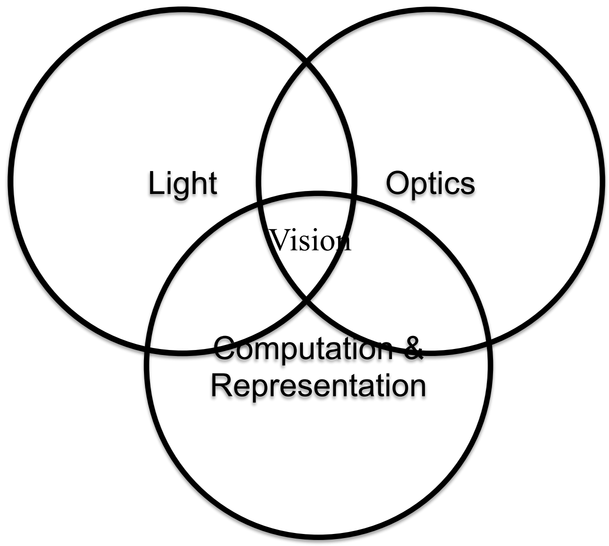
Biological Vision
a brief

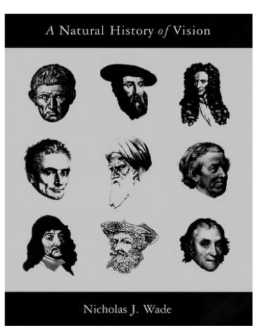
Ahmed Elgammal
Dept of Computer Science
Rutgers University

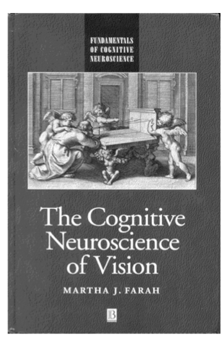
Outlines

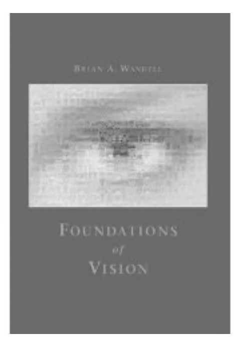
- How do we see: some historical theories of vision
- Biological vision: theories and results from psychology and cognitive neuroscience of vision.



Sources







- N. Wade “A Natural History of Vision” MIT press 1999
- Martha J. Farah “The Cognitive Neuroscience of Vision” Blackwell 2000
- Brian Wandell, Sinauer “Foundations of Vision”, Associates, Sunderland MA, 1995

How do we see

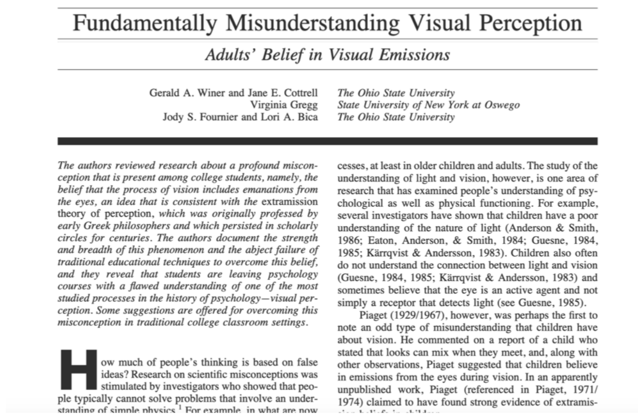
- Understanding Vision requires:
  - Understanding the physics of light and its interaction with objects
  - Understanding optics
  - Understanding how our brain works
- Historical: Two ancient opposing views of vision
  - Extramission
  - Intromission

How do we see - Historical view

- Extramission theories of vision (Euclid, Plato,…)
  - eye emits rays, and a person perceives the objects struck by these rays
  - Plato (350 B.C.) - from our eyes flows a light similar to the light of the sun
  - How that interact with luminance sources like, sun, etc ?
  - Ptolemy (ca. 90 -- ca. 168 AD): visual flux for our eyes + external light, study of refraction
  - "Therefore, when these three conditions concur, sight occurs, and the cause of sight is threefold: the light of the innate heat passing through the eyes, which is the principal cause, the exterior light kindred to our own light, which both acts and assists, and the light that flows from visible bodies, flame or color; without these the proposed effect [vision] cannot occur." [Chalcidius (ca. 300), middle ages].

How do we see - Historical view

- Extramission theories faced many difficulties
  - why do we see faraway objects instantaneously when we open our eyes?
    - the visual spirit that leaves the eyes is exceptionally swift
  - why don't the vision systems of different people looking at the same object interfere with each other?
    - they just don't
  - what if the eyes are closed when the visual spirit returns?
    - the soul has things timed perfectly - this never happens
  - Other non-material theories (spiritual, the "evil eye")



Fundamentally Misunderstanding Visual Perception

Adults' Belief in Visual Emissions

Gerald A. Winer and Jane E. Cottrell — *The Ohio State University*
Virginia Gregg — *State University of New York at Oswego*
Jody S. Fournier and Lori A. Bica — *The Ohio State University*

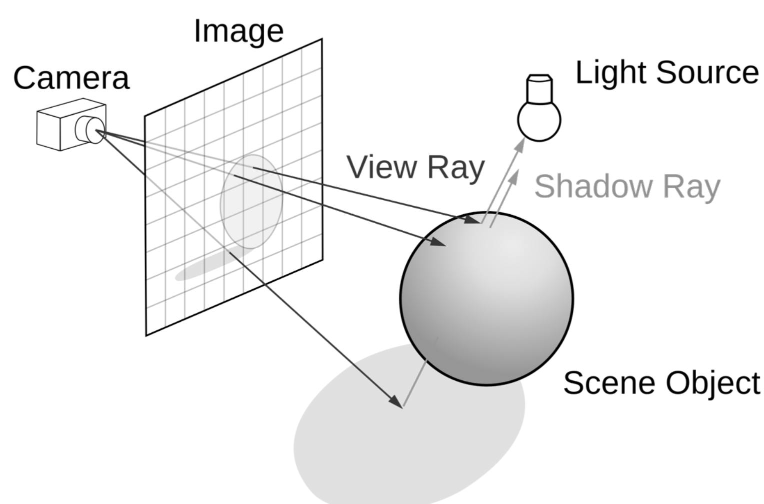
The authors reviewed research about a profound misconception that is present among college students, namely, the belief that the process of vision includes emanations from the eyes, an idea that is consistent with the extramission theory of perception, *which was originally professed by early Greek philosophers and which persisted in scholarly circles for centuries. The authors document the strength and breadth of this phenomenon and the abject failure of traditional educational techniques to overcome this belief, and they reveal that students are leaving psychology courses with a flawed understanding of one of the most studied processes in the history of psychology—visual perception. Some suggestions are offered for overcoming this misconception in traditional college classroom settings.*

How much of people's thinking is based on false ideas? Research on scientific misconceptions was stimulated by investigators who showed that people typically cannot solve problems that involve an understanding of simple physics.[1] For example, in what are now

cesses, at least in older children and adults. The study of the understanding of light and vision, however, is one area of research that has examined people's understanding of psychological as well as physical functioning. For example, several investigators have shown that children have a poor understanding of the nature of light (Anderson & Smith, 1986; Eaton, Anderson, & Smith, 1984; Guesne, 1984, 1985; Kärrqvist & Andersson, 1983). Children also often do not understand the connection between light and vision (Guesne, 1984, 1985; Kärrqvist & Andersson, 1983) and sometimes believe that the eye is an active agent and not simply a receptor that detects light (see Guesne, 1985).

Piaget (1929/1967), however, was perhaps the first to note an odd type of misunderstanding that children have about vision. He commented on a report of a child who stated that looks can mix when they meet, and, along with other observations, Piaget suggested that children believe in emissions from the eyes during vision. In an apparently unpublished work, Piaget (referenced in Piaget, 1971/1974) claimed to have found strong evidence of extramis-

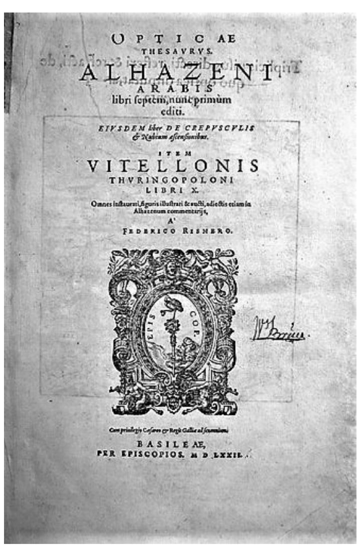
Winer et al. (2002) have found recent evidence that as many as 50% of American college students believe in emission theory.[3]

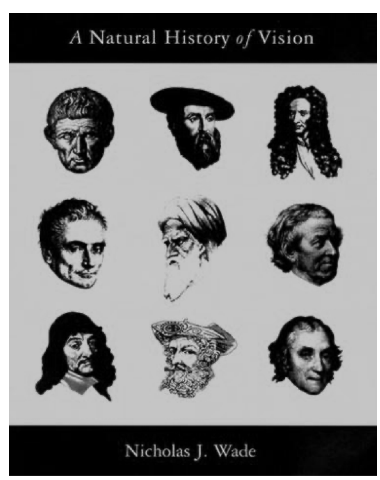


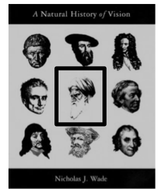
Ray Tracing in Computer Graphic

How do we see - Historical view

- Intromission theories of vision (Aristotle, Democritus, …)
  - Atomists: objects create “material images” (copies) that are transported through the atmosphere and enter the eye (Aristotle 330 B.C.)
    - but how do the material images of large objects enter the eye?
    - why don’t the material images of different objects interfere?
  - “light” travels from an object to the observer's eye, that’s why we see reflection in the eye pupil!



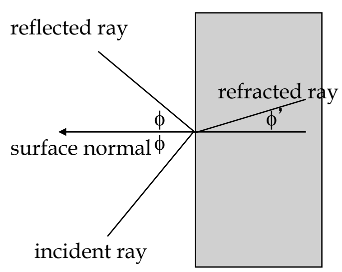




- Abu Ali al-Hassan ibn al-Hasan ibn al-Haytham (965-1040)
  - mercifully shortened to **Alhazen**
  - greatest optical scientist of the middle ages,
  - revolutionized the theory of optics "*Book of Optics" (7 volumes, translated to Latin in 1270)*
    - Light is a physical phenomenon (independent of vision)
    - Light radiates from self luminous bodies: sun, moon, light
    - Lights travel in straight lines
    - Concept of medium: transparent and opaque.
    - Light is refracted between two transparent medium
    - When light hits an object it irradiates in all directions.
    - pointillist theory of vision - we see a collection of points on the surfaces of objects
    - geometric theory to explain the 1-1 correspondence between the world and the image formed in our eyes

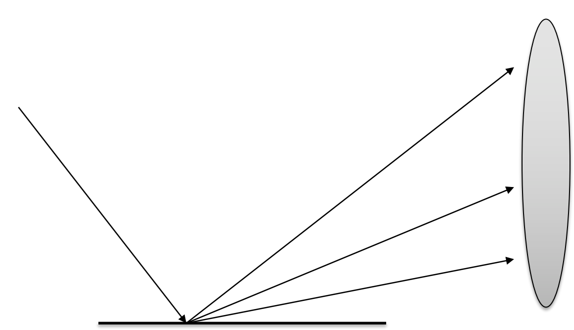
Lens and image formation

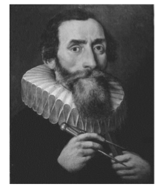
- Ray of light leaves the light source, and travels along a straight line
- Light hits an object and is
  - reflected and/or
  - refracted
- If the object is our lens, then the useful light for imaging is the refracted light



Ptolemy, Alhazen and refraction

- The phenomena of refraction was known to Ptolemy
- Alhazen's problem - since light from a surface point reaches the entire surface of the eye, how is it that we see only a single image of a point?
  - he assumed that only the ray that enters perpendicular to the eye affects vision
  - the other rays are more refracted, and therefore "weakened"
  - but in fact, the optical properties of the lens combine all of these rays into a single "focused" point under favorable conditions



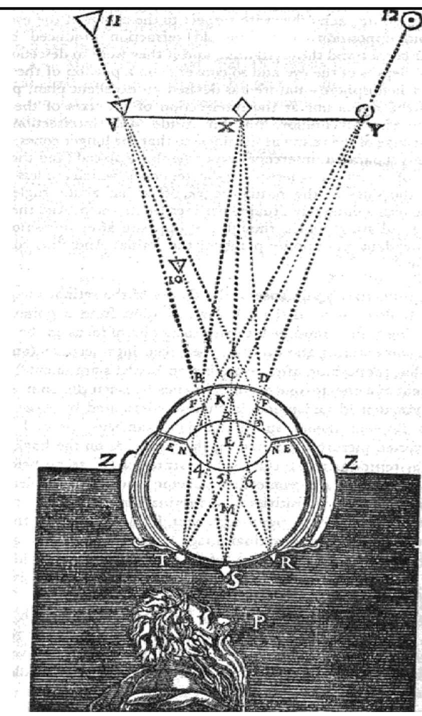


- Johannes Kepler (1571-1630)
- Founder of modern theories about optics and light.
  - Light has the property of flowing or being emitted by its source towards a distance place
  - From any point the flow of light takes place according to an infinite number of straight line.
  - Light itself is capable of advancing to the infinite
  - The lines of these emissions are straight and are called rays.

Kepler's retinal theory

Even though light rays from "many" surface points hit the same point on the lens, they approach the lens from different directions.

Therefore, they are refracted in different directions - separated by the lens

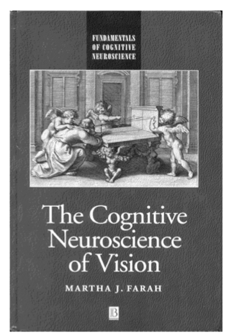


Modern theories of Vision

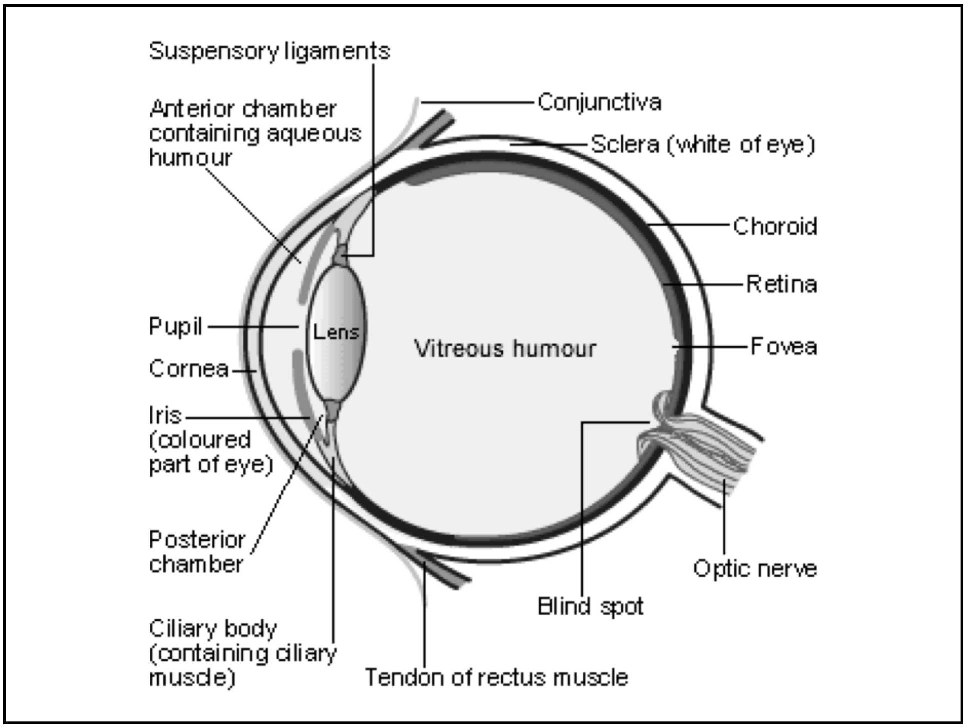
Three main streams contribute to our understanding of vision:

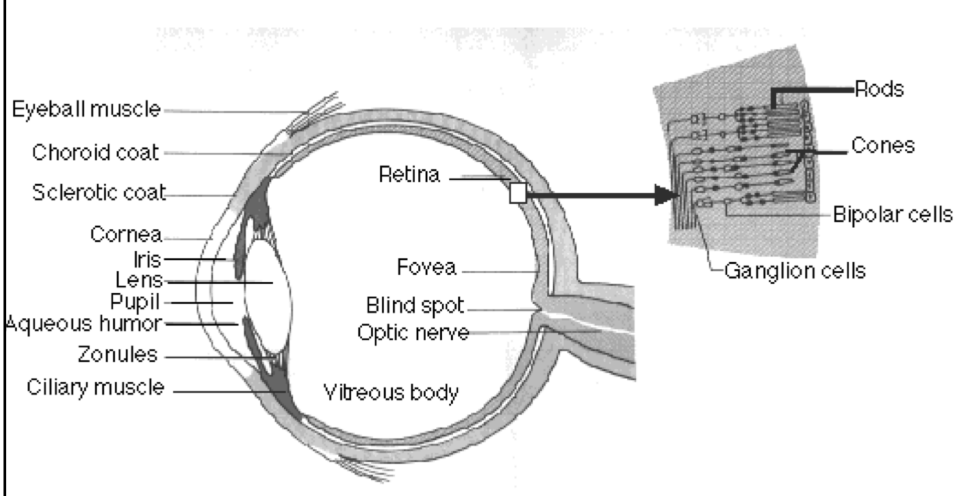
- Psychology of perception: functionalities
- Neurophysiology: explanations
- Computational vision: more problems

Biological Vision

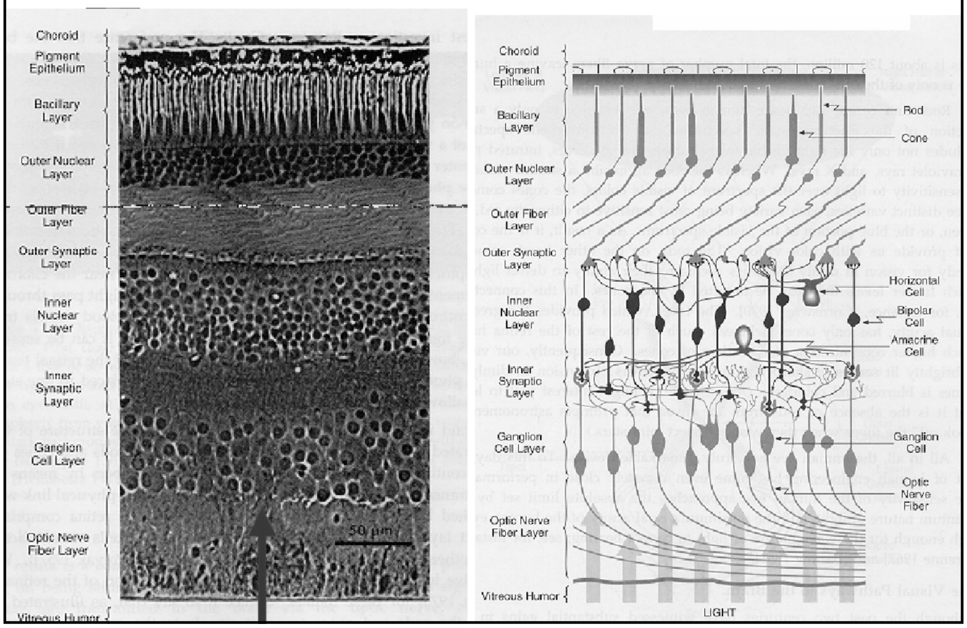


- Early vision:
  Parallelism. Multiplexing. Partitioning.
- High-level vision:
  Modularity.


Suspensory ligaments
Anterior chamber containing aqueous humour
Conjunctiva
Sclera (white of eye)
Choroid
Retina
Fovea
Pupil
Lens
Vitreous humour
Cornea
Iris (coloured part of eye)
Posterior chamber
Optic nerve
Blind spot
Ciliary body (containing ciliary muscle)
Tendon of rectus muscle


Eyeball muscle
Choroid coat
Sclerotic coat
Cornea
Iris
Lens
Pupil
Aqueous humor
Zonules
Ciliary muscle
Retina
Fovea
Blind spot
Optic nerve
Vitreous body
Rods
Cones
Bipolar cells
Ganglion cells

Retina



Retina

Three layers of cells:

- Receptor cells
- Collector cells
- Retinal ganglion cells

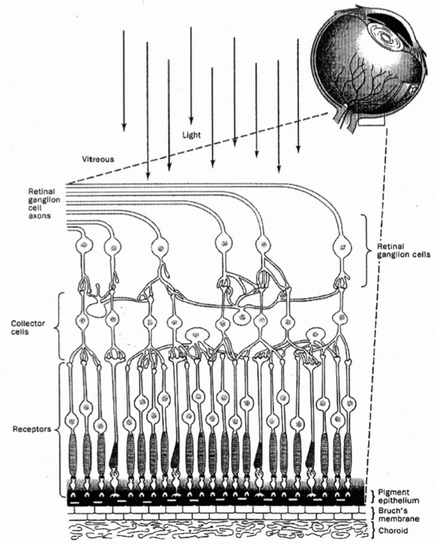
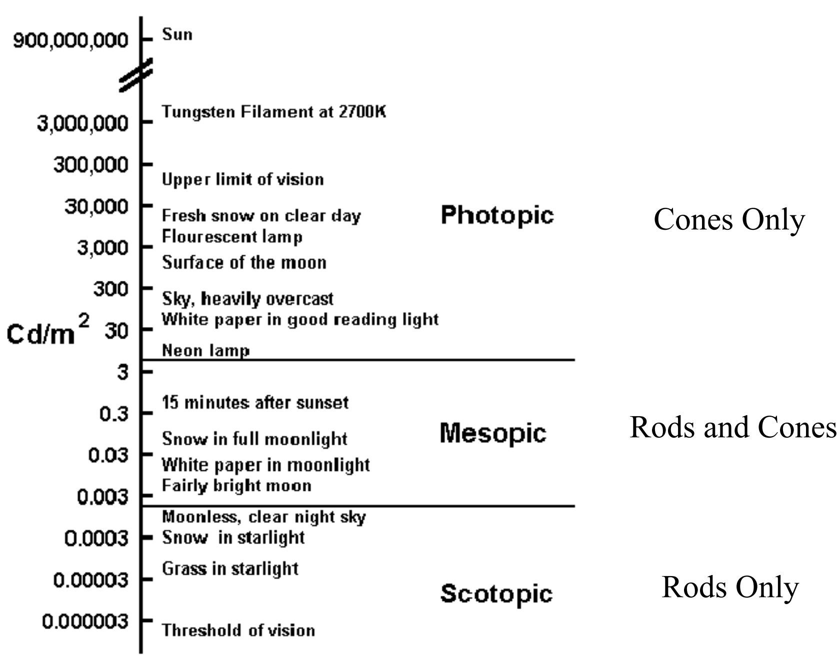


Figure 1.1 Cross-section through the retina, showing three functionally distinct layers of cells: receptor cells, collector cells, and ganglion cells.
From R. Sekuler and R. Blake, Perception, New York, McGraw-Hill, 1994.

Photoreceptor mosaics

- The retina is covered with a mosaic of photoreceptors
- Two different types of photoreceptors
- rods - approximately 100,000,000
- cones - approximately 5,000,000
- Rods
  - sensitive to low levels of light: scotopic light levels
- Cones
  - sensitive to higher levels of light: photopic light levels
- Mesopic light levels - both rods and cones active
- Difference in conversion to receptor cells.

Scotopic 10^{-2} to 10^{-6} cd/m2, Mesopic 10^{-2} to 1 cd/m², Photopic 1- to 10^{6} cd/m



https://www.visualexpert.com/Resources/nightvision.html

Rods & Cones: Conversion to Receptor Cells

Pooling (conversion) of the output of receptor cells:

Rods: several rods connect to each collector cell

Cones: limited pooling to collector cells

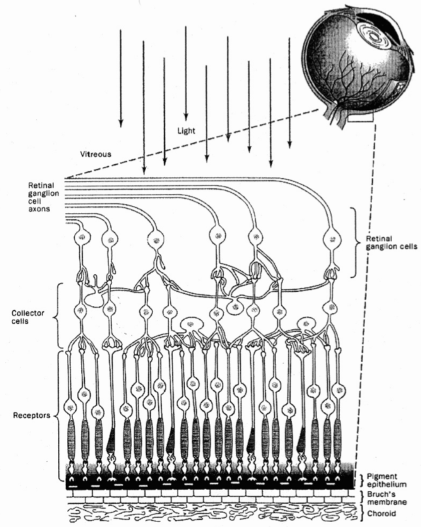


Figure 1.1 Cross-section through the retina, showing three functionally distinct layers of cells: receptor cells, collector cells, and ganglion cells.
From R. Sekuler and R. Blake, Perception, *New York, McGraw-Hill, 1994.*

Duplex retina

- Trade off: Sensitivity to light vs. spatial resolution.
- Two parallel systems:
  - One that favor sensitivity to light (Rods)
  - One that favor resolution (Cones)

Duplex Retina

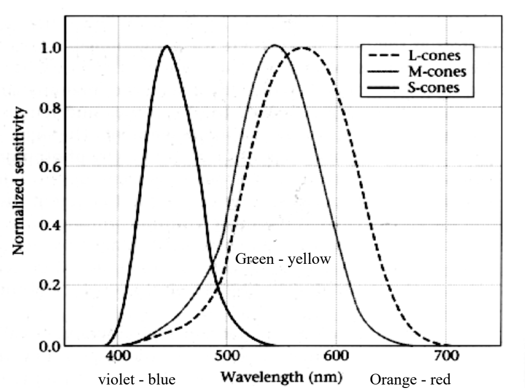
Trade off: Sensitivity to light vs. spatial resolution

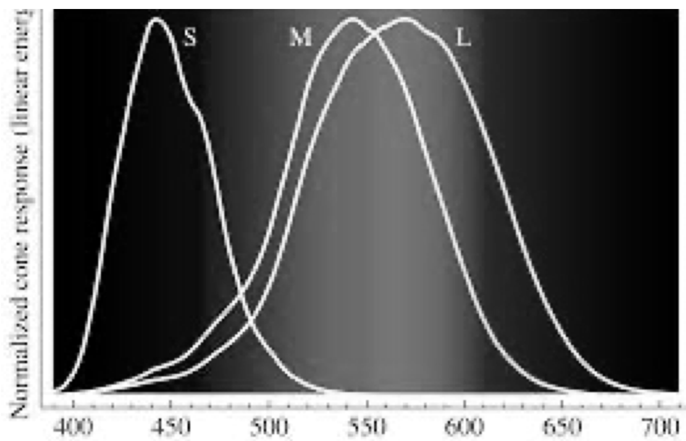
- Rods:
  - high sensitivity (sensitive to low levels of light: scotopic light levels)
  - extensive convergence onto collector & ganglion cells
  - ⇒ low resolution image of the world that persists even in low illumination condition
- Cones:
  - sensitive to higher levels of light: photopic light levels
  - much limited convergence
  - ⇒ High resolution image of the world in good illumination.

Cones and color

- Three different types of cones
  - they differ in their sensitivity to different wavelengths of light (blue-violet, green, yellow-red)

3.3 SPECTRAL SENSITIVITIES OF THE L-, M-, AND S-CONES in the human eye. The measurements are based on a light source at the cornea, so that the wavelength loss due to the cornea, lens, and other inert pigments of the eye plays a role in determining the sensitivity. Source: Stockman and MacLeod, 1993.



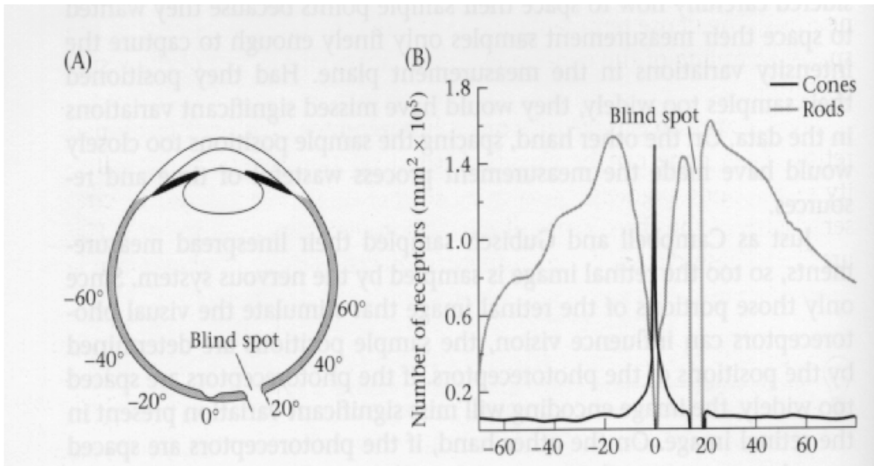


Cones and Color

- Example of a distributed representation
- Three different photopigments which absorbs different wavelengths of light to different degrees.
- Recall: Cones traded resolution for sensitivity (inactive in low light)

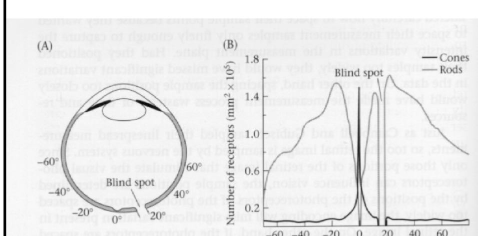
$\Rightarrow$ color blindness in low illumination

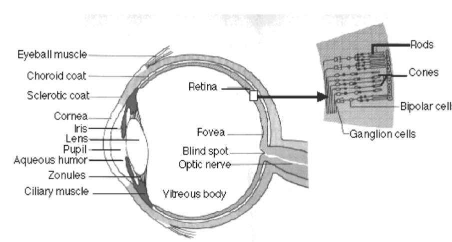
Photoreceptor mosaics



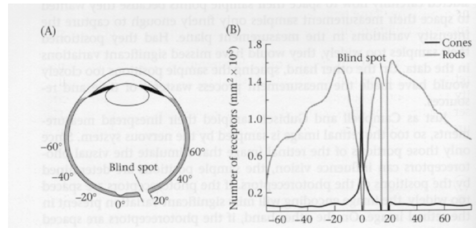
Photoreceptor mosaics

- Fovea is area of highest concentration of photoreceptors
- fovea contains no rods, just cones
- approximately 50,000 cones in the fovea
- cannot see dim light sources (like stars) when we look straight at them!





Blind spot



- Close left eye
- Look steadily at white cross
- Move head slowly toward and away from figure
- At a particular head position the white disk completely disappears from view

Retina

Three layers of cells:
- Receptor cells
- Collector cells
- Retinal ganglion cells

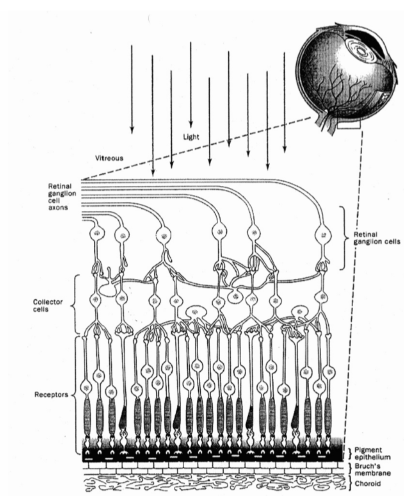
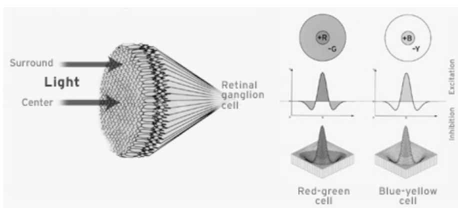
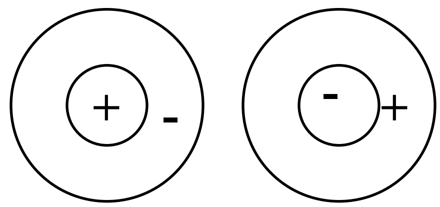


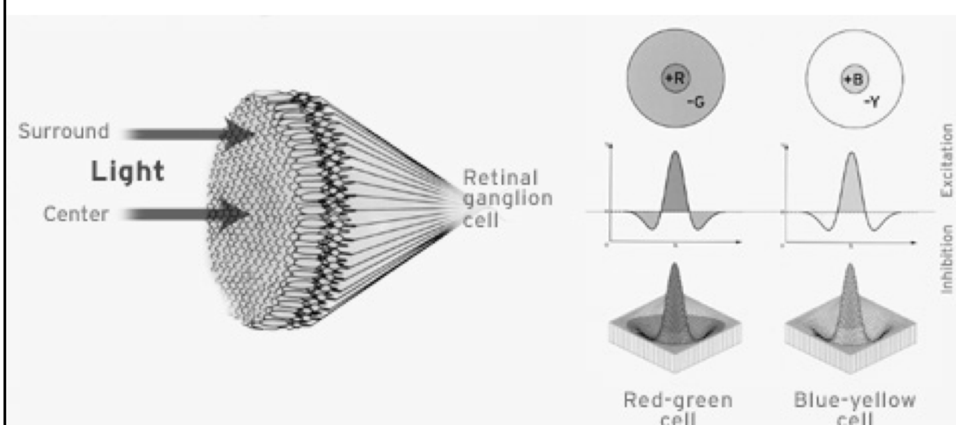
Figure 1.1 Cross-section through the retina, showing three functionally distinct layers of cells: receptor cells, collector cells, and ganglion cells.
From R. Sekuler and R. Blake, Perception, New York, McGraw-Hill, 1994.

Retinal Ganglion cells

- First stage of visual processing
- Function: Absolute levels of illumination is replaced by a retinotopic map of “differences”
- How: center-surrounding organization of their receptive fields:
  - on-center (off-surrounding) cells
  - off-center (on-surrounding) cells

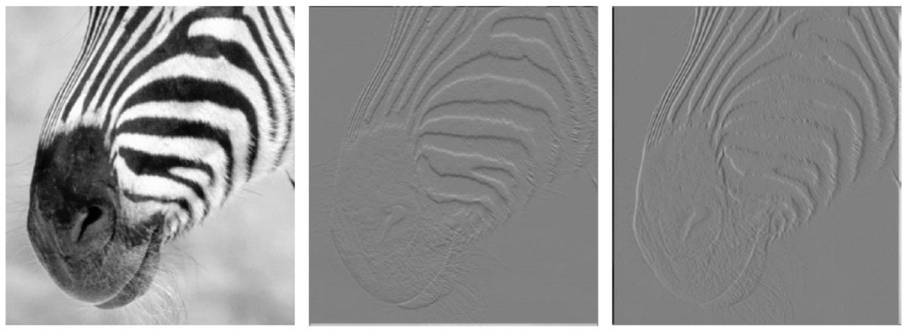




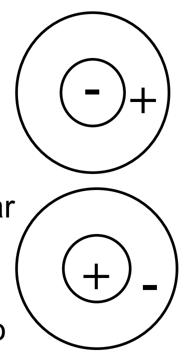


Retinal Ganglion cells

- How a spatial difference image might look like ?



Retinal Ganglion cells



- Why: objects are not associated with any particular brightness, but with differences in brightness between themselves and the background.
- The differences can be amplified without having to represent the enormous range of values that would result from the amplification of absolute values.
- ⇒ groundwork for perception of objects.

Retinal Ganglion cells

Another partition:
- M and P cells:
- Feeds into the M and P channels (magnocellular and parvocellular layers in LGN)
- Tradeoff: temporal vs. spatial resolution

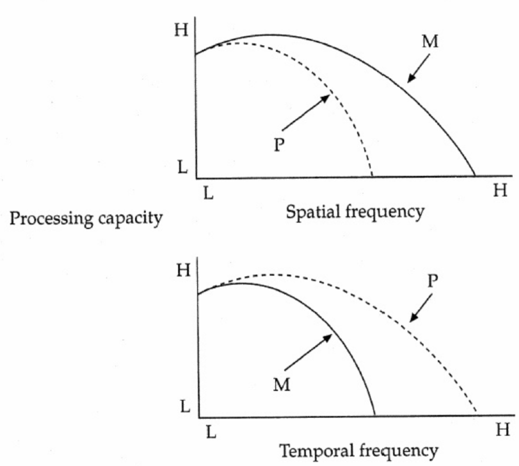


Figure 1.4 Schematic depiction of the complementarity of M and P channels for spatial and temporal information.
From P. H. Schiller, "Parallel pathways in the visual system," in B. Gulyas, D. Ottoson, and P. E. Roland (eds), Functional Organisation of the Human Visual Cortex, *Oxford, Pergamon Press, 1993.*

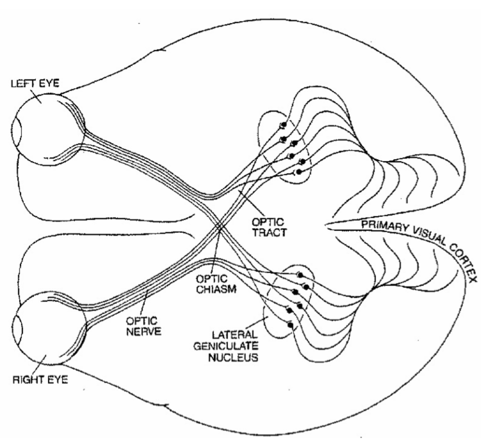
Retinal Ganglion cells

- Tradeoff: temporal vs. spatial resolution
- M cells: input from large number of photoreceptors ⇒ good light sensitivity, good temporal resolution (can sample easily from large input), low spatial resolution.
- P cells: input from small number of photoreceptors ⇒ good spatial resolution, poor temporal resolution.

- M cells are larger, faster nerve conduction velocities, responses are more transient.
- P cells show color sensitivity, M cells don't.

M cells: Temporal resolution, fast
⇒ motion perception, sudden stimulus.
P cells: Spatial resolution + color
⇒ Color, texture, patterns (major role in object perception).

- Bundle of axons leaving the eye: optic nerve
- Split into a number of pathways
- Retinotopy organization



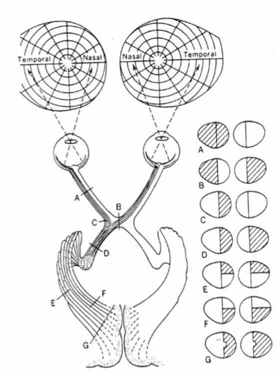
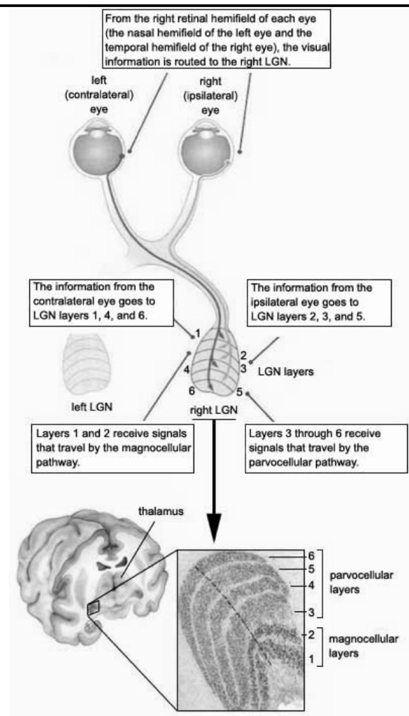


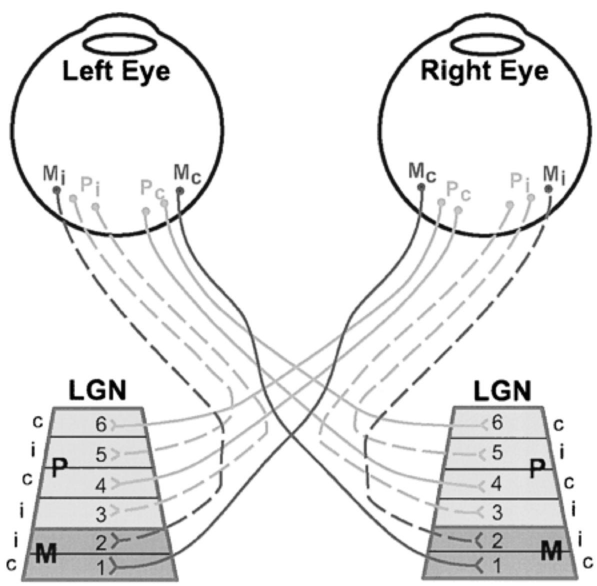
Figure 1.6 Correspondences between location of lesion within the visual system and pattern of visual field defects.
From J. Homans, A Textbook of Surgery, Springfield, IL, Charles C. Thomas, 1945.

The lateral geniculate nucleus (LGN):

- One LGN in each cerebral hemisphere
- Magnocellular layers (lower two) : feed from M-cells
  - Best temporal resolution
- Parvocellular layers (four) : feed from P-cells
  - Best spatial resolution, wavelength sensitivity
- Another example of division of labor and multiplexing
- Neurons in all layers show center-surrounding organization
- Retinotopy: all layers keep retinotopic organization of the image
- Feed back from visual cortex
- What is LGN for ? Gate or Amplify visual input, attention ?



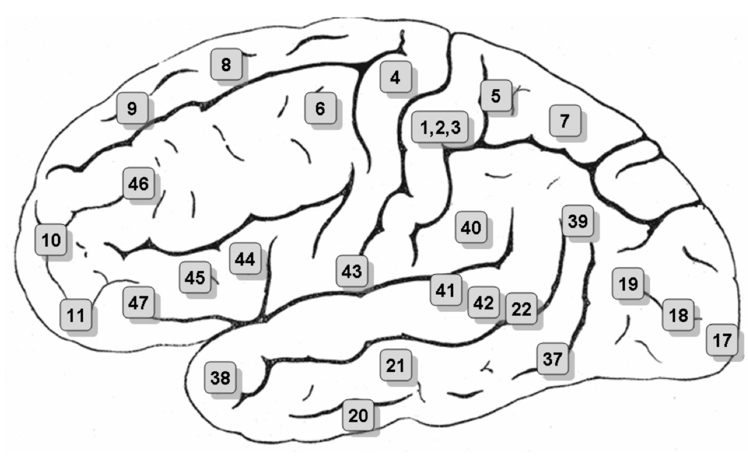
http://thebrain.mcgill.ca



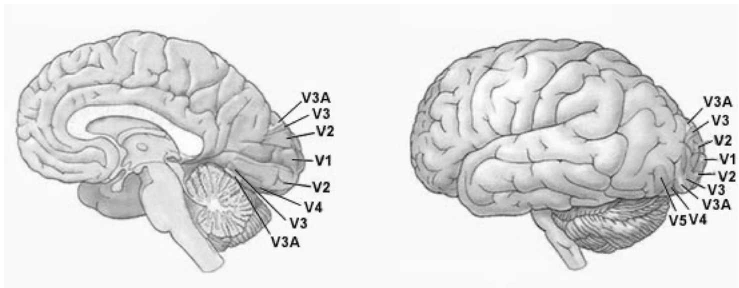
Claire Meissirel et al PNAS vol 94 No 11, 1997

The primary visual cortex

- Also known as area 17, Striate cortex, V1
- David H. Hubel & Torsten N. Wiesel : Nobel prize
- Three types of cells (1962):
- Center-surrounding
- Simple cells:
  - Like center-surrounding with elongated excitatory and inhibitory regions.
  - edges at particular location and orientation.
- Complex cells:
  - more abstract type of visual information. Partially independent of location within the visual field.



Brodmann numbering



http://thebrain.mcgill.ca

- David H. Hubel & Torsten N. Wiesel : Nobel prize 1981
  - Discovering of simple and complex cells, their functions and anatomical organization
  - Pioneering the technique for single cell recording in cortex

- Simple cells:
  - Like center-surrounding with elongated excitatory and inhibitory regions.
  - edges at particular location and orientation.

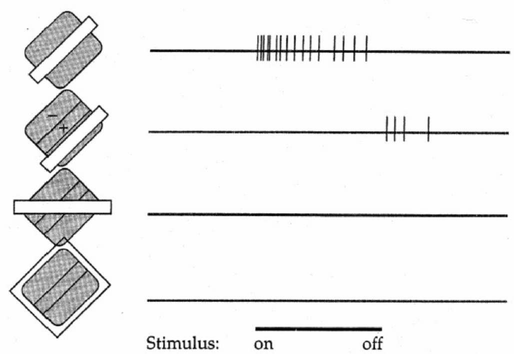


Figure 1.8 Bar stimuli of different orientations (left) and the responses they evoke from a simple cell in primary visual cortex (right).
From D. H. Hubel, Eye, Brain, and Vision, *New York, Scientific American Library, 1988.*

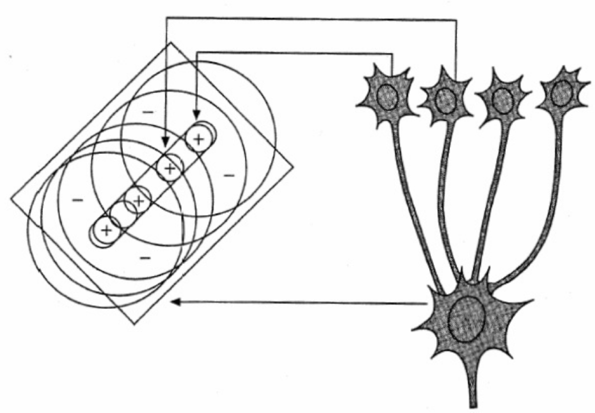


Figure 1.9 Illustration of the idea that simple cells result form the feedforward convergence of a set of center–surround cells.
Adapted from D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," Journal of Physiology, *160, 1962.*

The primary visual cortex

- Feed forward sequence or hierarchy of visual processing
  Center-surrounding → Simple → Complex
- Cells' responses become increasingly specific w.r.t the form of the stimulus (ex. oriented edges or bars)
- Increasingly general w.r.t viewing conditions (from just one location to a range of locations)
- These dual-trends are essential for object recognition
- can respond to specific form (like familiar face) generalized over changes in size, orientation, viewpoint
- More recent research: lateral interaction plays important role (Gilbert 1992)

- spatial arrangement of cells

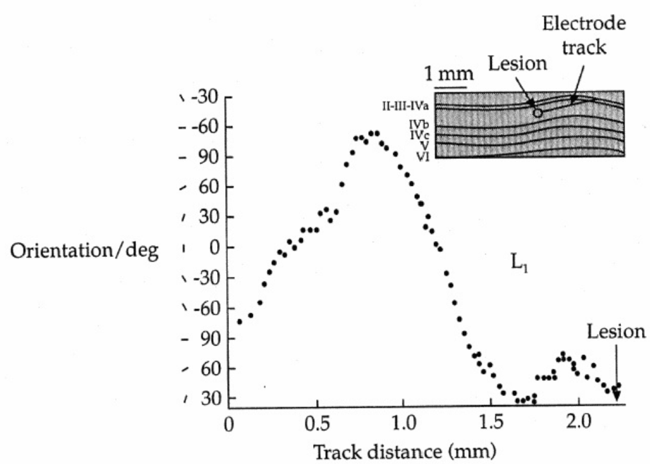


Figure 1.10 The orderly progression of orientation preference as a function of electrode position during an oblique penetration of primary visual cortex.
From R. C. Reid, "Vision," in M. J. Zigmond, F. E. Bloom, S. C. Landis, J. L. Roberts, and L. R. Squire (eds), Fundamental Neuroscience, *San Diego, Academic Press, 1999.*

- Organization and orientation selectivity (why and how ?):
- spatial arrangement of cells for minimizing the distance between neurons representing similar stimulus along three different stimulus dimensions:
  - Eye of origin
  - Orientation
  - Retinotopic location
- Hebb rule : neurons that fire together wire together.

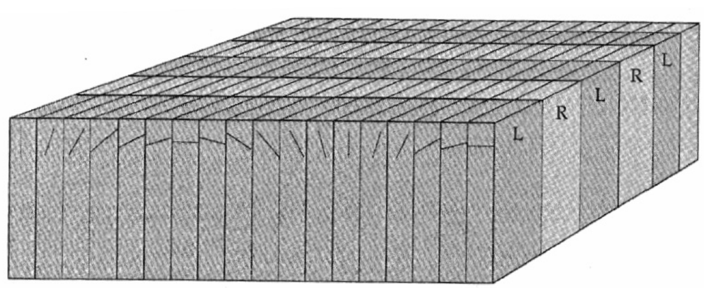


Figure 1.11 Idealized depiction of the organization of orientation selectivity and ocular dominance in primary visual cortex.
Adapted from D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," Journal of Physiology, *160, 1962.*

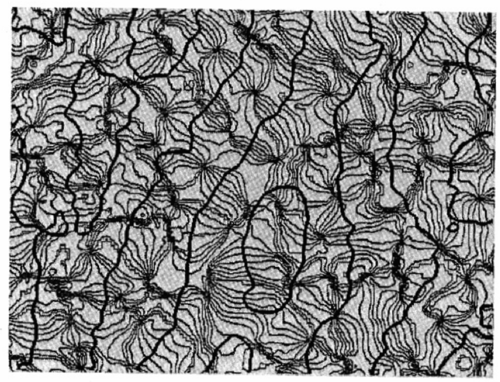


Figure 1.12 Reconstruction of the actual relations between orientation columns (gray) and ocular dominance columns (black) in primary visual cortex.
From K. Obermayer and G. G. Blasdel, "Geometry of orientation and ocular dominance columns in monkey striate cortex," Journal of Neuroscience, *13, 1993.*

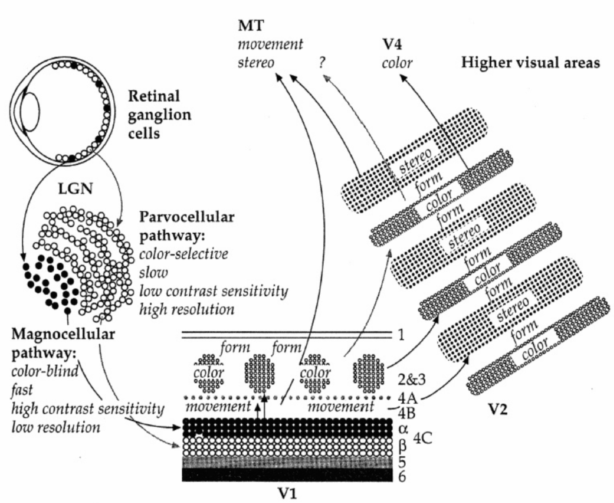
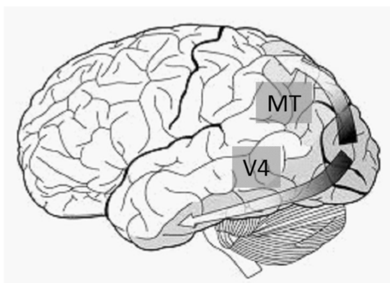
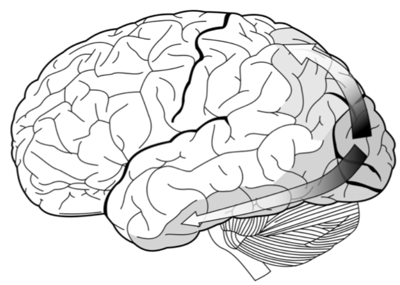


Figure 1.15 Original version of Livingstone and Hubel's hypothesis, according to which anatomically and physiologically defined subdivisions of the visual system formed independent streams of processing from the retina through extrastriate cortex, with each stream responsible for distinct perceptual functions.
From M. S. Livingstone and D. H. Hubel, "Segregation of form, color, movement, and depth: anatomy, physiology, and perception," Science, *240, 1988. Copyright 1988 American Association for the Advancement of Science.*

- Ventral-Dorsal Streams
- Ventral Stream: “What” Pathway: Object Recognition
- Dorsal Stream: “Where” & “How” Pathway: Localization, Motion, Actions





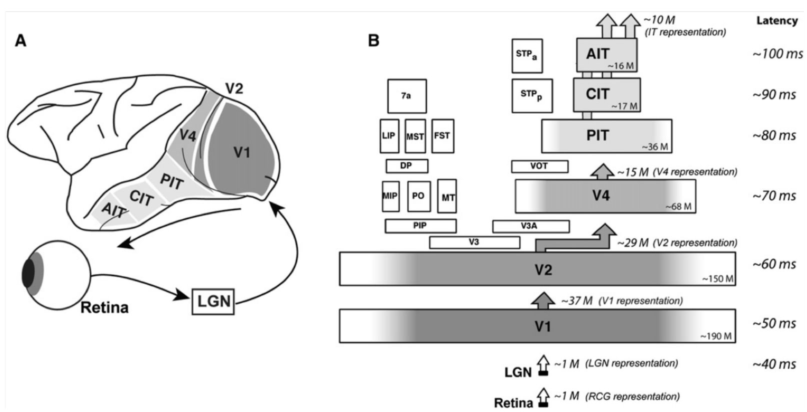


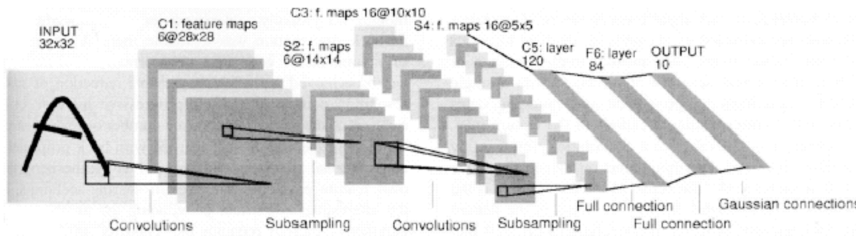
Figure 3. The Ventral Visual Pathway
(A) Ventral stream cortical area locations in the macaque monkey brain, and flow of visual information from the retina.
(B) Each area is plotted so that its size is proportional to its cortical surface area (Felleman and Van Essen, 1991). Approximate total number of neurons (both hemispheres) is shown in the corner of each area (M = million). The approximate dimensionality of each representation (number of projection neurons) is shown above each area, based on neuronal densities (Collins et al., 2010), layer 2/3 neuronal fraction (O’Kusky and Colonnier, 1982), and portion (color) dedicated to processing the central 10 deg of the visual field (Brewer et al., 2002). Approximate median response latency is listed on the right (Nowak and Bullier, 1997; Schmolesky et al., 1998).

From: DiCarlo et al 2011 “How Does the Brain Solve Visual Object Recognition?”

- IT (Inferior temporal cortex)
  - PIT (posterior IT): crude retinotopy exists
  - CIT (Central IT): retinotopy not reported
  - AIT (anterior IT): retinotopy not reported
- IT representation: some degree of invariance to position, scale, clutter

Modern theories of vision

- Reconstructionist: Marr
  - Internal reconstruction of the 3D world as the central representation.
- Hierarchy of feature detectors: Edelman
  - “Bug detector” in the frog retina (Lettvin 1959)

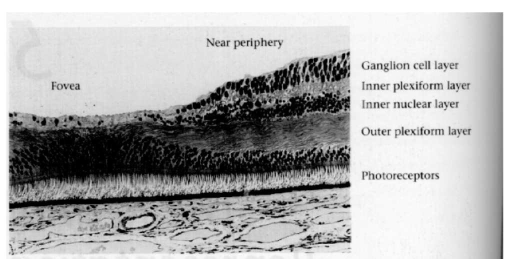


Example of a biologically motivated recognition system
A convolutional neural network, LeNet; the layers filter, subsample, filter, subsample, and finally classify based on outputs of this process.

Figure from "Gradient-Based Learning Applied to Document Recognition", Y. Lecun et al Proc. IEEE, 1998 copyright 1998, IEEE

The human eye

- Limitations of human vision
  - Blood vessels and other cells in front of photoreceptors
  - shadows cast on photoreceptors
  - non-uniform brightness



5.1 THE HUMAN RETINA, seen here in cross-section, is a thin sheet of neural tissue that lines the back of the eye. In the periphery and near periphery, the retina is a multilayered structure. The cornea and lens would be at the top of this picture, so in the periphery light must pass through the several retinal layers before being absorbed by the photoreceptors. In the fovea, the retina consists of only a single layer of photoreceptors, as the neurons responsible for carrying the responses of the foveal cones are displaced to the side, out of the light path. Micrograph courtesy of Anita Hendrickson.

The human eye

- Limitations of human vision
  - the image is upside-down!
  - high resolution vision only in the fovea
    - only one small fovea in man
    - other animals (birds, cheetas) have different foveal organizations
  - blind spot