Outlines

• How do we see: some historical theories of vision
• Human vision: results from cognitive neuroscience of vision.
• Some optics.
Sources

- N. Wade "A Natural History of Vision" MIT press 1999
- Slides by Prof Larry Davis at UMD

How do we see?

- Two opposing views of vision
  - Intromissive
  - Extromissive
- Intromissive theories of vision
  - 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?
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 (ca. 300), 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?

- 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
 Biological Vision

- Early vision:
- High-level vision:
  Modularity.
Retina

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

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

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, 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 $1.47 \times 10^{-4}$ 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

Cones and color

- Three different types of cones
  - they differ in their sensitivity to different wavelengths of light (blue-violet, green, yellow-red)
**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)
  \[ \Rightarrow \text{color blindness in low illumination} \]

**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

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

- Bundle of axons leaving the eye: optic nerve
- Split into a number of pathways
- Retinotopy organization
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: all layers keep retinotopic organization of the image
- Feed back from visual cortex
- What is LGN for? Gate or Amplify visual input?

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

- David H. Hubel & Torsten N. Wiesel: Nobel prize
  - 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 2.8) Bar stimuli of different orientations (left) and the responses they evoke from a single cell in primary visual cortex (right).


(Figure 2.9) Illustration of the idea that simple cells result from the feedfoward convergence of a set of center-surround cells.

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)

- spatial arrangement of cells

Figure 3.10  The orderly progression of orientation preference as a function of electrode position during an oblique penetration of primary visual cortex.
• 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 3.22 Reconstruction of the actual relations between orientation columns (grey) and ocular dominance columns (black) in primary visual cortex.


Figure 3.25 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 accessory cortex, with each stream responsible for distinct perceptual functions.

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
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
**CCD cameras**

- Charge-coupled device (CCD)
- Image is read one row at a time
- This process is repeated several times per second (frame rate)

**Color cameras**

- **CCD cameras:** Charge-coupled-device (CCD)
- 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

Snell’s law

- If $\phi$ is the angle of incidence and $\phi'$ is the angle of refraction then

$$\frac{\sin \phi}{\sin \phi'} = \frac{\nu}{\nu'}$$

where $\nu$ and $\nu'$ are the refractive indices of the two media.

- Refractive index is the ratio of speed of light in a vacuum to speed of light in the medium.

Refractive indices:
- Glass - 1.52
- Water - 1.333
- Air - 1.000 - mercifully, or there would be no physical basis for vision!
Applying Snell's Law twice

- Pass light into and out of a prism (symmetric piece of glass)
  - By combining many infinitesimally small prisms we form a convex lens that will bring all of the refracted rays incident from a given surface point into coincidence at a point behind the lens
  - If the image or film plane is placed that distance behind the lens, then that point will be in focus
  - If the image plane is in front of or behind that ideal location, the image of that point will be out of focus

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
The structure of eyes - compound eyes

- Many (small) animals have compound eyes
  - each photoreceptor has its own lens
  - images seen by these eyes are equally sharp in all directions
  - images seen by these eyes are equally “bright” in all directions when viewing a field of constant brightness
  - examples: flies and other insects
- But these eyes do not “scale” well biologically

Thin lenses

- The optical behavior is determined by:
  - optical axis going through the lens center \( O \) and perpendicular to the lens center plane
  - the left and right focus \((F_l, F_r)\) located a distance \( f \), called the focal length, from the lens center
Thin lenses

• The shape of the lens is designed so that all rays parallel to the optical axis on one side are focused by the lens on the other side:
  - Any ray entering the lens parallel to the optical axis on one side goes through the focus on the other side.
  - Any ray entering the lens from the focus on one side, emerges parallel to the optical axis on the other side.

Lensmaker's equation

• Lens imperfections might cause these rays not to intersect at a point
  - deviations in shape from the ideal lens
  - material imperfections that might cause the refractive index to vary within the lens
• Lensmaker's equation
  - relates the distance between the point being viewed and the lens to the distance between the lens and the ideal image (where the rays from that point are brought into focus by the lens)
  - Let $P$ be a point being viewed that is not too far from the optical axis.
    • $Z + f$ is the distance of $P$ from the lens along the optical axis
    • The thin lens focuses all the rays from $P$ onto the same point, the image point, $p$
Lensmakers equation

- $p$ can be determined by intersecting two known rays, $PQ$ and $PR$.
  - $PQ$ is parallel to the optical axis, so it must be refracted to pass through $F_r$.
  - $PR$ passes through the left focus, so emerges parallel to the optical axis.
- Note two pairs of similar triangles
  - $PS/FR = Z/f$ and $PS/F_r = f/z$
  - But $PS/FR = f/z$ and $F_r = f^2$
  - Let $Z' = Z + f$ and $z' = z + f$

\[
\frac{1}{Z'} + \frac{1}{z'} = \frac{1}{f}
\]

Lensmaker's equation

- Notice that the distance behind the lens, $z'$, at which a point, $P$, is brought into focus depends on $Z'$, the distance of that point from the lens
  - familiar to us from rotating the focus ring of any camera.
- In real lenses, system is designed so that all points within a given range of distances [$Z_1, Z_2$] are brought into focus at the same distance behind the lens center.
  - This is called the depth of field of the lens.
Lensmaker’s equation

\[ \frac{1}{Z'} + \frac{1}{z'} = \frac{1}{f} \]

- Note:
  - as \( Z' \) gets large, \( z' \) approaches \( f \)
  - as \( Z' \) approaches \( f \), \( z' \) approaches infinity

Optical power and accommodation

- Optical power of a lens - how strongly the lens bends the incoming rays
  - short focal length lens bends rays significantly
  - it images a point source at infinity (large \( d_s \)) at distance \( d \) behind the lens. The smaller \( f \), the more the rays must be bent to bring them into focus sooner.
  - optical power is \( 1/f \), measured in meters. The unit is called the diopter
  - Human vision: when viewing faraway objects the distance from the lens to the retina is \( .017m \). So the optical power of the eye is 58.8 diopters
Accommodation

• From the lensmaker's equation and the known optical power of the eye \(1/f\) we can calculate the ideal image distance of a point being viewed as a function of the distance to that point.

  - Sources at > 1 meter are imaged at the same distance.
  - Sources closer than 1 m are imaged at different distances.

Accommodation

• How does the human eye bring nearby points into focus on the retina?
  - By increasing the power of the lens.
  - Muscles attached to the lens change its shape to change the lens power.
  - Accommodation: adjusting the focal length of the lens.
  - Bringing points that are nearby into focus causes faraway points to go out of focus.
  - Depth-of-field: range of distances in focus.

• Physical cameras - mechanically change the distance between the lens and the image plane.
Lensless imaging systems - pinhole optics

- Pinhole optics focuses images
  - without lens
  - with infinite depth of field
- Smaller the pinhole
  - better the focus
  - less the light energy from any single point

Diffraction

- Two disadvantages to pinhole systems
  - light collecting power
  - diffraction
- Diffraction
  - when light passes through a small aperture it does not travel in a straight line
  - it is scattered in many directions
  - process is called diffraction and is a quantum effect
- Human vision
  - at high light levels, pupil (aperture) is small and blurring is due to diffraction
  - at low light levels, pupil is open and blurring is due to lens imperfections
Diffraction and pinhole optics

2.18 DIFFRACTION LIMITS THE QUALITY OF PINHOLE OPTICS. These three images of a bulb filament were made using pinholes with decreasing size. (A) When the pinhole is relatively large, the image rays are not properly converged, and the image is blurred. (B) Reducing the size of the pinhole improves the focus. (C) Reducing the size of the pinhole further worsens the focus, due to diffraction. From Ruechart, 1958.

Complications of color

• Spectral composition of light
  - Newton’s original prism experiment
  - Light decomposed into its spectral components

4.1 NEWTON’S SUMMARY DRAWING of his experiments with light. Using a point source of light and a prism, Newton separated sunlight into its fundamental components. By reconverging the rays, he also showed that the decomposition is reversible.
Complications of color

• Why does the prism separate the light into its spectral components?
  - prism bends different wavelengths of light by different amounts
    • refractive index is a function of wavelength
    • shorter wavelengths are refracted more strongly than longer wavelengths

<table>
<thead>
<tr>
<th>Wavelength</th>
<th>Color (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>700</td>
<td>Red</td>
</tr>
<tr>
<td>610</td>
<td>Orange</td>
</tr>
<tr>
<td>580</td>
<td>Yellow</td>
</tr>
<tr>
<td>540</td>
<td>Green</td>
</tr>
<tr>
<td>480</td>
<td>Blue</td>
</tr>
<tr>
<td>400</td>
<td>Violet</td>
</tr>
</tbody>
</table>

* - viewed in isolation

4.4 THE SPECTRAL POWER DISTRIBUTION of two important light sources are shown: (left) blue skylight and (right) a tungsten bulb.
Chromatic aberration

- Chromatic aberration
  - The prism effect of focusing different wavelengths of light from the same point source at different distances behind the lens
  - when incident light is a mixture of wavelengths, we can observe a chromatic fringe at edges
  - accommodation can bring any wavelength into good focus, but not all simultaneously
  - human visual system has other mechanisms for reducing chromatic aberration
  - color cameras have similar problems

![Chromatic Aberration Graph](image)