COLLEGE OF ENGINEERING
ELECTRICAL ENGINEERING & COMPUTER SCIENCE
UNIVERSITY OF MICHIGAN

Foundations of Computer Vision
Introduction

EECS 598-08 Fall 2014

http://web.eecs.umich.edu/~jjcorso/t/598F14

Instructor: Jason Corso
jjcorso@eecs.umich.edu

Materials on these slides have come from many sources in addition to myself; I am infinitely grateful to these, especially Greg Hager and Silvio Savarese.

Source: http://www.ted.com/talks/sebastian_thrun_google_s_driverless_car?

Your turn!

- Form groups of 4.
- You are a new "Google Computer Vision Design Team"!
- Select one of the three tasks.
- Design a model / algorithm for the task using the knowledge you currently have about computer vision, if any.
- Discuss for 5 minutes.


Lane Following


Pedestrian Avoidance


Stoplight Detection

What is Computer Vision?

- **Ballard and Brown (1982)**
  - The construction of explicit, meaningful description of physical objects from images.
- **Trucco and Verri (1998)**
  - computing properties of the 3D world from one or more digital images.
- **Stockman and Shapiro (2001)**
  - To make useful decisions about real physical objects and scenes based on sensed images.
- **Forsyth and Ponce (2003)**
  - ...extracting descriptions of the world from pictures or sequences of pictures.
- **Extraction of information from visual content.**

Some related terms

- **Image Processing:** the study of the properties of operators that produce images from other images
  - we will touch on image filtering and related operators from image processing

- **Machine Vision:** a somewhat outdated term which now tends to refer to industrial vision applications where (usually) a single camera is used to solve a structured inspection task
  - the "reverse CAD" model

- **Pattern Recognition:** typically refers to the recognition of structures in 2D images (usually without reference to any underlying 3D information).

- **Photogrammetry:** the science of measurement though non-contact sensing, e.g. terrain maps from satellite images. Usually is more focused on accuracy issues than interpretation.

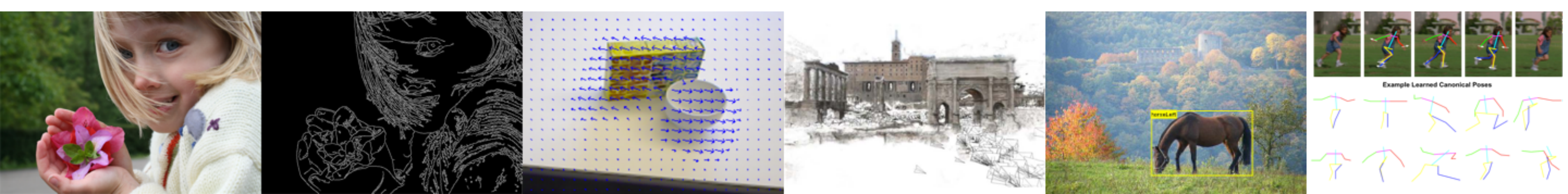
Source: G.D. Hager Slides.

What information is in images?

Source: GD Hager slides.

Five classes of information to be extracted

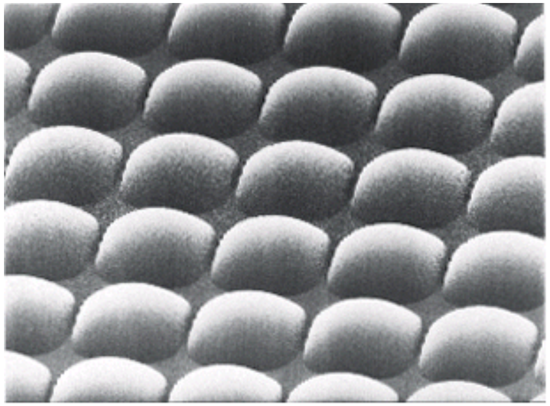
1. **Early Processes**
   – extracting basic features, edges, contours, and segmentation.
2. **Motion Tracking**
   – extracting movement, optical flow, tracking, and filtering.
3. **Shape**
   – extracting 3D structure, epipolar geometry, stereo, SFM, shape from X.
4. **Objects**
   – extracting objects, detection, recognition, and matching.
5. **Actions**
   – extracting actions, space-time localization, detection

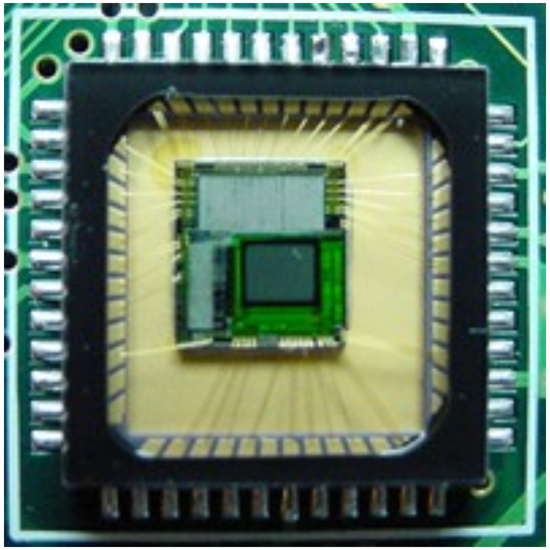


Extracting that information is harder than you think

Sensors and image formation

- Basic image sensing process:
  - photons hit a detector
  - the detector becomes charged
  - the charge is read out as brightness
- Sensor types:
  - CCD (charge-coupled device)
    - most common
    - high sensitivity
    - high power
    - cannot be individually addressed
    - blooming
  - CMOS
    - simple to fabricate (cheap)
    - lower sensitivity, lower power
    - can be individually addressed





Source: G.D. Hager Slides.

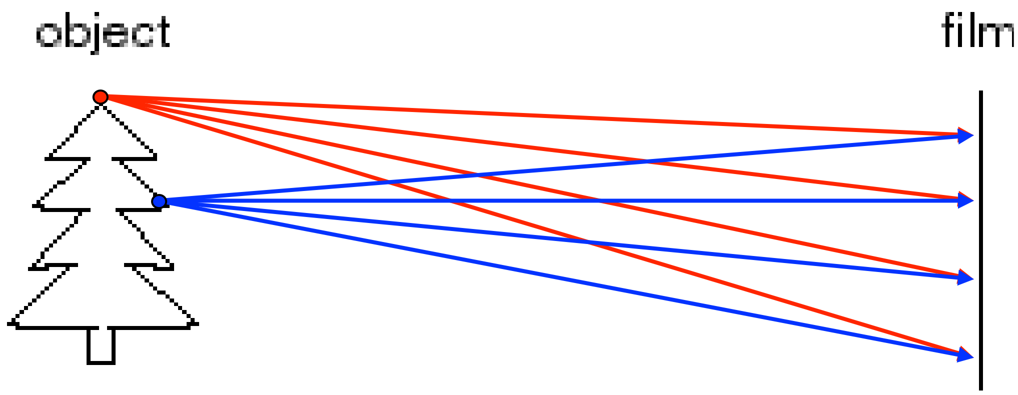
Sensors and image formation

62	70	31	47	100	125	164	166
62	63	40	112	159	140	160	161
50	50	100	143	167	153	150	148
43	73	142	152	165	167	115	114
57	134	170	164	155	114	106	93
111	163	187	144	61	45	50	62
143	180	166	89	51	60	81	176
141	163	105	120	112	99	123	154
167	91	113	135	140	135	135	139

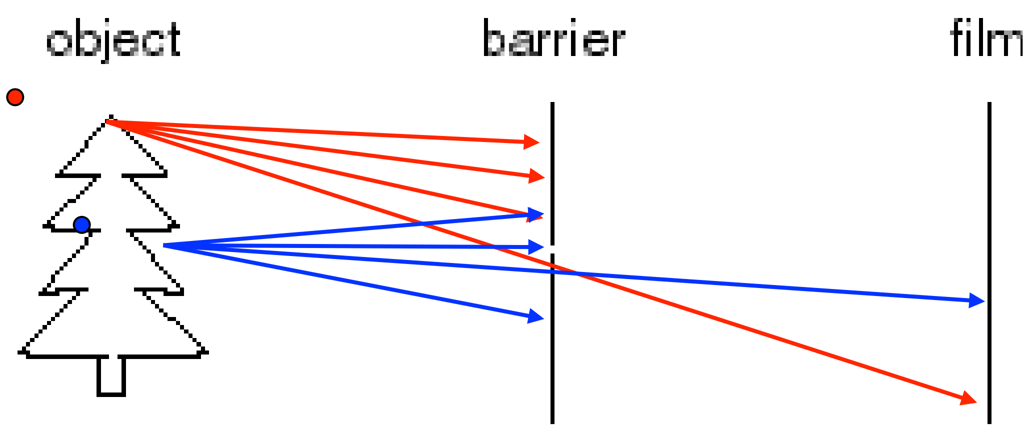
Each pixel is a measure of the brightness (intensity of light) that falls on an area of an sensor (typically a CCD chip)

Sensors and image formation

- Getting the light to the sensor.



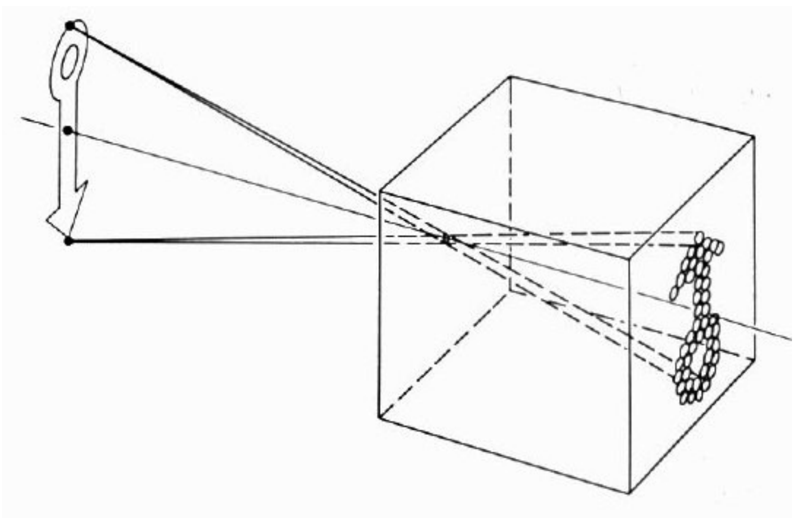
- Add a barrier to block most of the light rays: **aperture.**



Source: GD Hager and S Seitz slides.

The pinhole camera

- Camera obscura
  - First camera; known to Aristotle.
  - Aperture size impacts image.



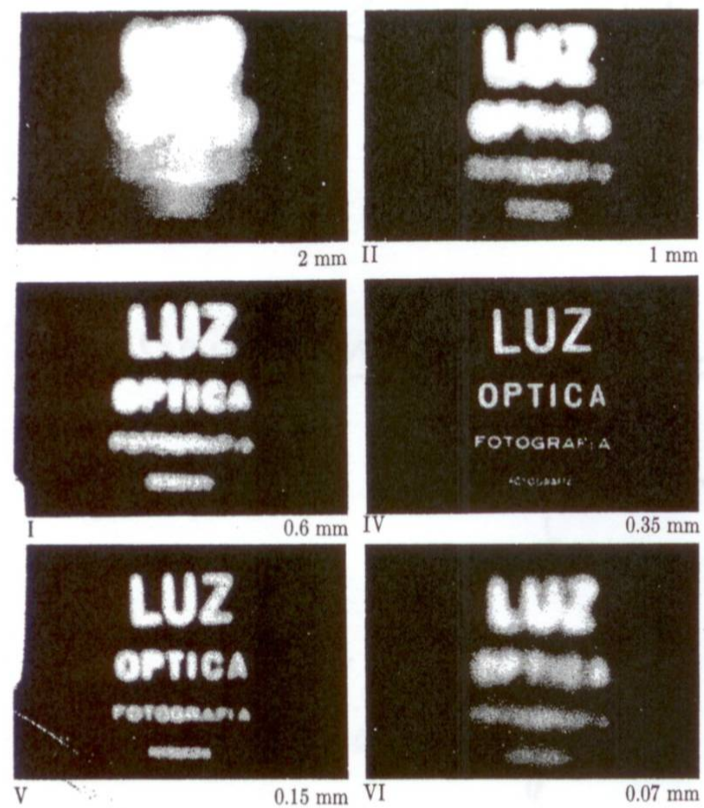


Fig. 5.96 The pinhole camera. Note the variation in image clarity as the hole diameter decreases. [Photos courtesy Dr. N. Joel, UNESCO.]

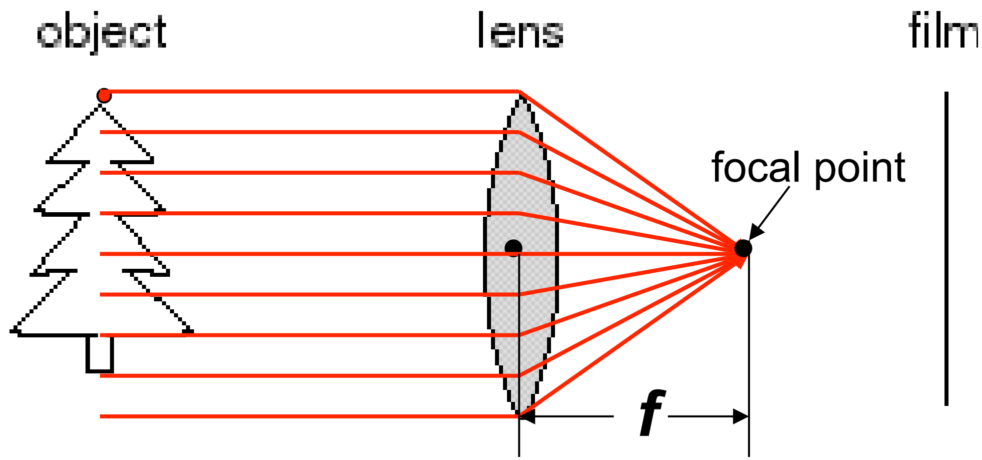
Source: GD Hager and S Seitz slides.

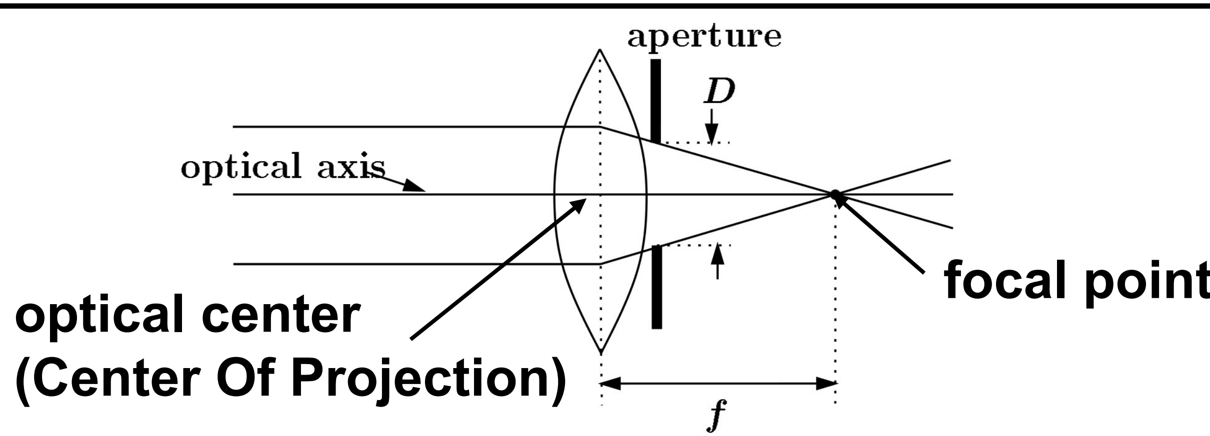
The pinhole camera

Source: https://youtube.com/watch?v=m5VVl0EhoJY

Adding a lens

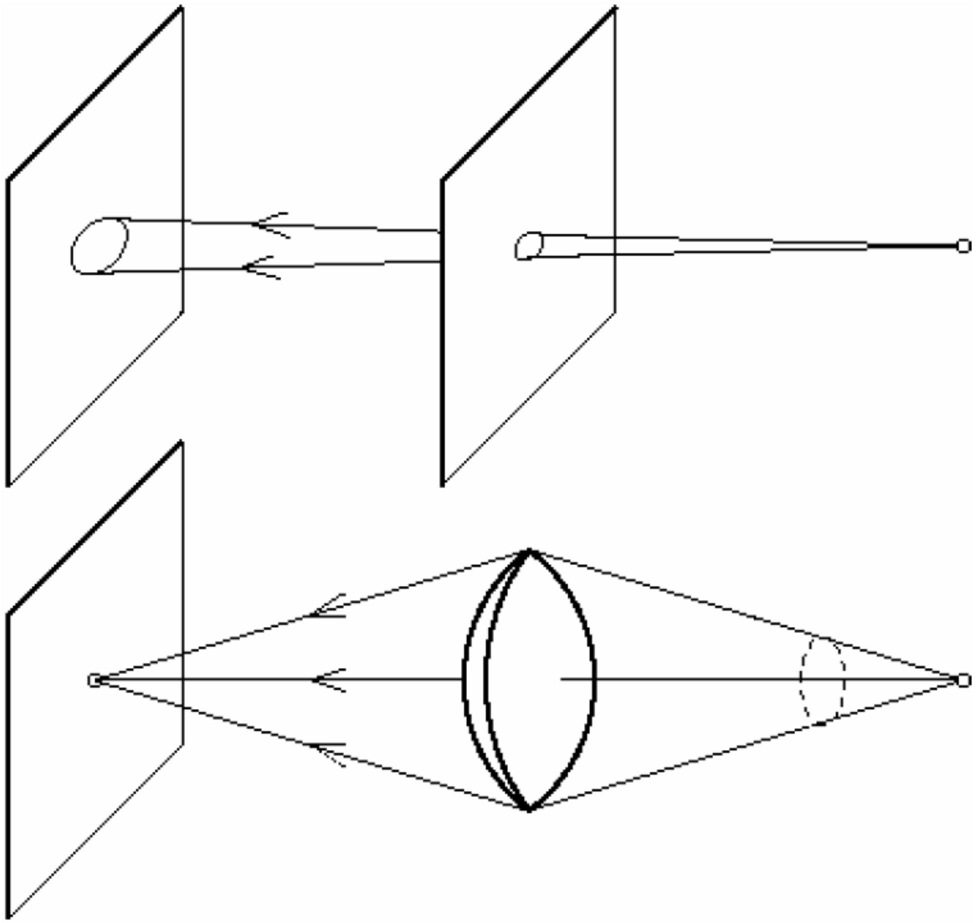
- A lens focuses the light onto the film/CCD.
- Rays passing through the center are not deviated.
- All parallel rays converge to one point on a plane located at the **focal length** *f.*





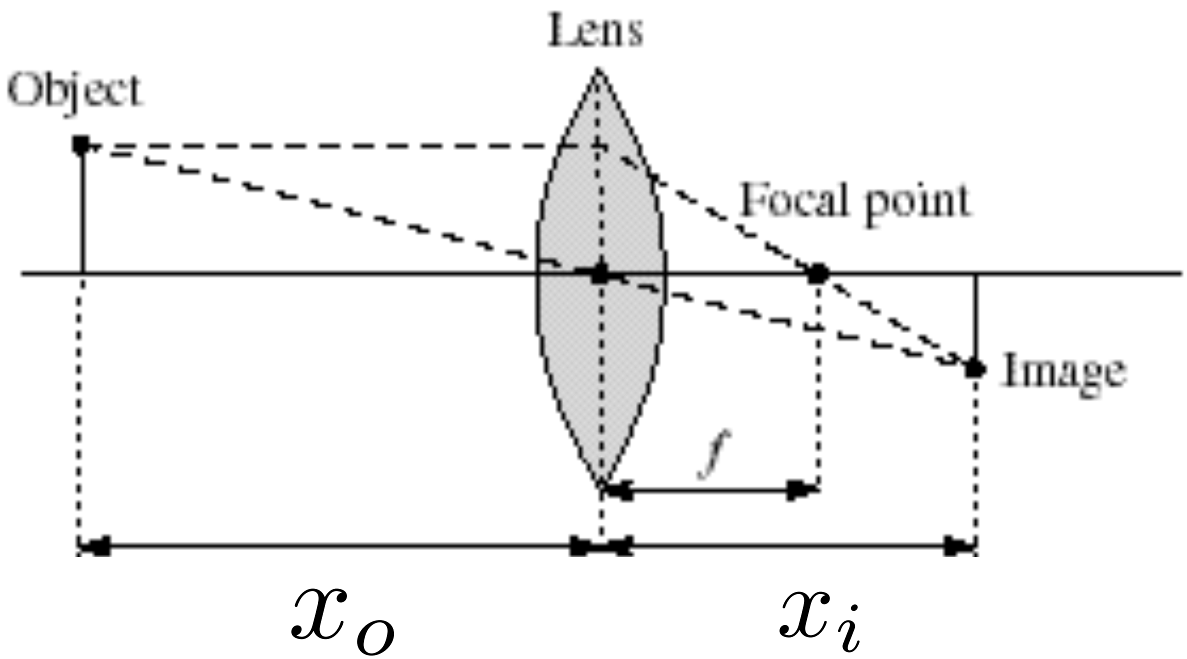
Source: GD Hager and S Seitz slides.

Pinhole vs. lens



Source: GD Hager and S Seitz slides.

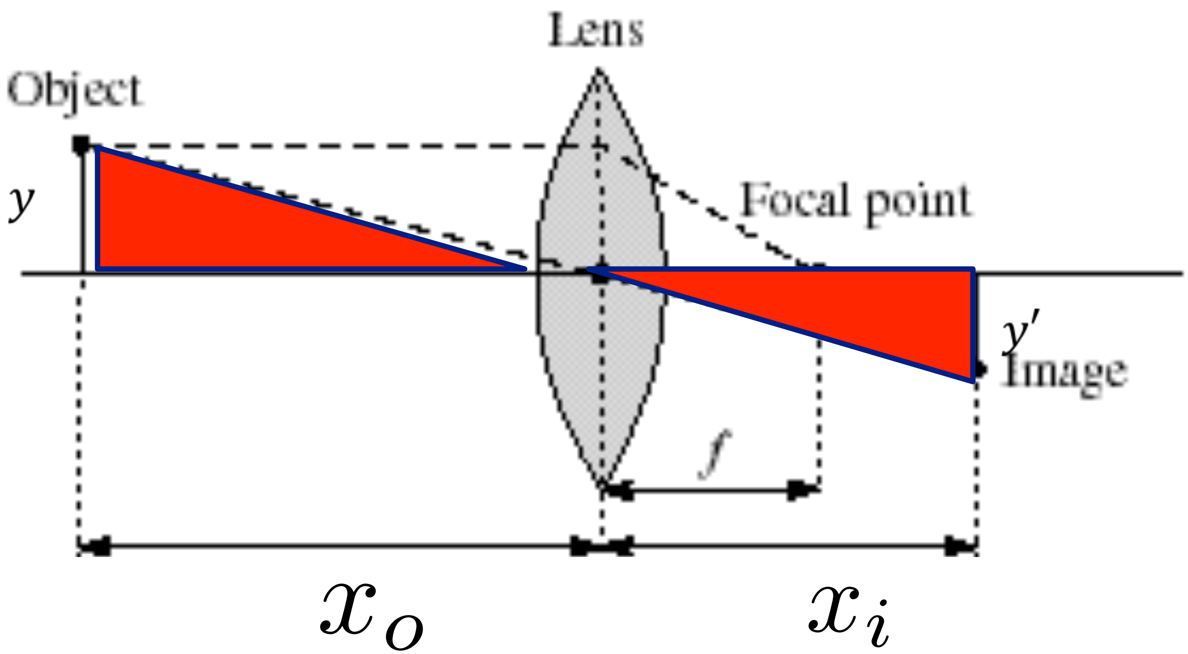
Thin lens equation



- How to relate distance of object from optical center (x_o) to the distance at which it will be in focus (x_i), given focal length f?

Source: GD Hager and S Seitz slides.

Thin lens equation

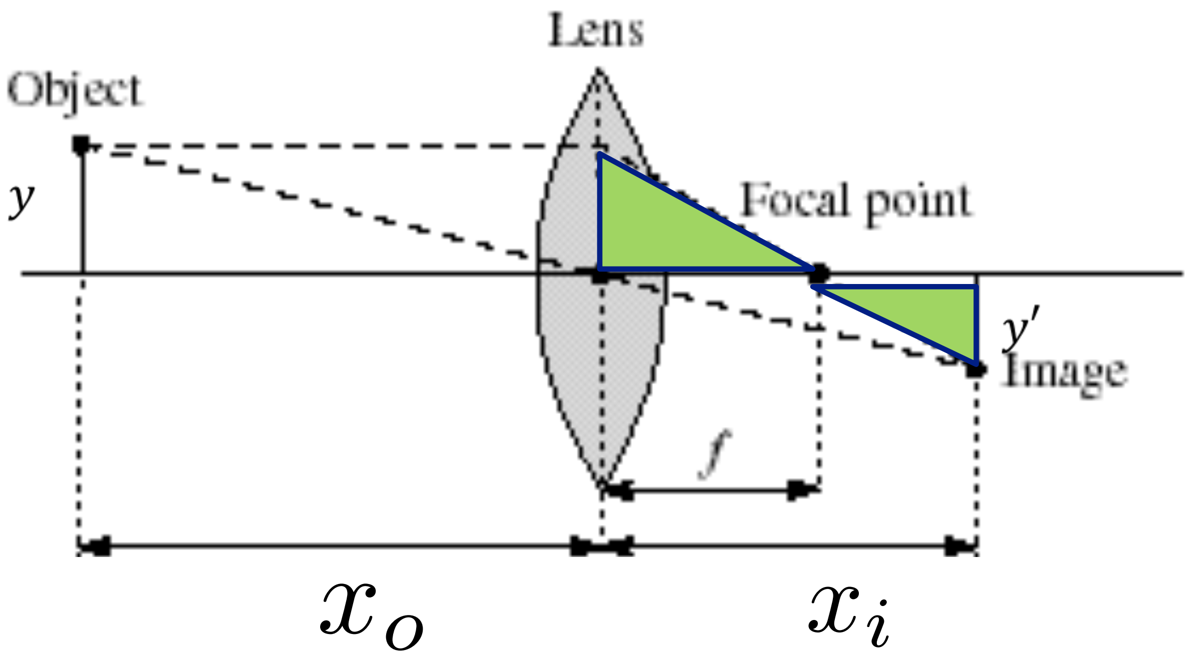


$$\frac{y'}{y} = \frac{x_i}{x_o}$$

- How to relate distance of object from optical center (x_o) to the distance at which it will be in focus (x_i), given focal length f ?

Source: GD Hager and S Seitz slides.

Thin lens equation



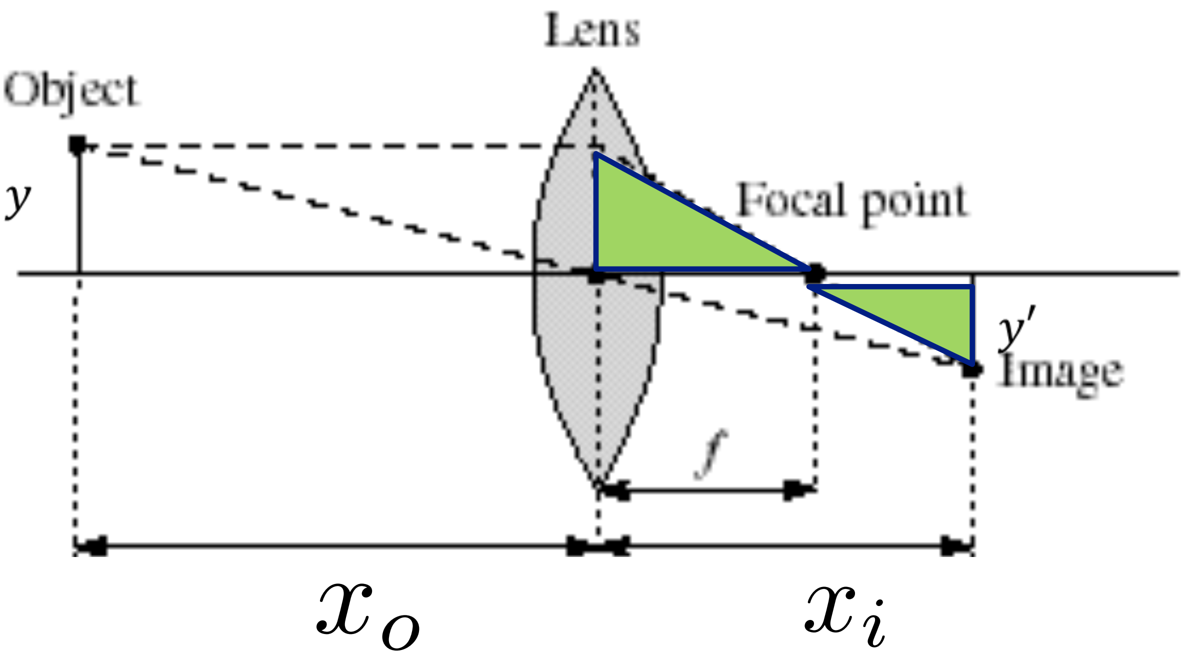
$$\frac{y'}{y} = \frac{x_i}{x_o}$$

$$\frac{y'}{y} = \frac{x_i - f}{f}$$

- How to relate distance of object from optical center (x_o) to the distance at which it will be in focus (x_i), given focal length f ?

Source: GD Hager and S Seitz slides.

Thin lens equation



$$\frac{y'}{y} = \frac{x_i}{x_o}$$

$$\frac{y'}{y} = \frac{x_i - f}{f}$$

$$\frac{1}{f} = \frac{1}{x_o} + \frac{1}{x_i}$$

- Any object point satisfying this equation is in focus

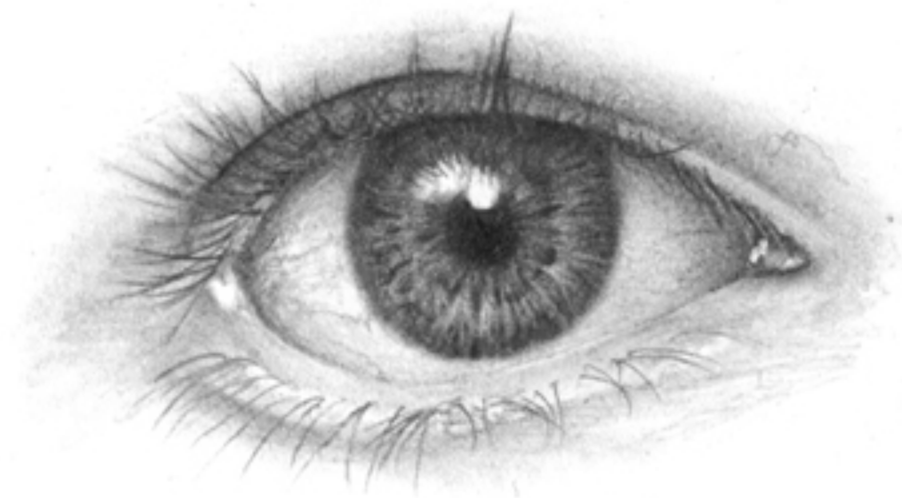
Source: GD Hager and S Seitz slides.

And we have our image...

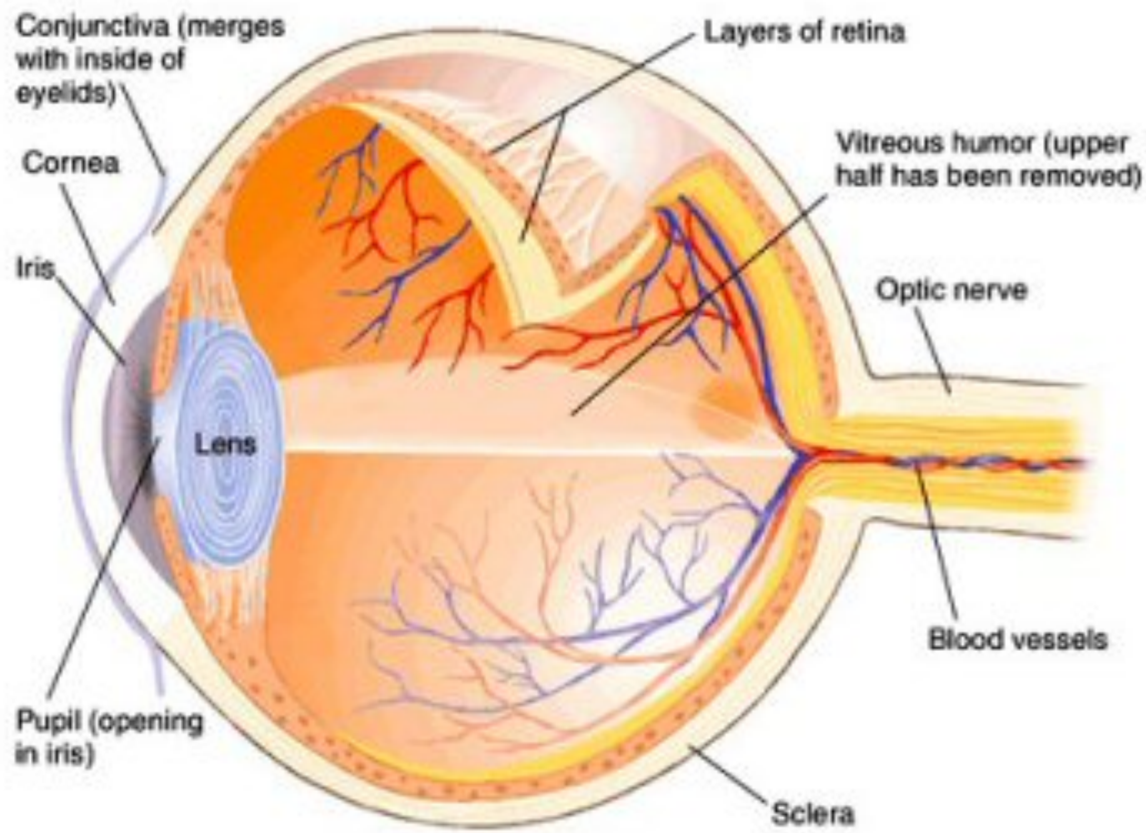
62	70	31	47	100	125	164	166
62	63	40	112	159	140	160	161
50	50	100	143	167	153	150	148
43	73	142	152	165	167	115	114
57	134	170	164	155	114	106	93
111	163	187	144	61	45	50	62
143	180	166	89	51	60	81	176
141	163	105	120	112	99	123	154
167	91	113	135	140	135	135	139

Human vision: our benchmark

- Is a human eye just that much better than a CCD camera?



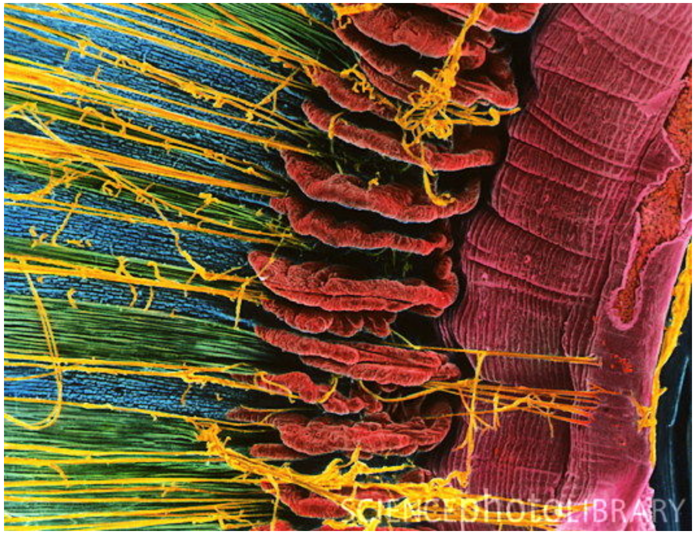
Structure of the eye



The iris is roughly equivalent to the diaphragm in a camera, the cornea and the lens are both lens-like objects, and the retina is where the image is recorded, similar to a CCD sensor or film.

Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

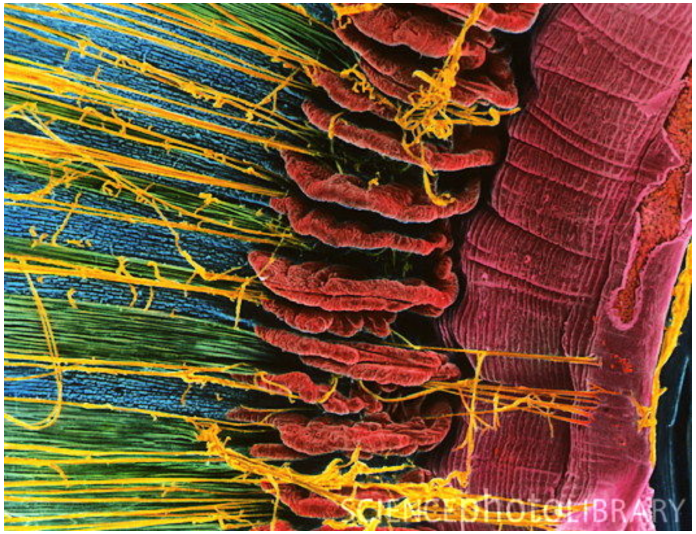
Structure of the eye: Iris

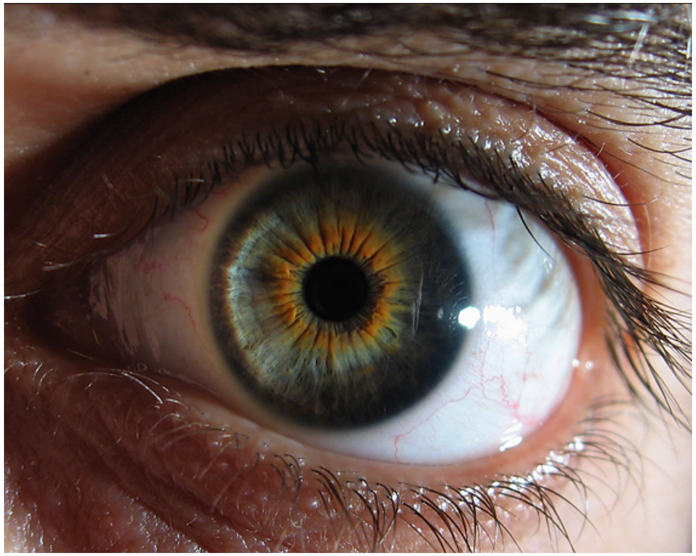


- The iris is similar to the diaphragm/aperture in a camera
- Your iris *widens* in dim light and *narrows* in bright light
- The f-number of your eye varies from f/2 (large opening) to f/8 (small opening)
- Compare this to the range of an average camera lens, which may have f-numbers from f/2.8 to f/22.

Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Structure of the eye: iris

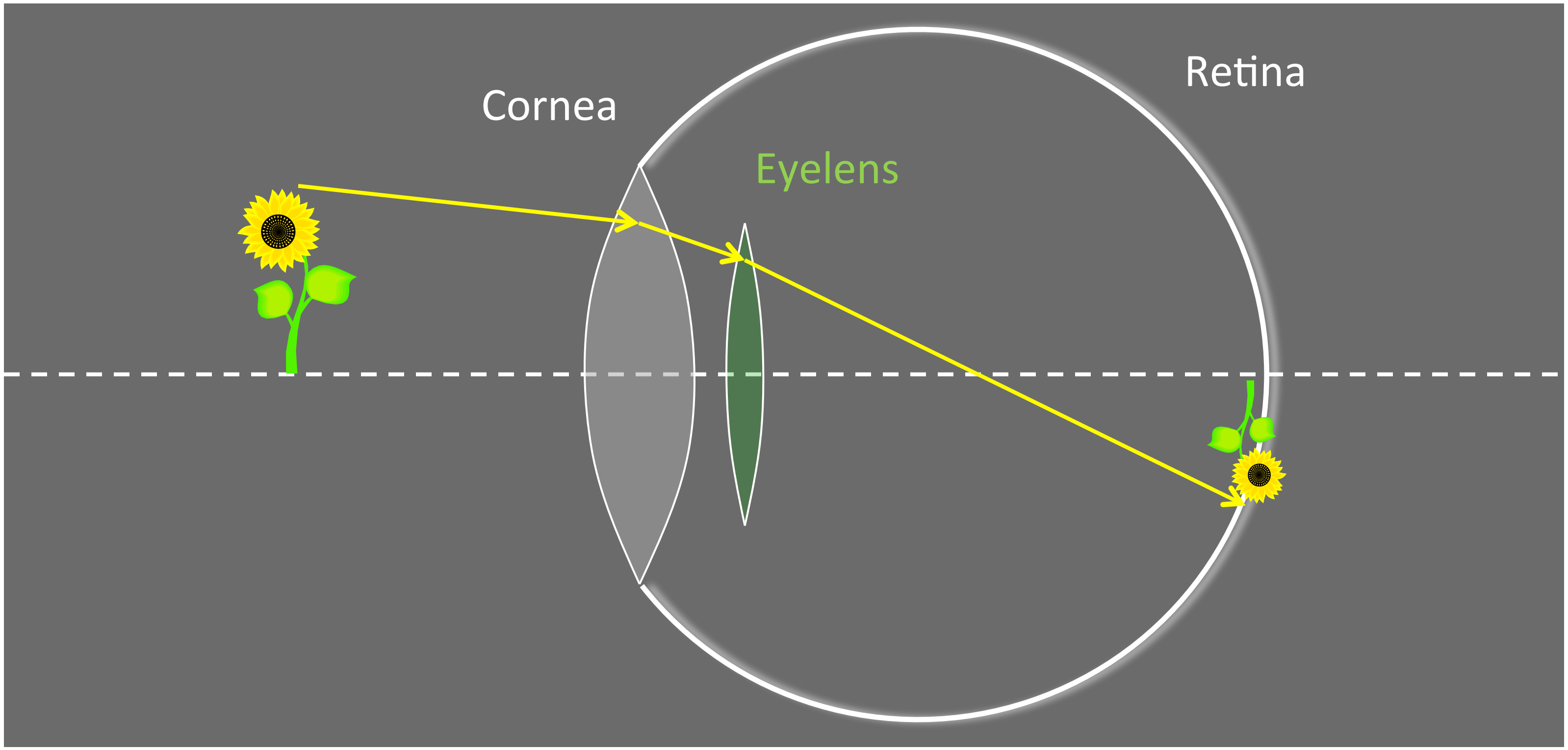




- With a range of only f/2 – f/8, your iris can only reduce the light coming into your eye by a factor of 20.
- The range of intensities that your eye can respond to is a factor of 10^{13}
- The main function of the iris is ***not*** to control the intensity of light coming into your eye
- Main functions of iris
  - Reduce aberrations, sharpen image
  - Increase depth of field

Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Structure of the eye: cornea and lens



- There are two lenses in your eye, the cornea and the eyelens.
- The cornea, the front surface of the eye, does most of the focusing in your eye.
- The eyelens (crystalline lens) provides adjustable fine-tuning of the focus.

Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

How a camera lens focuses

- A **camera** is focused by changing the **distance**, x_i ,from the lens to the image at the back on the CCD as the distance to the object, x_o , changes.

Focal length (fixed)

Object distance (varies)

$$\frac{1}{x_o} + \frac{1}{x_i} = \frac{1}{f}$$

Image distance (changes to satisfy equation when x_o changes)

Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

How a human eye focuses

- The **eyelens** is a fixed distance, x_i, from the retina at the back of the eyeball where the image is created.

Focal length (changes to satisfy the equation)

Object distance (varies)

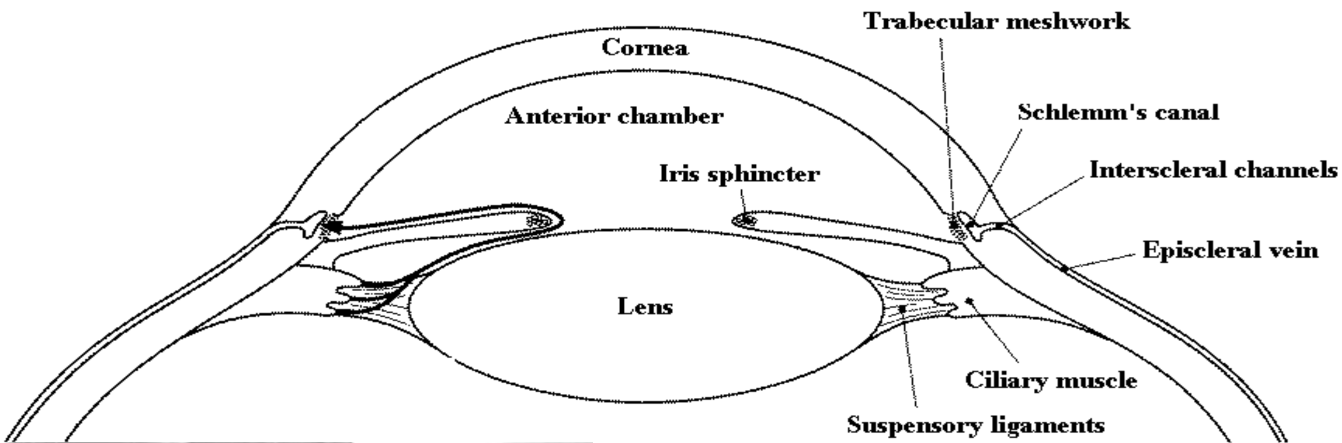
$$\frac{1}{x_o} + \frac{1}{x_i} = \frac{1}{f}$$

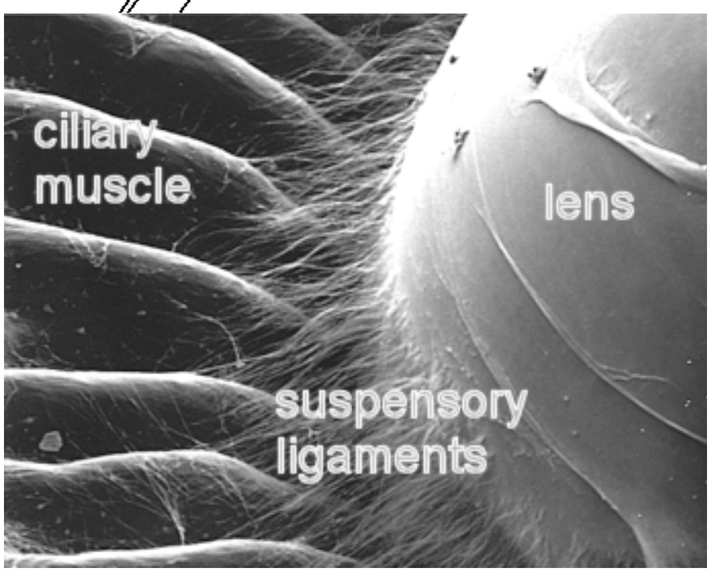
Image distance (fixed)

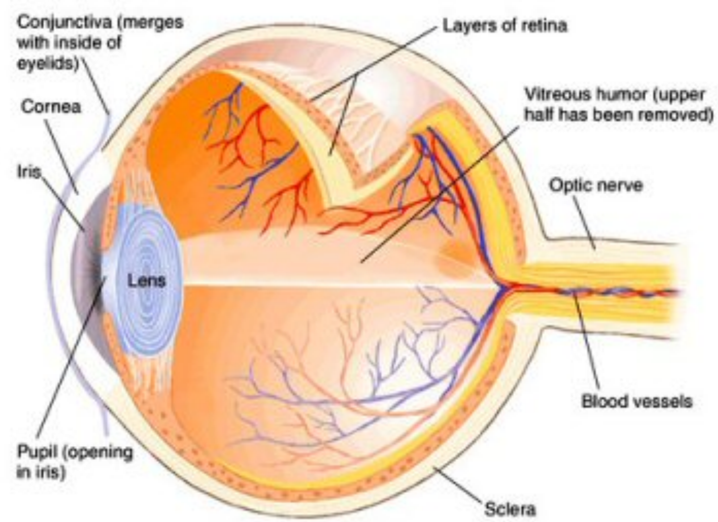
Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Eyelens: focusing and accommodation

- The eyelens changes its focal length by changing its shape. Ligaments pull on the lens to change the amount of "bulge".





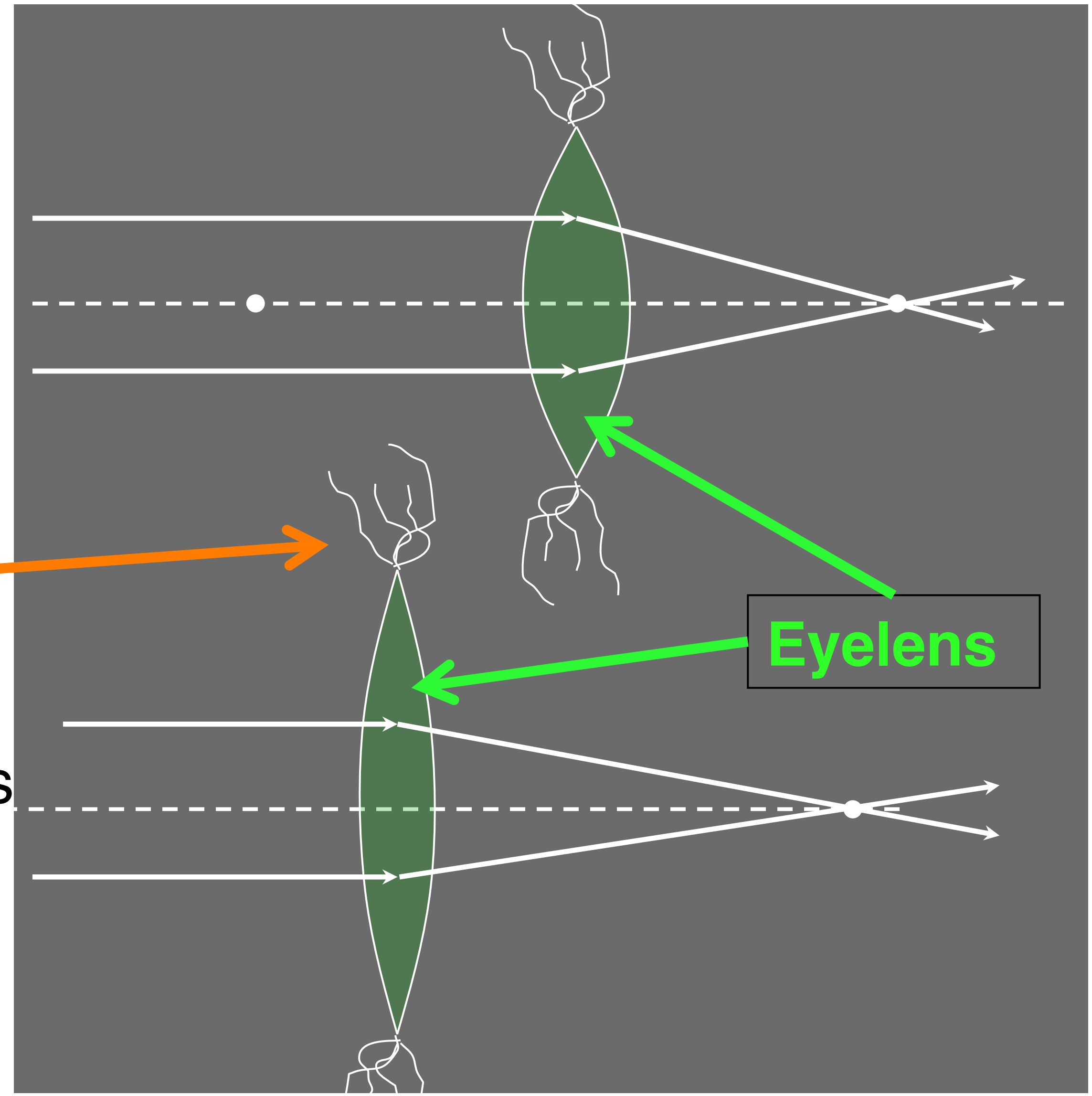


Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Eyelens: focusing and accommodation

Muscles contract, ligaments relax, more bulge, more bending power, shorter focal length

Ligaments



Muscles relax, ligaments contract, less bulge, less bending power, longer focal length

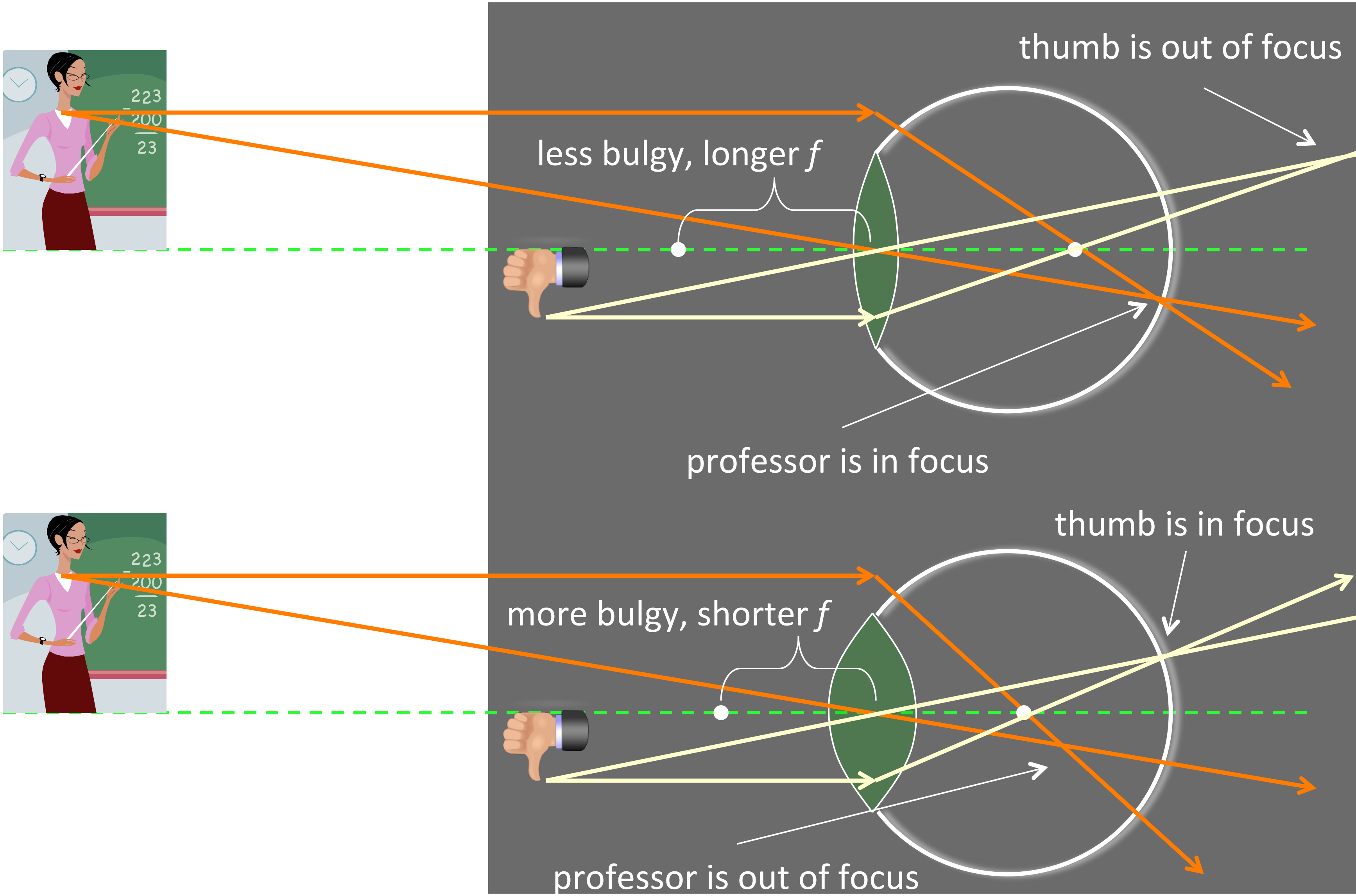
Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Eyelens: focusing and accommodation

- Your eyelens has a ***small depth of field***
  - You can't see something close and far with both objects in focus at the same time.
- Hold out your thumb about a foot away from your eye
  - Then, alternately focus on thumb and me (right above your thumb).
- Note that you cannot see ***both*** me ***and*** your thumb sharply (in focus) at the same time
  - You focus on one or the other by changing the bulge of your eyelens.

Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Eyelens: focusing and accommodation



Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Concept questions on focusing

You can't see me and your thumb clearly *at the same time*

a) because your *pupil* is too small
b) because your iris can't change fast enough
c) because your eye cannot *accommodate*
d) because your eye does not have enough depth of field

Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Concept questions on focusing

When you see someone *out-of-focus*

a) There is no image anywhere
b) There is an *in-focus* image on your fovea
c) There is an *in-focus* image on your retina
d) There is an image *in-focus* either in front or in back of your retina

Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Concept questions on focusing

In order to *focus* on close objects

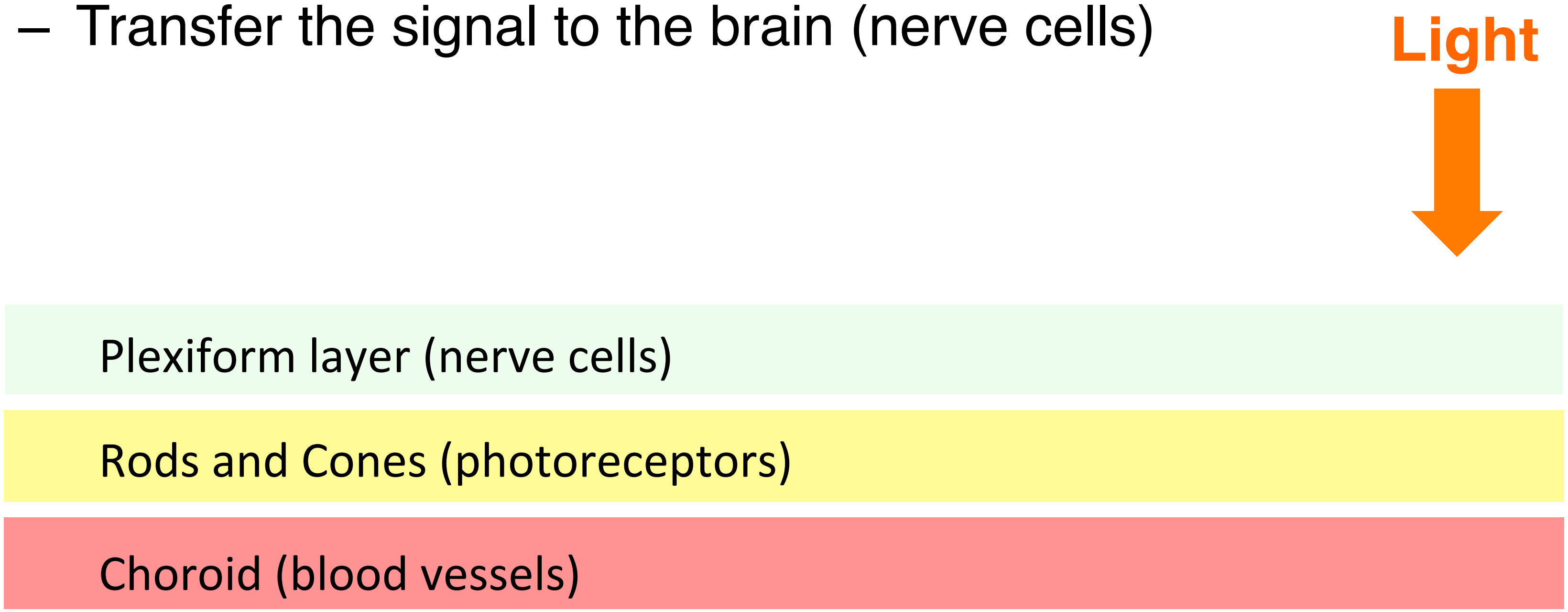
a) your eyelens needs to bulge
b) your eyelens needs to flatten
c) your cornea needs to bulge
d) your cornea needs to flatten
e) the distance (x_i) between your eyelens and retina needs to change

Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Structure of the eye: retina

- The retina is the sensor or film of your eye.

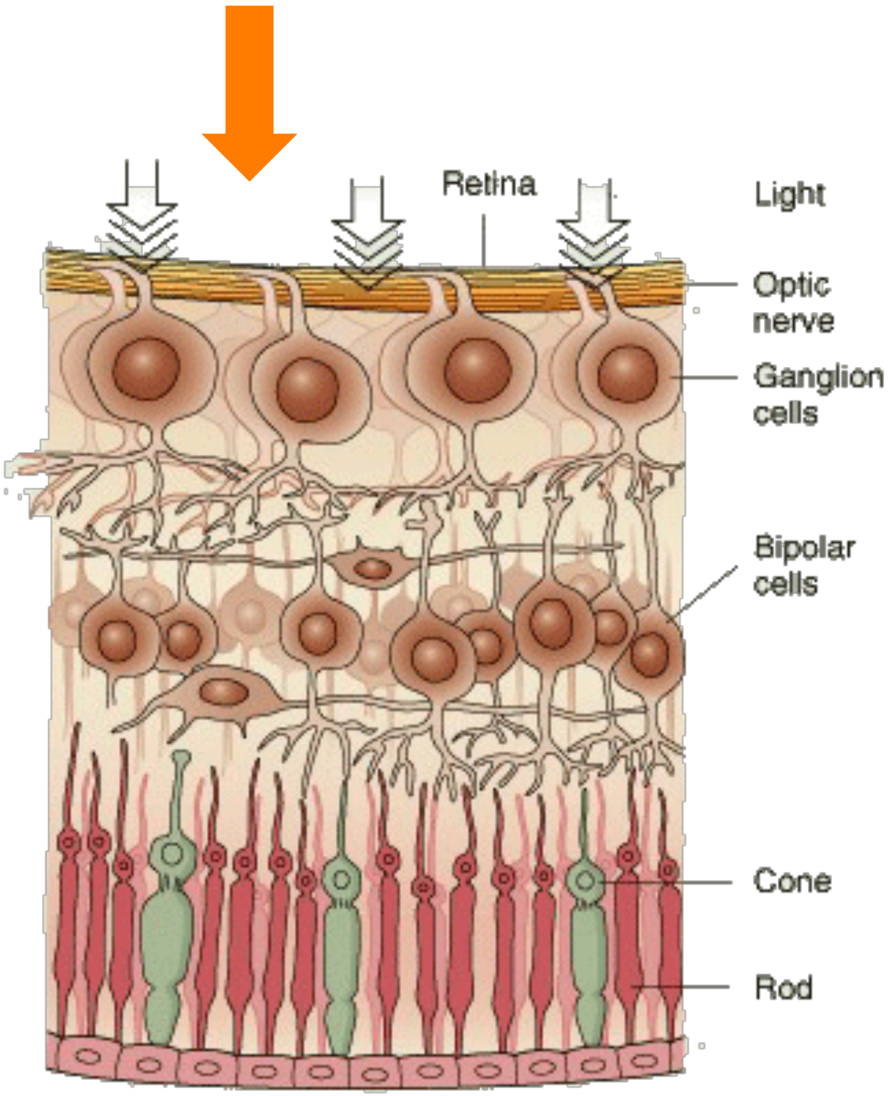
- Its layers do three things
  - Provide blood and nutrients (choroid)
  - Absorb light and convert to an electrical signal (photoreceptors)
  - Transfer the signal to the brain (nerve cells)



Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Structure of the retina

Light



← Nerve cells

← Photoreceptors

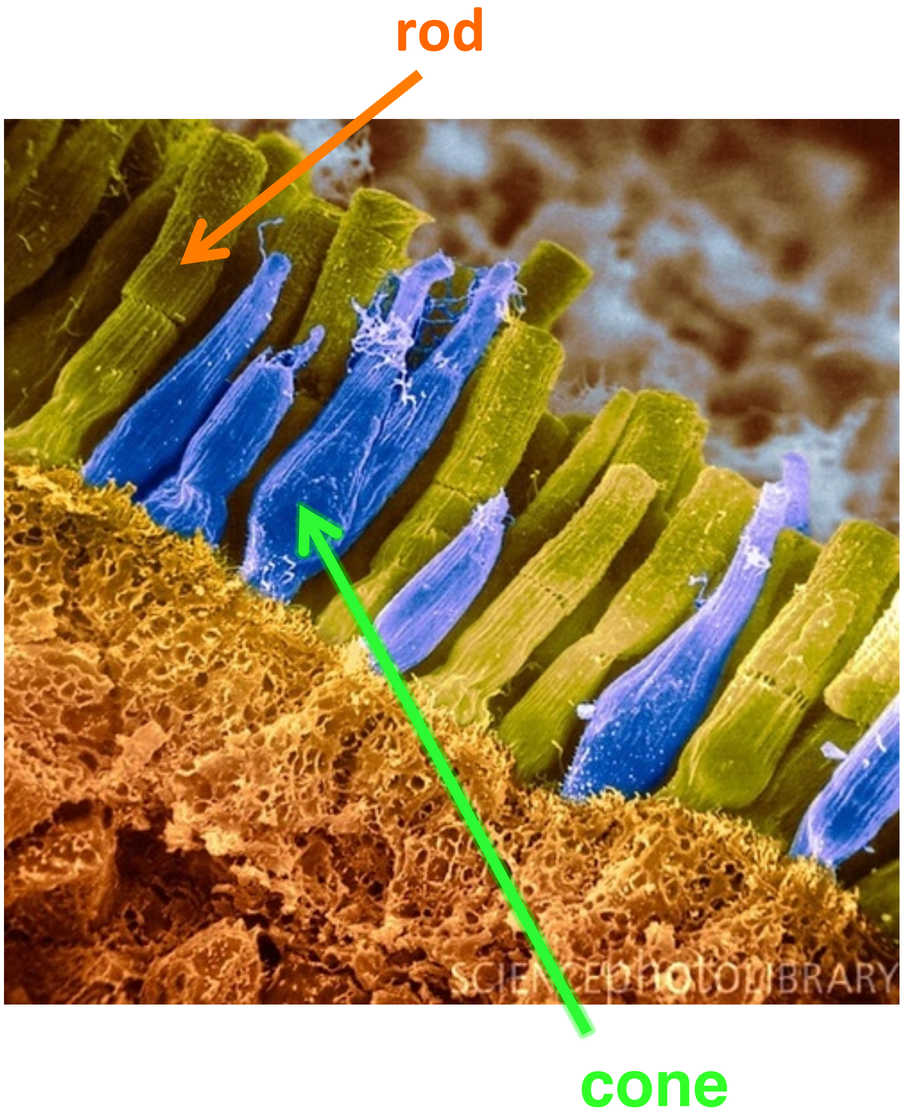
← Choroid

Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Photoreceptors: rods and cones

- Light is detected and converted to an electrical signal by the photoreceptors in the retina. There are two main kinds of receptors, rods and cones.

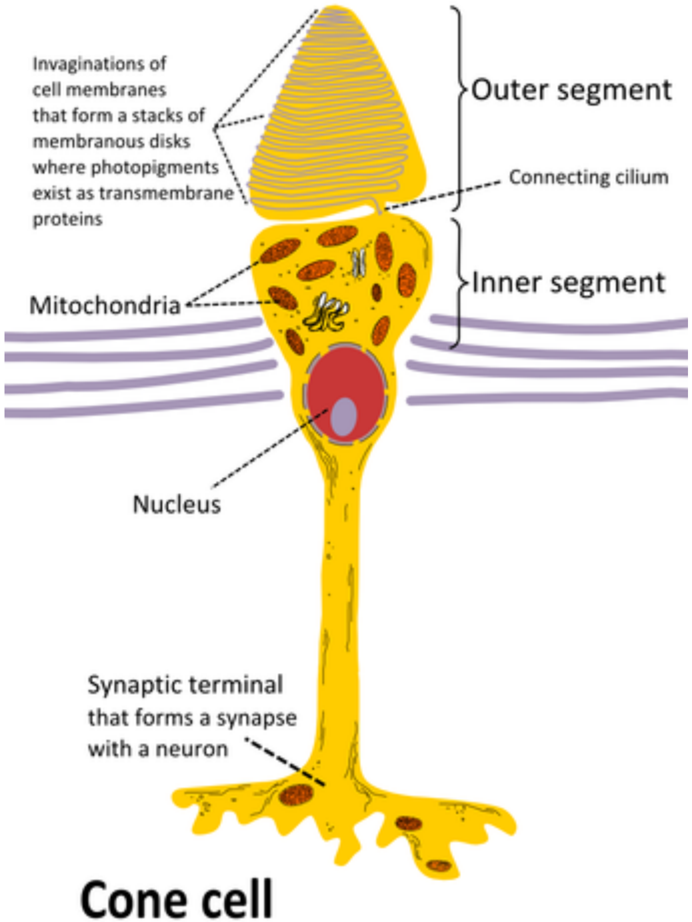
- This is a false color image, rods and cones are not actually different colors.



Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Photoreceptors: cones

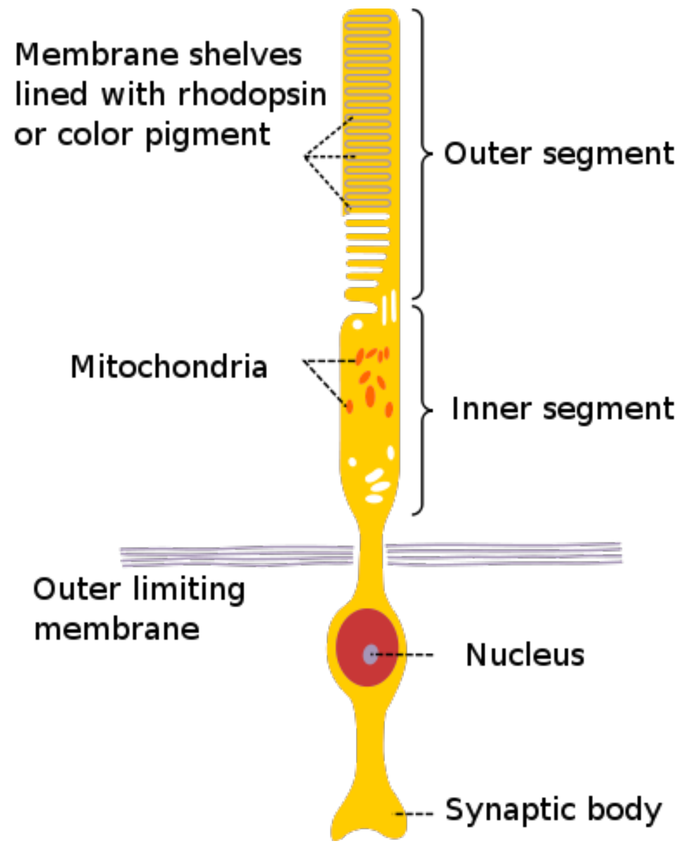
- Cones are responsible for our fine detailed and color vision
- Cones are clustered near the center of your retina, called the *fovea*
- There are 5 million cones in the average retina



Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Photoreceptors: rods

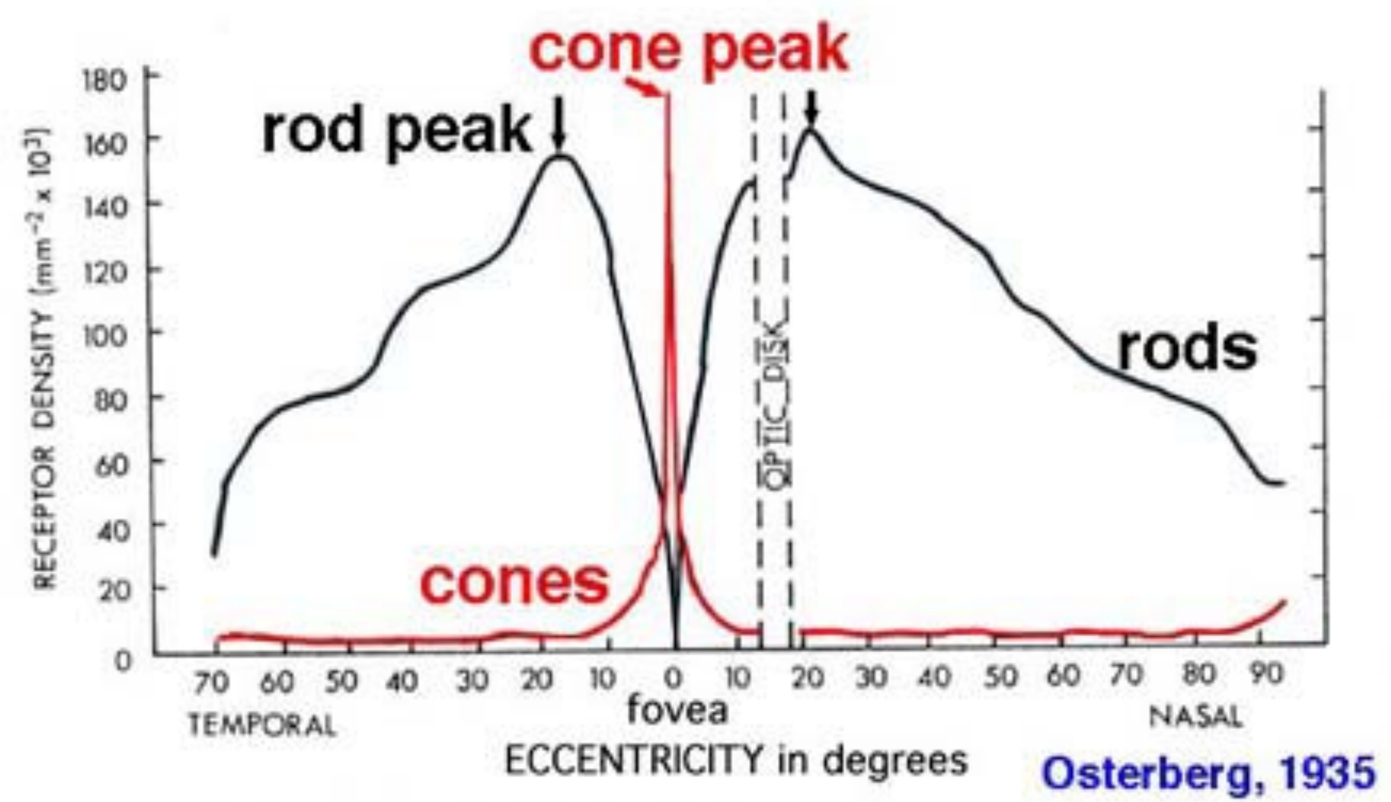
- Rods are responsible for low light and peripheral vision
- They are present everywhere in the retina except the fovea
- There are 125 million rods in the average retina



Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

Rods and cones

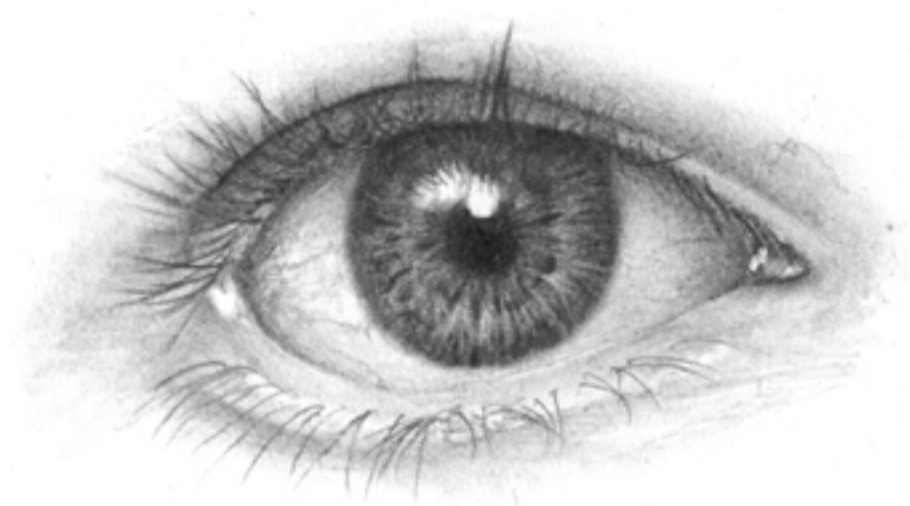
- Because of their different functions, rods and cones are present in varying densities in the retina. The blind spot is due to the connection of the optic nerve



Source: Dr. Ellen Keister, Physics 1230, U Colorado Fall 2012 slides and *Seeing the Light* by Falk, Brill and Stork (Wiley, 1986).

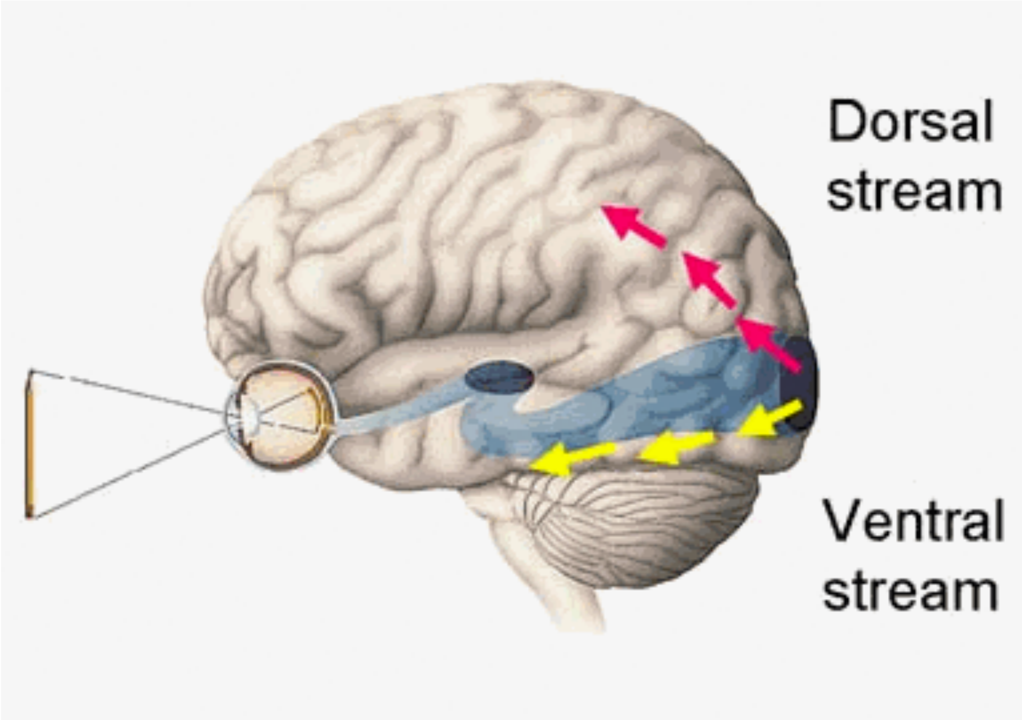
Human vision: our benchmark

- Is a human eye just that much better than a CCD camera?



Human vision: our benchmark

- Is a human eye just that much better than a CCD camera?

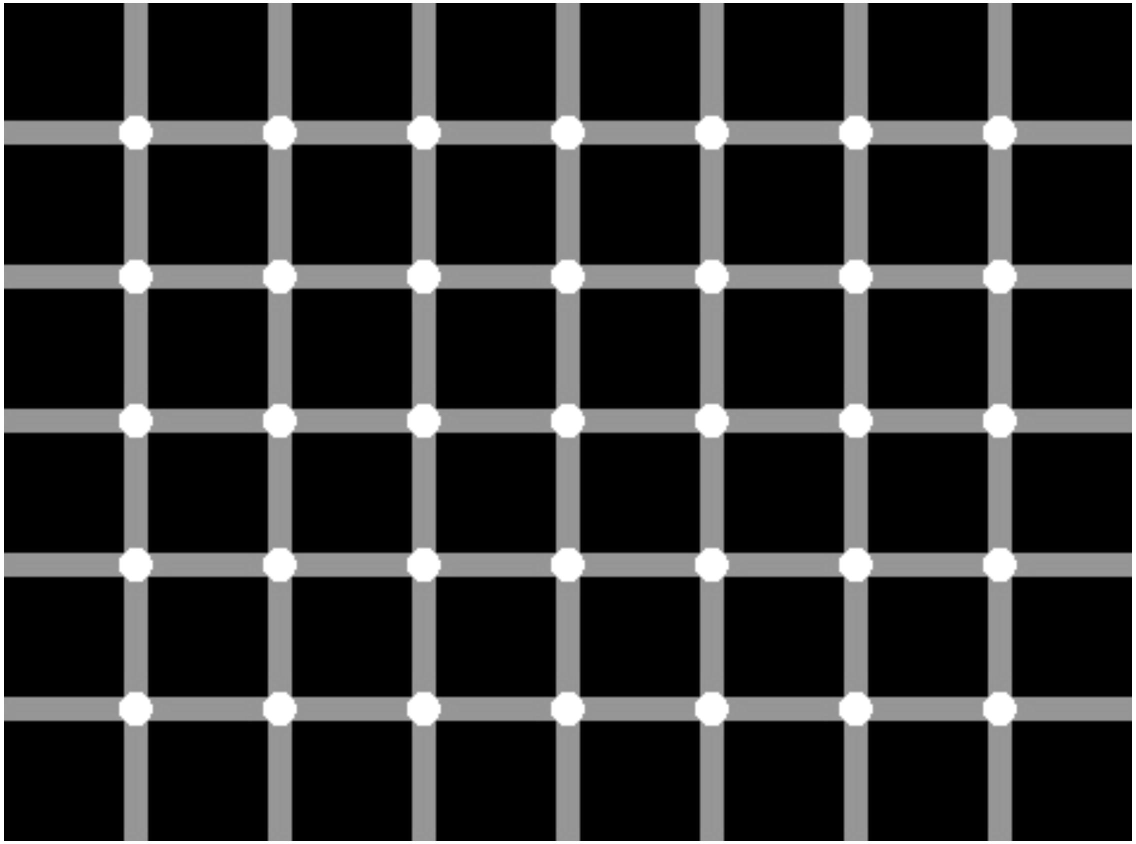


- Then how do humans ***see*** so well?

Human vision: our benchmark

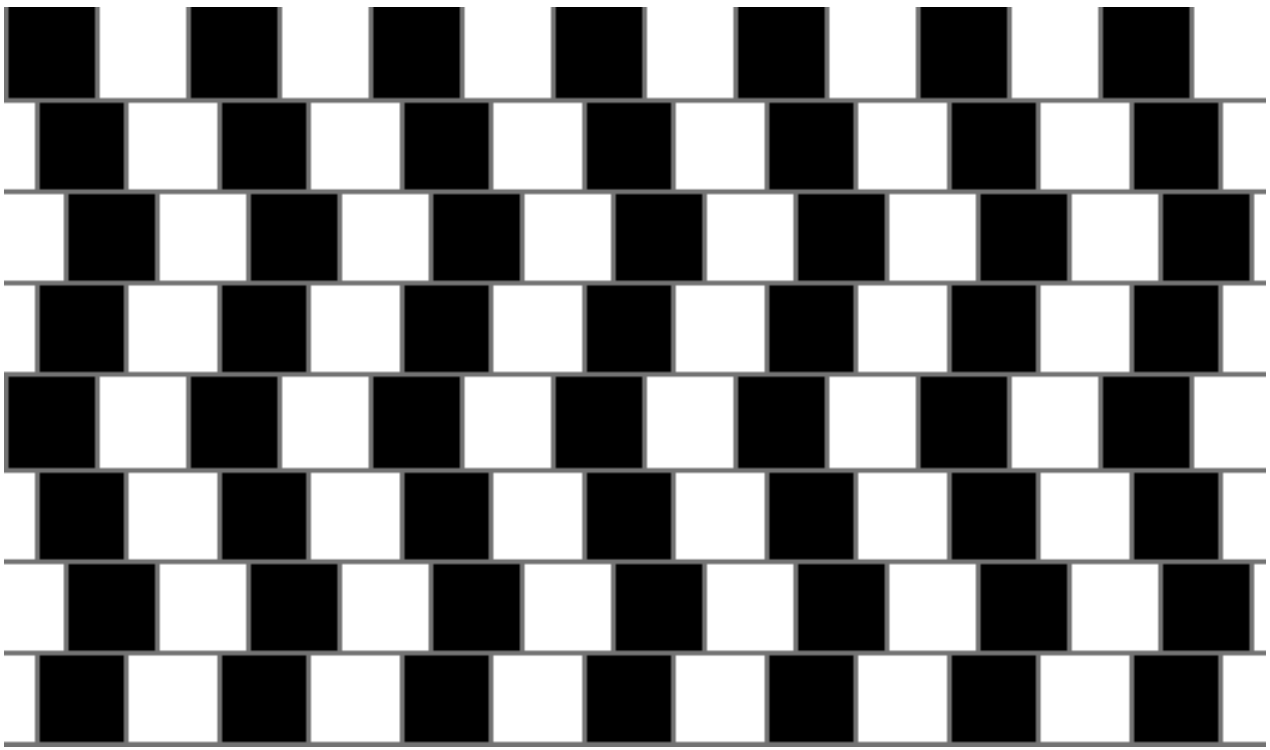
Source: G.D. Hager Slides.

Illusions: what do they tell us?



Source: G.D. Hager Slides.

Illusions: what do they tell us?



Source: G.D. Hager Slides.

Why is computer vision hard?

Source: G.D. Hager Slides.

Why is computer vision hard?

Source: G.D. Hager Slides.

Why is computer vision hard?

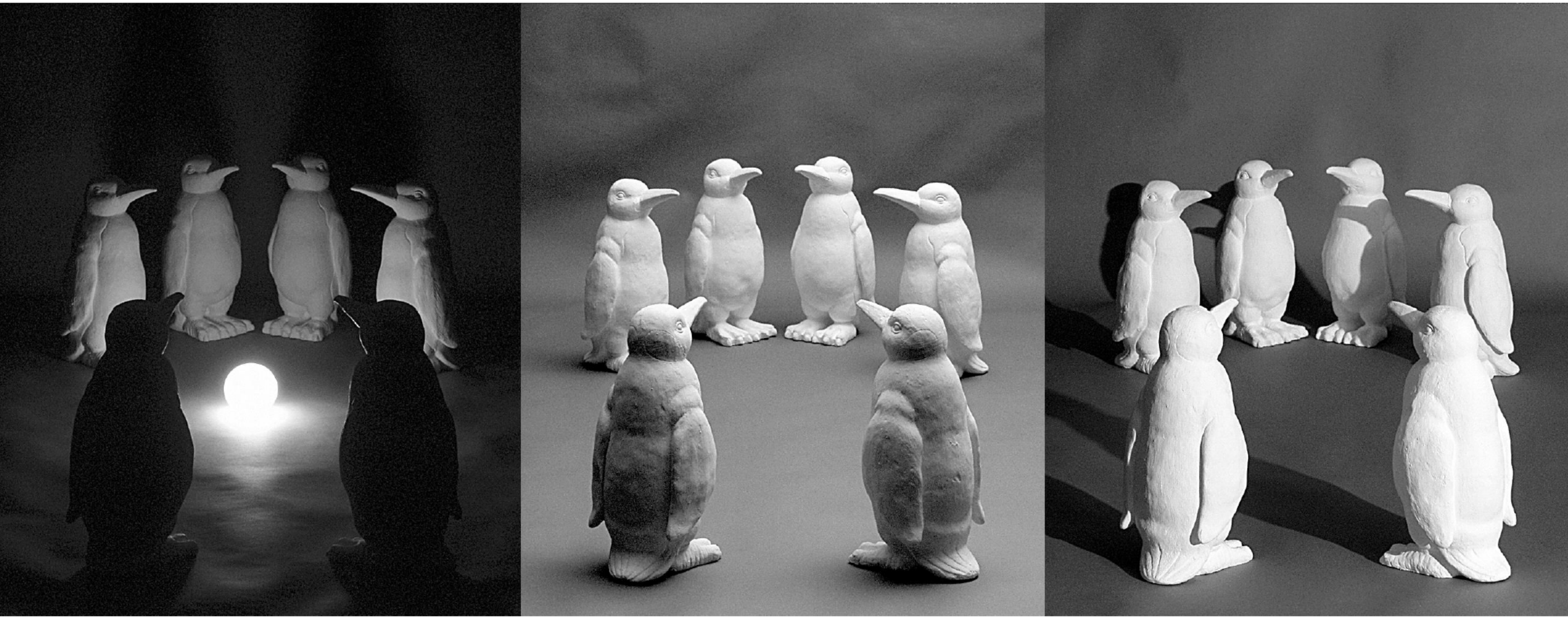
- “Context” counts for as much as appearance.
- Huge amounts of prior knowledge (learned and innate).
- AI complete (hard to solve a cleanly defined “simple” problem without invoking unrealistic assumptions).
- Lack of a clear metric for success (indeed, we are often completely wrong as you’ve seen).
- The diversity of the natural world.

Source: G.D. Hager Slides.

Challenges: viewpoint variation

slide credit: Fei-Fei, Fergus & Torralba

Challenges: illumination



Source: Saverese slides; Koenderink image.

Challenges: scale

Source: Saverese, Fei-Fei, Fergus, and Torralba.

Challenges: deformation

Source: Saverese slides.

Challenges: occlusion

Source: Magrit 1957; Saverese, Fei-Fei, Fergus, and Torralba.

Challenges: background clutter

Source: Kilmeny Niland 1995; Saverese slides.

Challenges: intra-class variation



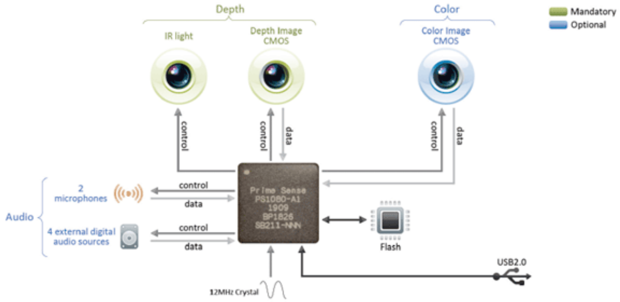



Source: Saverese, Fei-Fei, Fergus, and Torralba slides.

Challenges: class variation

Source: Saverese slides.

Computer Vision is everywhere



From D. Gavrila.

Applications of Computer Vision

- Link to Prof. Savarese application slides (roughly 16-40).
  - http://cvgl.stanford.edu/teaching/cs231a_winter1314/lectures/lecture1_introduction.pdf

Course Information

- EECS 598-08 Foundations of Computer Vision
  - Essentially an introduction to computer vision targeted to graduate student explicitly rather than undergrad/grad mix.
- Syllabus
  - Contains all necessary information for course structure and grading.
  - Piazza for all Q&A (except when privacy is a concern).

Coming up in Week 2

- No class on Monday 9/8 as I will be in Zurich at ECCV.
  - This is the only planned cancellation of the term.
- Wednesday 9/10 we begin with image formation and representation, covering in more detail the geometry of image formation. Read FP Ch. 1 and I will add supplemental materials.