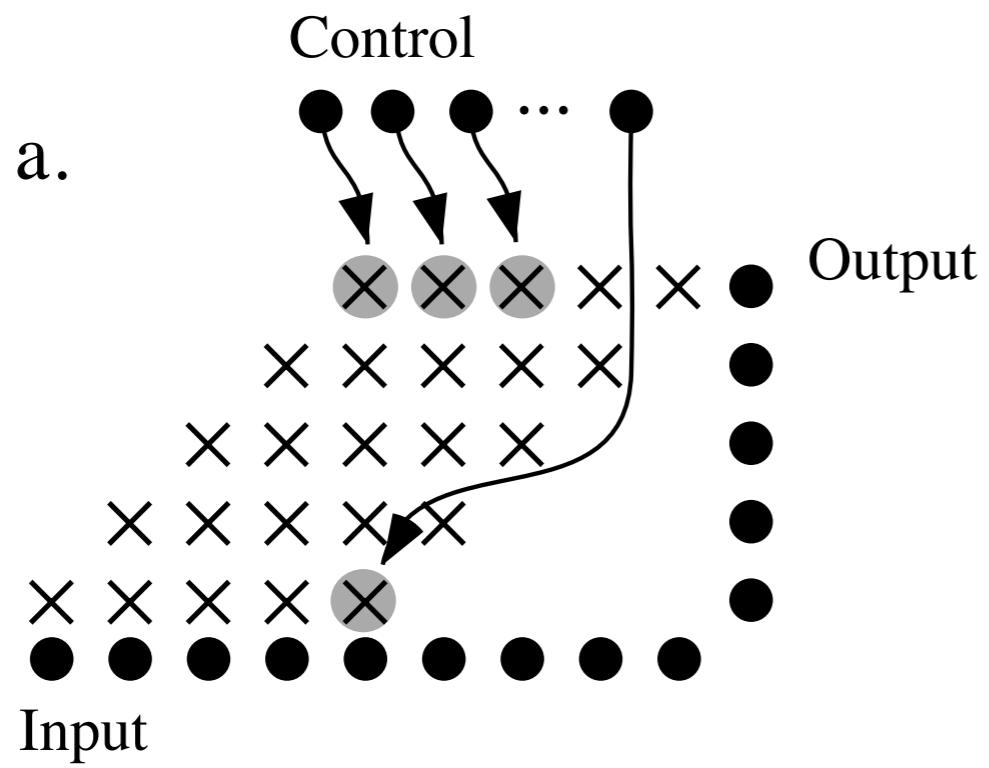


c.

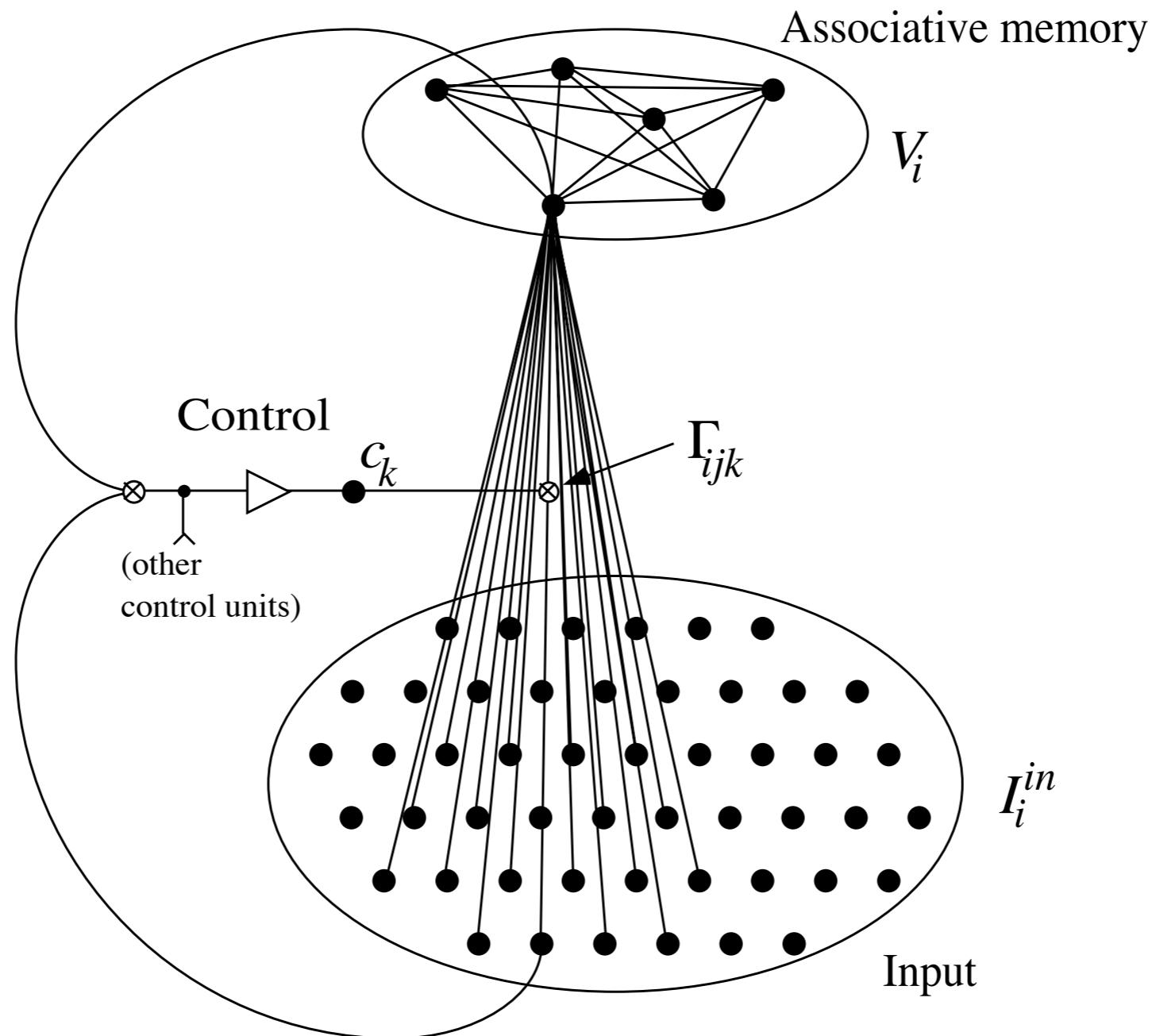


d.

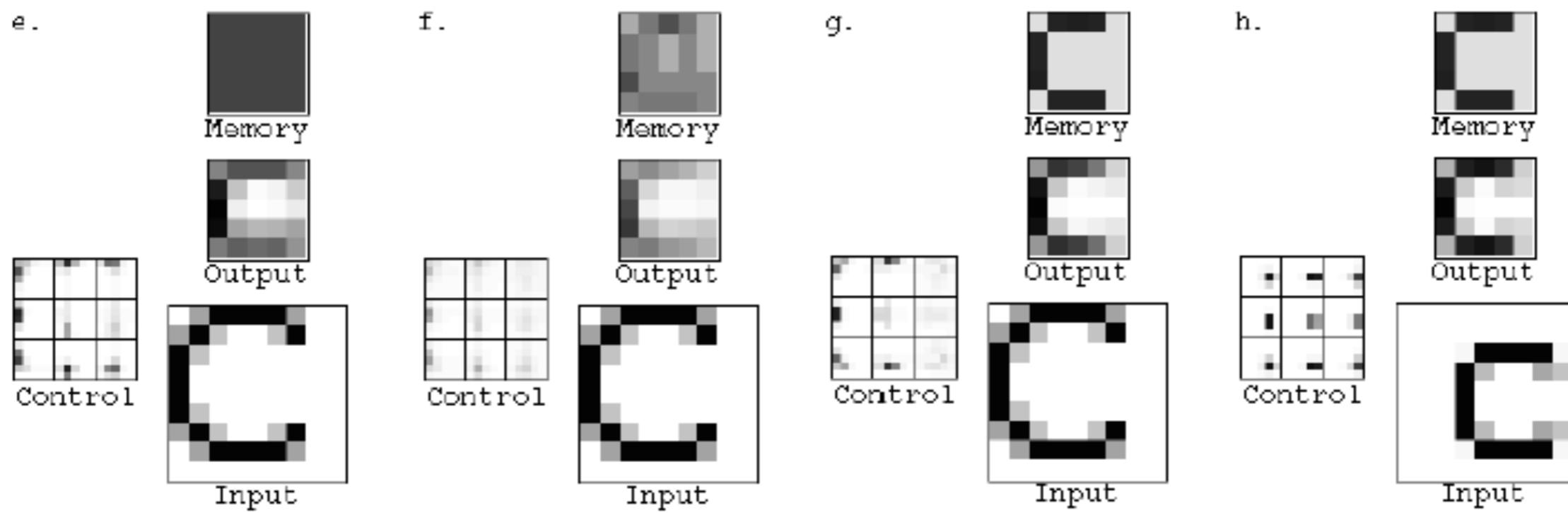
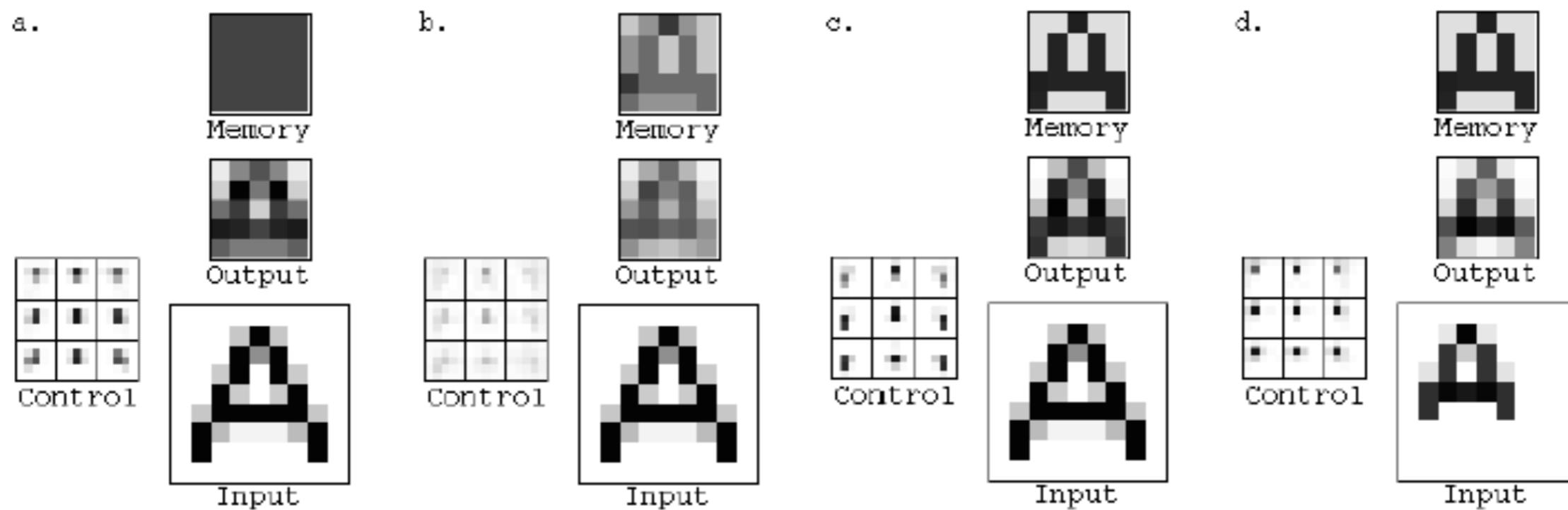




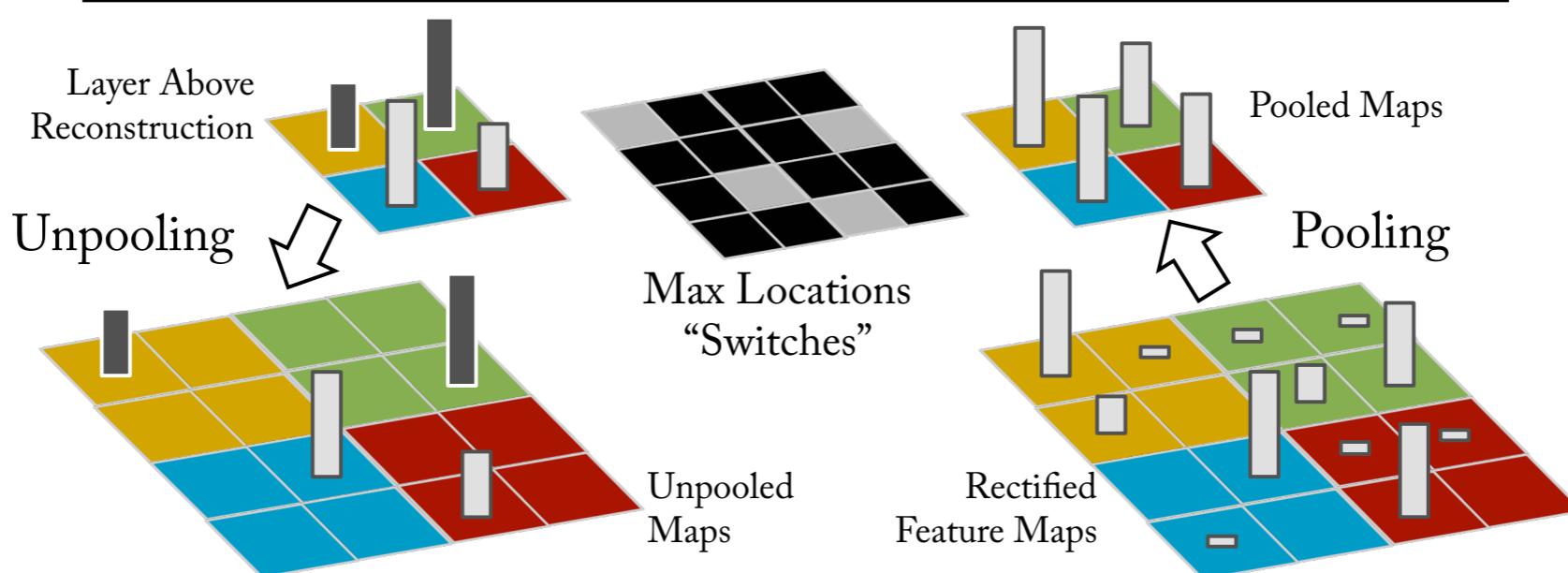
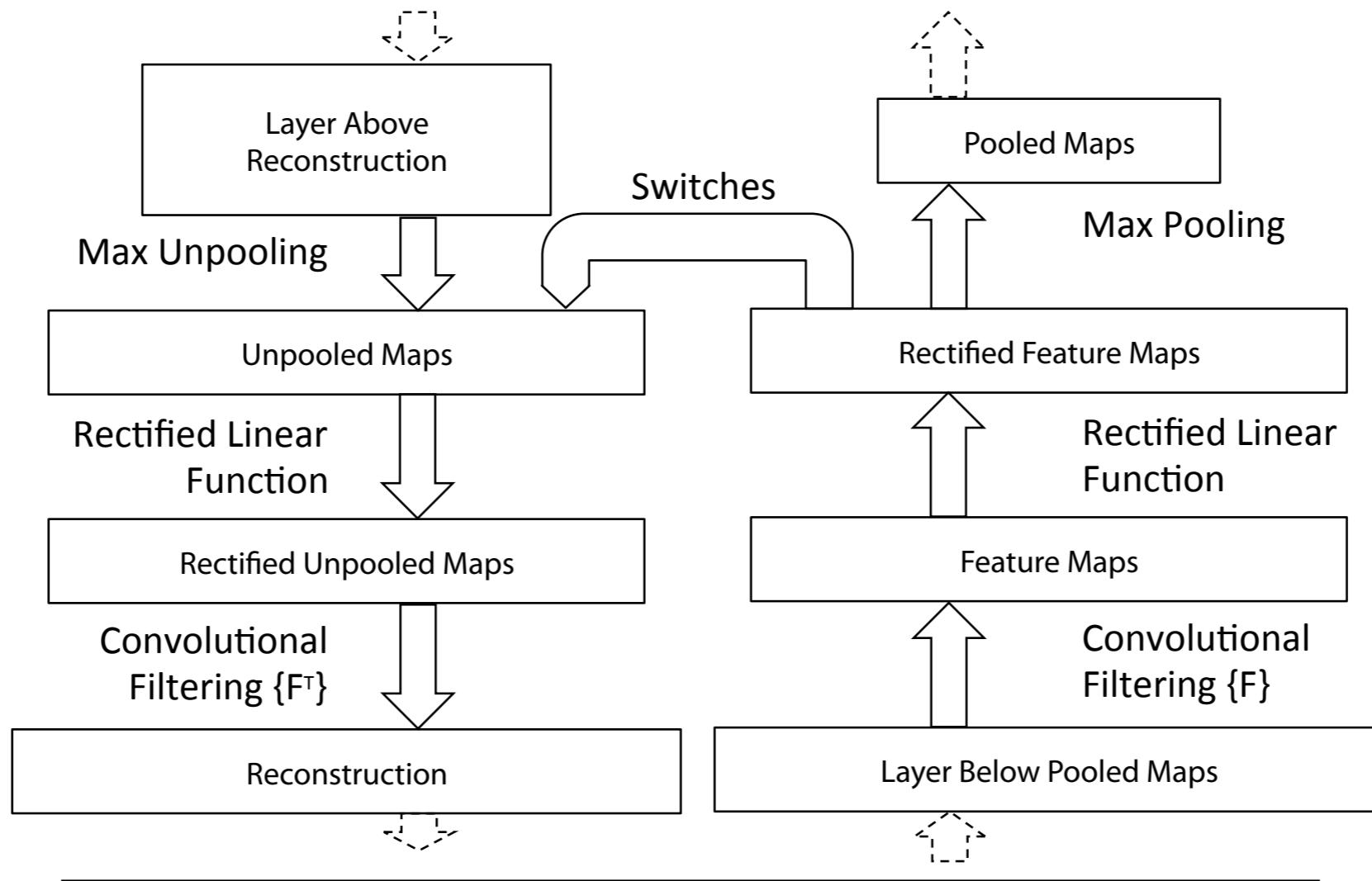
Pattern matching via dynamic routing



Pattern matching via dynamic routing

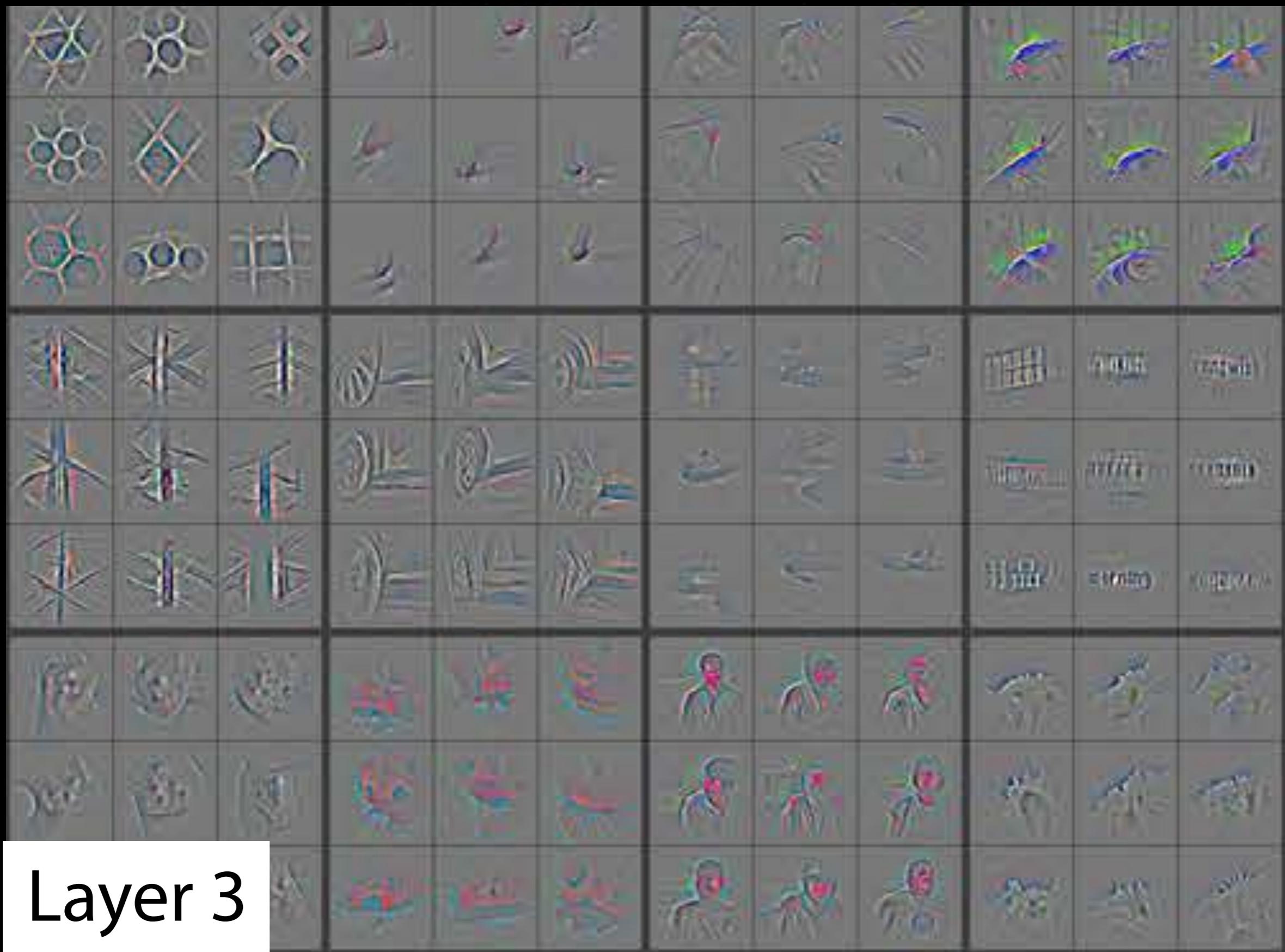


Dynamic routing in deep networks



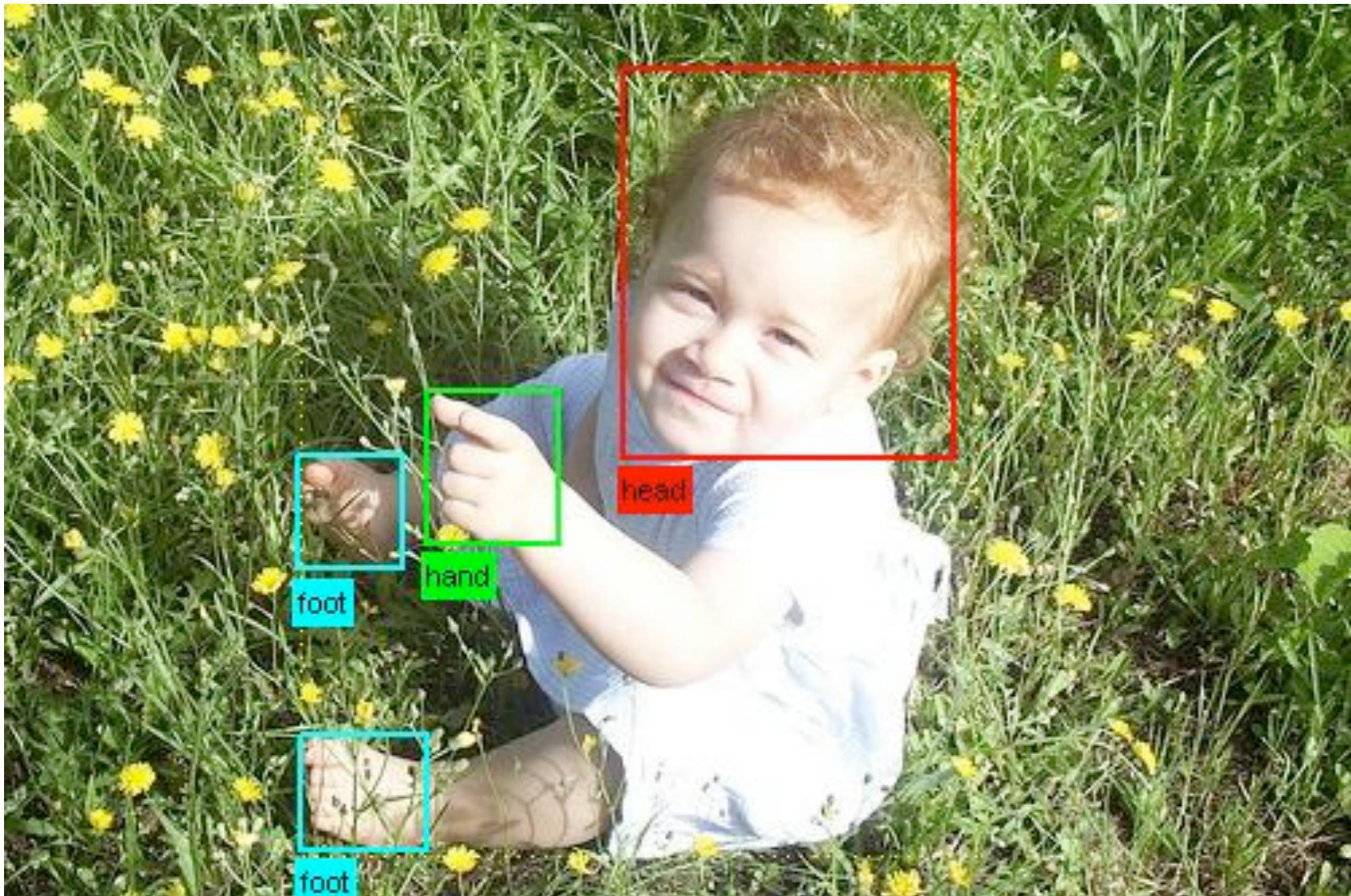
(Zeiler & Fergus, 2013)

Visualization of filters learned at intermediate layers (Zeiler & Fergus 2013)



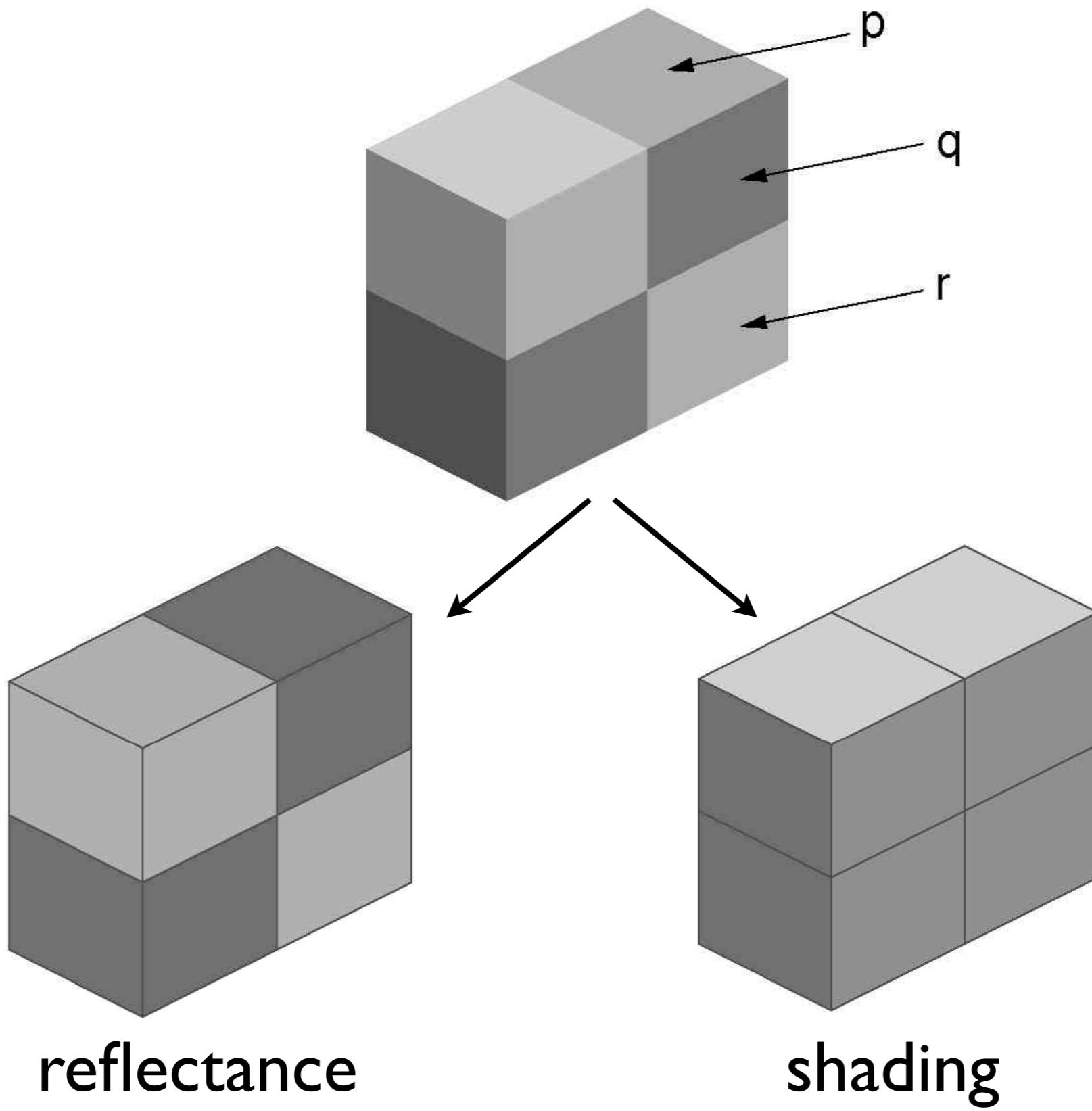
Perception as inference

Is this the goal of perception?

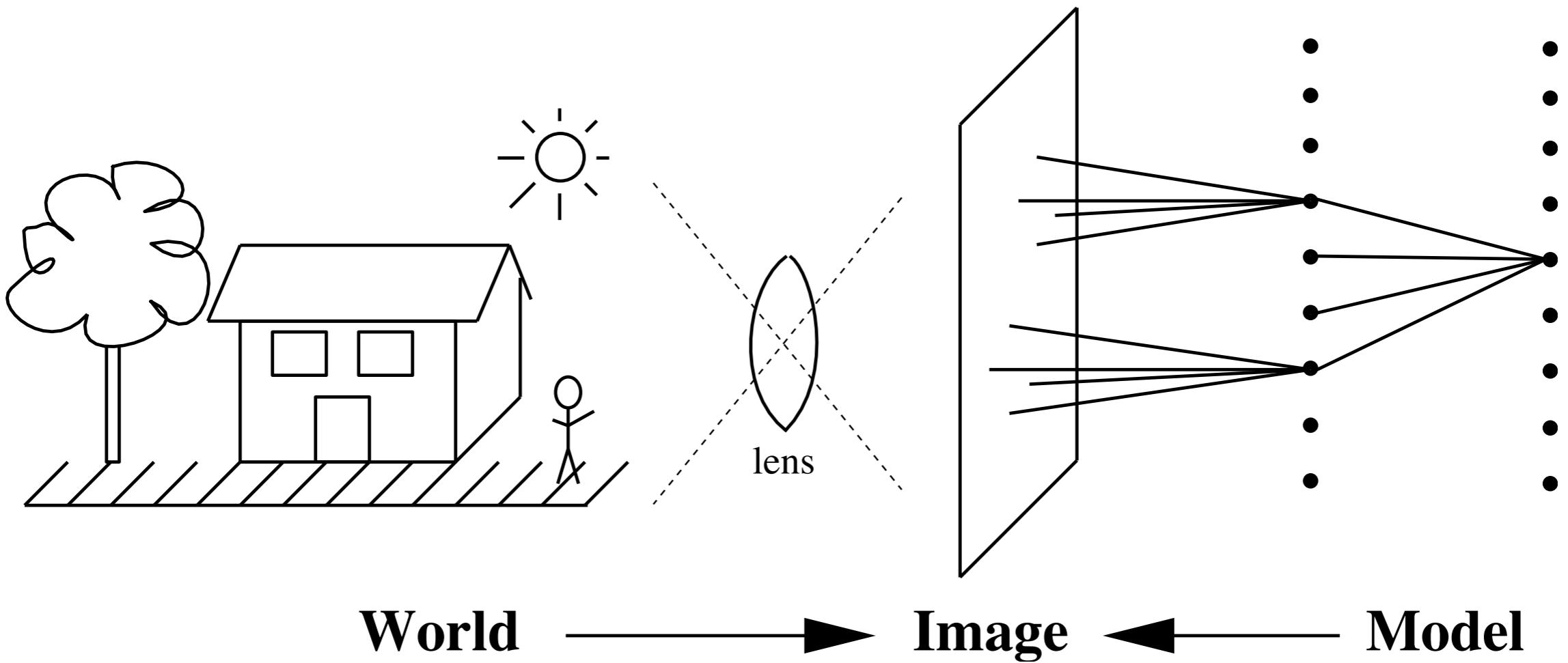




What do these edges mean?

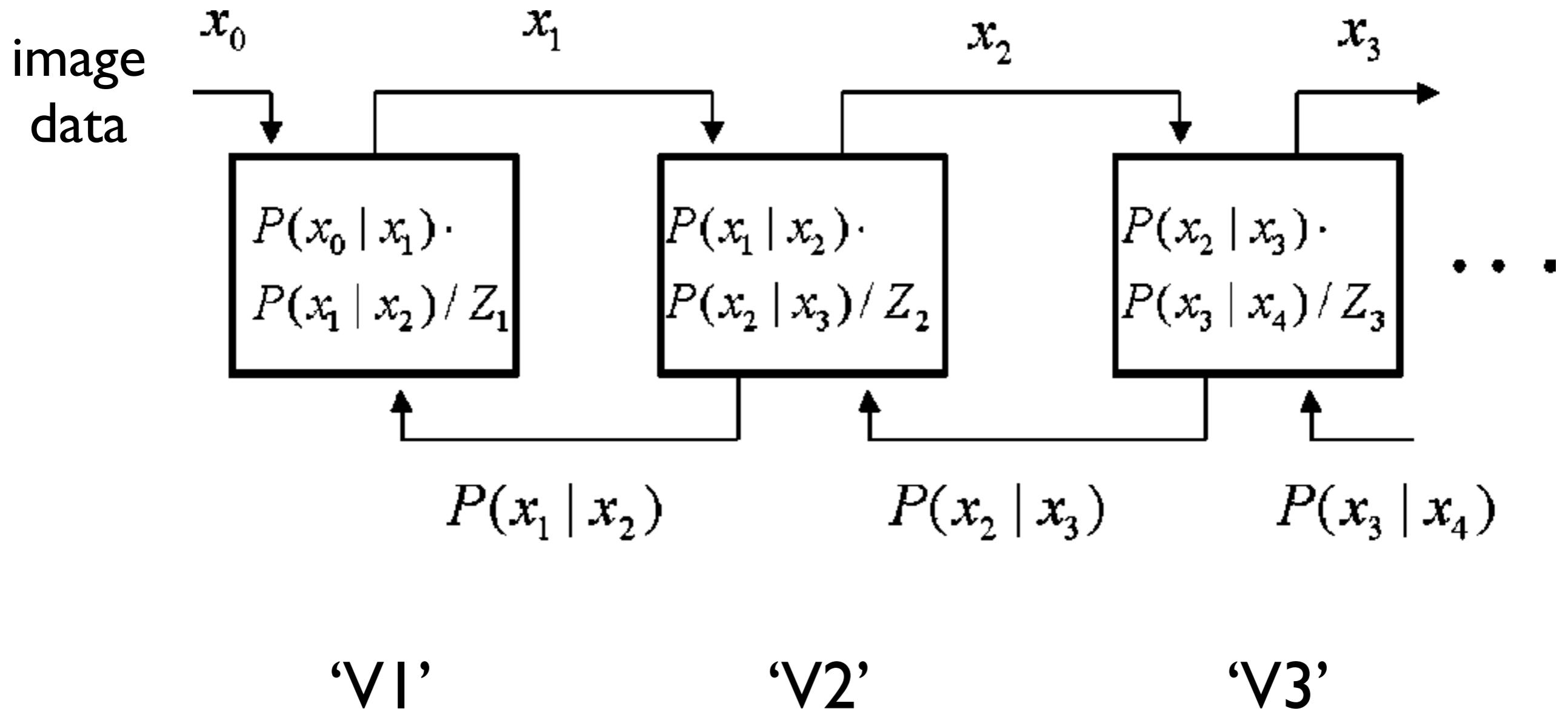


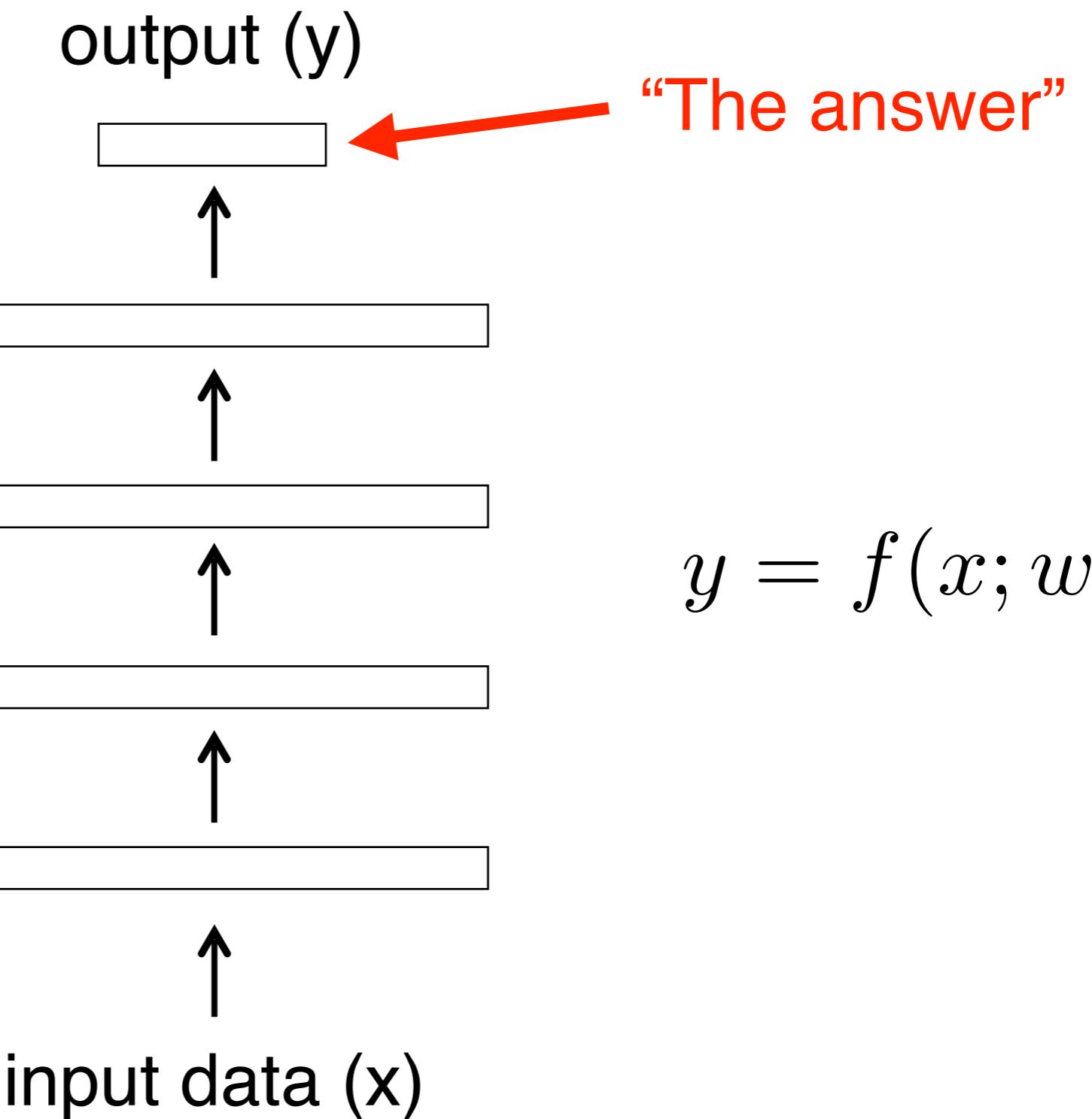
Vision as inference

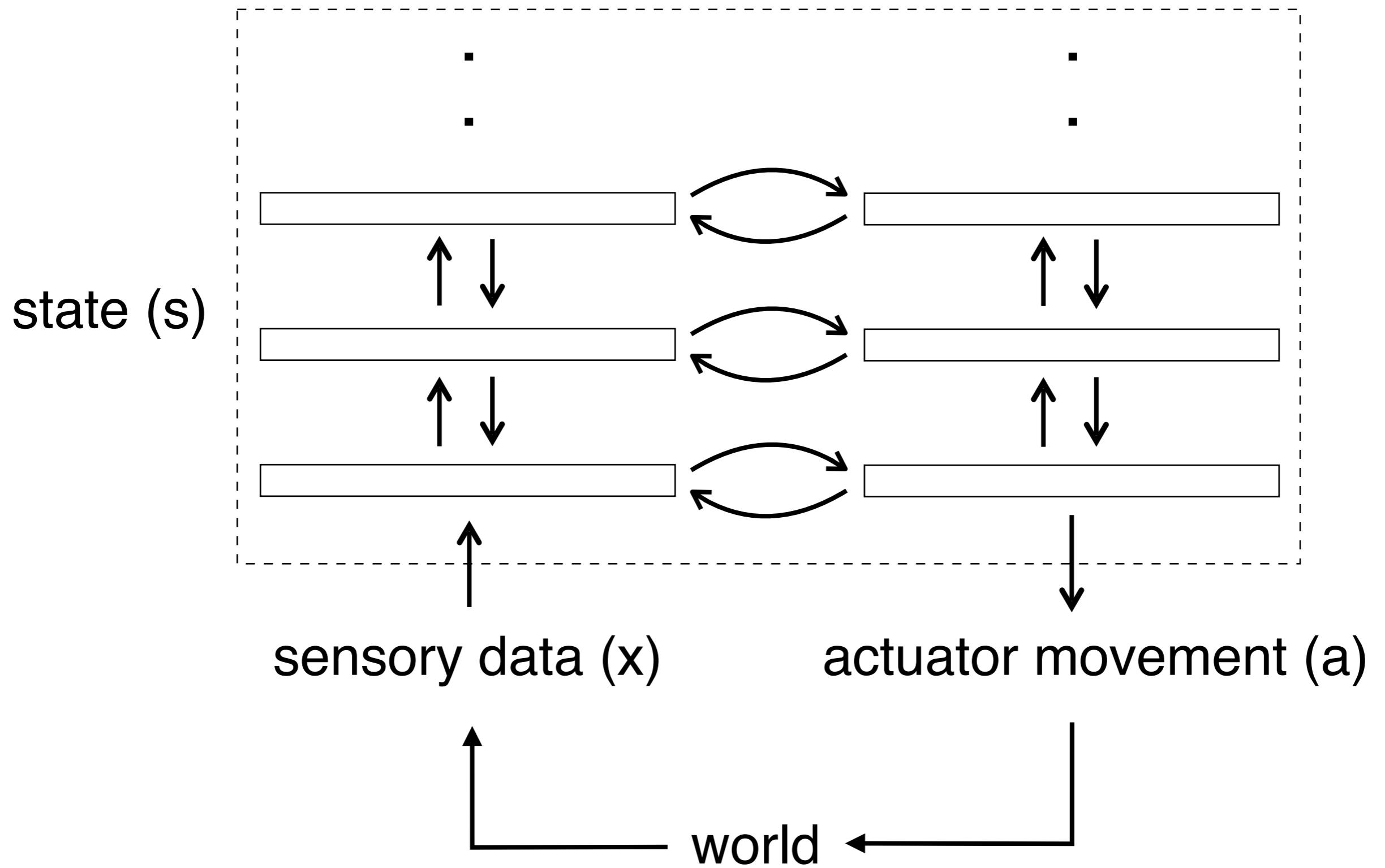


Hierarchical Bayesian inference in visual cortex

(Lee & Mumford, 2003)







Main points

- Multilayer perceptrons were a good idea in 1960's
- Neocognitron was a good idea in 1980's
- The way forward:
 - identify the right problems to be solved
 - exploit the computational richness offered by real neurons and cortical circuits
- Two examples:
 - Dynamic routing
 - Hierarchical Bayesian inference