Visual information processing and visual perception

Erhardt Barth

Institute for Neuro- and Bioinformatics, University of Lübeck
http://www.inb.mu-luebeck.de

barth@inb.mu-luebeck.de
Please send me email to request reprints!

Scope of the lecture

- Learn about the basic principles of visual information processing.
- Obtain an intuition for the fact that we and what we perceive depend upon things going on in our brains.
- Why learn that? Because the technical systems that we build today are much too simple relative to the complexity of the biological system they are build for (technical design must be more human centered).
- Present a current project (***Itap***) as an example of such human-centered design of a communication systems.



Why study biological vision systems?

Learn how to build technical systems, e.g.

- Machine-vision systems
- Robots
- Prostheses (e.g. electronic eyes)
- Image coders



Why study biological vision systems?

Human-Computer Interaction (HCI)

- Optimal Coding
- Ergonomics

An important trend in the last few years was to move *"From Computing to Interaction Design"*
(Terry Winograd, Stanford Univ.)

Interactive systems can only be build on the basis of interdisciplinary research.

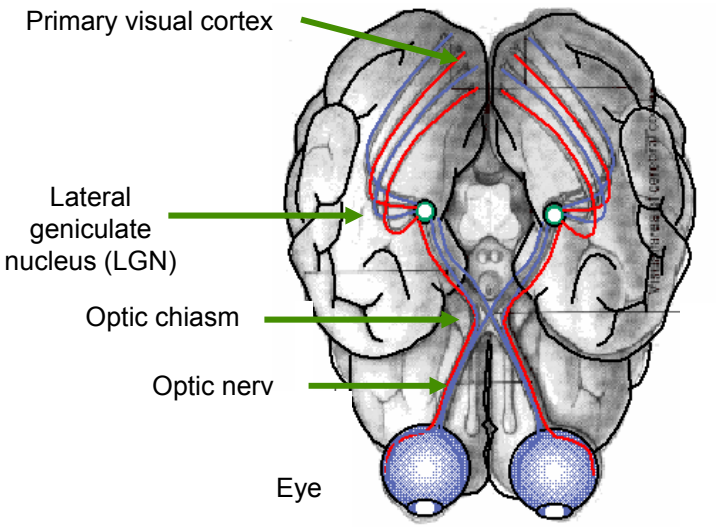


Vision-research paradigms

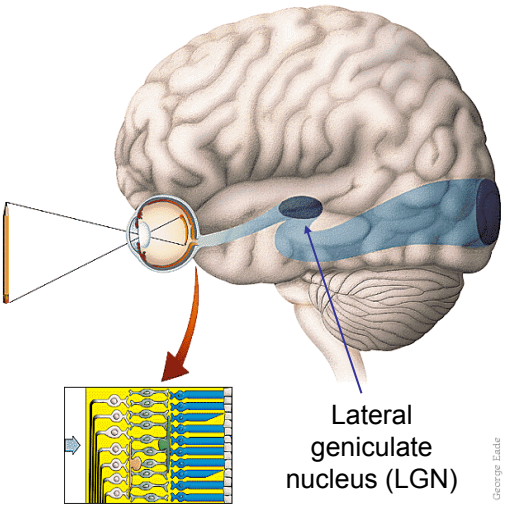
- Passive, bottom-up, information processing
  - Information processing in the retina
  - Information processing in the visual cortex
- Active vision and top-down processes
  - selection of information
  - senso-motoric loop



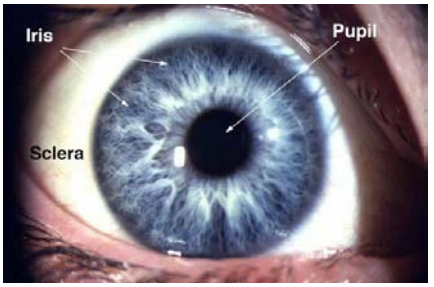
Visual pathway, bottom view

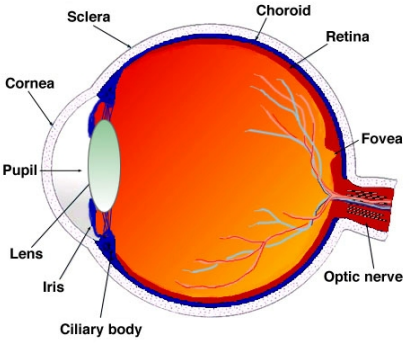


Visual pathway, lateral view



The eye





The eye: main functionality

Eye position is controlled by six muscles

Cornea and lens project the image on the retina

Ciliary muscles adjust the shape of the lens for optimal focus

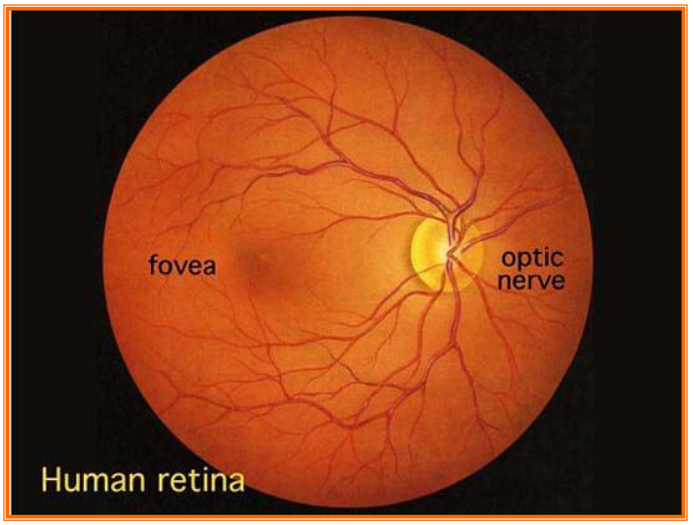
The iris (a circular muscle) controls the size of the pupil so that more or less light is allowed to enter the eye

The retina holds the light-sensitive receptors that have the highest density in the fovea

Retinal information leaves the eye through the optic nerve



Retina as seen through an opthalmoscope

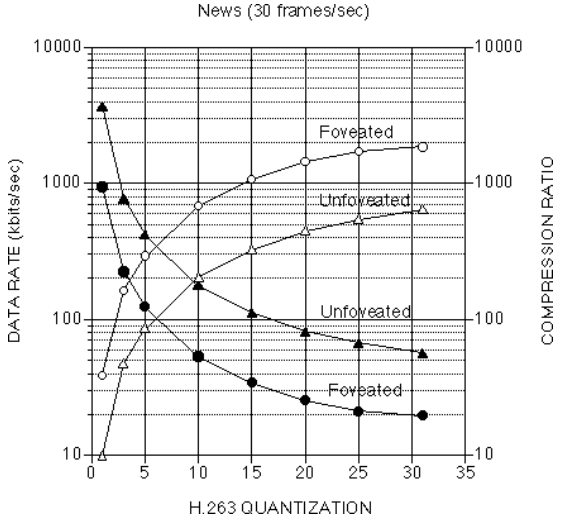




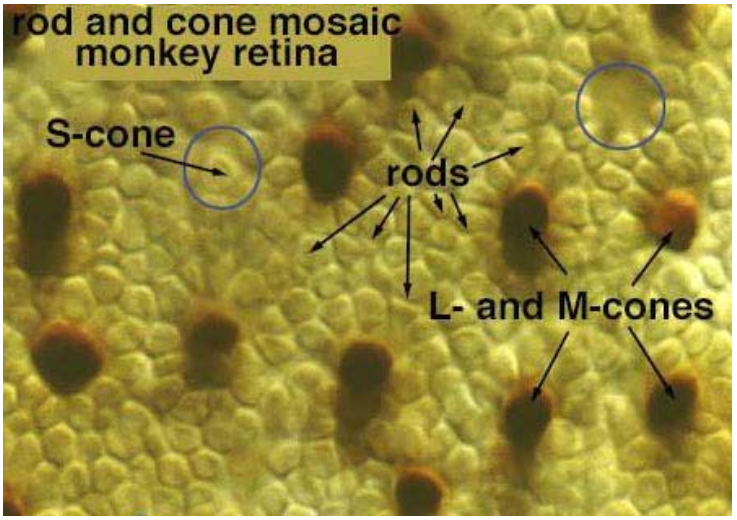
Foveated MPEG

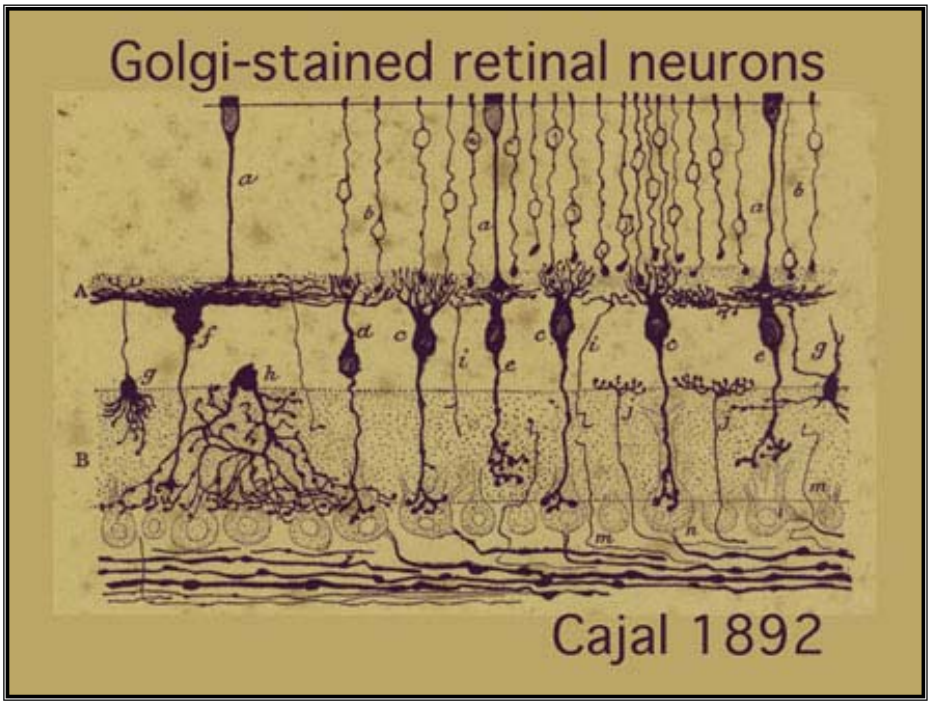


Foveated MPEG



Retina as seen through a microscope



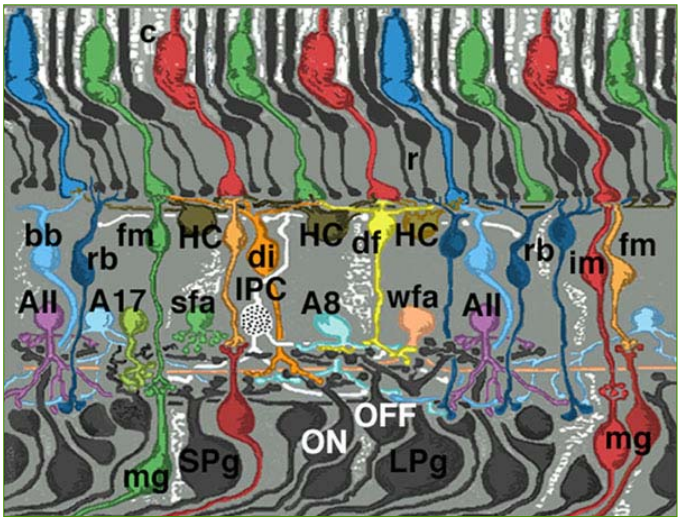


Neurons as discrete cells

- **Camillo Golgi** developed a histological technique that was able to stain individual neurons
- **Ramon y Cajal** used the Golgi method to draw incredibly beautiful and accurate pictures of neurons. He realized that *neurons communicate* with each other *specifically*:
  - Connections between different neurons are made with specific other cells and at specific sites, rather than just randomly
  - The information flow from one neuron to another is usually unidirectional
  - There are parts of neurons that are specialized for receiving information (e.g. dendrites) and other parts that are specialized for communicating information (e.g. axons)

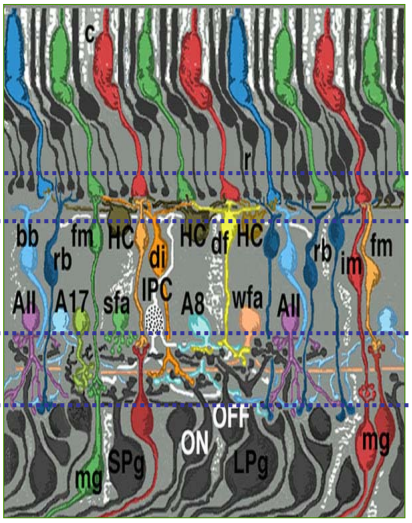


Schematic drawing of the retina





Organization of the retina



Receptor layer

Outer plexiform layer

Horizontal, bipolar, and amacrine cells

Inner plexiform layer

Ganglion cells



Cell types in the retina

- Receptors
- Horizontal cells
- Bipolar cells
- Amacrine cells
- Ganglion cells

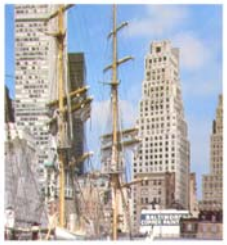


Receptors are of two types

- **Rods**
  - Responsible for night vision
  - High sensitivity (one single photon is sufficient)
  - Low spatial and temporal resolution
- **Cones**
  - Responsible for daylight vision
  - High spatial and temporal resolution
  - Color vision (3 types: L-, M-, S-cones)
  - Low sensitivity (need hundreds of photons)



Optimal color coding



RGB principal components

$$\begin{pmatrix} 0.44 & 0.55 & 0.71 \\ -0.77 & -0.16 & 0.61 \\ 0.44 & -0.82 & 0.36 \end{pmatrix}$$

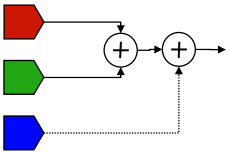


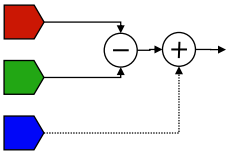


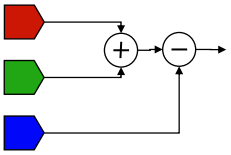
R+G+B -R-G+B R-G+B



Visual color coding







R+G+B

achromatic channel

R-G+B

red-green channel

R+G-B

yellow-blue c.

The three channels carry different spatial frequencies



Horizontal cells

(1) Adjust photoreceptor response (gain control)

Effect of background light is to increase gain of feedback onto photoreceptors.

(2) Provide color opponency (some species)

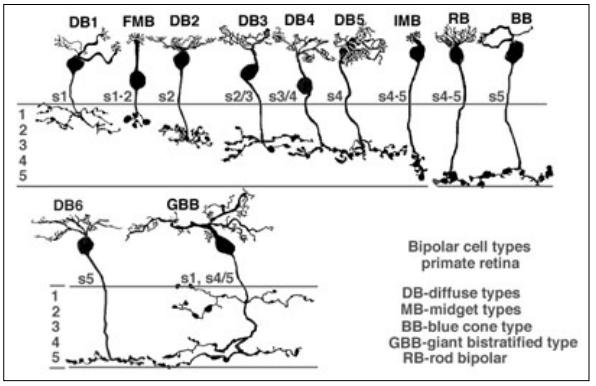
(3) Provide surround for subsequent neurons

Collect and summate photoreceptor responses over a large area due to their interconnectivity via gap junctions



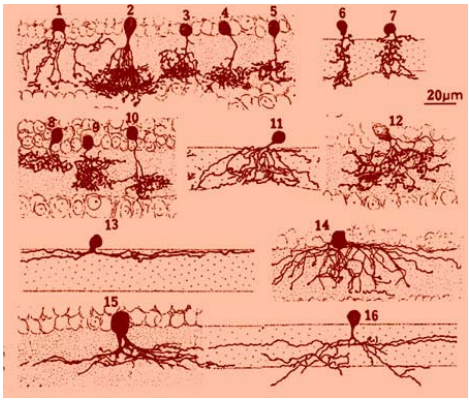
Bipolar cells

Bipolar cells collect the signals from the receptors and the horizontal cells, enhance edges by lateral inhibition, and transmit information to the ganglion cells.



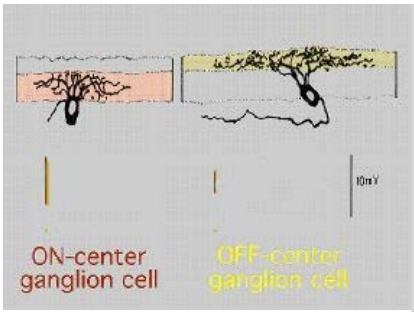
Amacrine cells

Amacrine cells make connections with bipolar and ganglion cells laterally across the retina. Their function is not yet well understood.



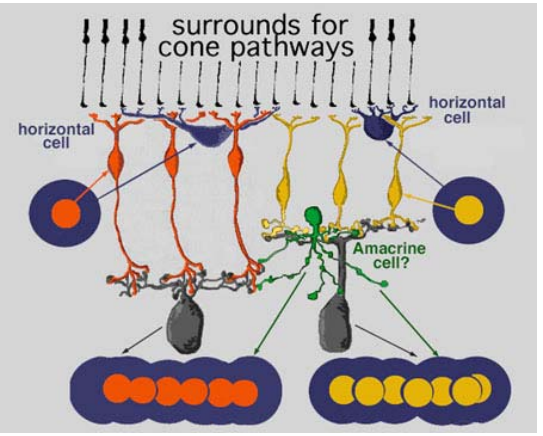


“ON” and “OFF” ganglion cells



Representation of a real number:

$$a = Max(0, a) - Max(0, -a)$$



Retinal output: parvo- and magno streams

Parvo stream:
- beta-type ganglion cells
- small
- slow
- sustained
- high spatial acuity
- color and form

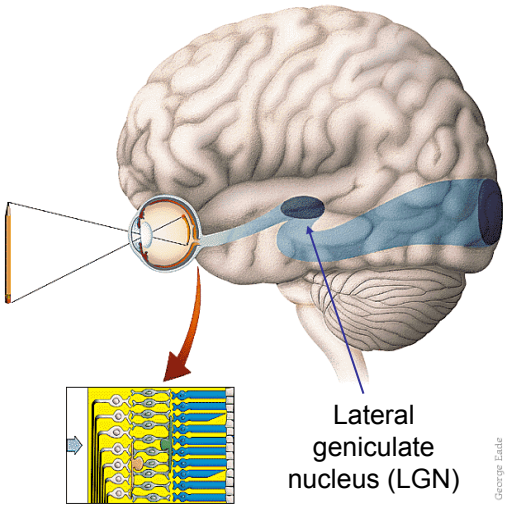
Magno stream:
- alpha-type ganglion cells
- large
- fast
- transient
- low spatial acuity
- motion

Retina: useful numbers

- area: 5x5 cm
- 1 deg = 0.3 mm
- no. of cones: $5x10^6$
- no. of rods: 10^8
- fovea: 1.5 mm (5.2 deg)
- foveola: 0.3 mm (1 deg)
- peak cone density: $1.6x10^5/mm^2$
- no. of optic nerve fibers: $1.6x10^6$ (human)
- receptor:ganglion cell ratio: 1:3 (foveal), 125:1

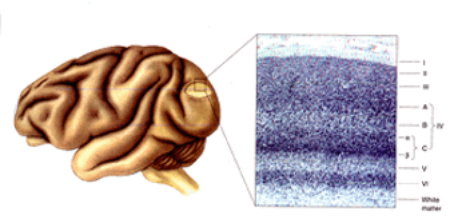


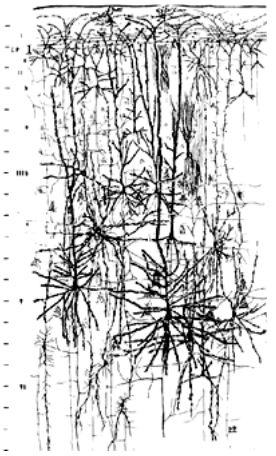
Visual pathway, lateral view





Primary visual cortex (V1)





Basic types of neurons in the primate V1:
- Spiny pyramidal cells
- Spiny stellate cells
- Smooth or sparsely spinous interneurons

Cortex : some numbers

- Total number of cortical neurons: 10^10;
- Density: 10^5 neurons / mm^3
- Number of macaque visual areas: 30
- Size of each area V1: 3cm by 8 cm
- Half of area V1 represents the central 10 deg (2% of the visual field)

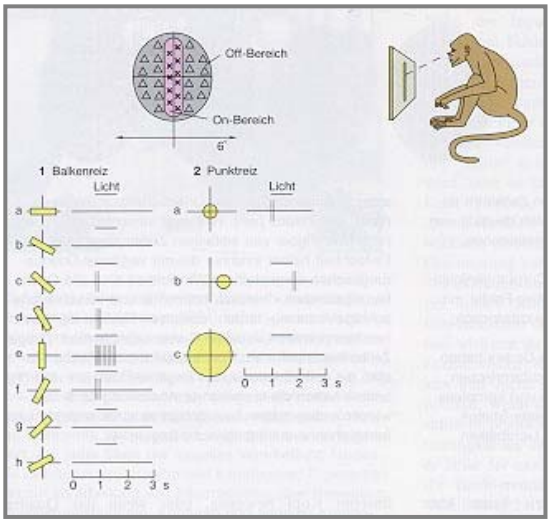
Visual cortex: overview of selected topics

- Orientation-selective neurons in V1
- End-stopped neurons in V1 and V2
- More complex features in IT
- Segregation of form, color, and motion



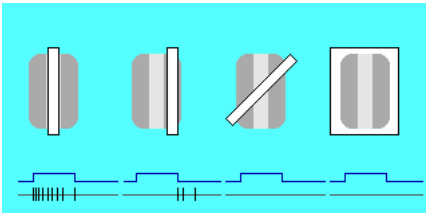
Visual cortex: orientation selectivity

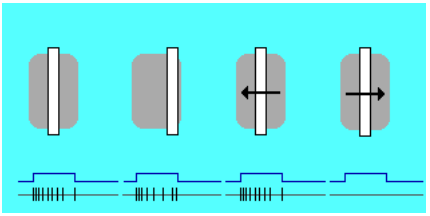
Hubel & Wiesel, 1962





Basic spatial properties of V1 neurons

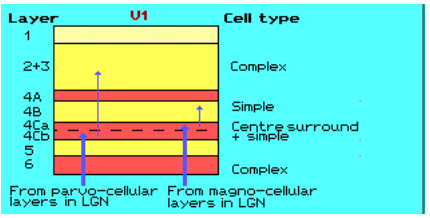




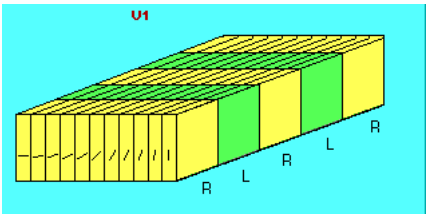
Simple cells:
- selective only to bars and edges and not to uniform illumination
- selective to orietation and position (spatial phase)

Complex cells:
- as above, but
- insensitive to exact position (spatial phase)

Basic organization of V1



6 **layers,** with layer 4C receiving the input from the LGN



Orientation colums with ocular dominance for the left (L) and right (R) eye respectively

Information and entropy

- Information (I)
  - Messages f_k (e.g. image grey levels), generated with probabilities p_k, $k=1,\ldots,N$
  - Information:

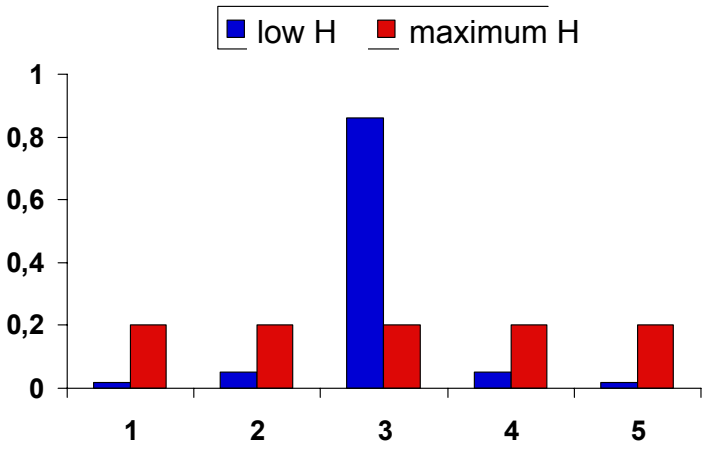
$$I_k = -\log_2 p_k \qquad [bits]$$

- Entropy (H)

$$H = -\sum_{k=1}^{N} p_k \log_2 p_k \qquad [bits / message]$$

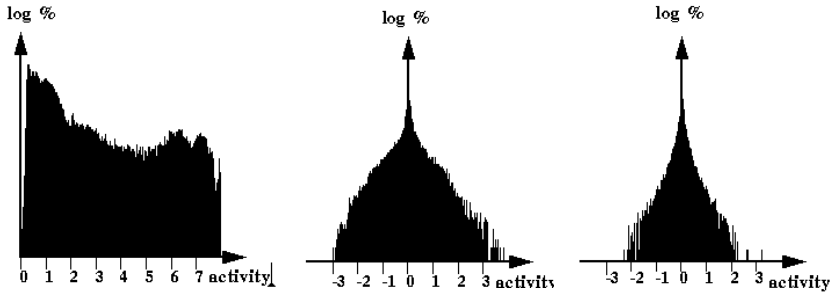


Examples for low and maximum entropy





The entropy is reduced along the visual pathway



Input

Retina
(lateral inhibition)

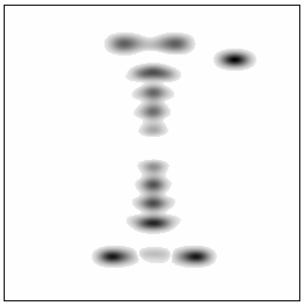
Cortical area V1
(orientation)

Visual cortex: endstopping

Input image



Simulated response



A significant number of V1 and V2 neurons are endstopped, i.e., they do not respond to straight edges or lines, but only to corners and line ends.

Visual cortex: endstopping

Endstopped neurons are, again, of two types:
- simple endstopped
- complex endstopped (or hypercomplex)

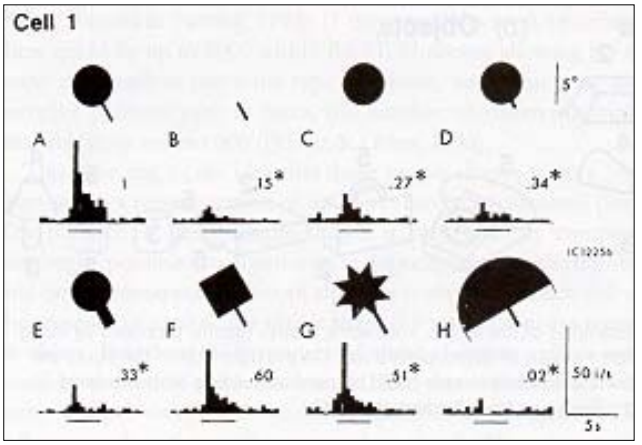
Endstopped representations further reduce the entropy and it has been shown that an endstopped representation can nevertheless be unique.

Models of endstopped neurons must be nonlinear and thus a theory of endstopping is more difficult and currently a topic of active research.



IT cortex of wake monkeys

Tanaka et al., 1991

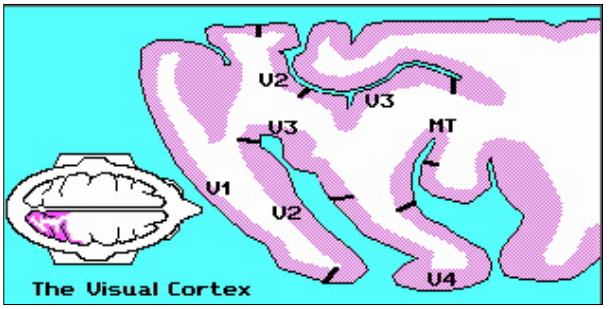


The selectivity of the neurons becomes increasingly specific.



Functional specialization

- Segregation of
  - form (IT)
  - motion (MT)
  - color (V4)



Paul Broca (1861)
• damage to a specific part of the brain disrupts language
↖ idea of functional specialization: different parts of the brain do different things

Cortical organization: specialized visual areas

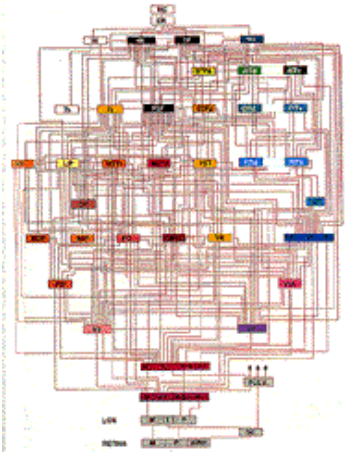
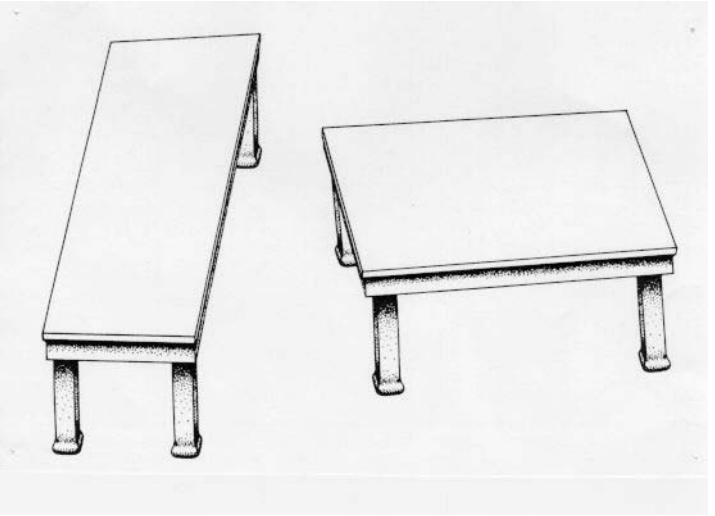


Diagram of visual pathway and different visual areas (Van Essen et.al.).

Note the massive connections that are both feed-forward and feed-back (bottom-up and top-down pathways).

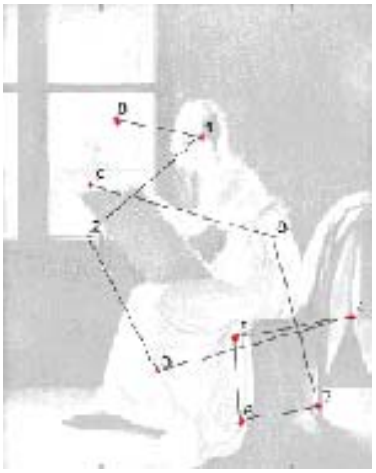
Example of top-down processing



Brief summary (passive vision)

- The retina is isotropically band-pass, but nonlinear due to gain control.
- Primary area V1 is orientation selective like a wavelet coder.
- In V2 many neurons are i2D selective, i.e., they do not respond to straight features.
- All of the above representations reduce the entropy of the visual input in case of natural images.
- In higher cortical areas selectivity becomes more and more specific.
- The brain is very silent, i.e. only few neurons are active.

Scan path



Types of eye movements:
- saccades (micro saccades)
- smooth pursuit

Saccades are fast (700 deg/sec) changes in gaze direction. Fixations between saccades last for about 300 ms.



Scan-path theory

- Human visual perception involves top-down processes, with a cognitive model actively driving foveated vision (from Stark and Ellis, 1981).

- Warren McCulloch:
  “The mind in the brain is like the grasp in the hand.”



Support for scan-path theory

- scan-path patterns are repetitive, and
- they vary more with subjects than with inputs
- the scan-path changes when the mental image of an ambiguous figure changes
- scan-paths are similar for real images and imagination



Visual attention and change blindness

- Experiments on change blindness have shown that what we see is very much limited by attention (remember the demonstrations Beppe gave).
- These experiments are therefore further evidence for top-down (as opposed to bottom-up) processes in vision.



References

- Wandell, B. A. (1995). Foundations of Vision . Sinauer Press.
- Hubel, D. H. (1988). Eye, Brain, and Vision . New York: W. H. Freeman.
- **http://www.visionscience.com**
- http://webvision.med.utah.edu

