CS 534: Computer Vision

Spring 2004
Ahmed Elgammal
Dept of Computer Science
Rutgers University

Human Vision

Outline

- How do we see: some historical theories of vision
- Human vision: results from 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 "Foundations of Vision", Associates, Sunderland MA, 1995
- Slides by Prof Larry Davis at UMD

How do we see?

- Extromissive theories of vision
  - Plato (350 B.C.) - from our eyes flows a light similar to the light of the sun
  - "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, middle ages].
  - Other non-material theories (spiritual, the "evil eye")

How do we see?

- Extromissive 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

How do we see?

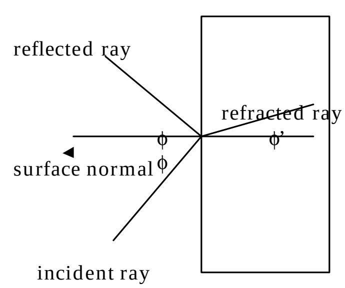
- Intromissive theories of vision
  - objects create "material images" that are transported through the atmosphere and enter the eye (Aristotle 330 B.C.)
    - but how do the material images of large objects enter the eye?

How do we see?

- Abu Ali al-Hassan ibn al-Hasan ibn al-Haytham (1040)
  - mercifully shortened to Alhazen
  - greatest optical scientist of the middle ages
    - Self luminous bodies: sun, moon, light
    - Lights travel in straight lines
    - When light hits an object it irradiates every place.
    - Concept of medium: transparent and opaque.
    - 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

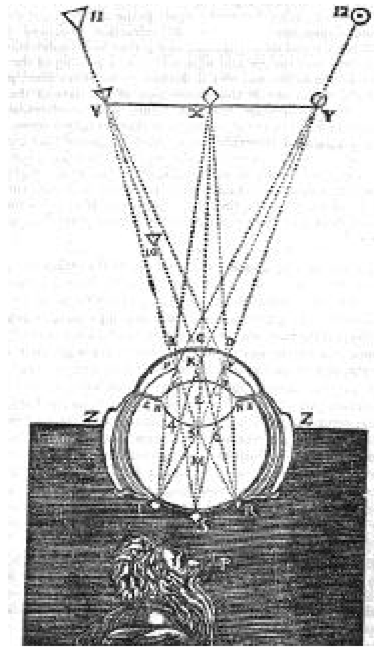
- The phenomena of refraction was known to Ptolemy (150 B.C.)
- Alhazen's problem - since light from a surface point reaches the entire surface of the eye, how is it that we see only a single image of a point?
  - he assumed that only the ray that enters perpendicular to the eye effects vision
  - the 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:

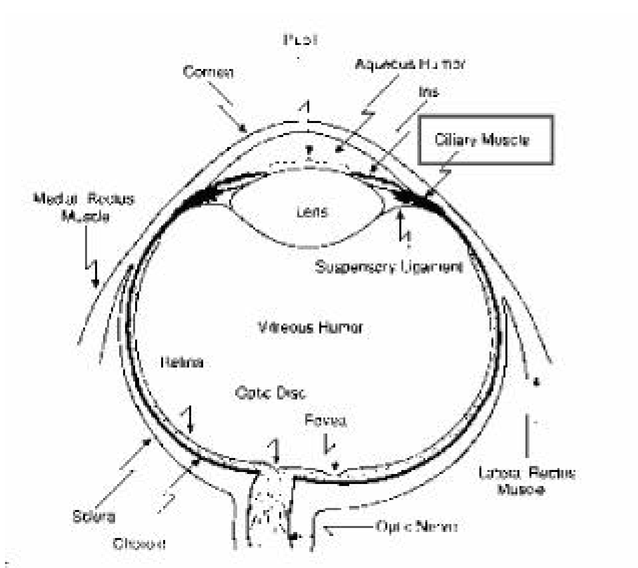
- Psychology of perception: functionalities
- Neurophysiology: explanations
- Computational vision: more problems

- Early vision:

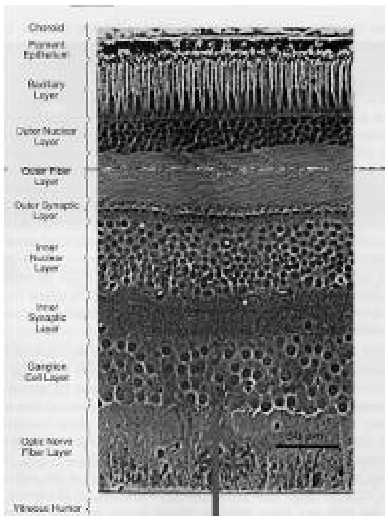
  Parallelism. Multiplexing. Partitioning.

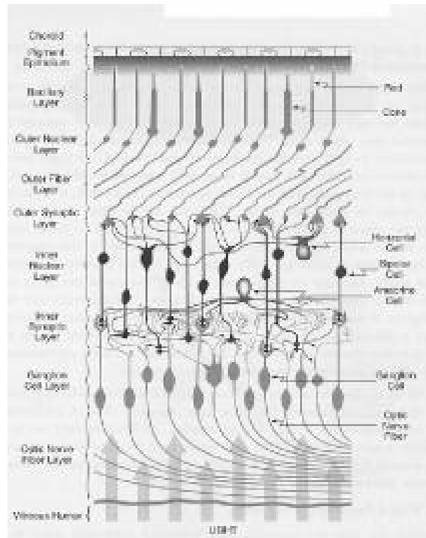
- High-level vision:

  Modularity.



Retina





Photoreceptor mosaics

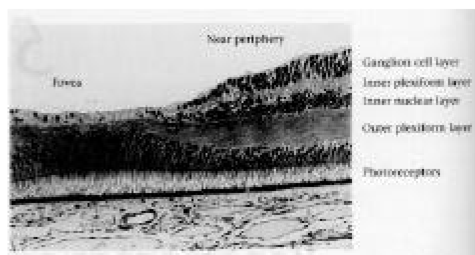
- The retina is covered with a mosaic of photoreceptors
- Two different types of photoreceptors
- rods - approximately 100,000,000
- cones - approximately 5,000,000
- Rods
  - sensitive to low levels of light: scotopic light levels
- Cones
  - sensitive to higher levels of light: photopic light levels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

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!
- TV camera photoreceptor mosaics
  - nearly square mosaic of approximately 800X640 elements for complete field of view

The human eye

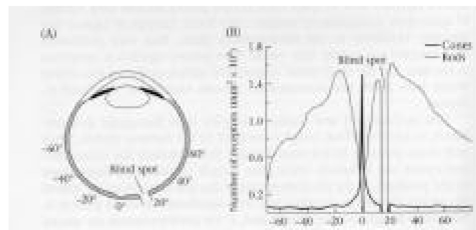
- Limitations of human vision
  - Blood vessels and other cells in front of photoreceptors
  - shadows cast on photoreceptors
  - non-uniform brightness



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



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

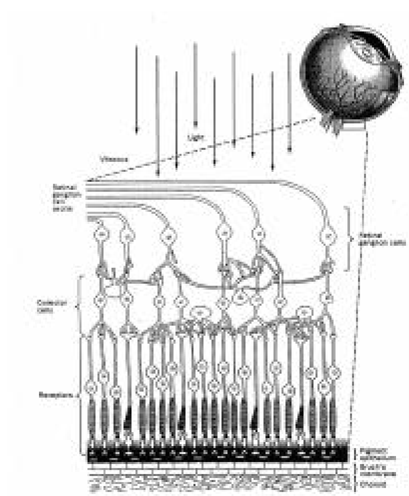
Cones, CCD's and space

- How much of the world does a cone see?
  - measured in terms of visual angle
  - the eye lens collects light over a total field of view of about 100°
  - each cone collects light over a visual angle of about 1.47×10^{-4} degrees, which is about 30 seconds of visual angle
- How much of the world does a single camera CCD see
  - example: 50° lens
  - 50/500 gives about 10^{-1} degrees per CCD

Retina

Three layers of cells:

- Receptor cells
- Collector cells
- Retinal ganglion cells



Cross-section through the retina, showing three functionally distinct layers of cells: receptor cells, collector cells, and ganglion cells.

Martha J. Farah "The Cognitive Neuroscience of Vision" Blackwell 2000

Duplex retina

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

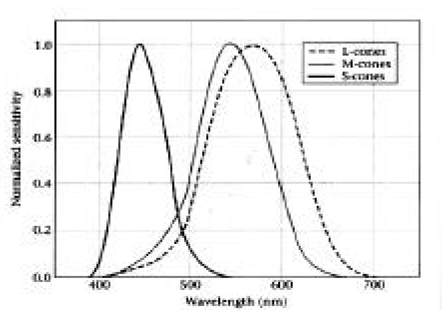
Duplex Retina

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.

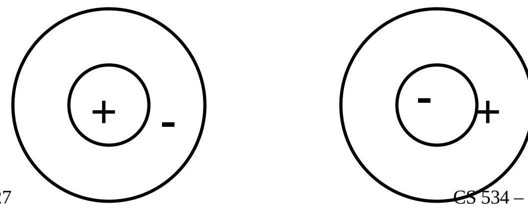
Cons and Color

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

⇒ color blindness in low illumination

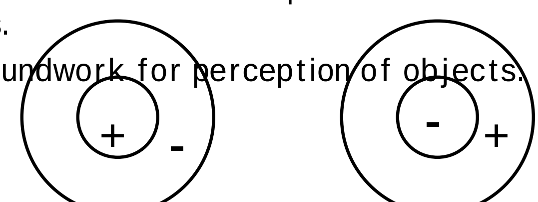
Retinal Ganglion cells

- 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

- 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
- 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.

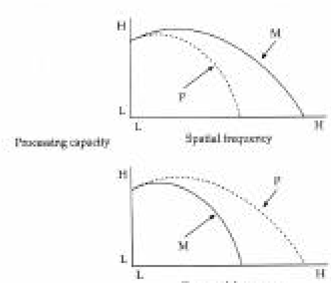


Figure 1.4 Schematic depiction of the complementarity of M and P channels for spatial and temporal information.
From P. H. Schiller, "Parallel pathways in the visual system," in B. Gulyas, D. Ottoson, and P. E. Roland (eds), Functional Organization of the Human Visual Cortex, Oxford, Pergamon Press,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

Martha J. Farah "The Cognitive Neuroscience of Vision" Blackwell 2000

Retinal Ganglion cells

Tradeoff: temporal vs. spatial 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

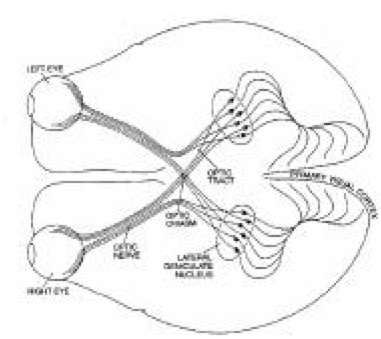
⇒ motion perception, sudden stimulus.

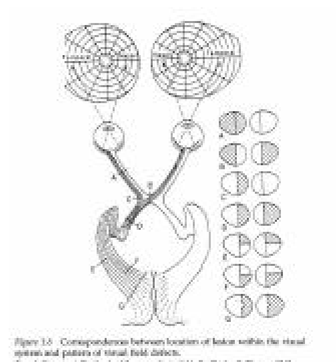
- P cells: Spatial resolution

⇒ Color, texture, patterns (major role in object perception).

Visual Path Ways

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





Martha J. Farah "The Cognitive Neuroscience of Vision" Blackwell 2000

The lateral geniculate nucleus (LGN):

- One LGN in each cerebral hemisphere
- Magnocellular layers (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 in LGN and beyond: all layers keep retiontopic organization of the image
- What is LGN for ? Amplify visual input ?

The primary visual cortex

- 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.

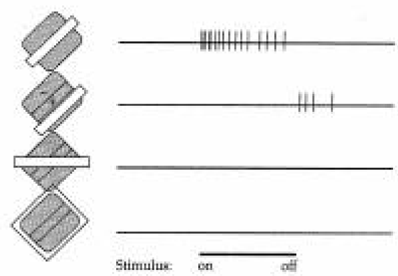


Figure 1.8 Bar stimuli of different orientations (left) and the responses they evoke from a simple cell in primary visual cortex (right).
From D. H. Hubel, Eye, Brain, and Vision, New York, Scientific American Library, 1988.

Martha J. Farah "The Cognitive Neuroscience of Vision" Blackwell 2000

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, view point)
- More recent research: lateral interaction plays important role (Gilbert 1992)

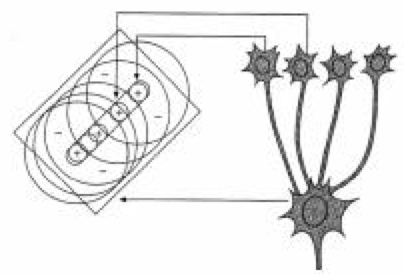


Figure 1.9 Illustration of the idea that simple cells result form the feedforward convergence of a set of center–surround cells.
Adapted from D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," Journal of Physiology, 160, 1962.

Martha J. Farah "The Cognitive Neuroscience of Vision" Blackwell 2000

- 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.

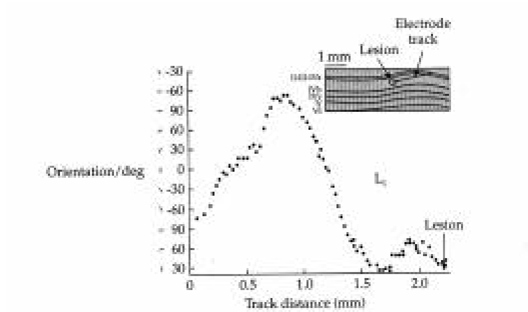


Figure 2.10 The orderly progression of orientation preference as a function of electrode position during an oblique penetration of primary visual cortex.
From R. C. Reid, "Vision," in M. J. Zigmond, F. E. Bloom, S. C. Landis, J. L. Roberts, and L. R. Squire (eds), Fundamental Neuroscience, *San Diego, Academic Press, 1999.*

Martha J. Farah "The Cognitive Neuroscience of Vision" Blackwell 2000

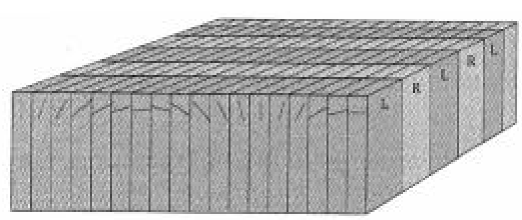


Figure 2.11 Idealized depiction of the organization of orientation selectivity and ocular dominance in primary visual cortex.
Adapted from D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," Journal of Physiology, 160, 1962.

Martha J. Farah "The Cognitive Neuroscience of Vision" Blackwell 2000

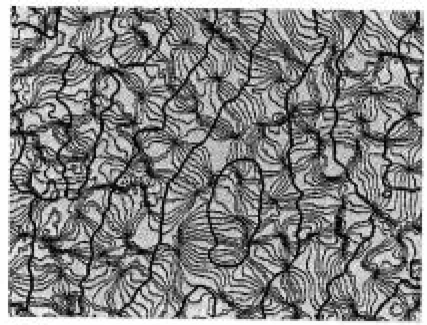


Figure 1.12 Reconstruction of the actual relations between orientation columns (gray) and ocular dominance columns (black) in primary visual cortex.
From K. Obermayer and G. G. Blasdel, "Geometry of orientation and ocular dominance columns in monkey striate cortex," Journal of Neuroscience, 13, 1993.

Martha J. Farah "The Cognitive Neuroscience of Vision" Blackwell 2000

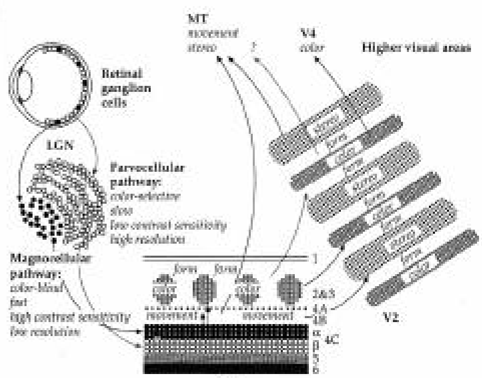


Figure 1.15 Original version of Livingstone and Hubel's hypothesis, according to which anatomically and physiologically defined subdivisions of the visual system formed independent streams of processing from the retina through extrastriate cortex, with each stream responsible for distinct perceptual 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.*

Martha J. Farah "The Cognitive Neuroscience of Vision" Blackwell 2000

Color cameras

- Two types of color cameras
  - Single CCD array
    - in front of each CCD element is a filter - red, green or blue
    - color values at each pixel are obtained by hardware interpolation
      - subject to artifacts
      - lower intensity quality than a monochromatic camera
      - similar to human vision
  - 3 CCD arrays packed together, each sensitive to different wavelengths of light

3 CCD cameras