TSBB15
Computer Vision

Lecture 9
Biological Vision Systems

Per-Erik Forssén

Two parts

1. Systems perspective

2. Visual perception

Two parts

1. Systems perspective

   Based on Michael Land's and Dan-Eric Nilsson's work

2. Visual perception

   Based on Slides from Gösta Granlund

Vision Systems

Camera vs. eye

Vision Systems

≠

Purpose:

Reproduce the world
as accurately as possible

Purpose:

Sensing device for
visual behaviours

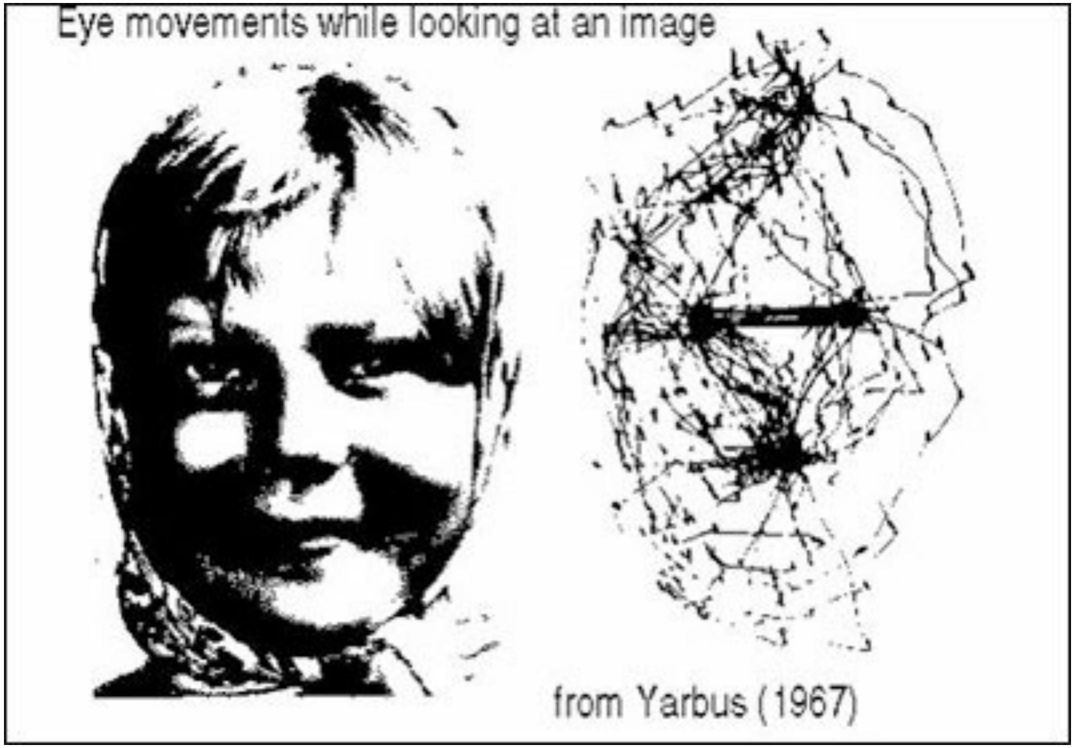
Vision Systems

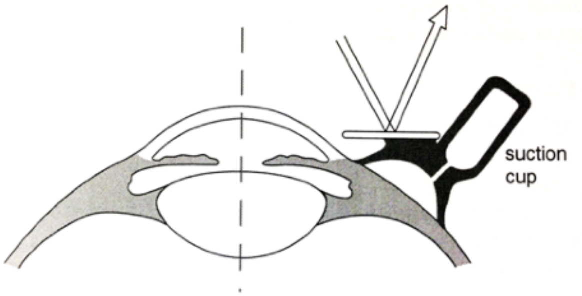
What a camera sees

Vision Systems

What the human eye sees

Vision Systems





Device used by Yarbus 1965
Illustration from:
M. F. Land ”Looking and Acting”

Vision Systems

≠

Uniform resolution
Smooth motion

just central 2° are sharp
saccadic motions (avg. 3Hz, around 700°/s)

Vision Systems

Peripheral view

Foveal view

What a robot sees

Karlsruhe robot ARMAR-III

Visual Behaviours

Saccadic motion is an example of a visual behaviour

Purpose?

Visual Behaviours

Other examples of visual behaviours:

1. **Fixate moving targets**
2. **Compensate for head and body movement**

3. **Change detection**
4. **Recognition**

Visual Behaviours

Other examples of visual behaviours:

1. **Fixate moving targets** - Optoknetic Reflex (OKR)
2. **Compensate for head and body movement** - Vestibulo - Ocular Reflex (VOR)
3. **Change detection**
4. **Recognition**

Visual Behaviours

Experiment:
Hold out your hand and raise a finger:

1. turn head while looking at finger (VOR)
2. move hand while looking at finger (OKR)

Which reflex is faster?

Visual Behaviours

Visual input for VOR (stabilization)?

Visual input for OKR (tracking)?

Visual Behaviours

Visual input for VOR (stabilization)?

- Optical flow (dense over entire visual field)

Visual input for OKR (tracking)?

- Tracking (region around fovea)

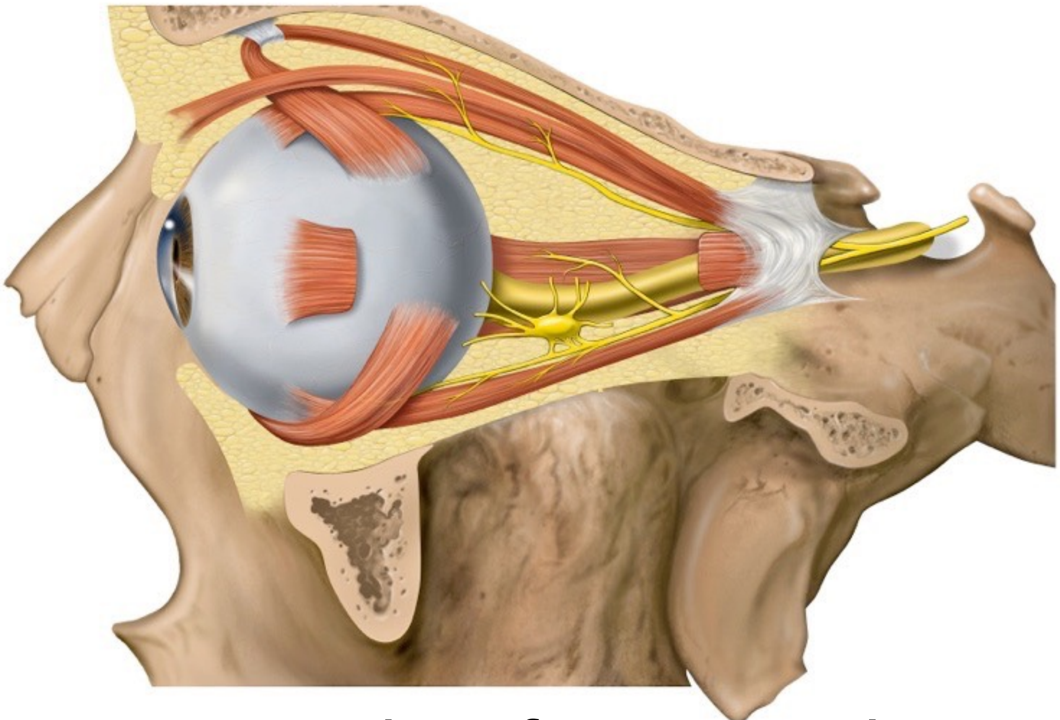
Visual Behaviours

Visual input for VOR (stabilization)?
- Optical flow (dense over entire visual field)

Visual input for OKR (tracking)?
- Tracking (region around fovea)

Note: VOR mainly uses input from the vestibular system. Optical flow is used as an error signal, for learning/adaptivity.

Visual Behaviours



- Three opponent pairs of eye muscles
- Whole neck-eye system is involved in gaze control

VCR in Weka bird

Whole head has to move in birds - Vestibulo-Collic Reflex

Weka VCR - YouTube

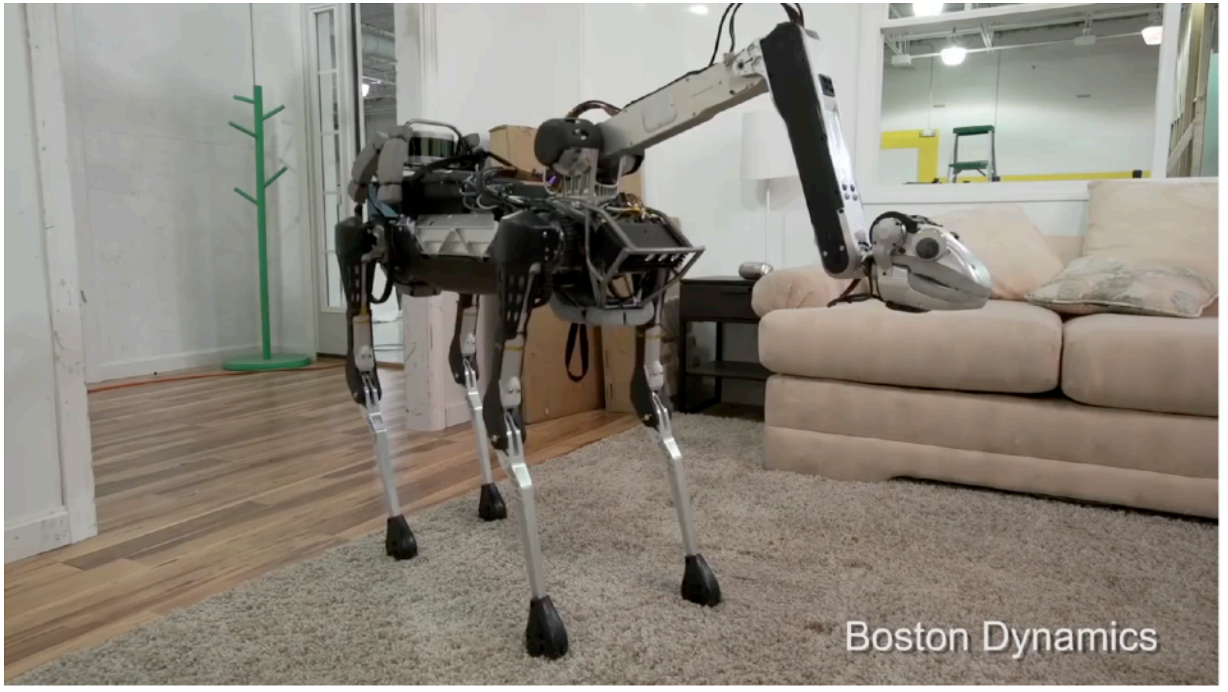
VCR in Chicken

Whole head has to move in birds - Vestibulo-Collic Reflex

Chicken VCR - YouTube

VCR on Robot

Boston Dynamics version of VCR



Boston Dynamics - YouTube

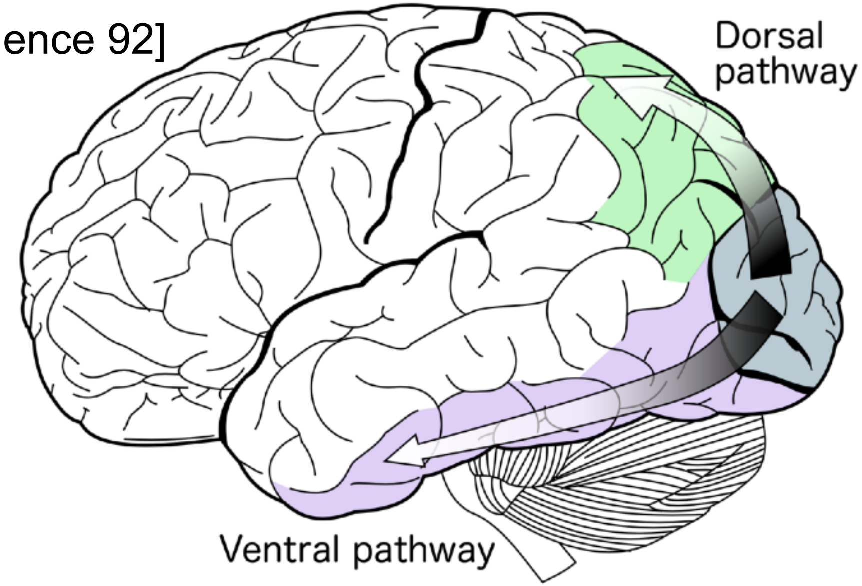
Visual Behaviours

Examples of visual behaviours:

1. **Fixate moving targets** - OKR
2. **Compensate for head and body movement** - VOR,VCR
3. **Change detection** - 1&2 + time difference
4. **Recognition** - Saccadic motions + 1&2 + Perceptual hierarchy

Visual Perception

- How and what separation
  [Godale & Milner, Trends Neuroscience 92]



- Dorsal pathway
  controls gaze
  and action
- Ventral pathway
  handles visual recognition

Complex problem

Recognition using direct matching to prototype images is untenable:

- Large number of objects
- Large number of variations

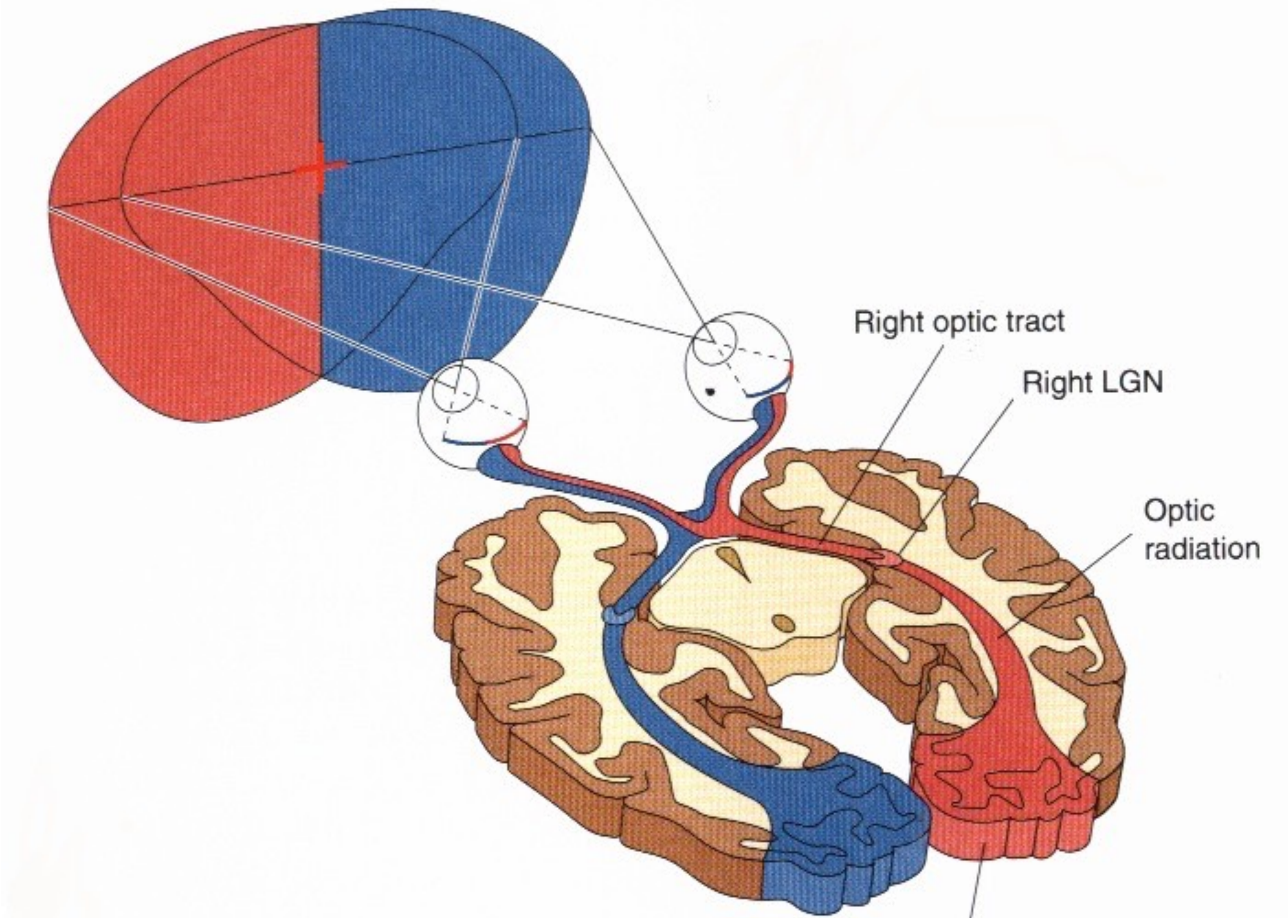
Complex problem

Recognition using direct matching to prototype images is untenable:

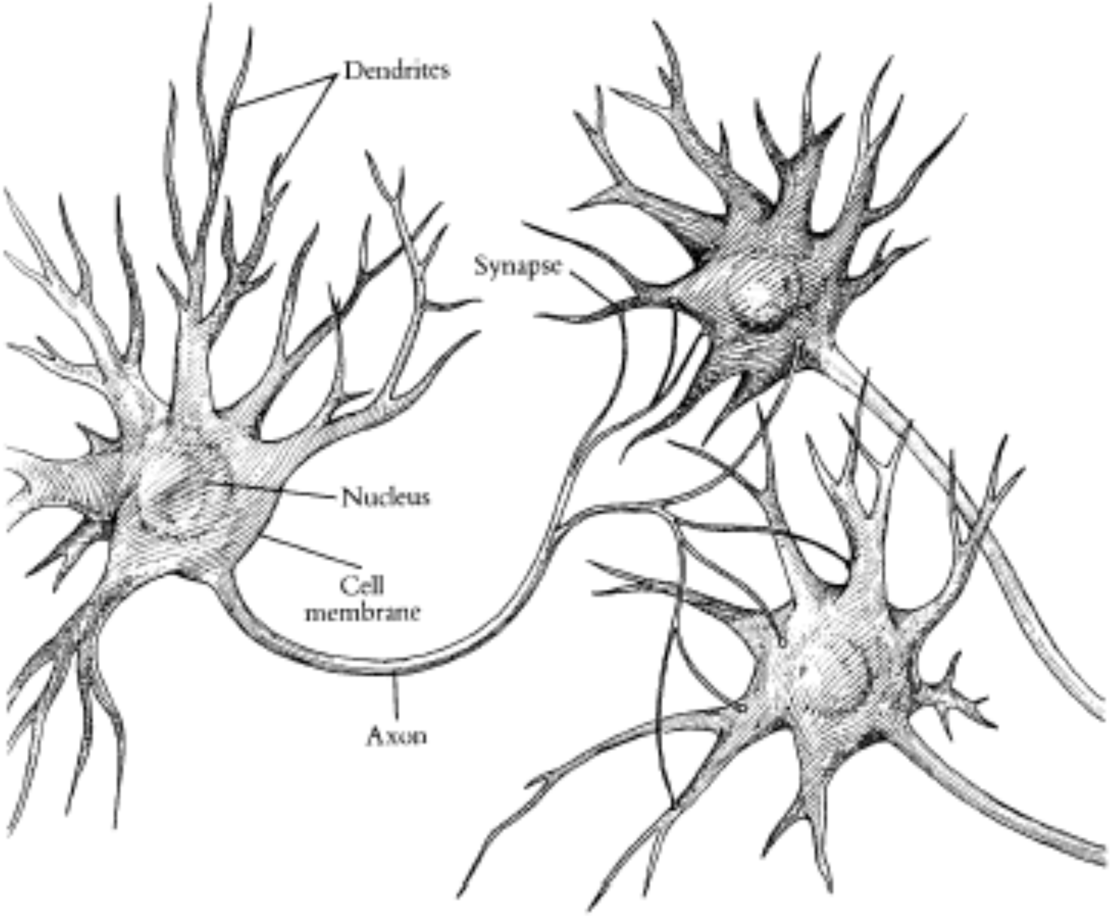
- Large number of objects
- Large number of variations

Abstraction is necessary!

The visual pathway



Principal parts of a nerve cell



Signals of neurons

- Carried through a chemical process
- Resting potential -70 mV inside axon
- Reversal to +40 mV inside axon
- Refractory time about 1 msek
- A few to > 1000 impulses per second
- Most neurons use **pulse frequency coding**
- A few types have **graded** signals

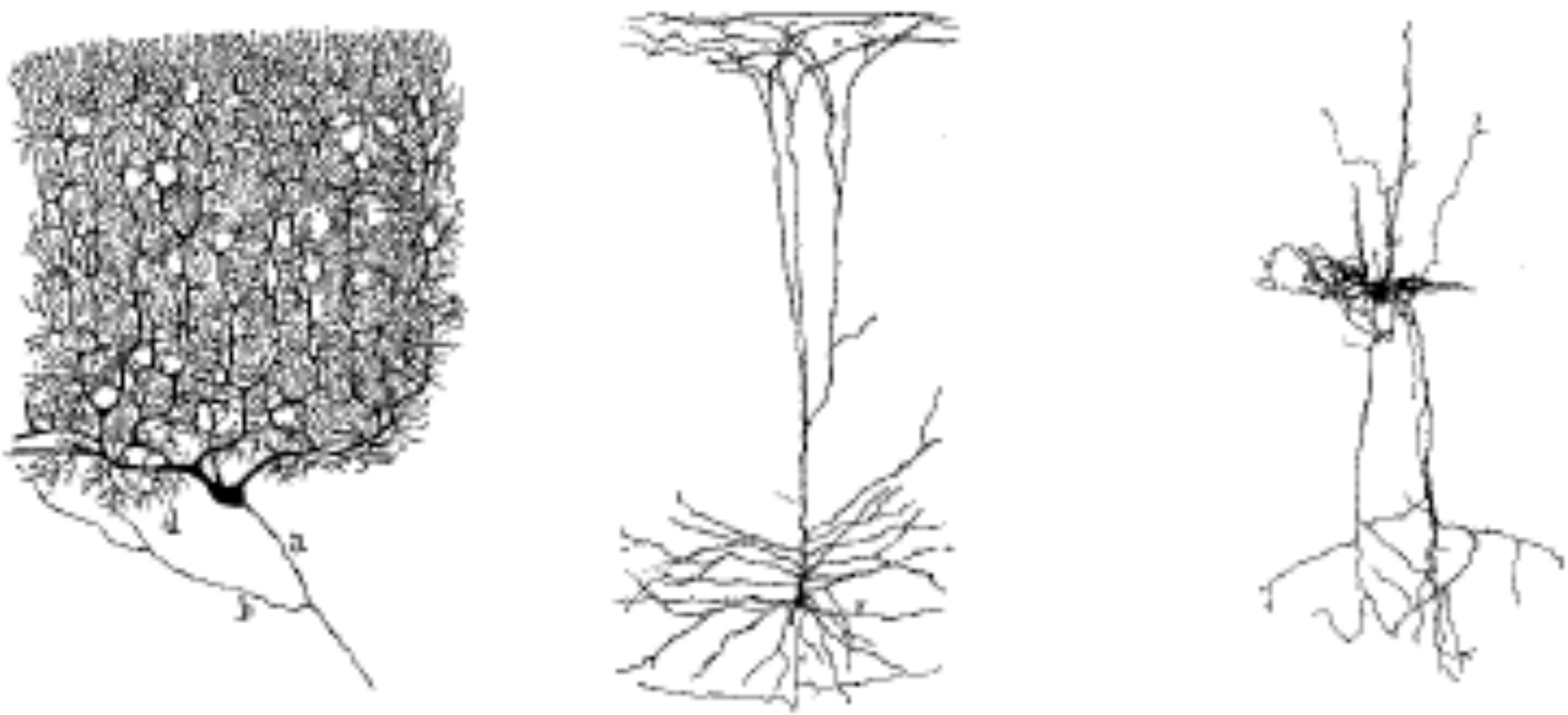
Neurons

Axons can be < 1 mm to > 1 m

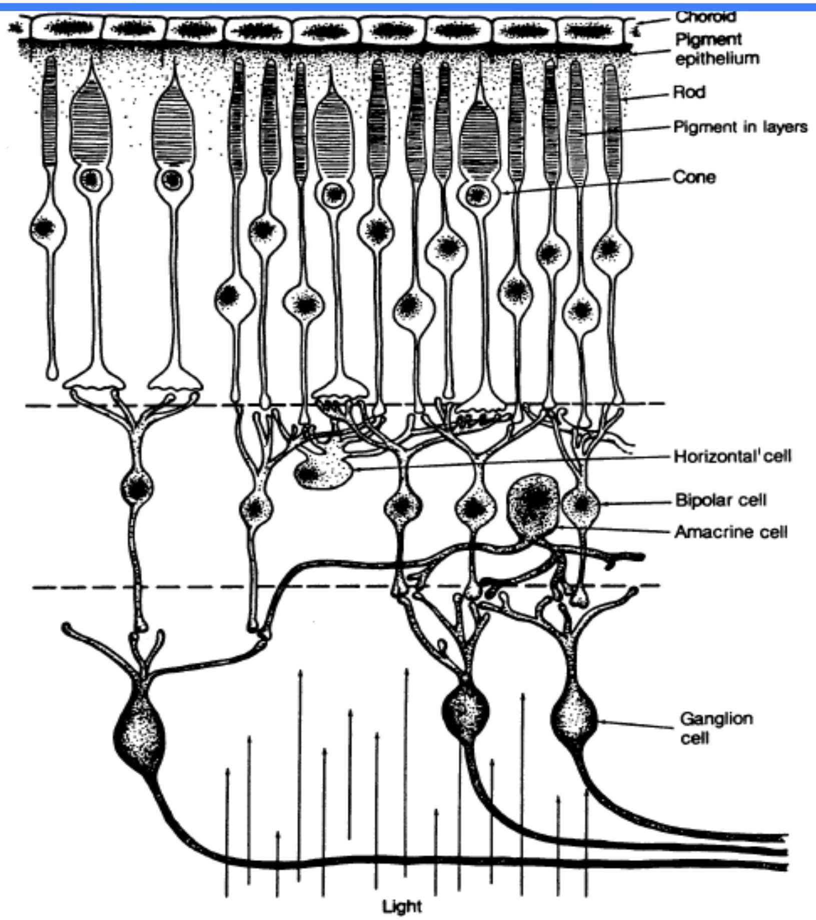
Synapses can be excitatory or inhibitory

50 – 100 neurotransmitters

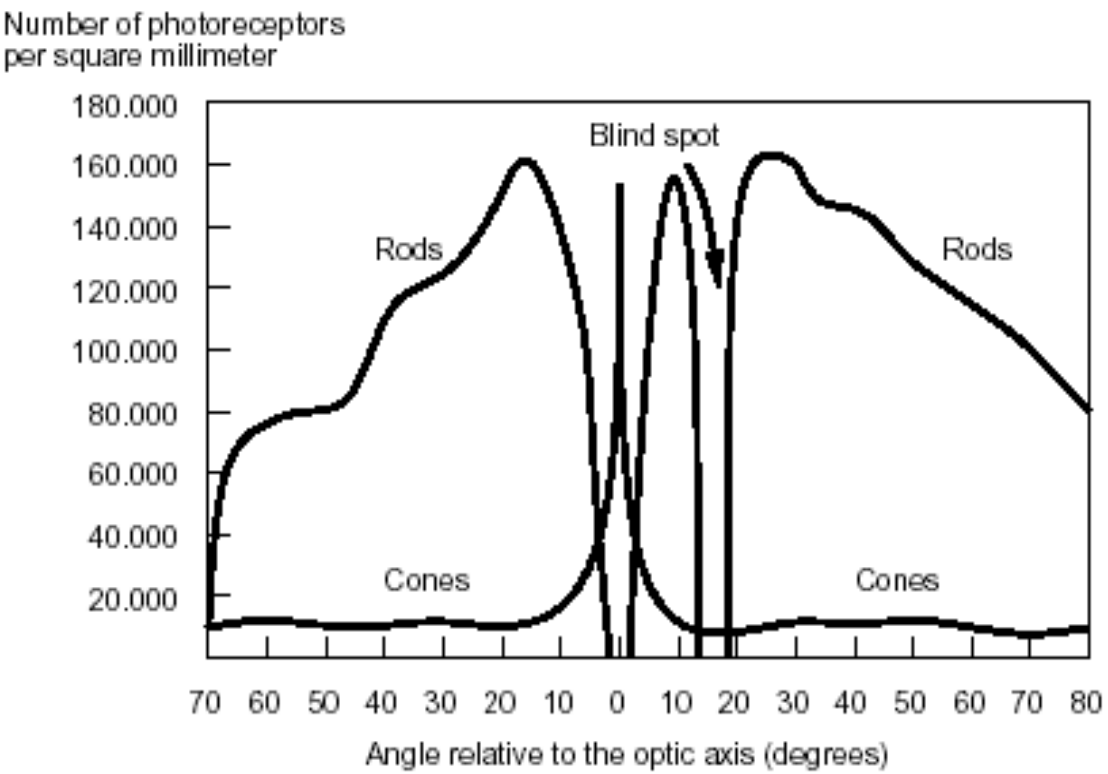
> 100 different types of neurons

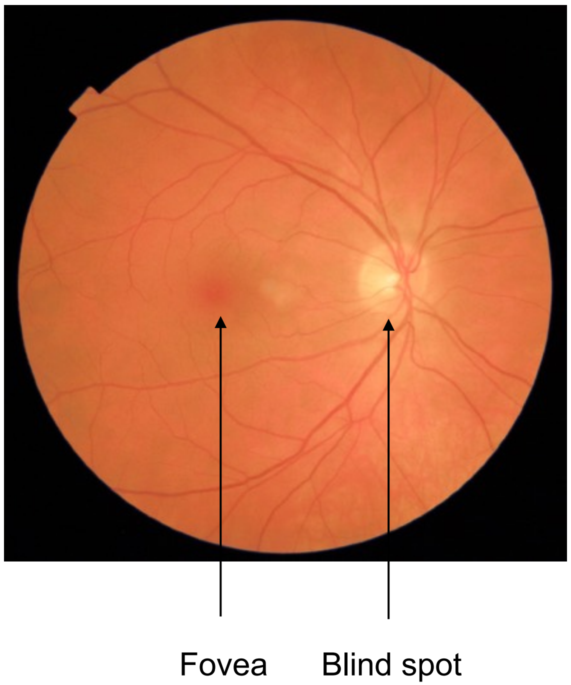


The retina

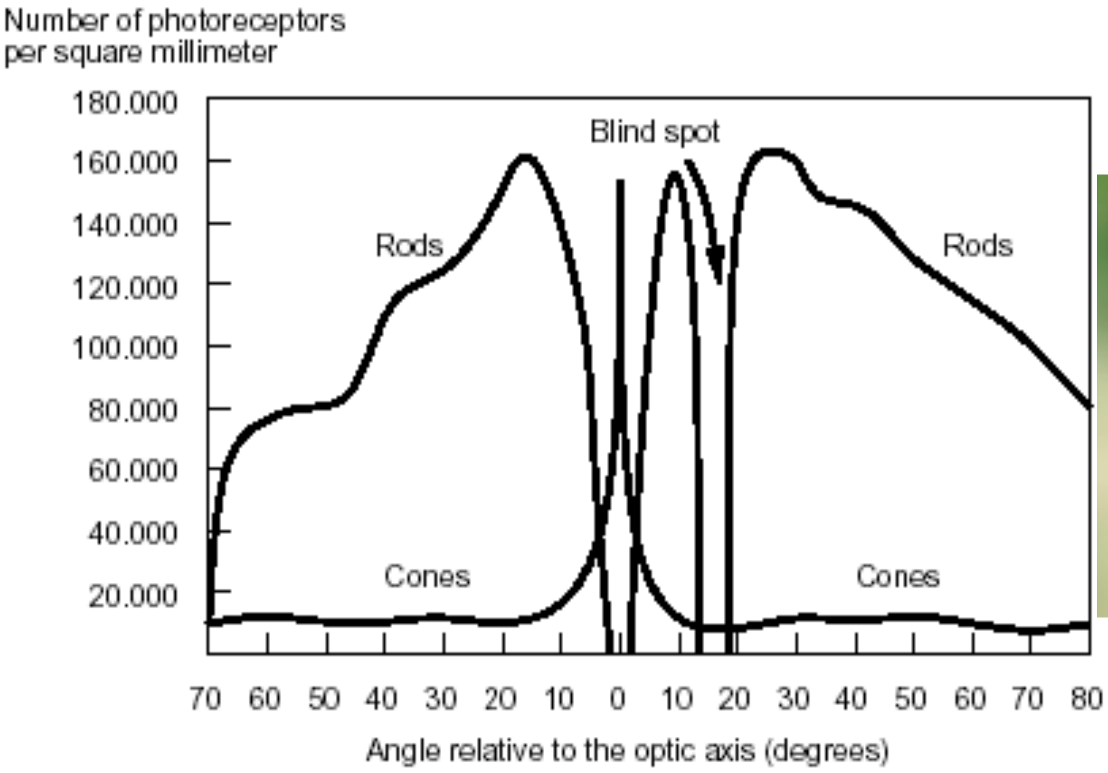

Choroid
Pigment epithelium
Rod
Pigment in layers
Cone
Horizontal cell
Bipolar cell
Amacrine cell
Ganglion cell
Light

Density of photoreceptors

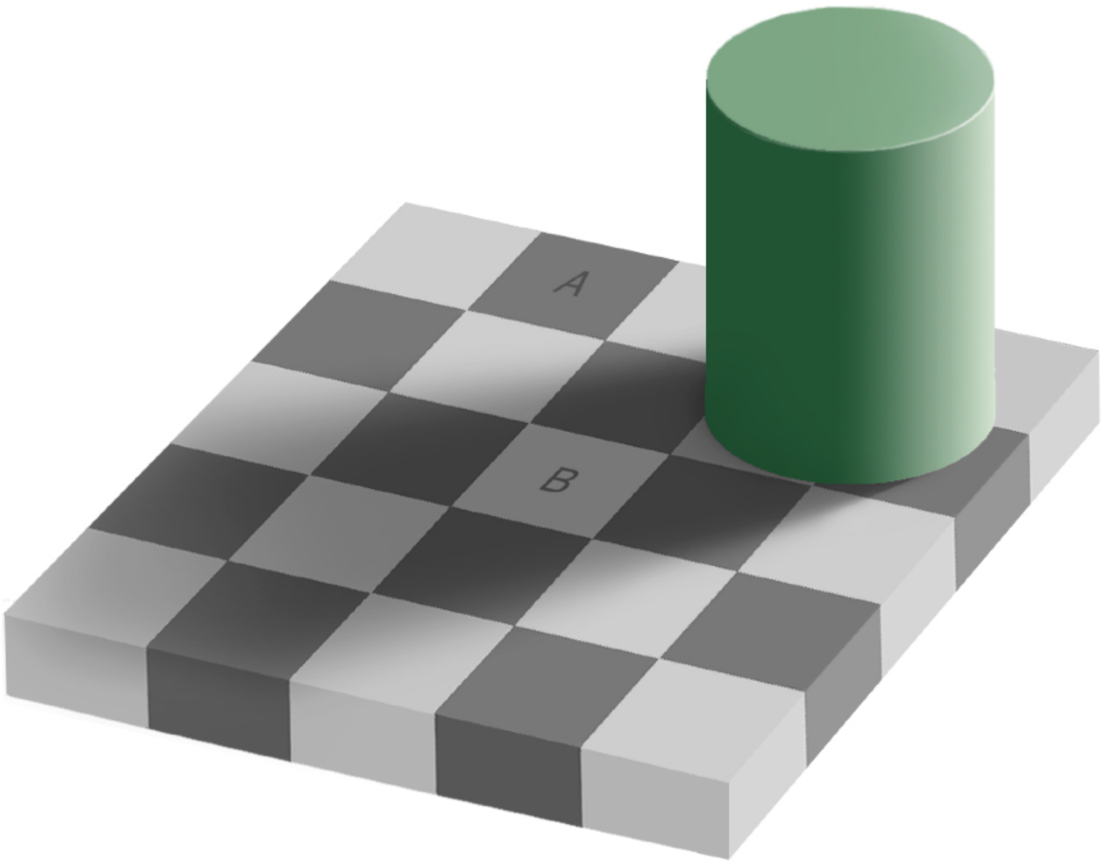




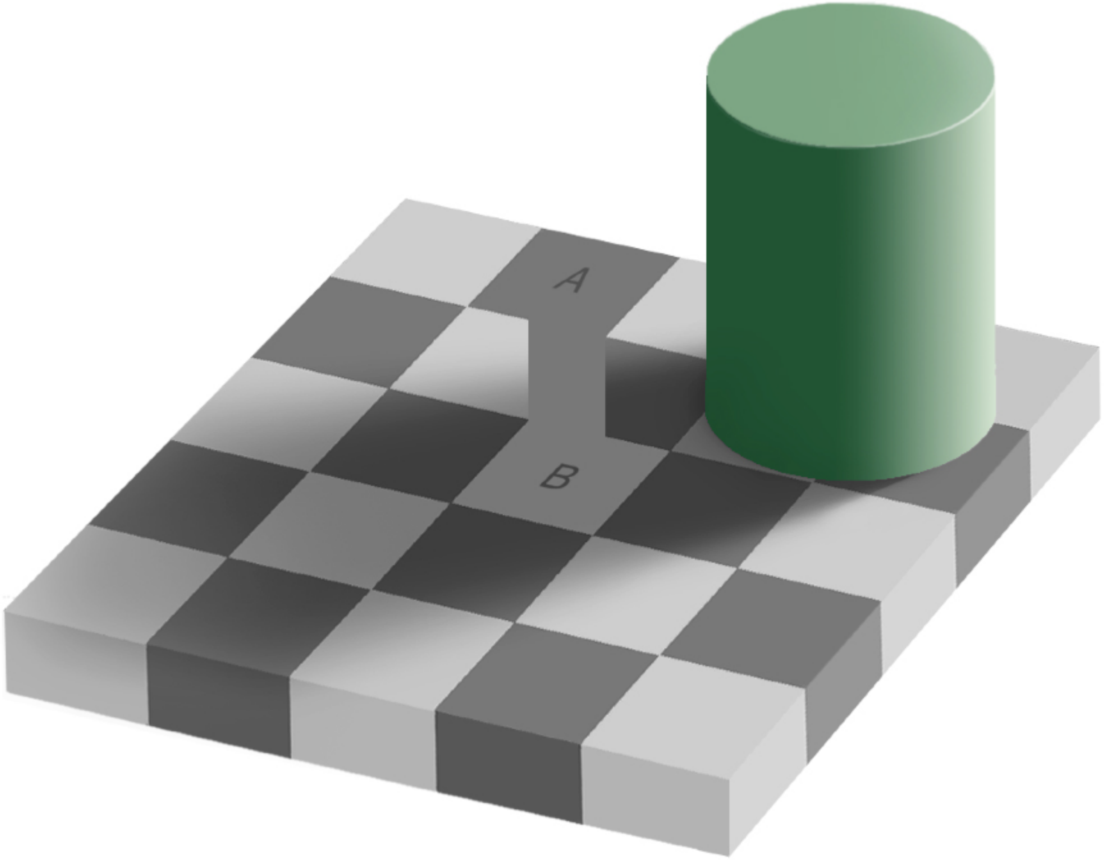
Density of photoreceptors



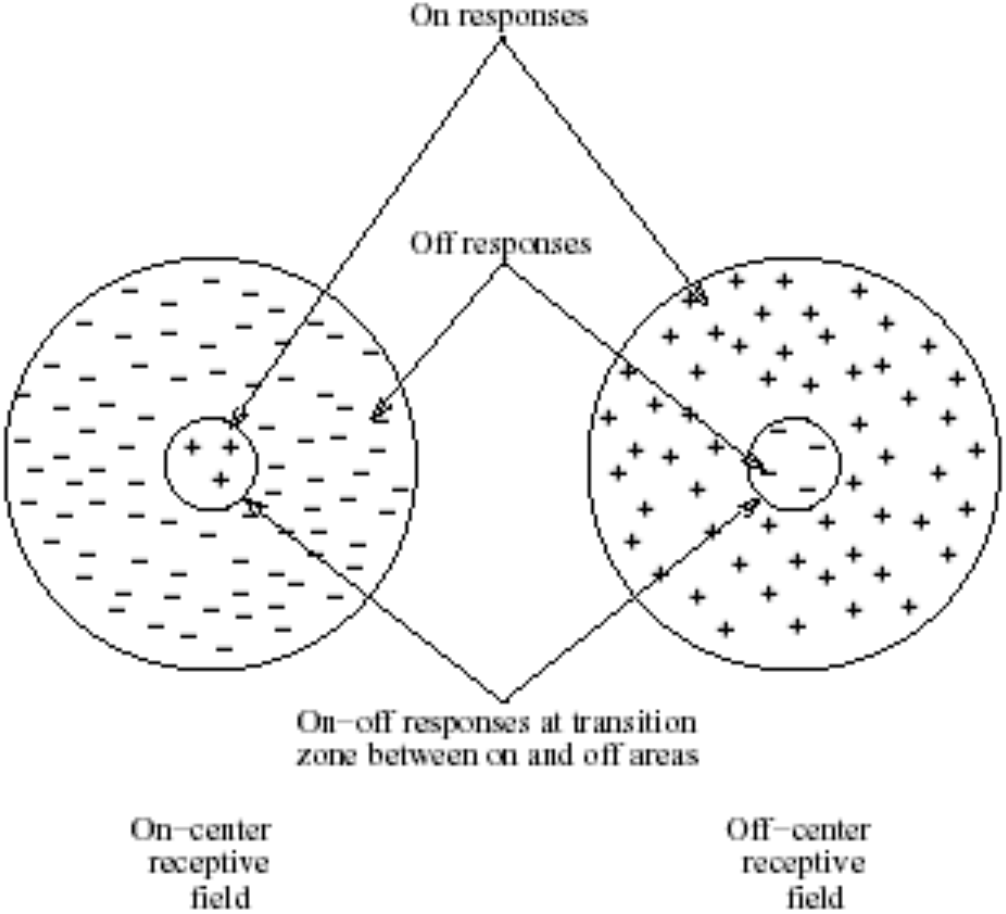
Stability with respect to illumination



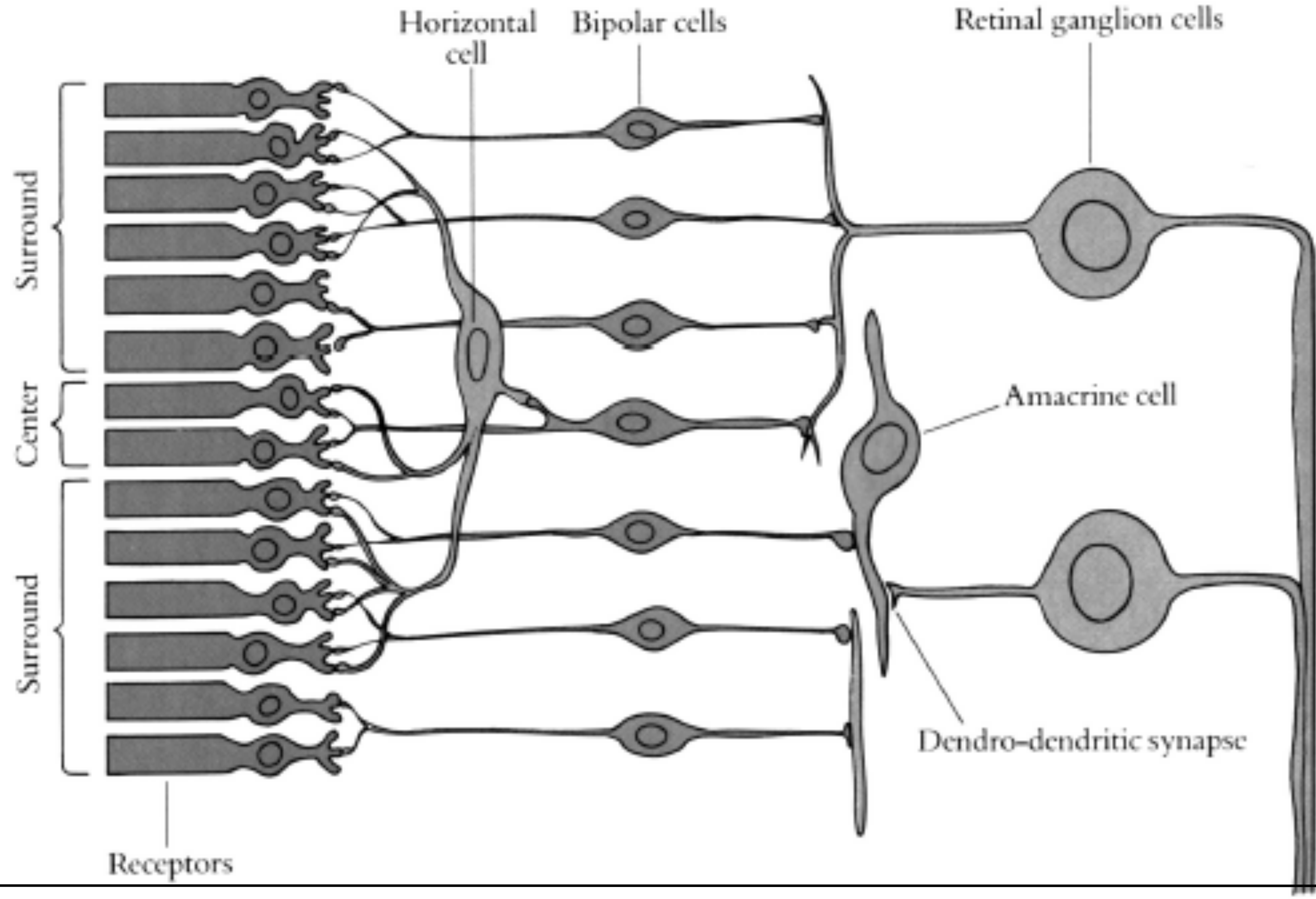
Stability with respect to illumination



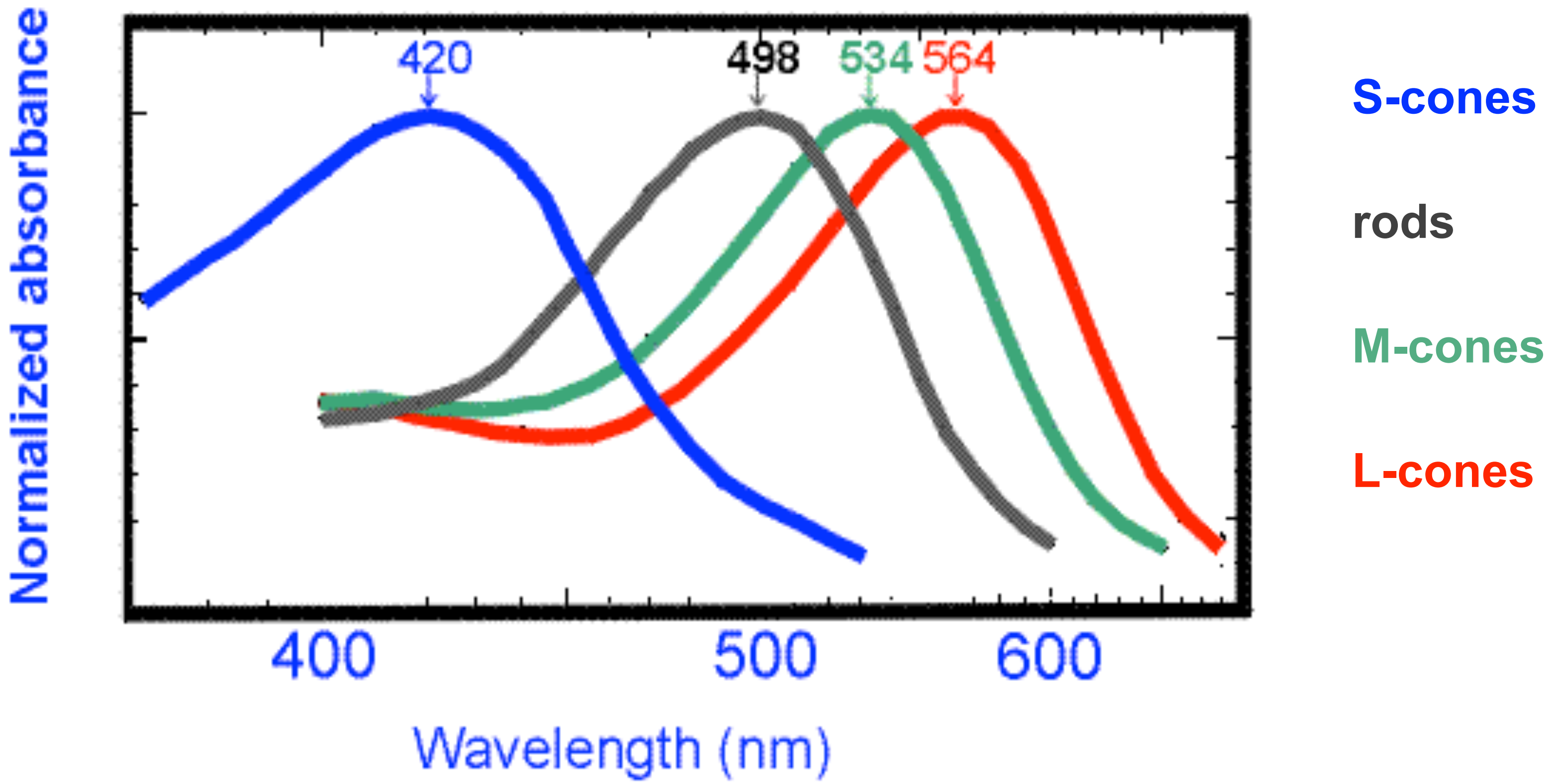
Center-surround receptive fields



Generation of center-surround fields



Colour vision



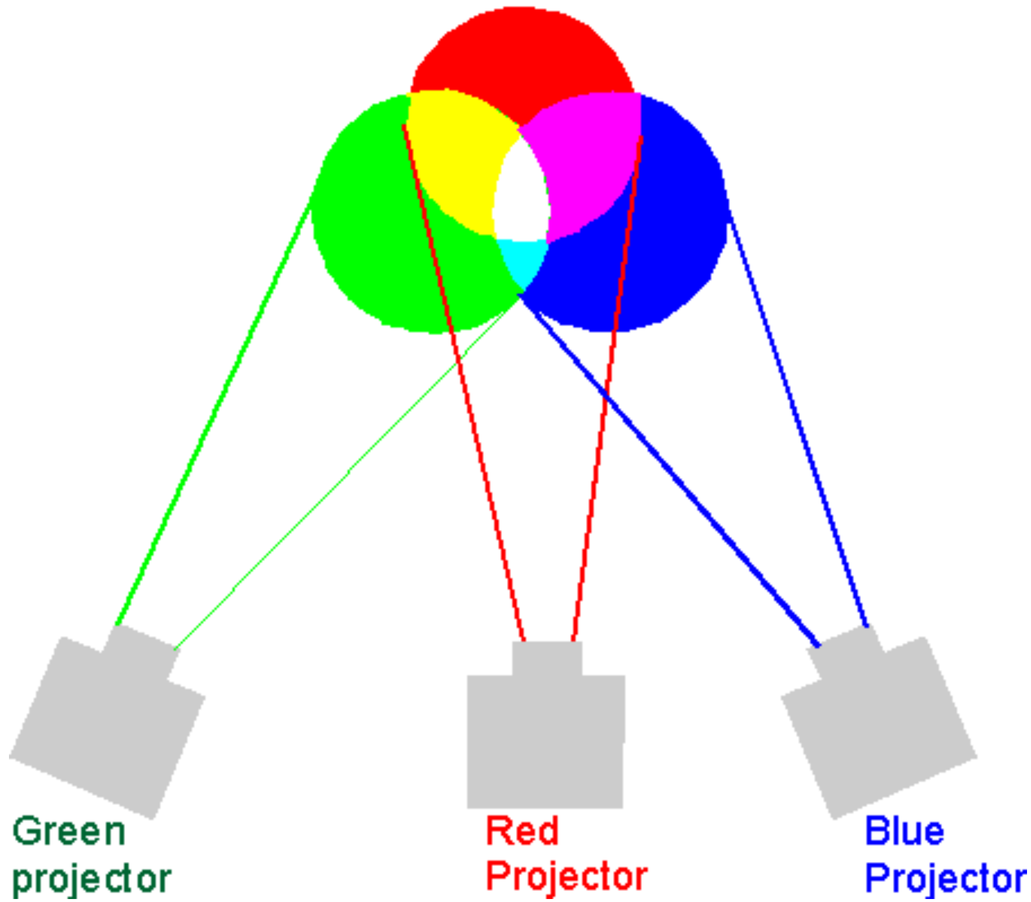
After Bowmaker & Dartnall, 1980

Colour vision theories

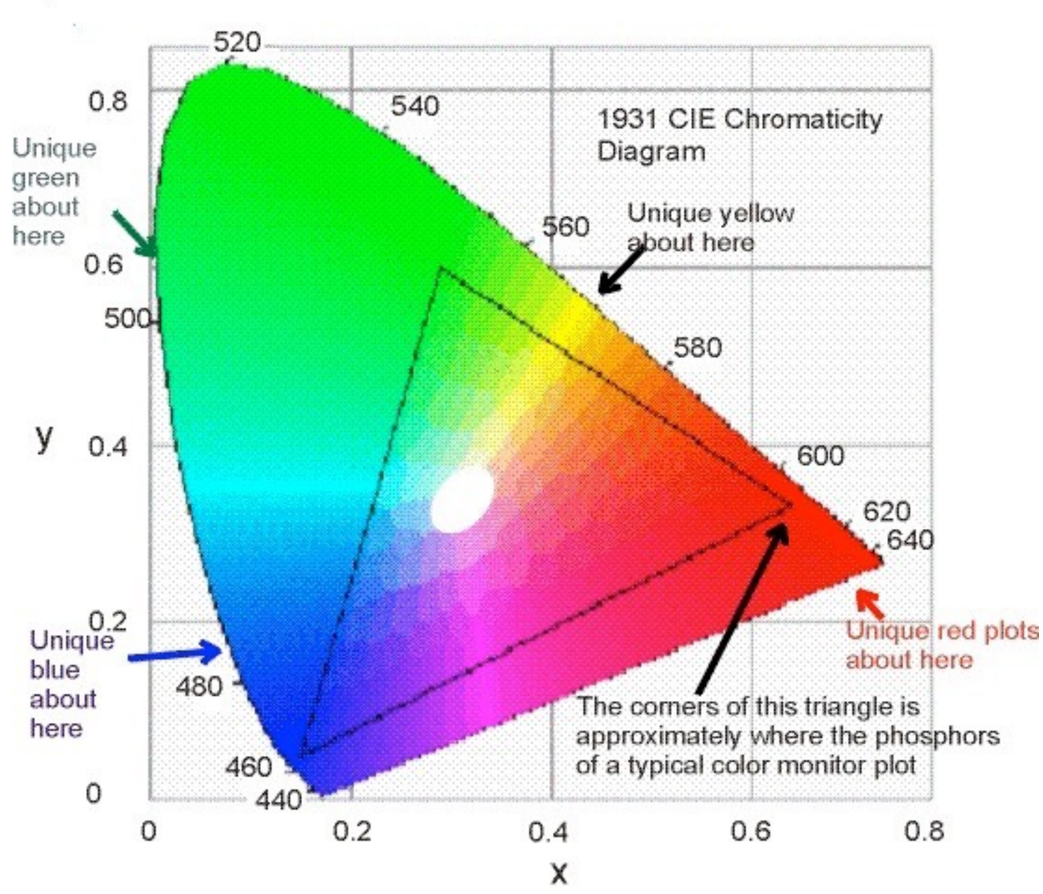
The **trichromatic** theory operates at the receptor level

The **opponent processes** theory applies to the subsequent neural level of colour vision processing

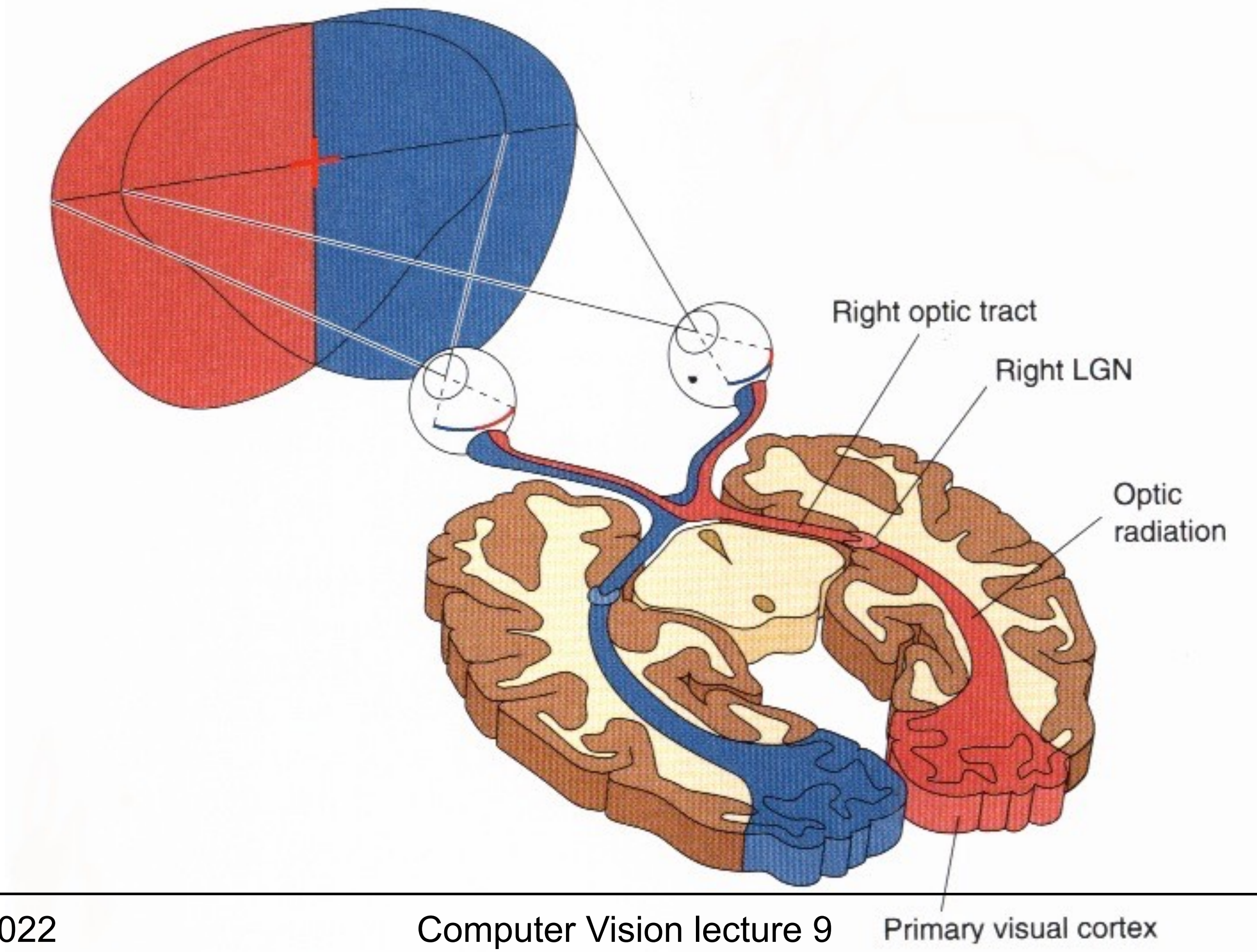
Additive colour mixing



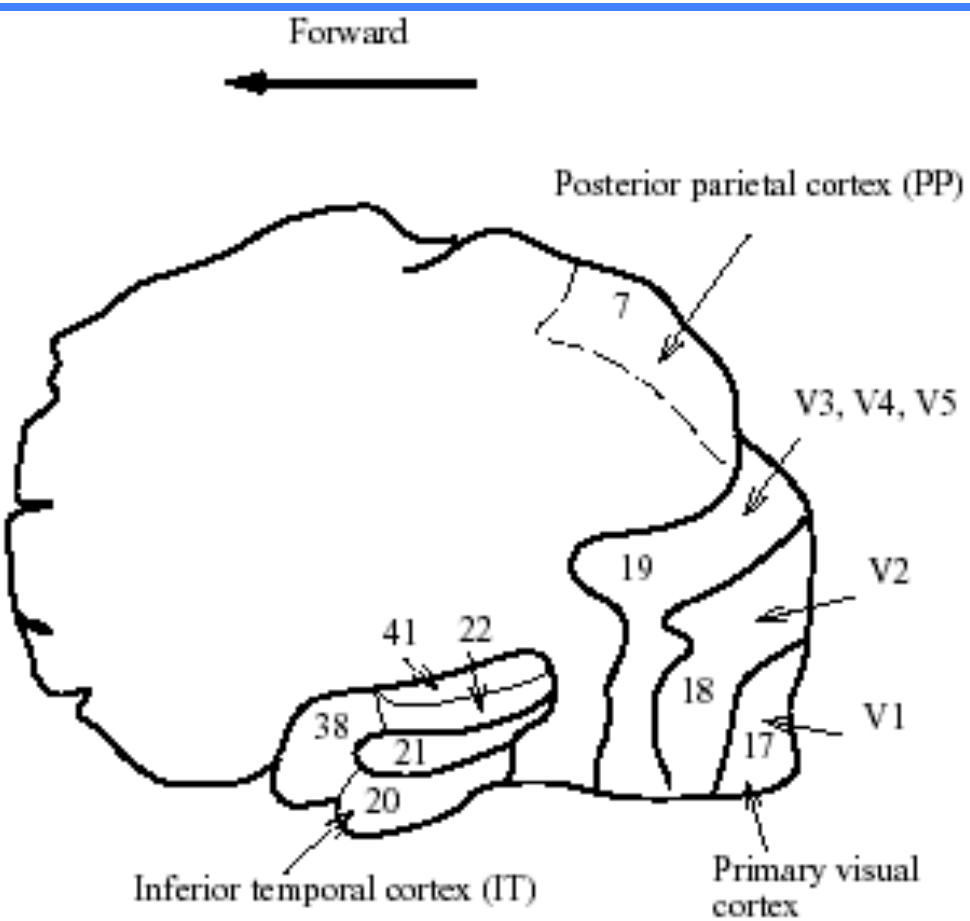
The CIE colour diagram



The visual pathway

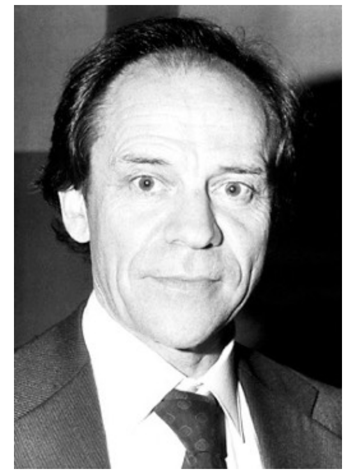


Cortical maps



1981 Nobel prize in Medicine

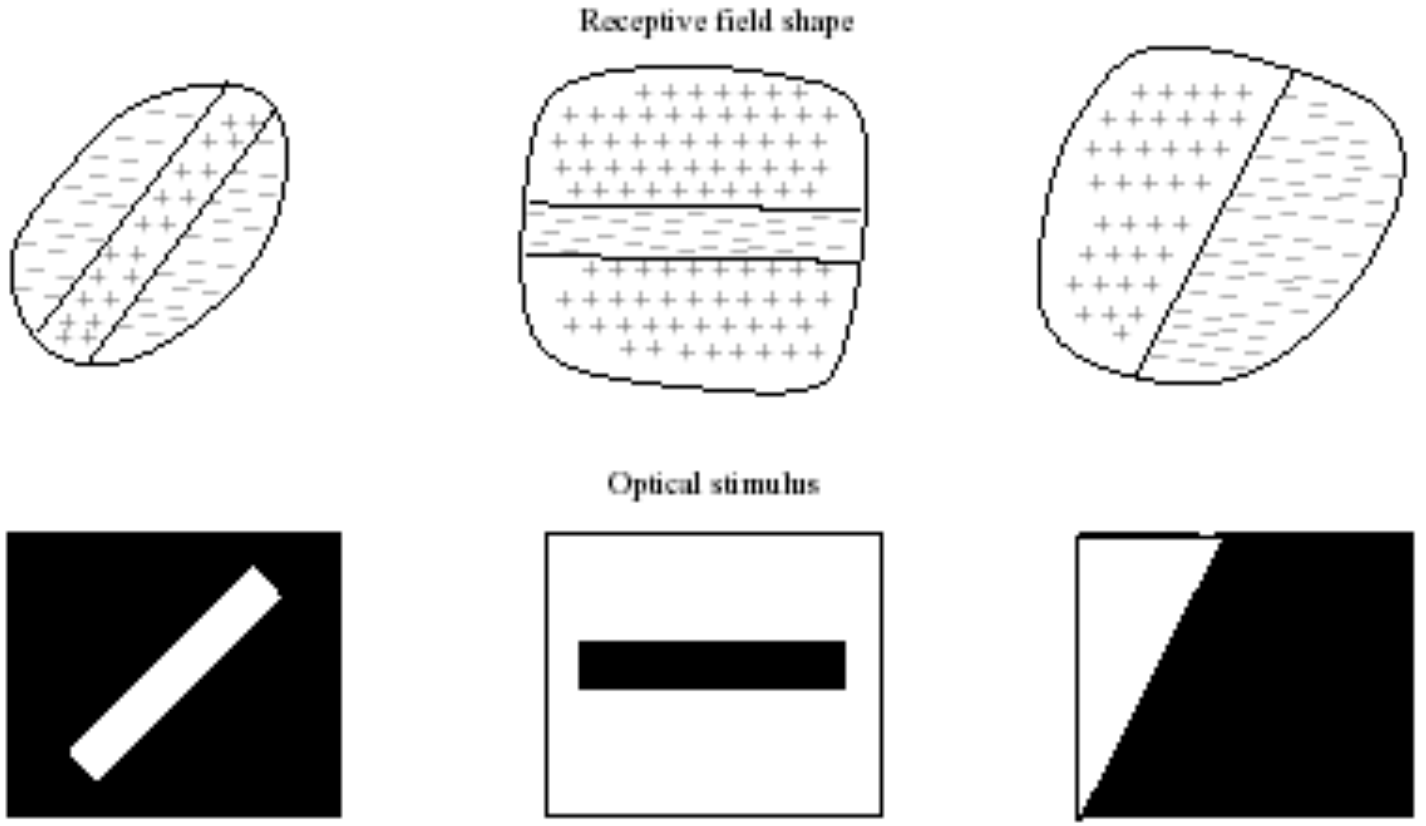
David Hubel, Harvard



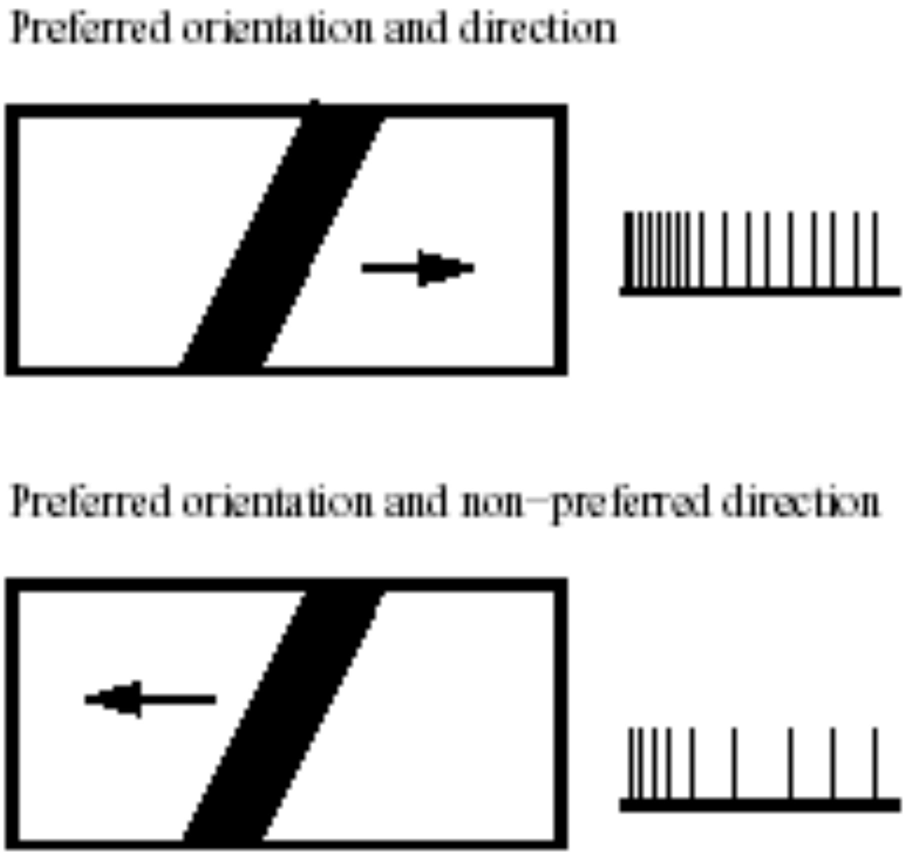
Torsten Wiesel, Harvard (initially KI)

- Microelectrodes in primary visual cortex of anasthesized cats
- What visual patterns are a particular cell sensitive to?

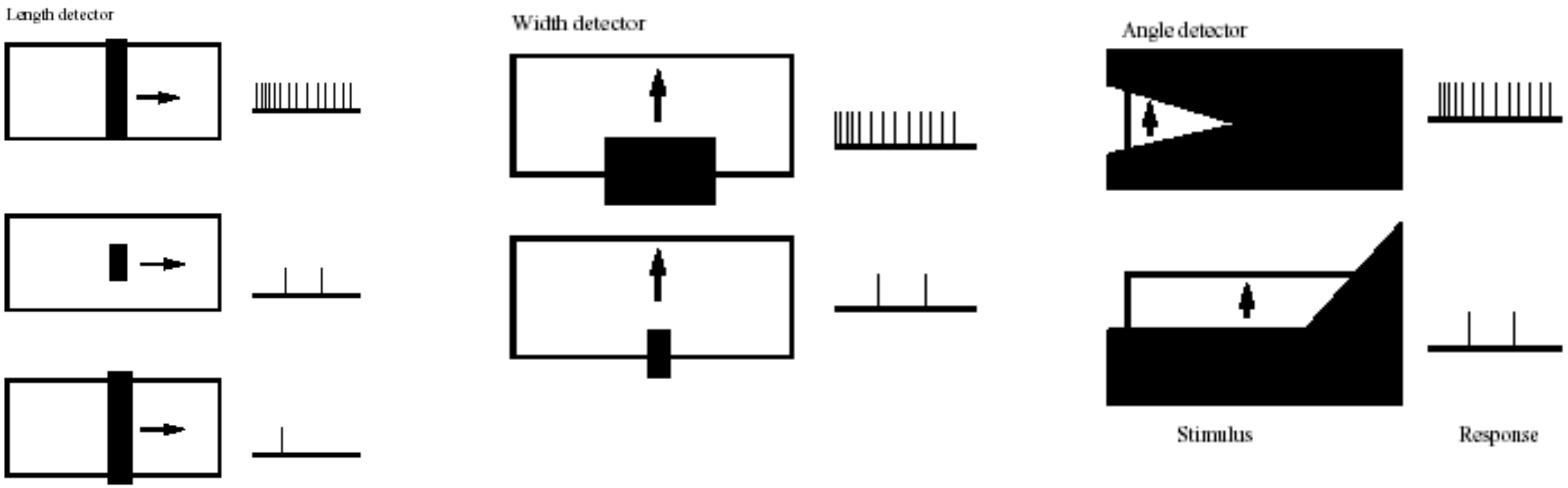
Receptive fields of simple cells



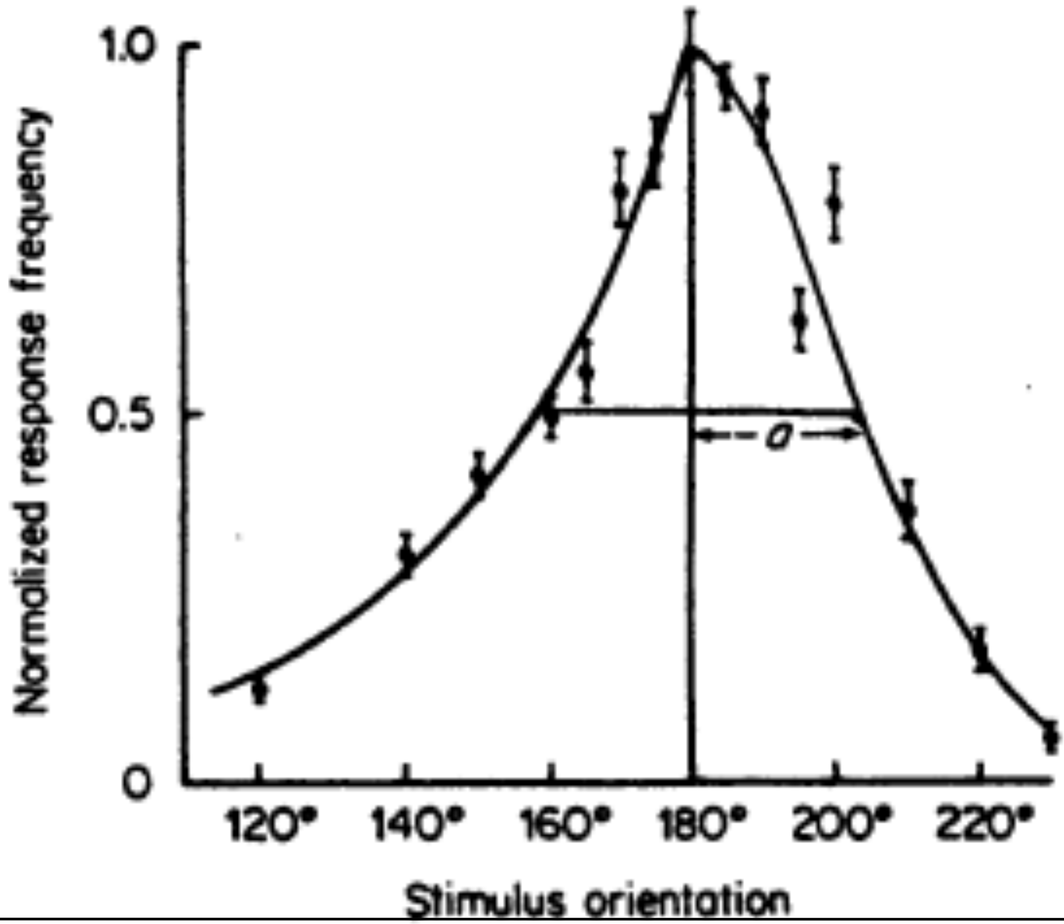
Preference of orientation and direction



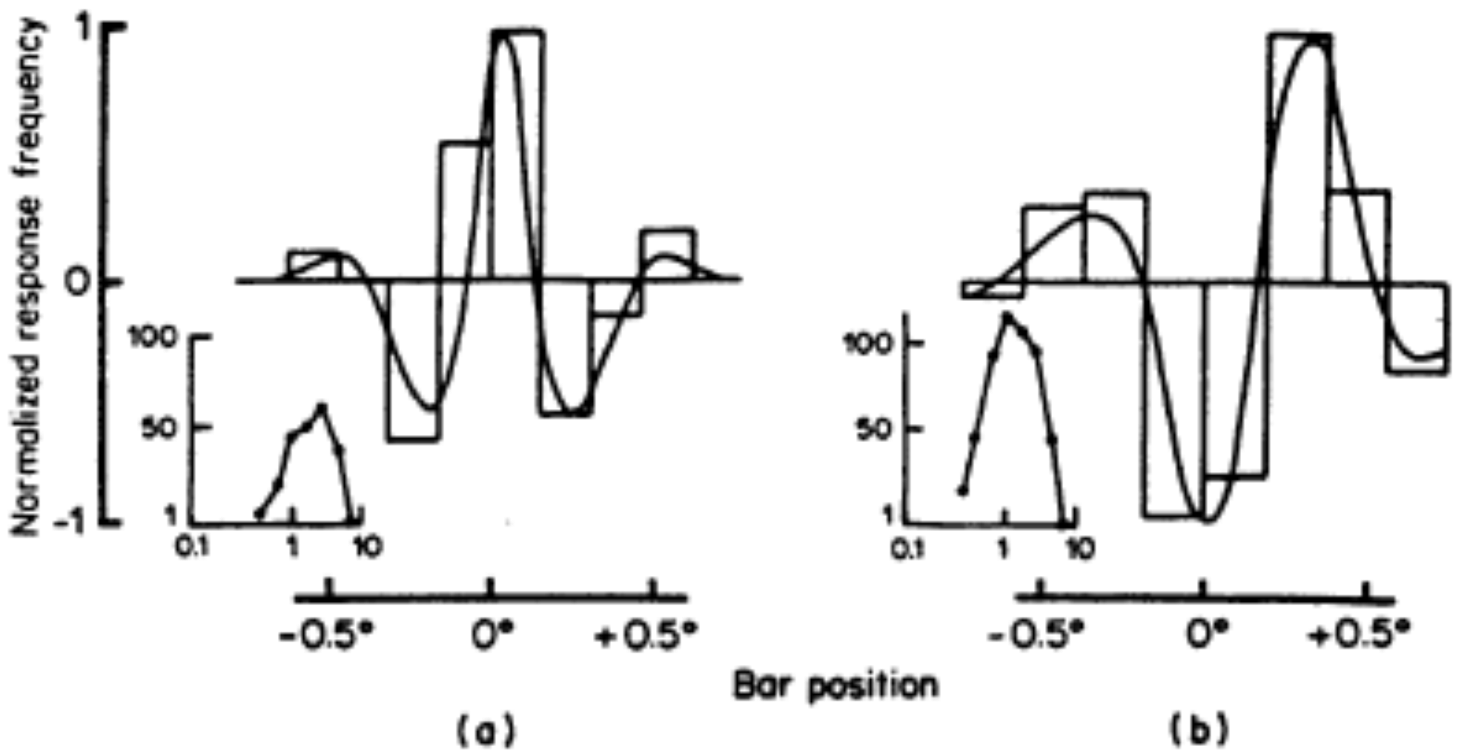
Other detectors: length, width, angle etc…



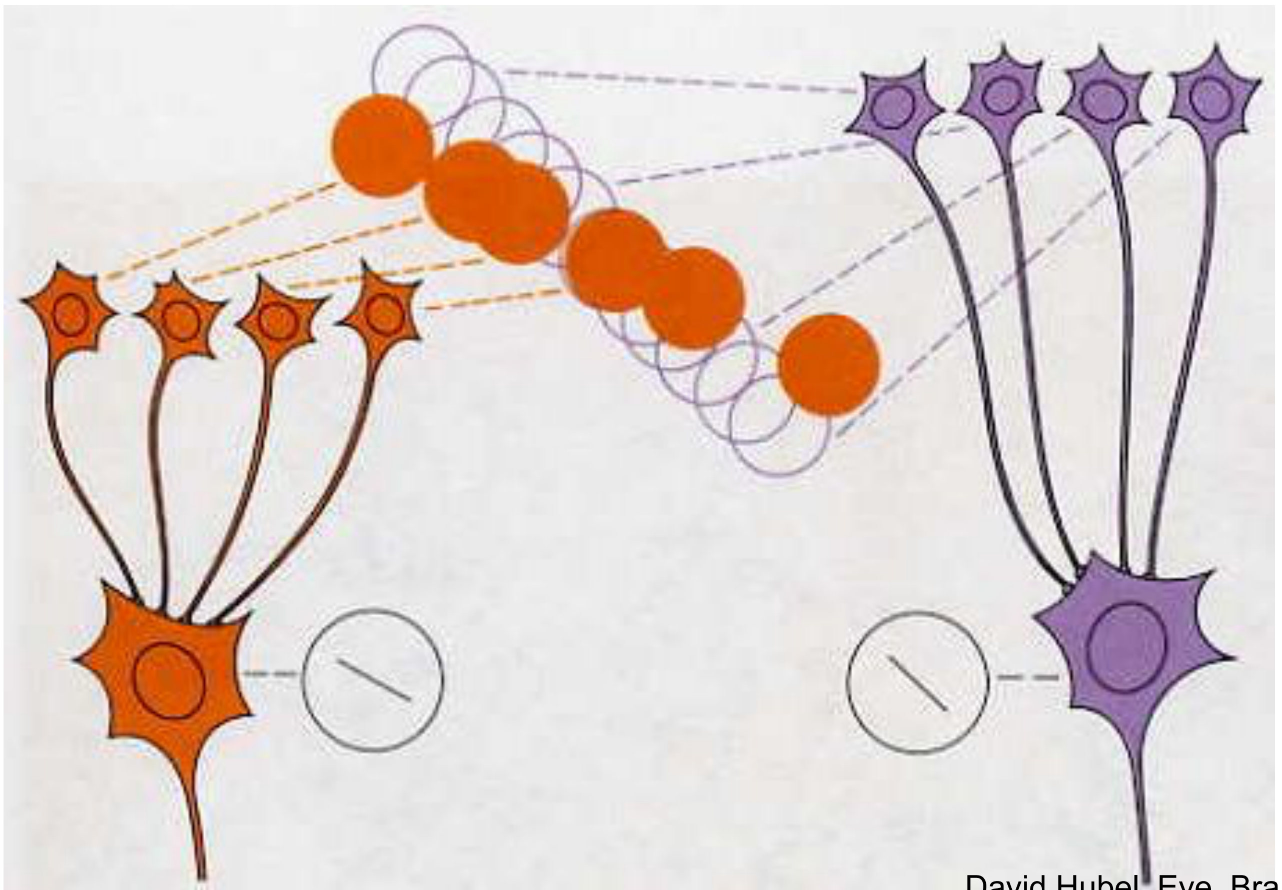
Orientation tuning
Simple cell of cat



Sensitivity profiles of simple cells a)Bisymmetrical b)Antisymmetrical

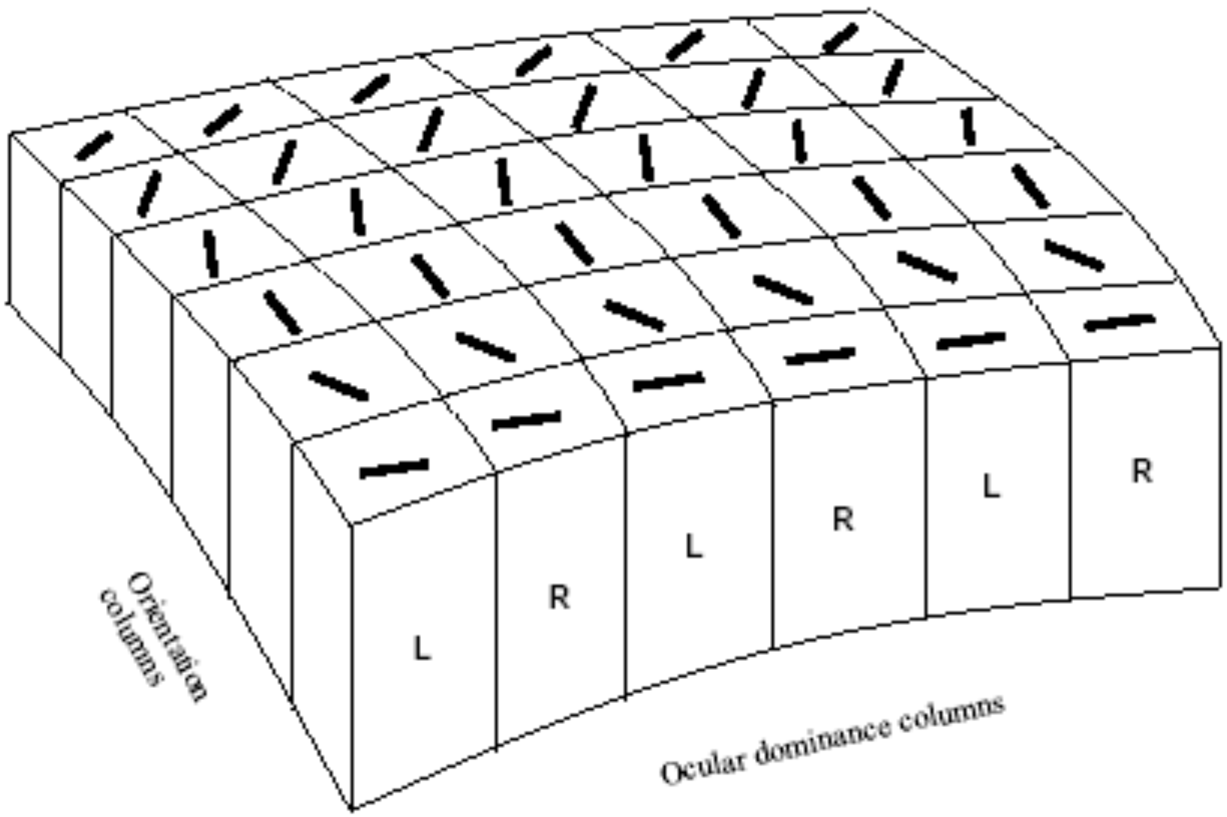


Implementation of simple cell receptive fields

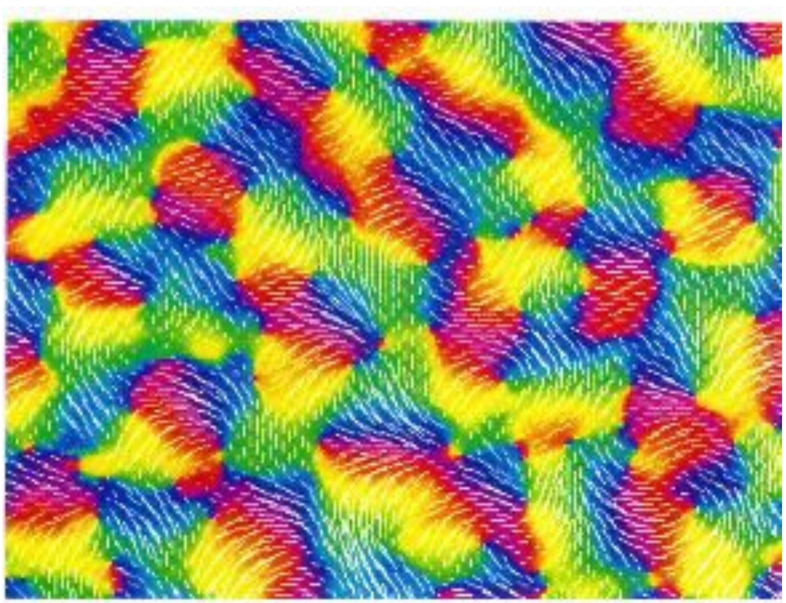


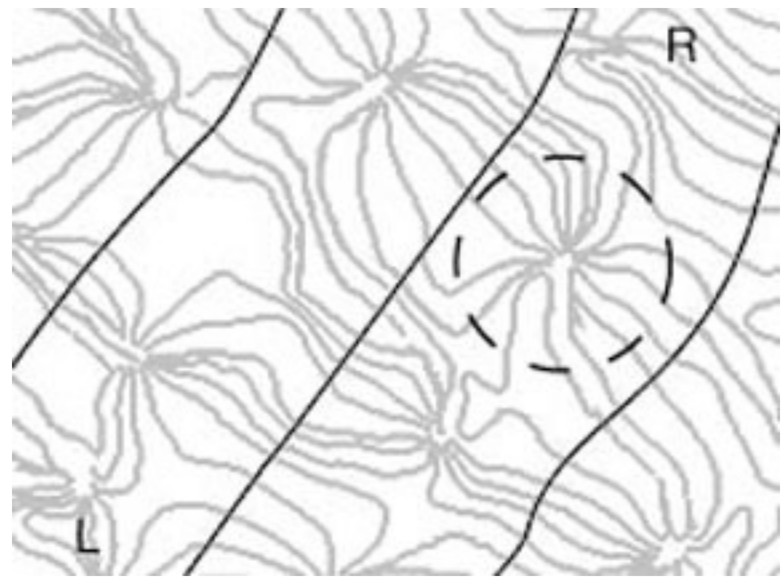
David Hubel, Eye, Brain and Vision

Orientation and ocular dominance columns

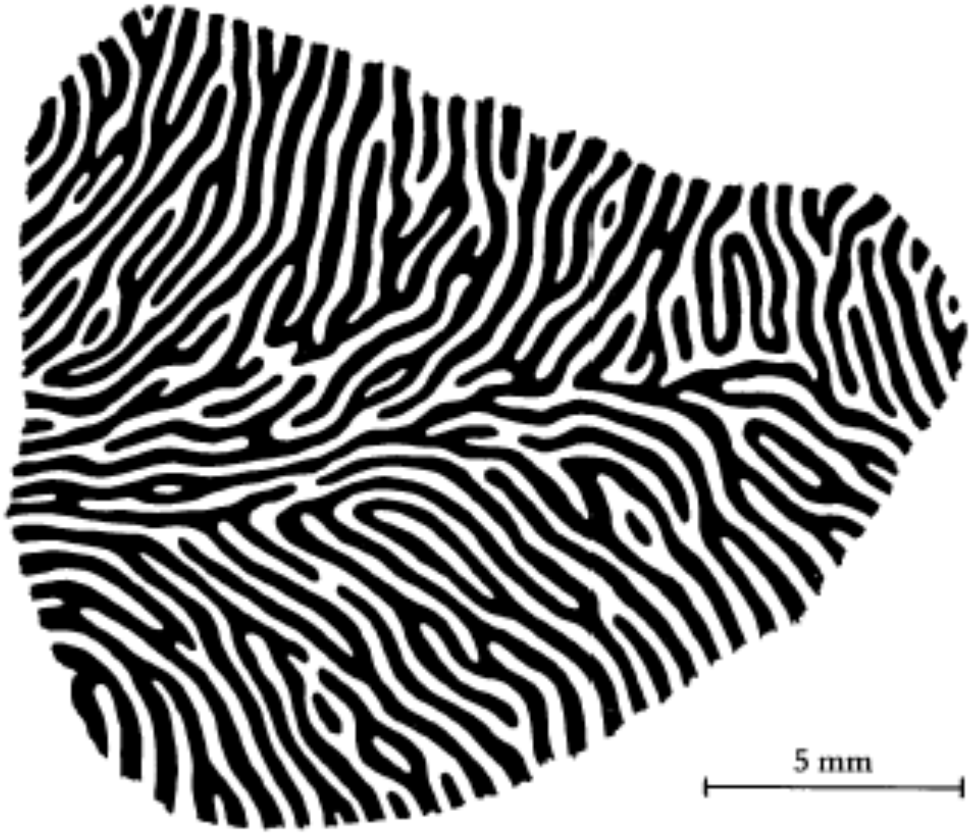


Orientation dominance

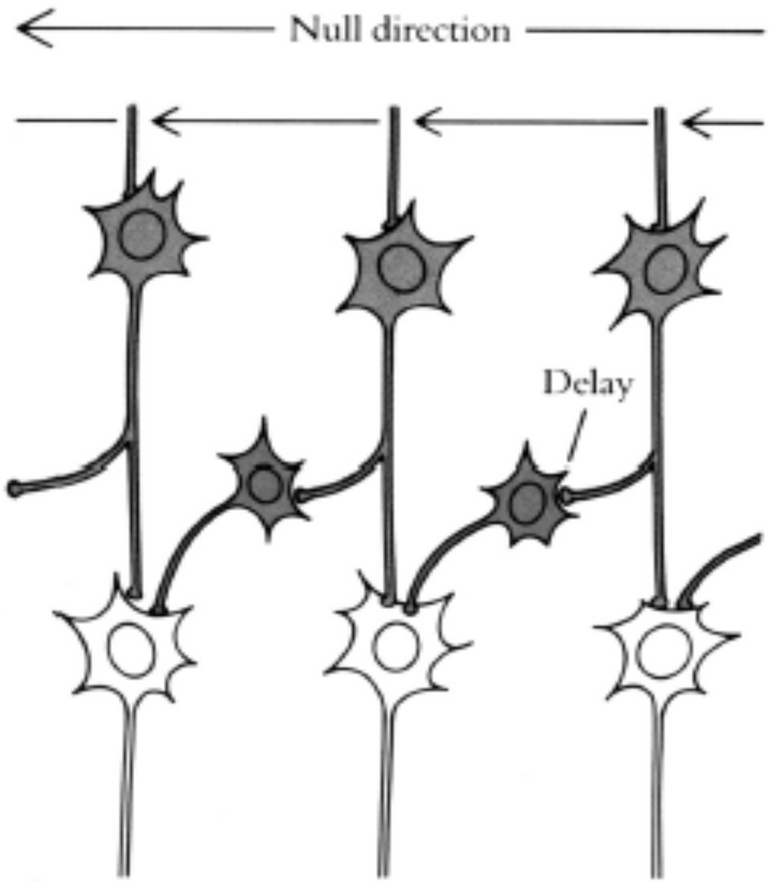




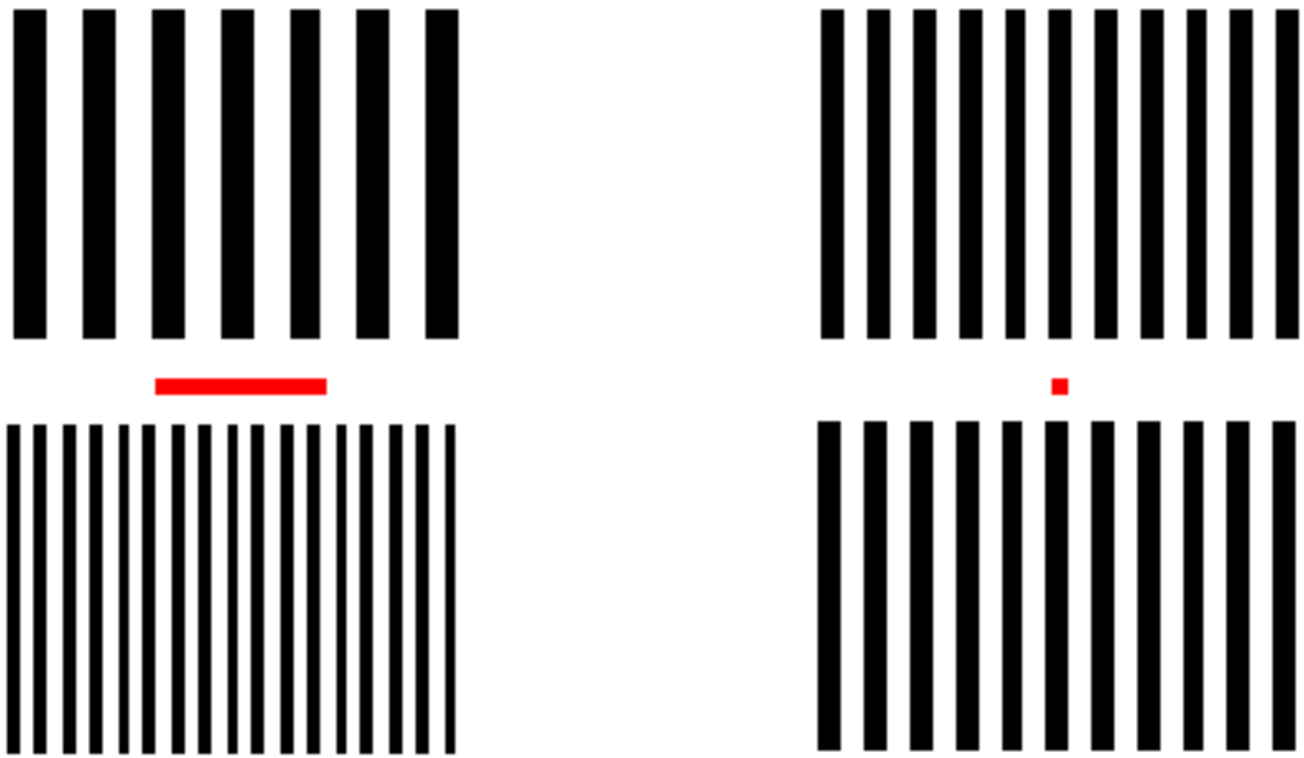
Ocular dominance map



Implementation of direction-sensitive cell

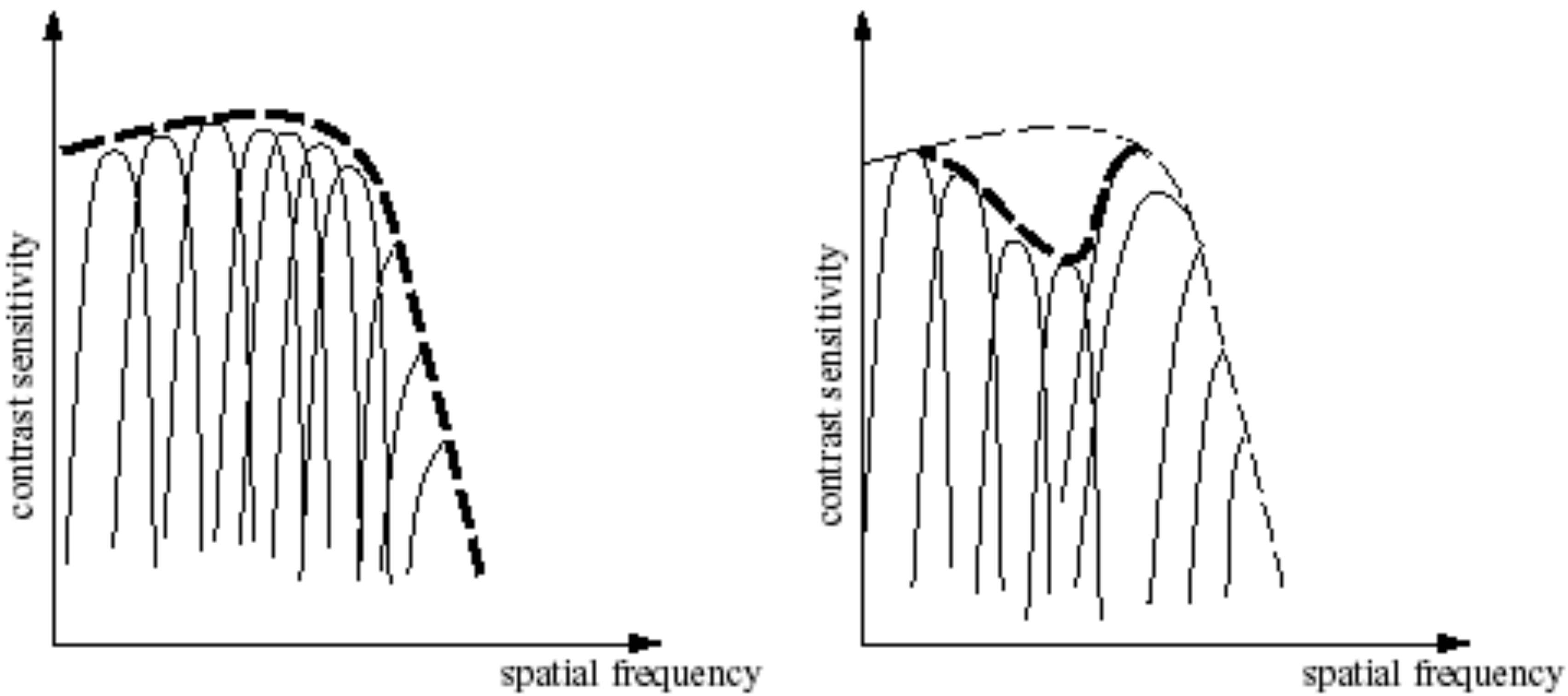


Spatial frequency adaptation

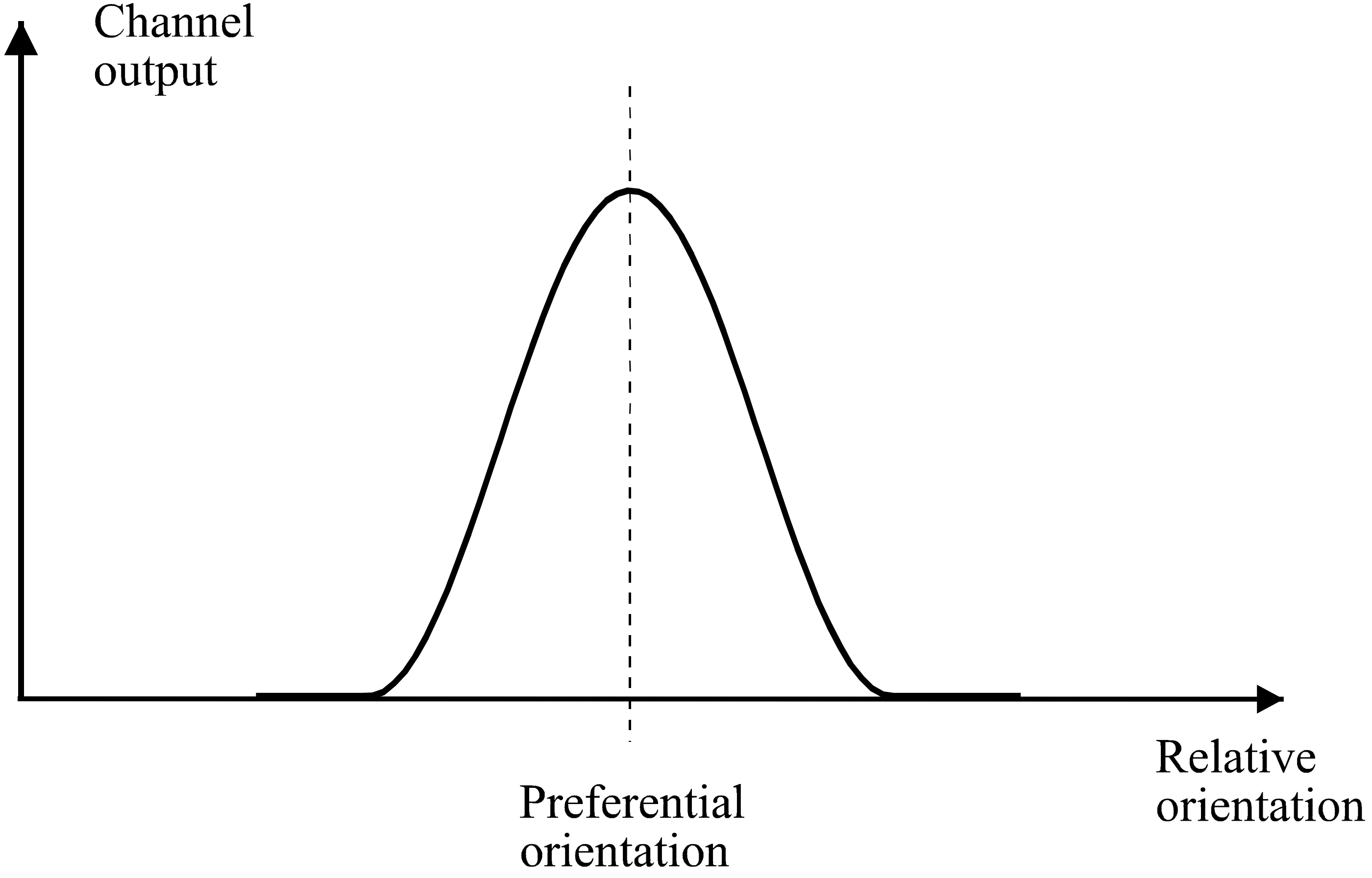


Adapted from Blakemore & Sutton, 1969

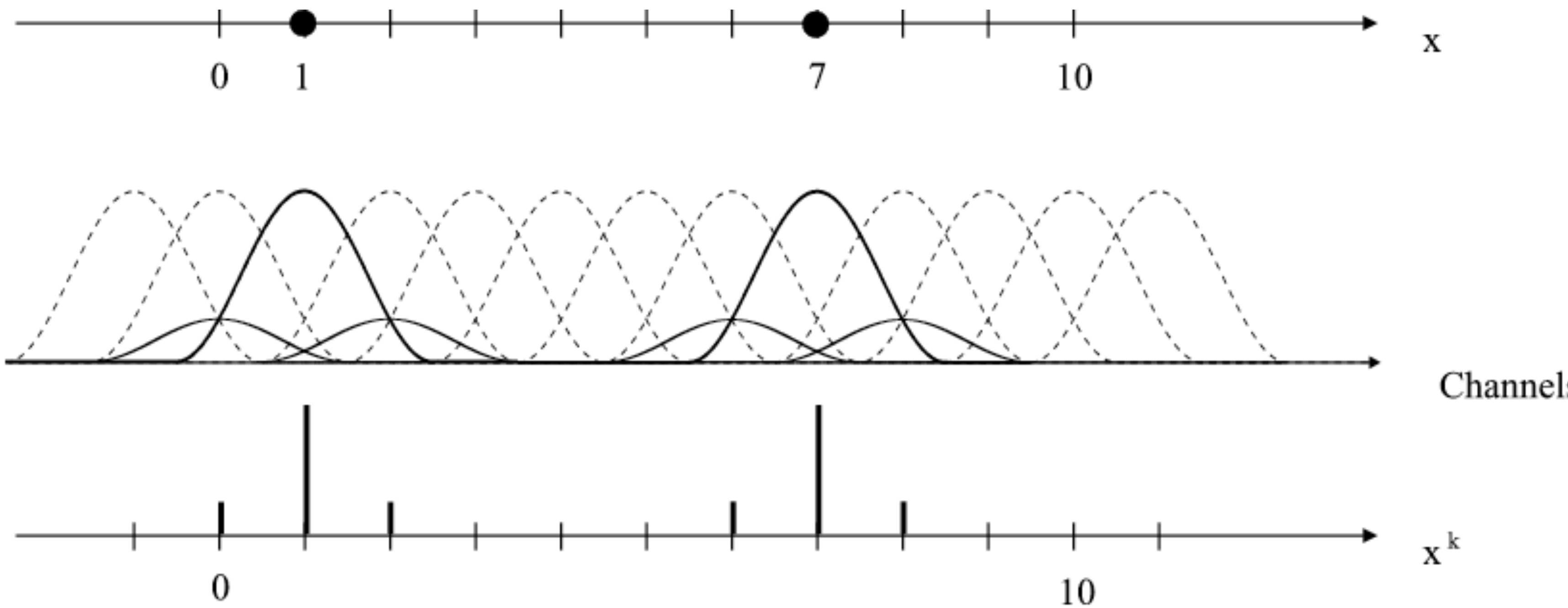
Build-up from separate channels Effect on sensitivity of channels



Channel representation



Channel Information Representation

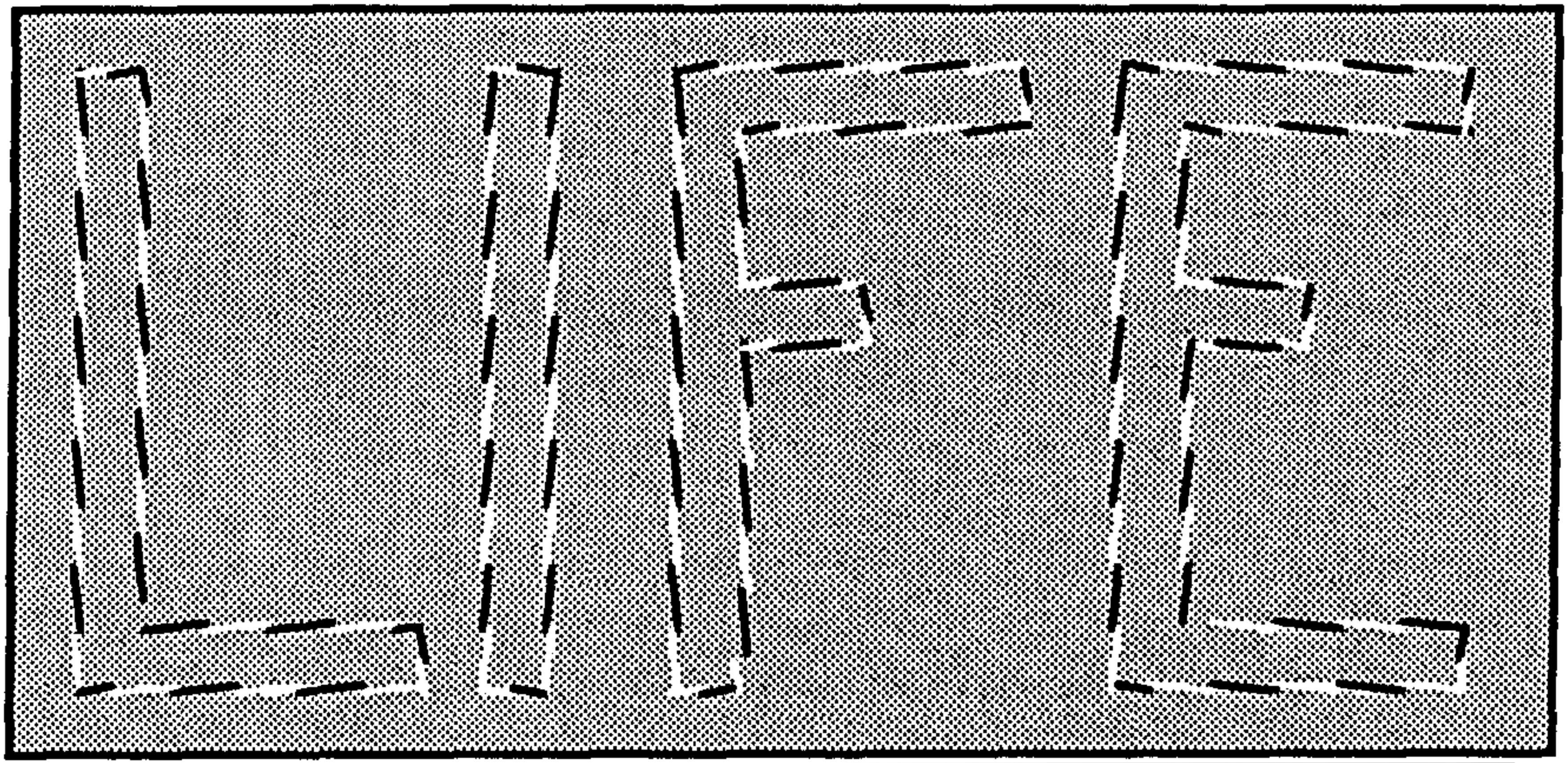


$$\mathbf{x} = [\ 0 \quad 0.25 \quad 1.0 \quad 0.25 \quad 0 \quad 0 \quad 0 \quad 0.25 \quad 1.0 \quad 0.25 \quad 0 \quad 0 \quad 0\ ]^T$$

Advantages of channel representation

- Several values can be represented for a variable, allowing support to alternative hypotheses
- Locality allows a fast optimization in learning
- Locality allows implementation of non-linear models using linear mappings
- Allows representation of confidence or certainty
- Monopolarity allows **zero** to represent **no information** (i.e. ”I don’t know”)
- Sparse representation

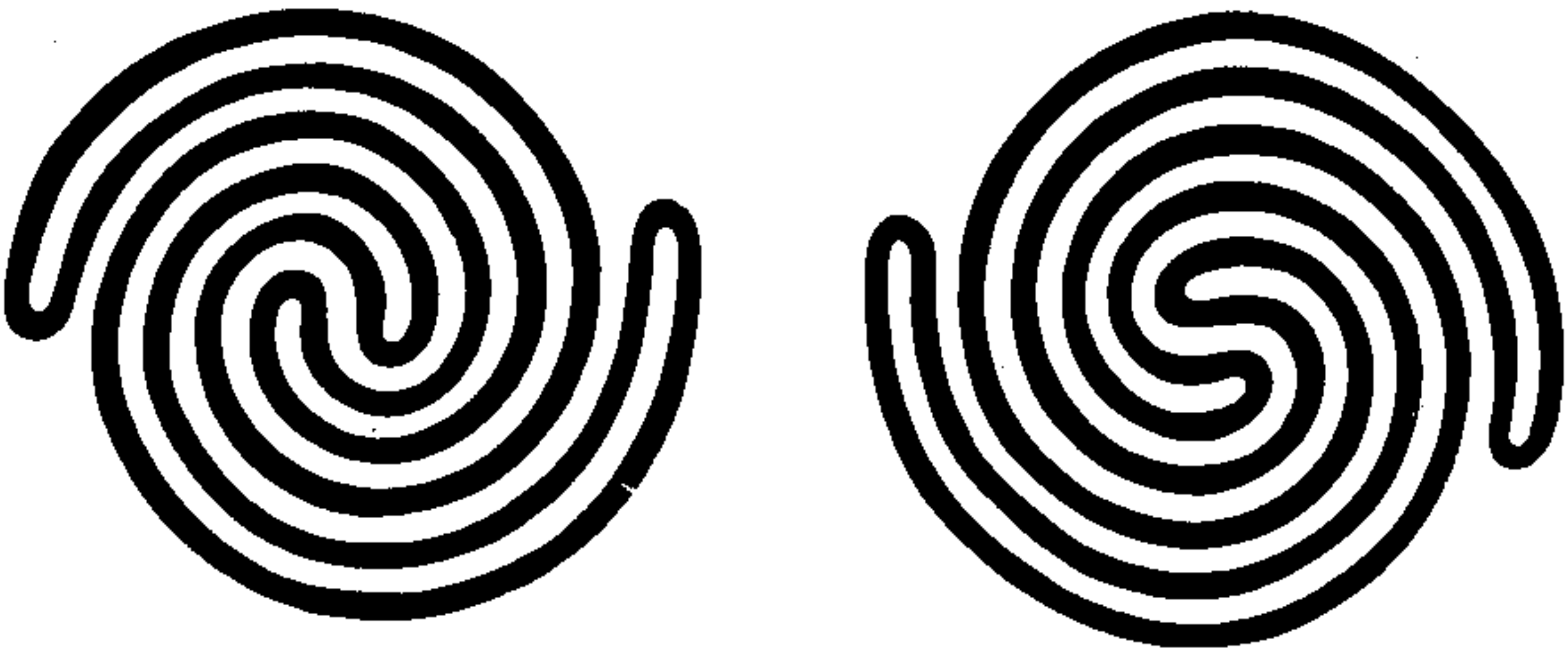
Local versus global properties



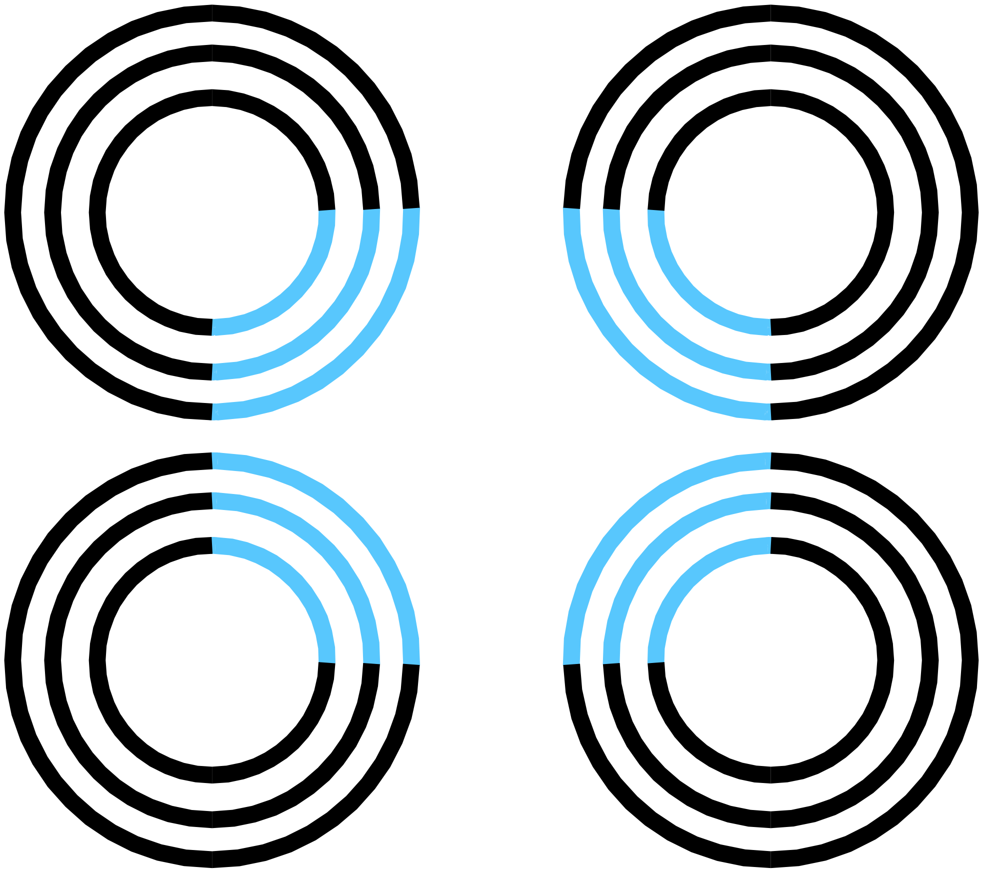
Conflicting interpretations

Parallel interpretation

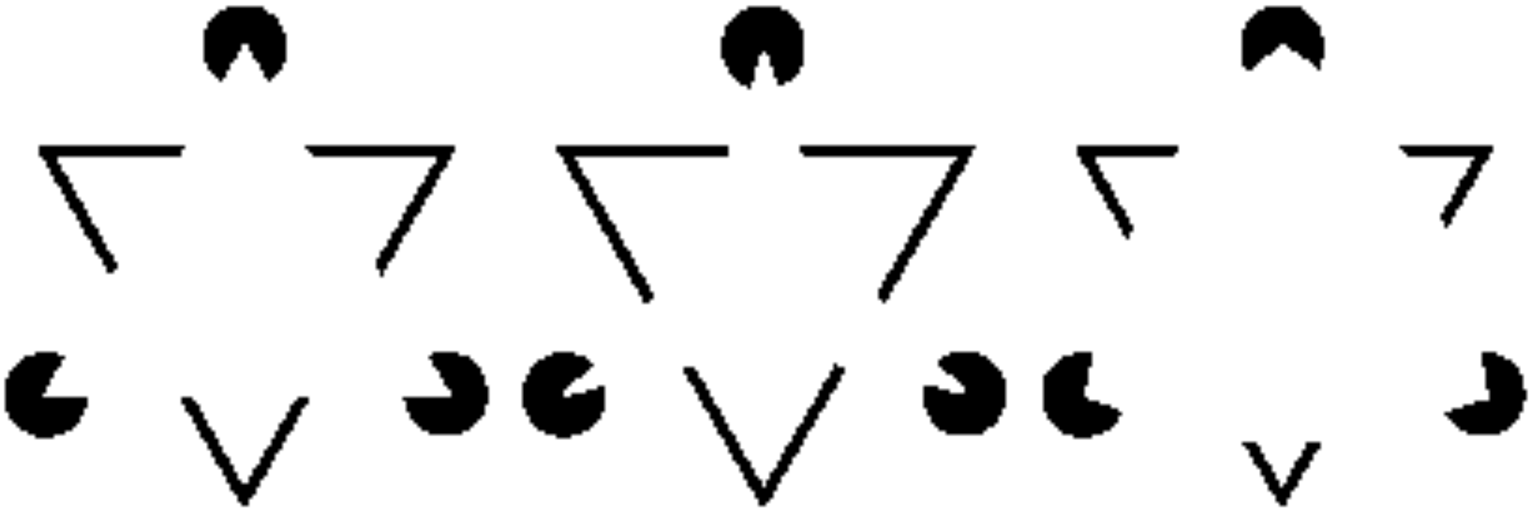
Sequential interpretation



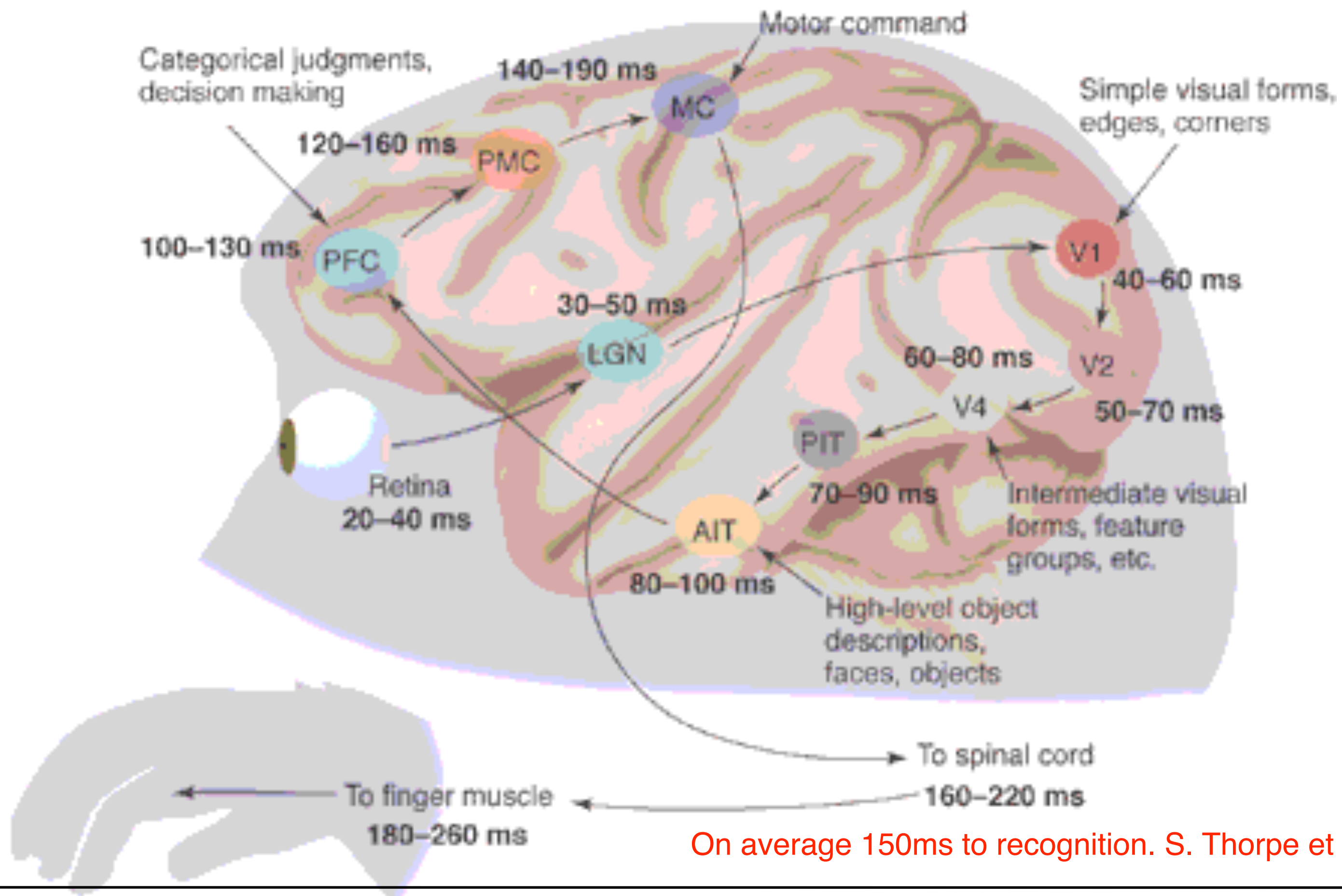
Extrapolations forming illusions



The Kanitza triangle

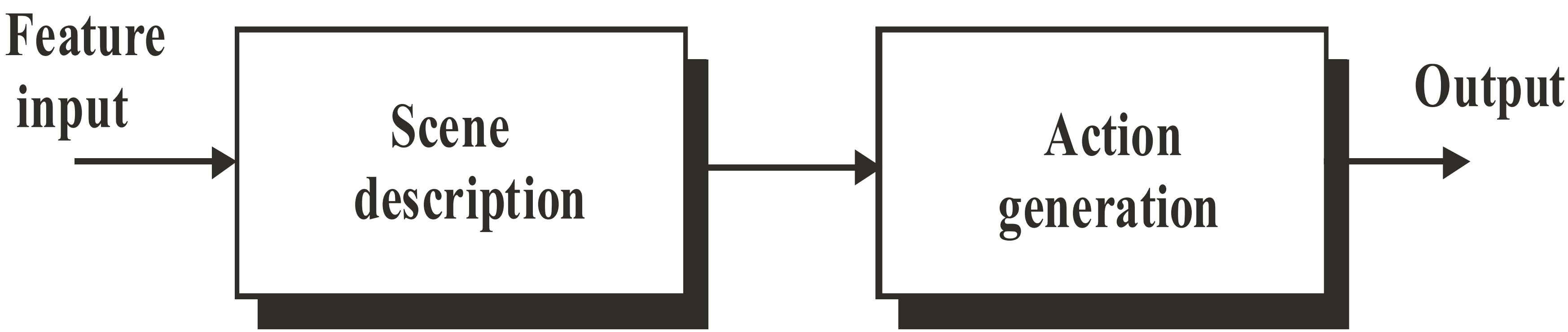


Computation times

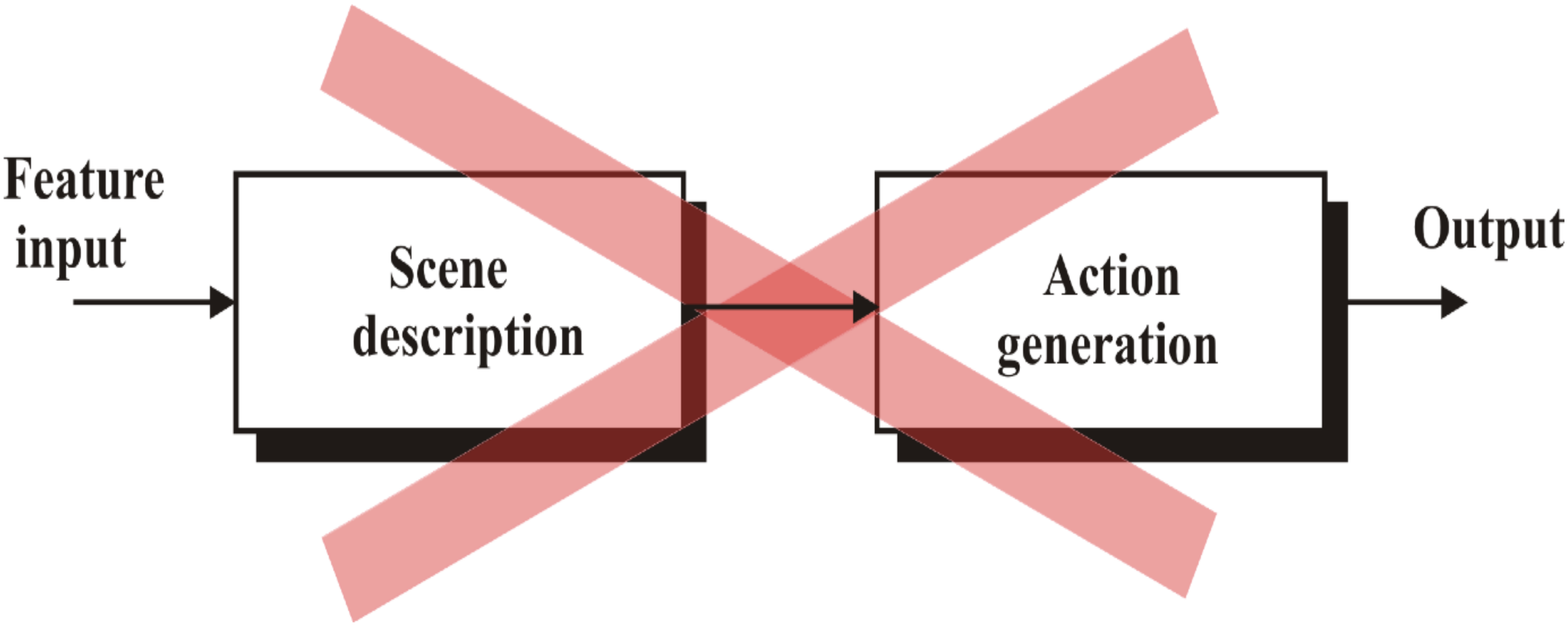


On average 150ms to recognition. S. Thorpe et al. 1996

A conventional robotics structure



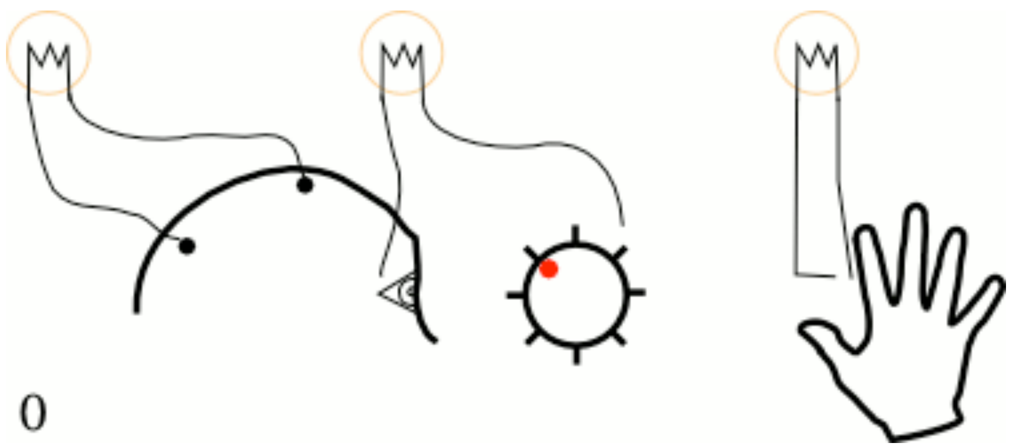
Not done in biological vision



Consciousness - an afterthought

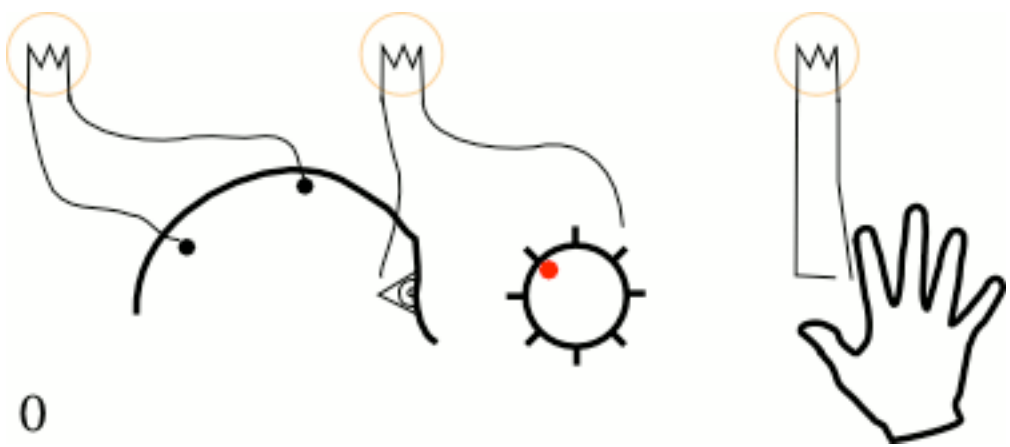
Experiments by Benjamin Libet show that:

Action is initiated before it reaches consciousness



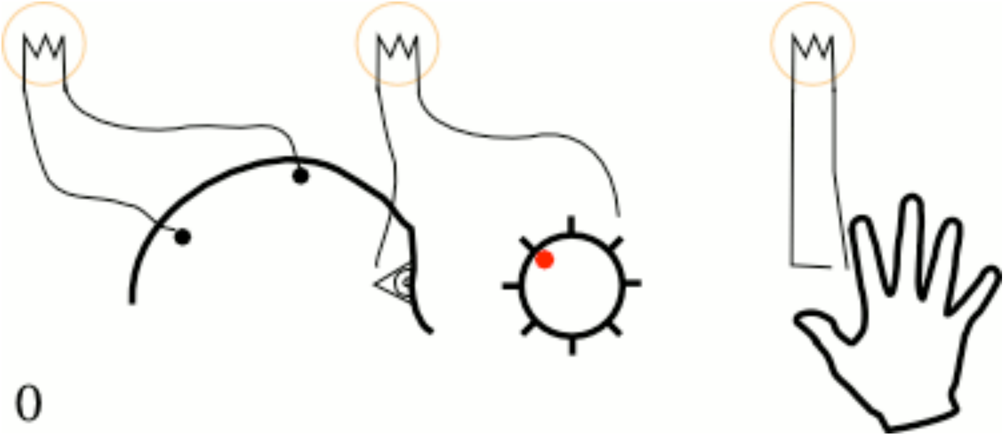
Consciousness - an afterthought

Synchronized EEG and rotating clock, subjects note position on timer when ”they were first aware of the wish or urge to act”



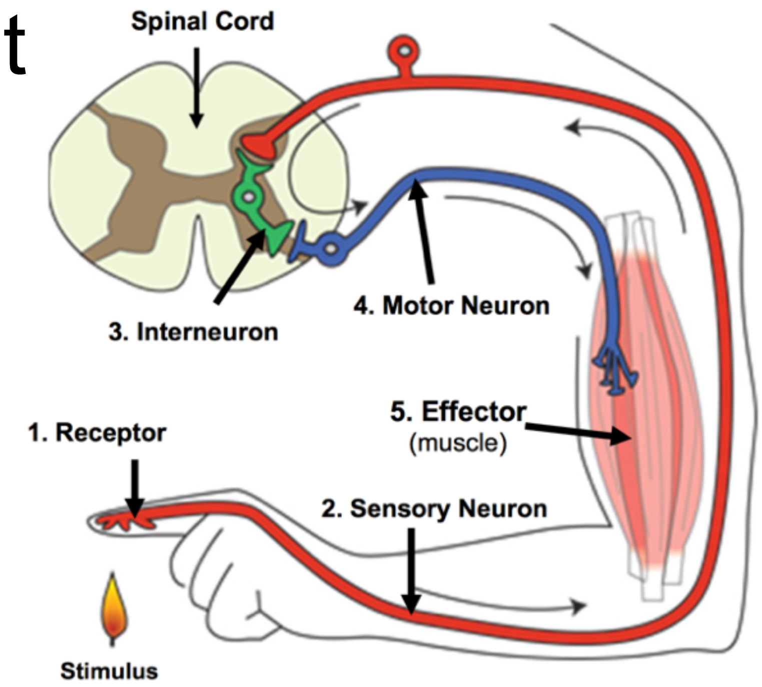
Consciousness - an afterthought

T-500ms: Readiness potential is measured by EEG
T-200ms: Observed time is registered by consciousness
by looking at synchronised clock
T: Action takes place

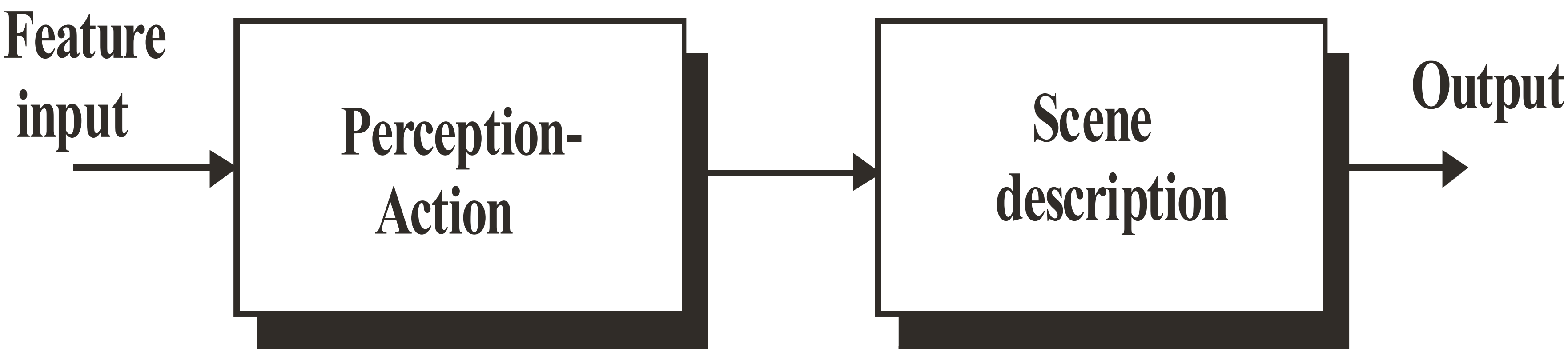


Other examples

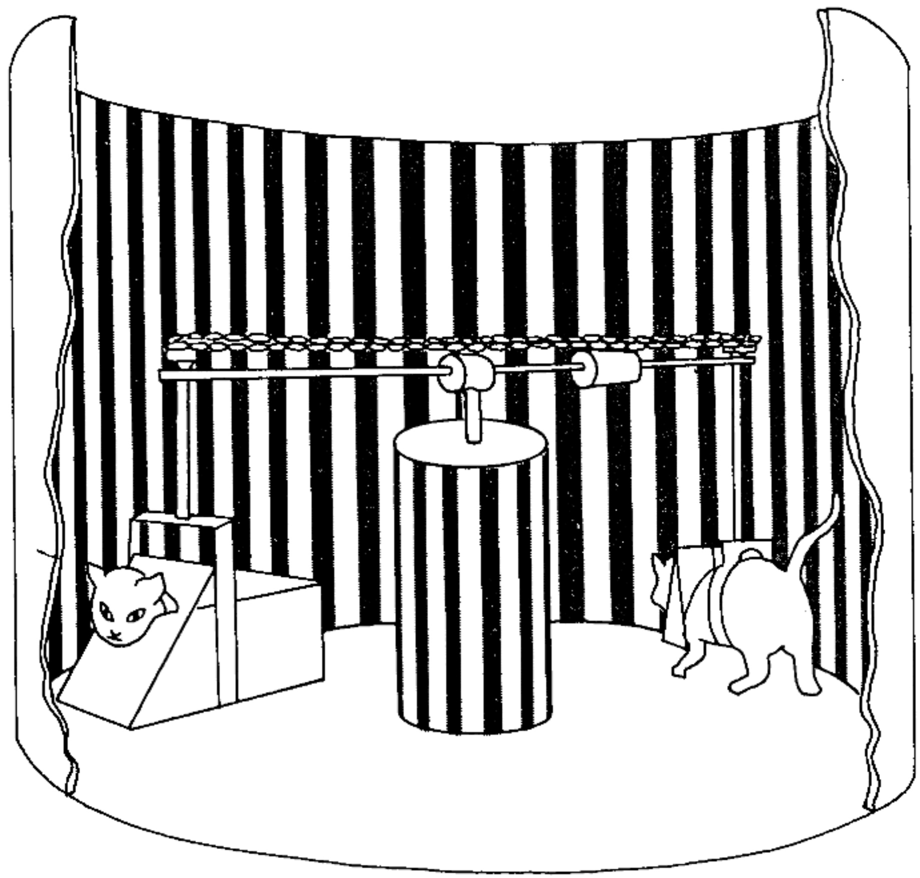
1. It is well known that reflex actions are pre-conscious
2. You do not consciously plan all details of e.g. walking pattern



Order is the opposite!



Active versus passive exposure



Active versus passive exposure

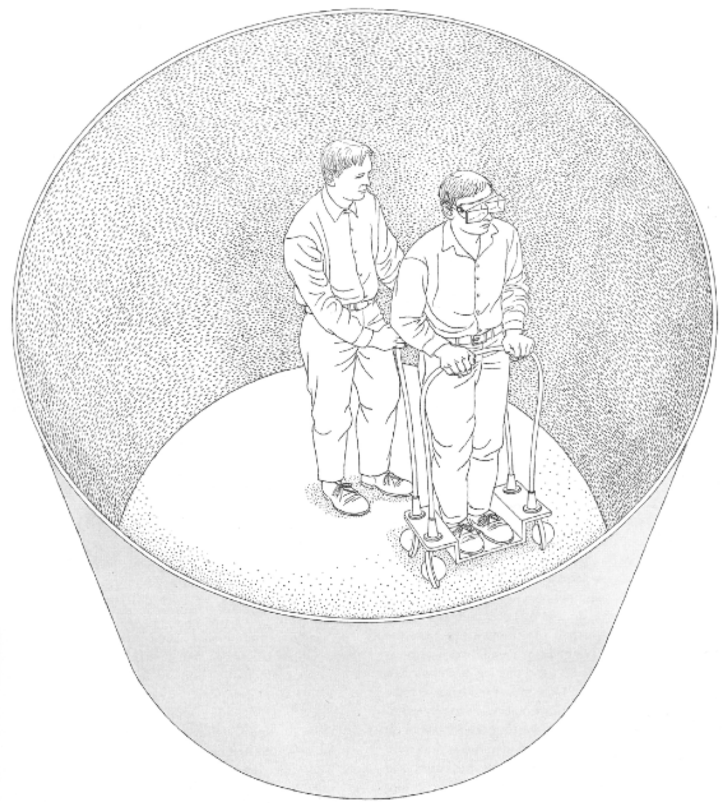


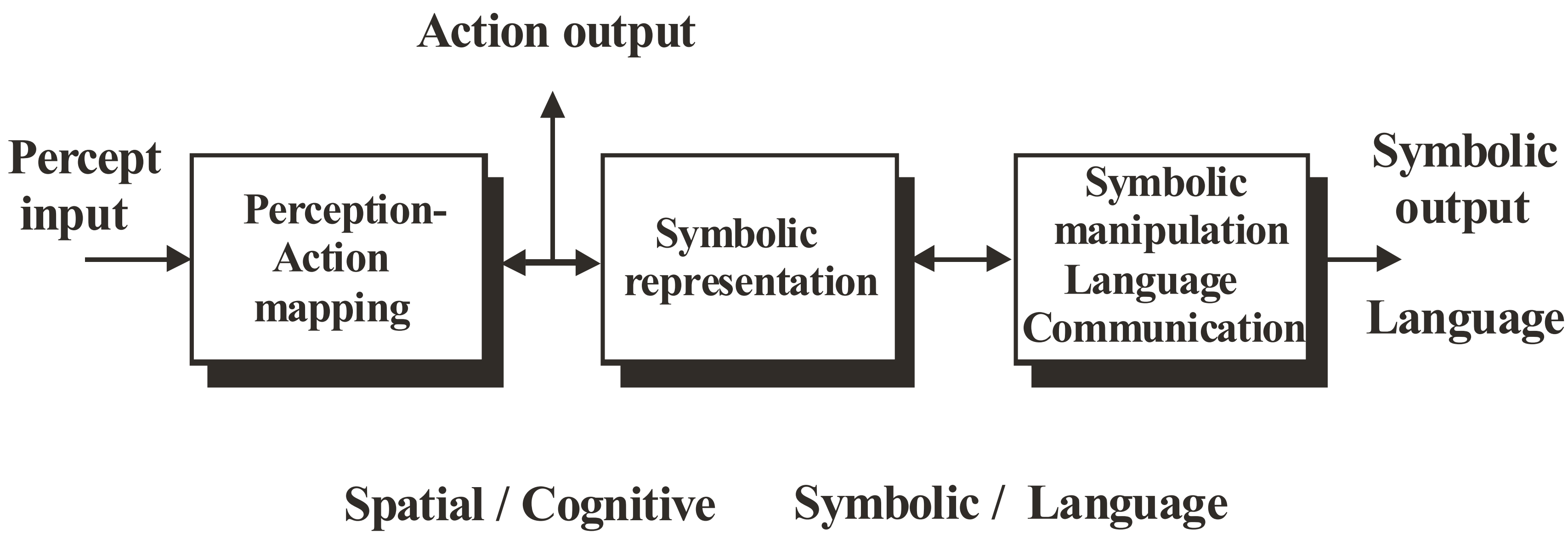
Figure 7.5.2 Adaptation during active versus passive movement. Some observers were allowed to actively explore their environment while wearing distorting goggles whereas others received similar stimulation during passive movement as shown here. The active observers adapted more fully to the optical transformation produced by the goggles than the passive observers did. (From Held, 1965.)

Why active learning?

Act-perceive-learn cycle

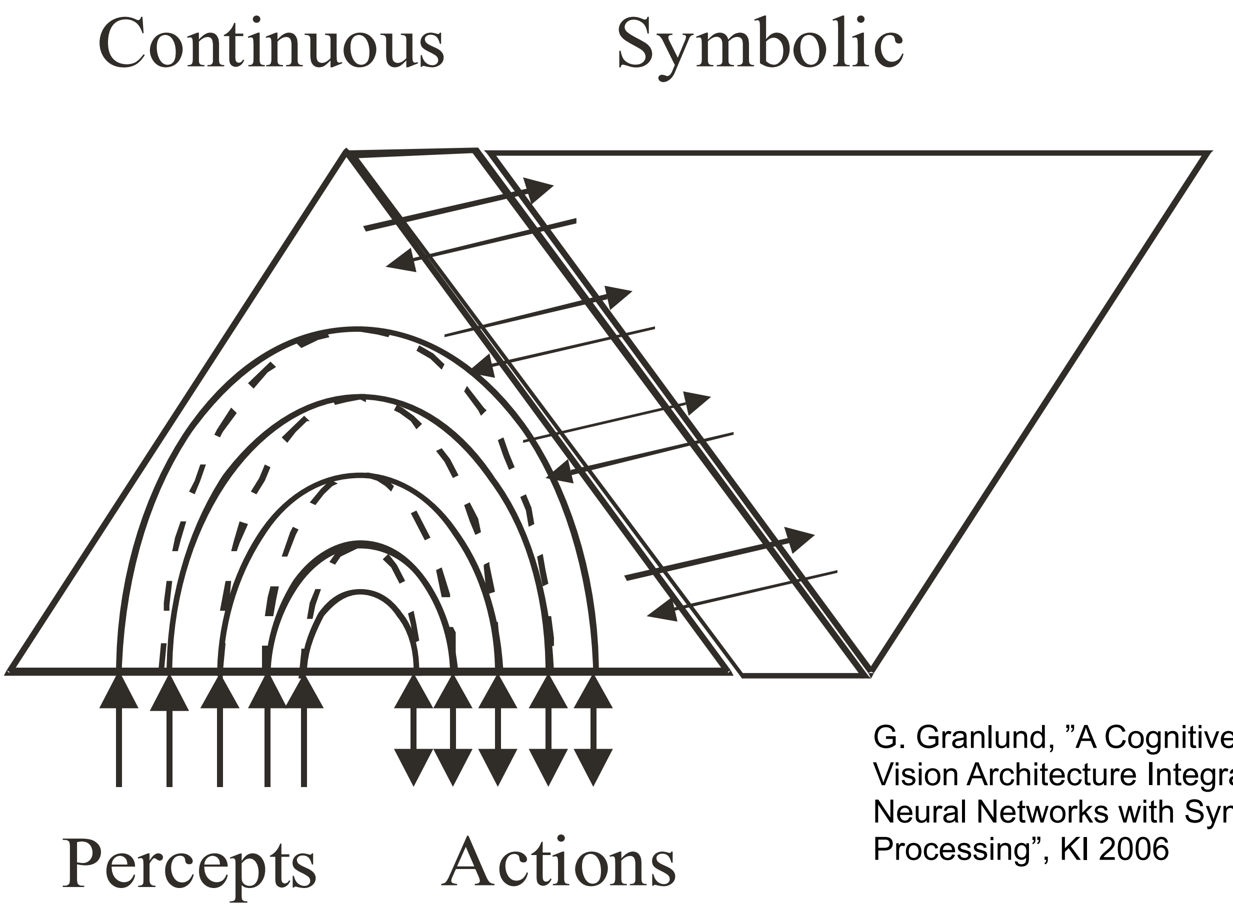
- Only features that change are related to the action or state change
- The action or state space is much less complex than the percept space
- Does not require consciousness (other forms of learning do)

Extended Cognitive Structure



G. Granlund, "A Cognitive Vision Architecture Integrating Neural Networks with Symbolic Processing", KI 2006

Pyramid version



G. Granlund, "A Cognitive Vision Architecture Integrating Neural Networks with Symbolic Processing", KI 2006

Summary

- Biological vision systems are not monolithic, but a collection of **visual behaviours**
- Visual perception is done in **cortical maps**, for e.g. colour, edges, and faces
- Much of visual learning is **active**, and pre-conscious