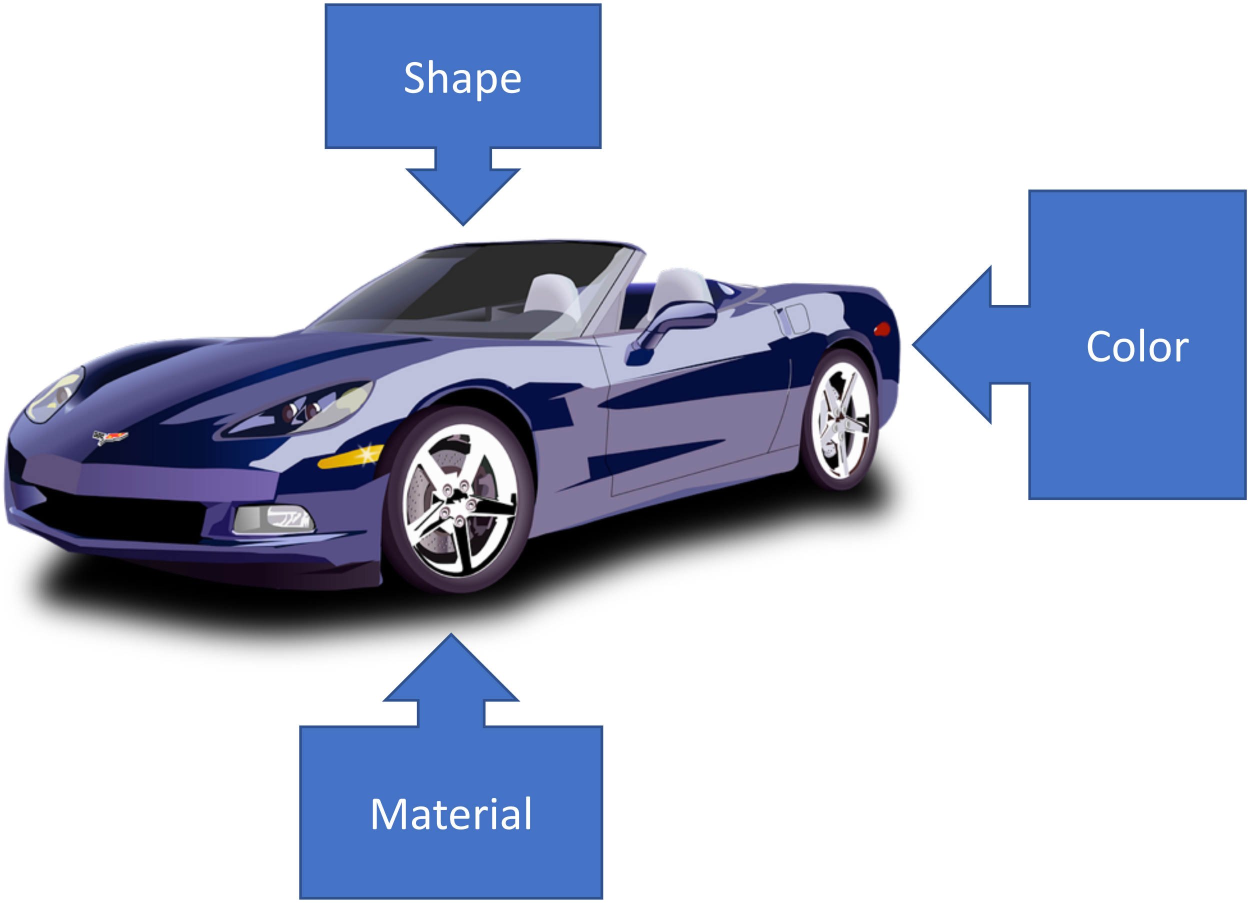
The visual world

What do images look like?

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

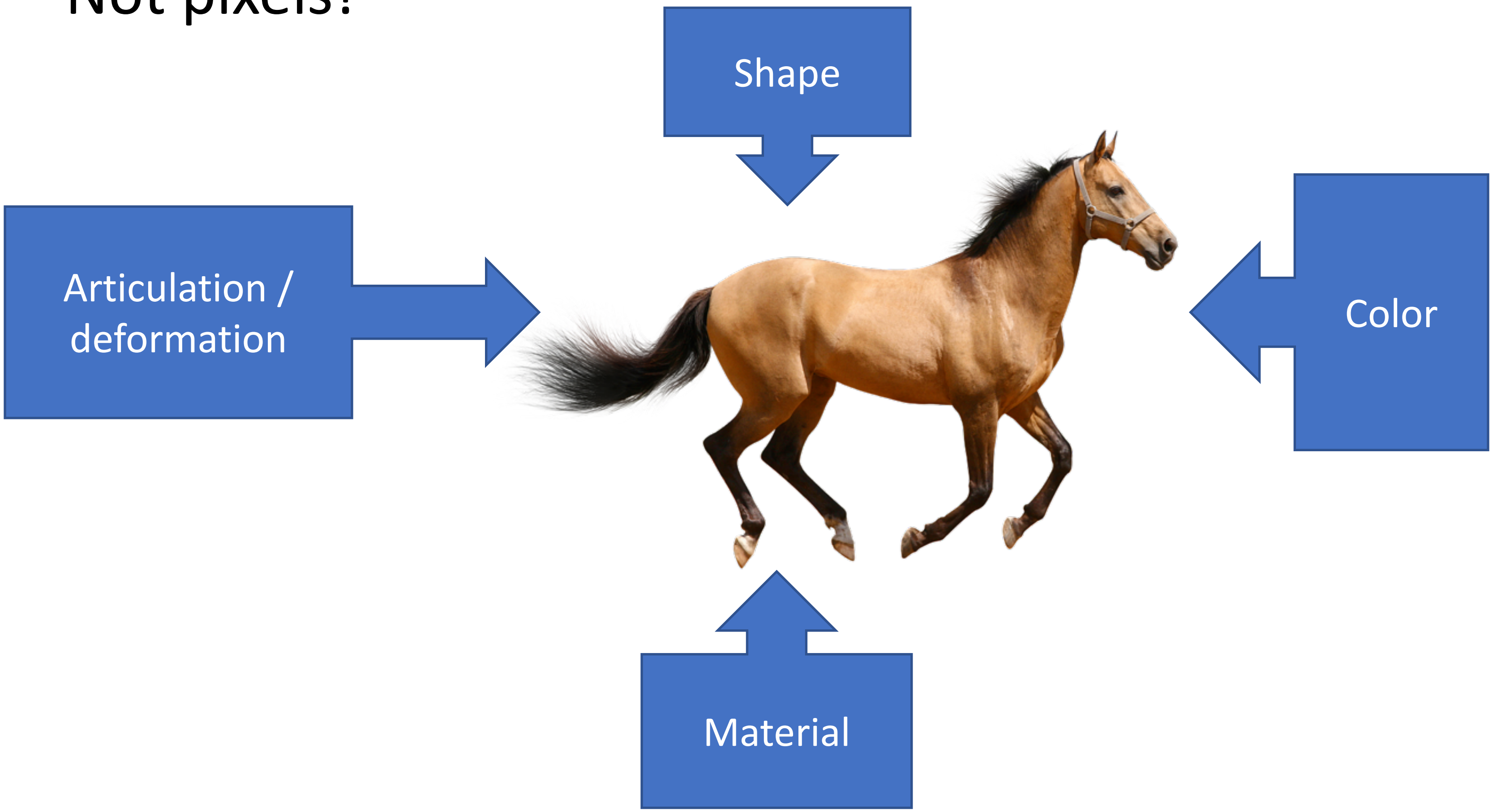
The *physical* world

- Not pixels!



The *physical* world

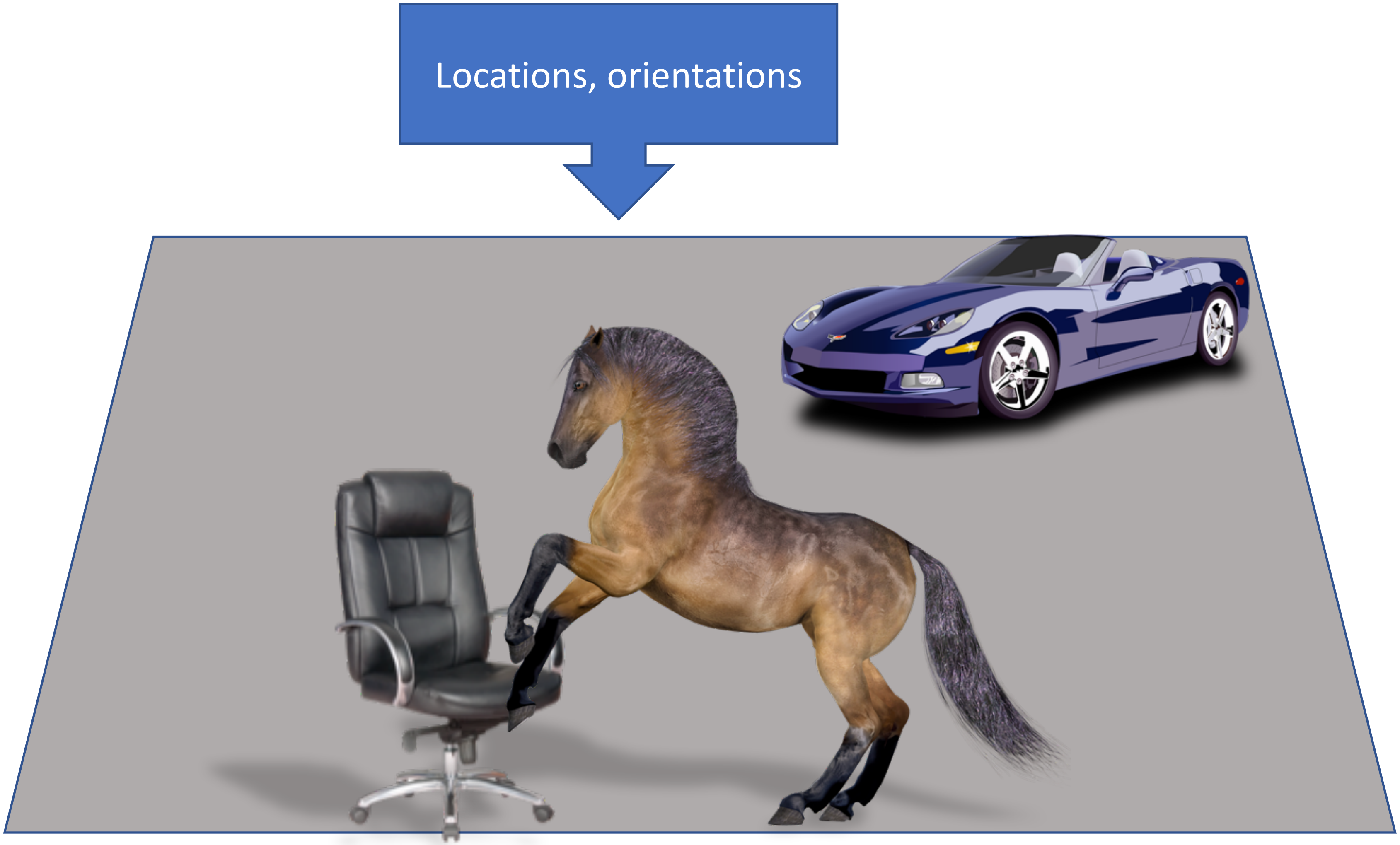
- Not pixels!



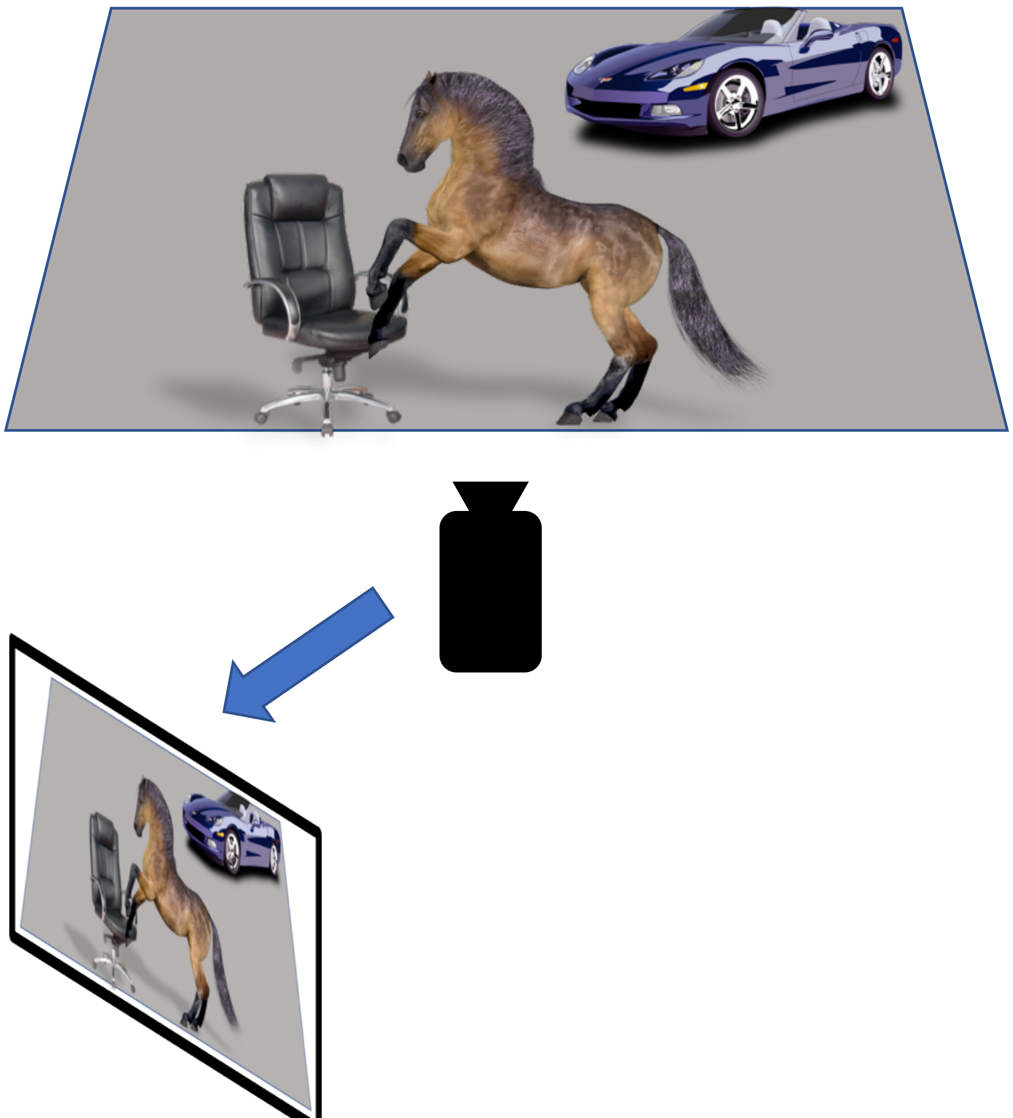
The *physical* world

- Not pixels!

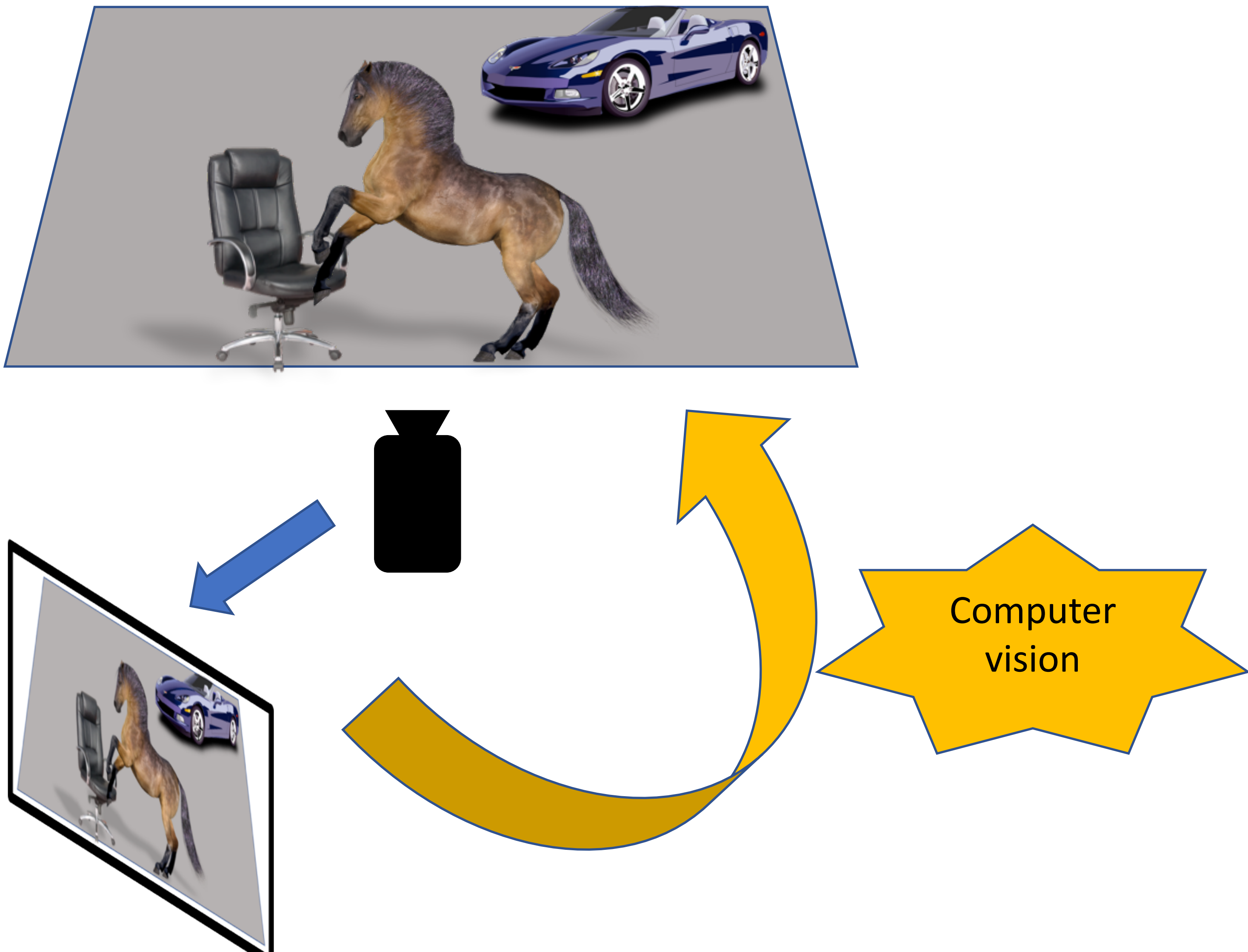
The *physical* world



World → Images

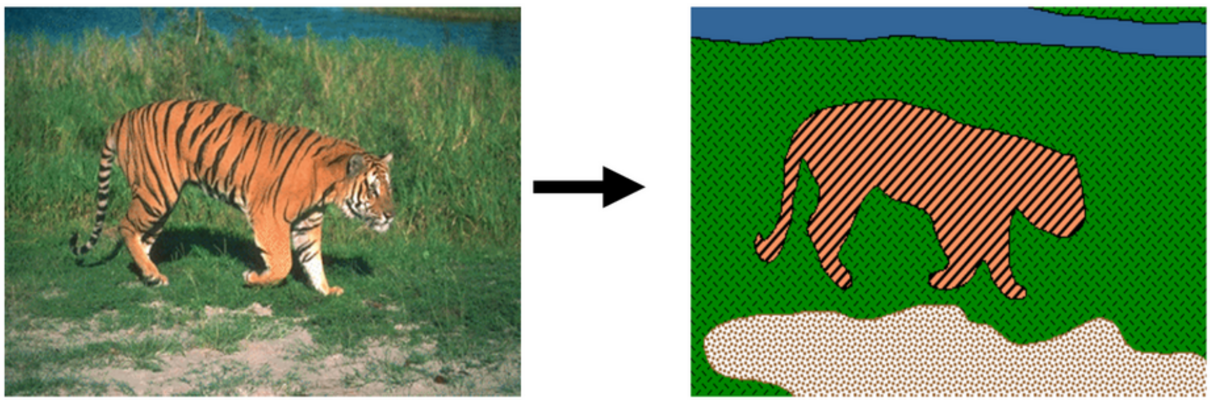


The goal of computer vision



The goal(s) of computer vision

- *Grouping ("Reorganization")*
  - Convert from "pixels" to "objects": which groups of pixels correspond to objects?



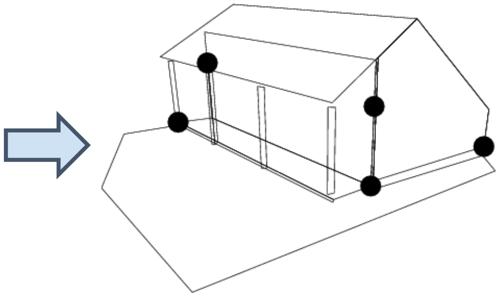
[Blobworld]

The goal(s) of computer vision

- *Reconstruction*
  - Go from 2D arrays to 3D: what does every pixel correspond to in 3D

Left View

Right View



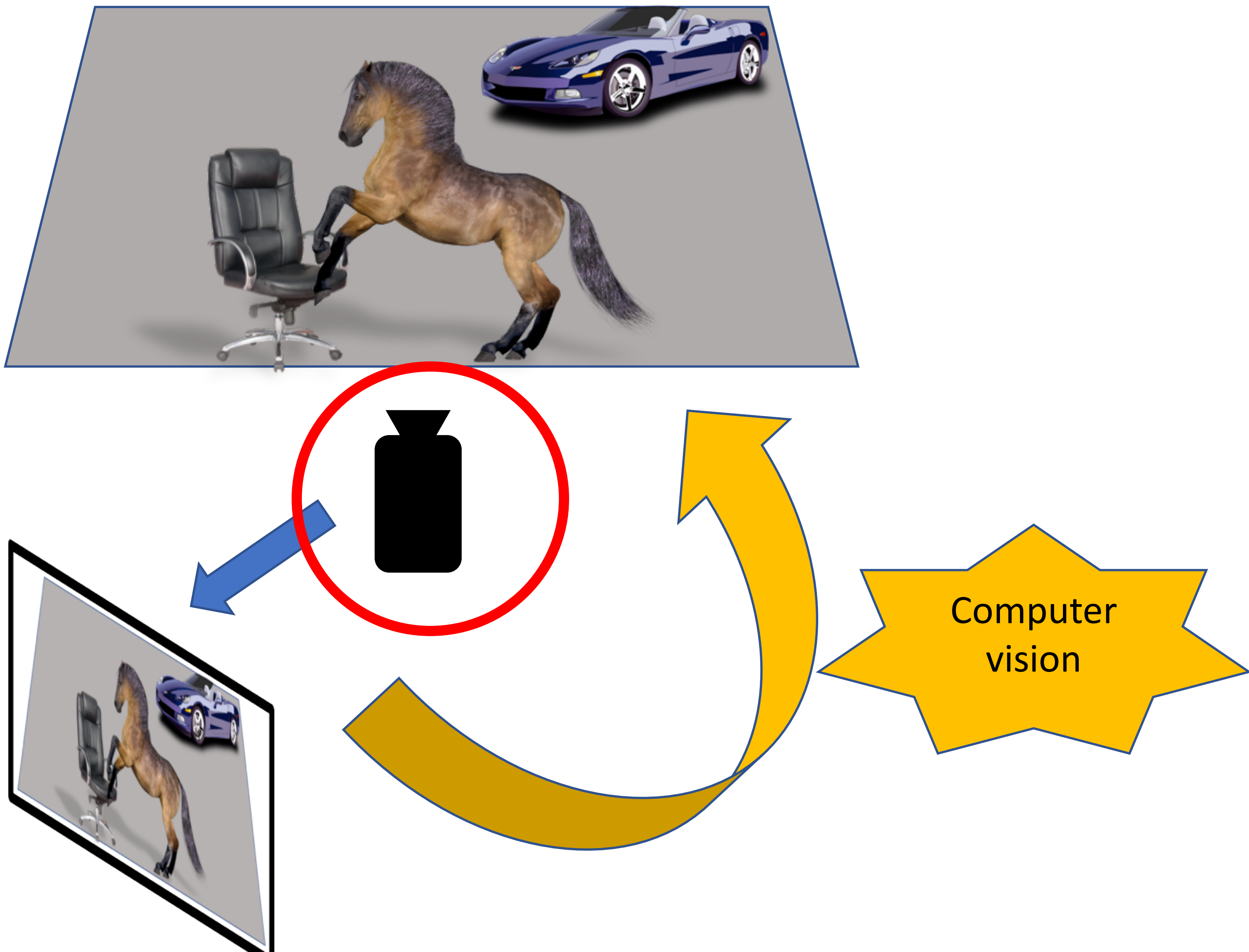
[Hartley & Zisserman]

The goal(s) of computer vision

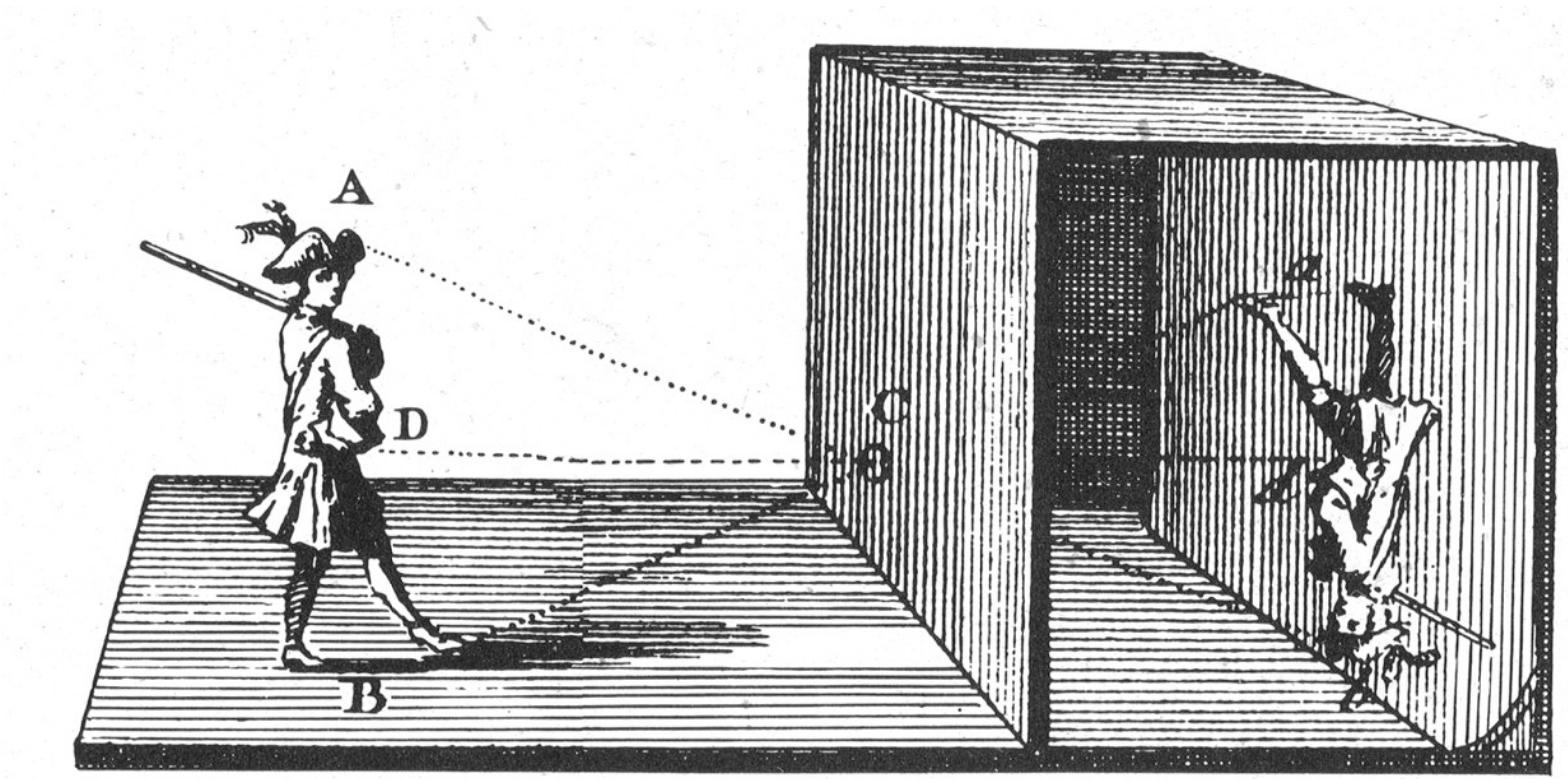
- *Recognition*
  - “Name” the object: what class does it belong to?

Picture credit: Magritte , Jon Barron

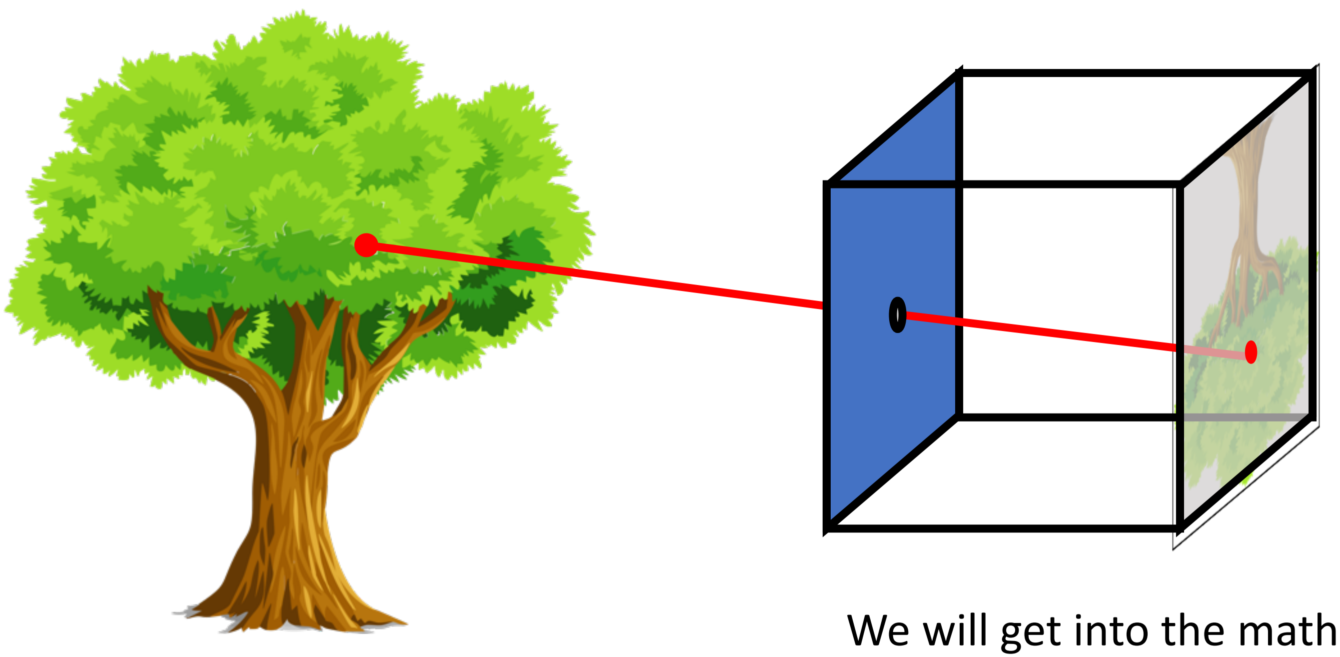
How do we do this?



The pinhole camera - *Camera Obscura*



The pinhole camera

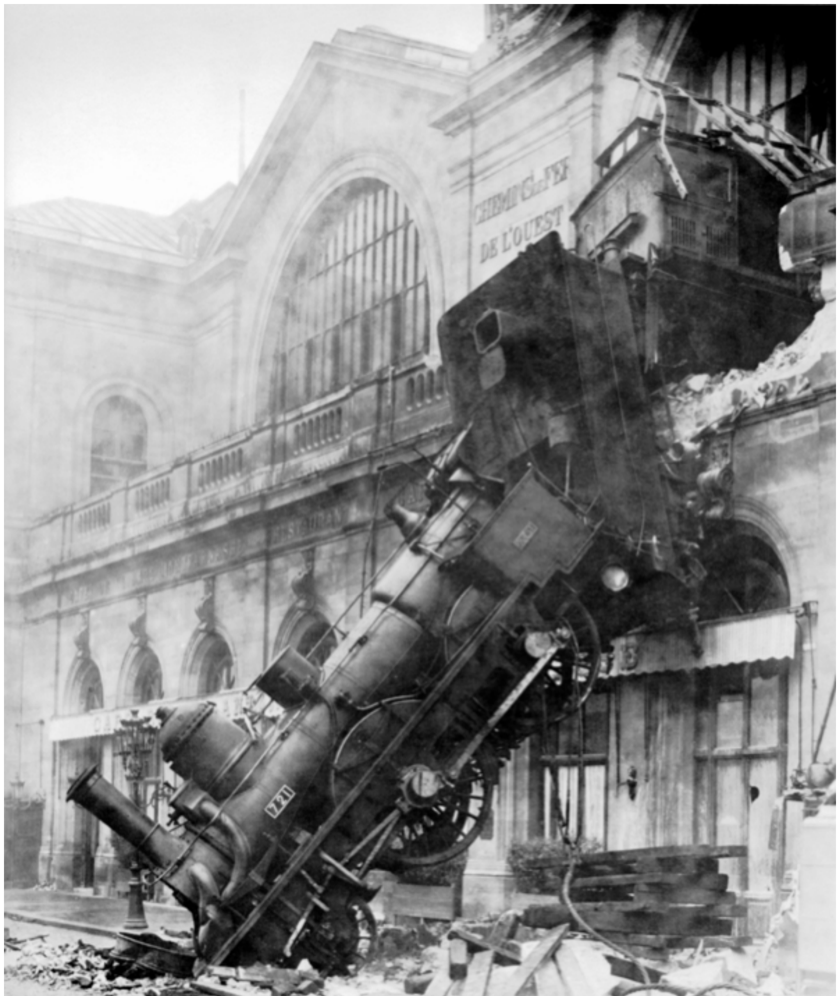


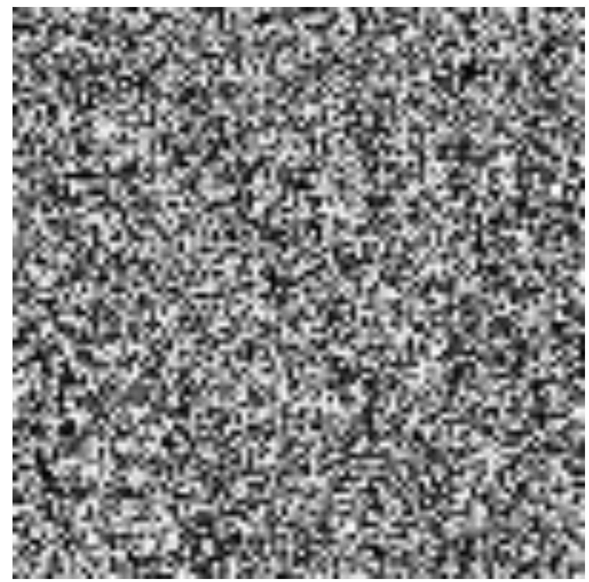
We will get into the math later

The pinhole camera

Not all 2D arrays are images

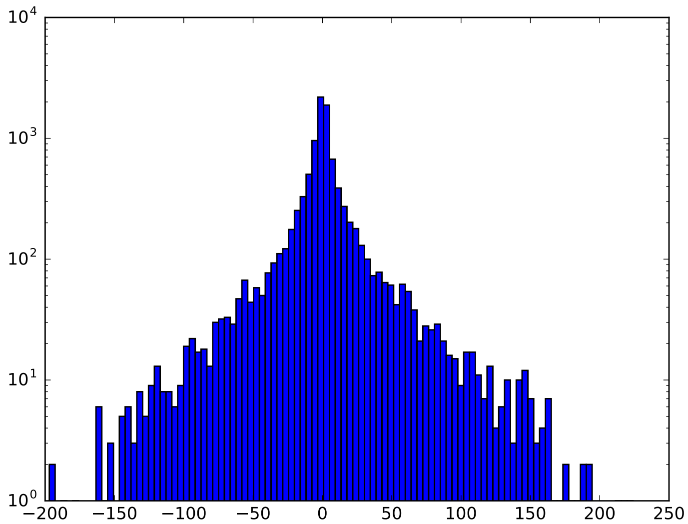
Consequence 1: nearby pixels are similar



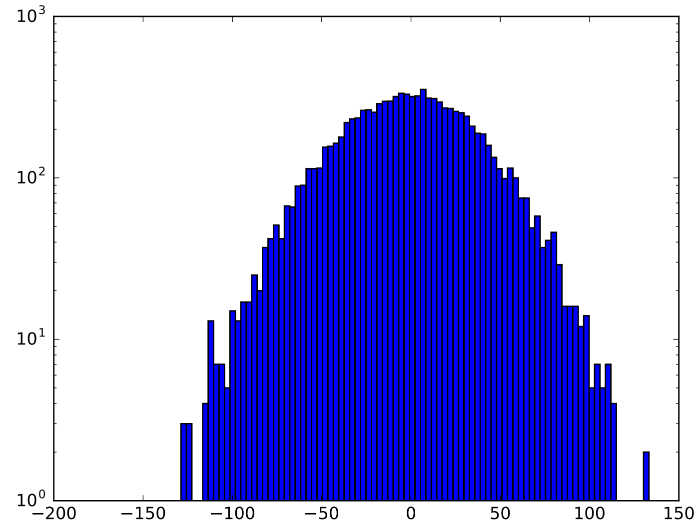


Consequence 1: nearby pixels are similar

Log histogram of differences between adjacent pixels



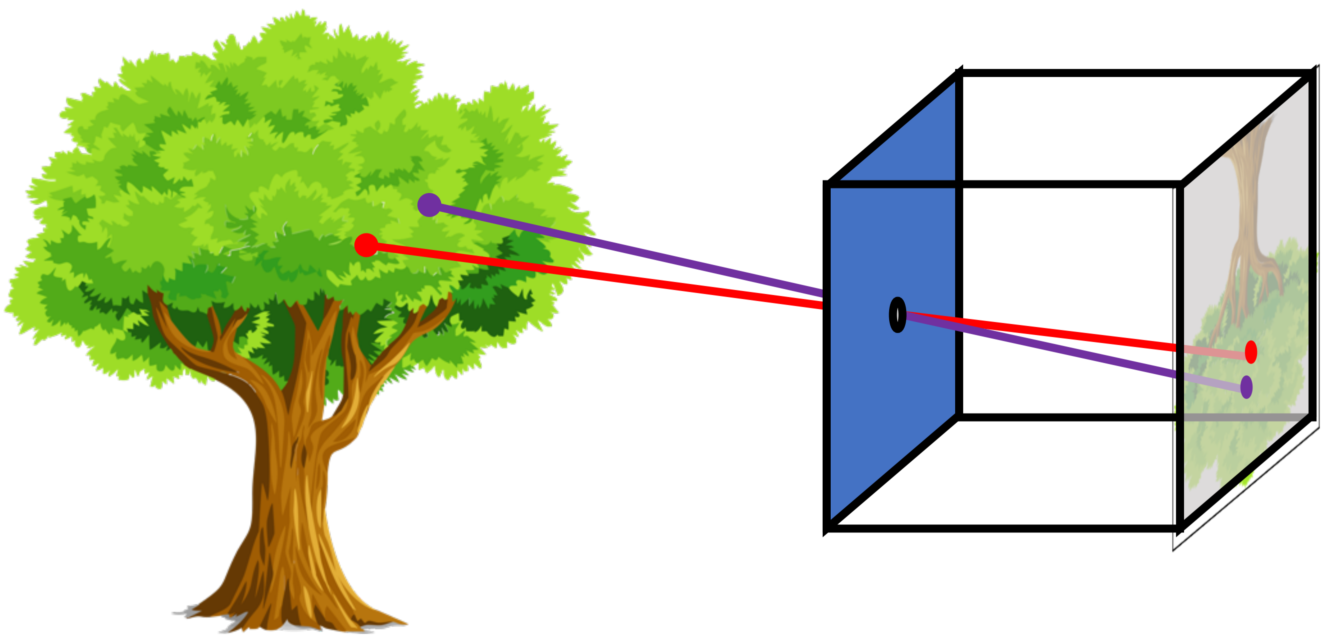
Natural images



Random arrays

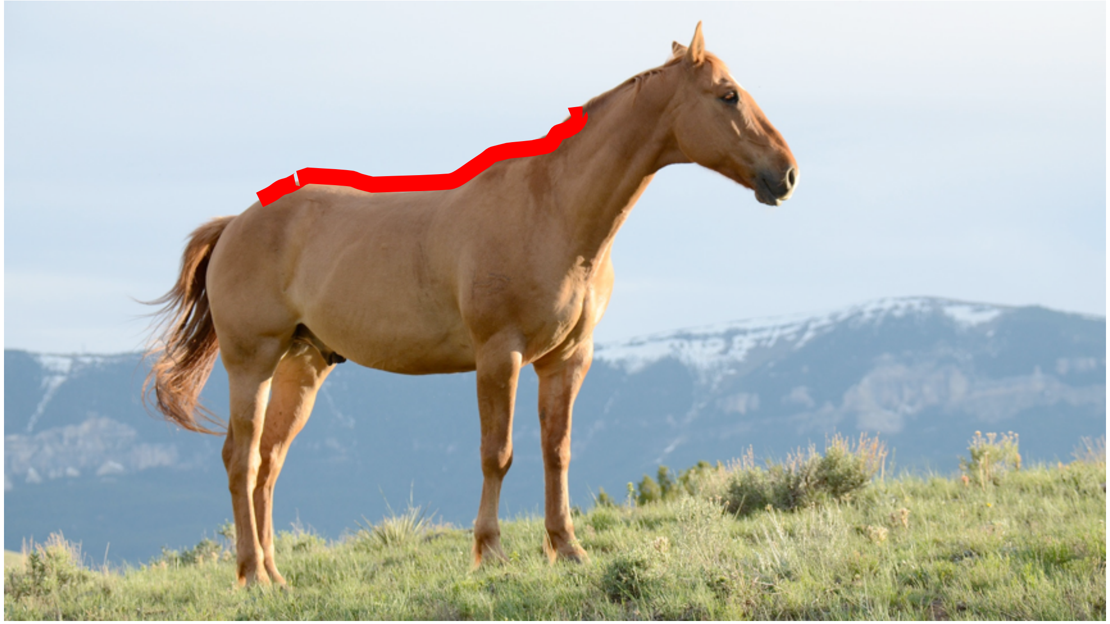
Consequence 1: nearby pixels are similar

- Why?
- Nearby pixels in pinhole camera lead to nearby rays
- Nearby rays *mostly* fall on the same object



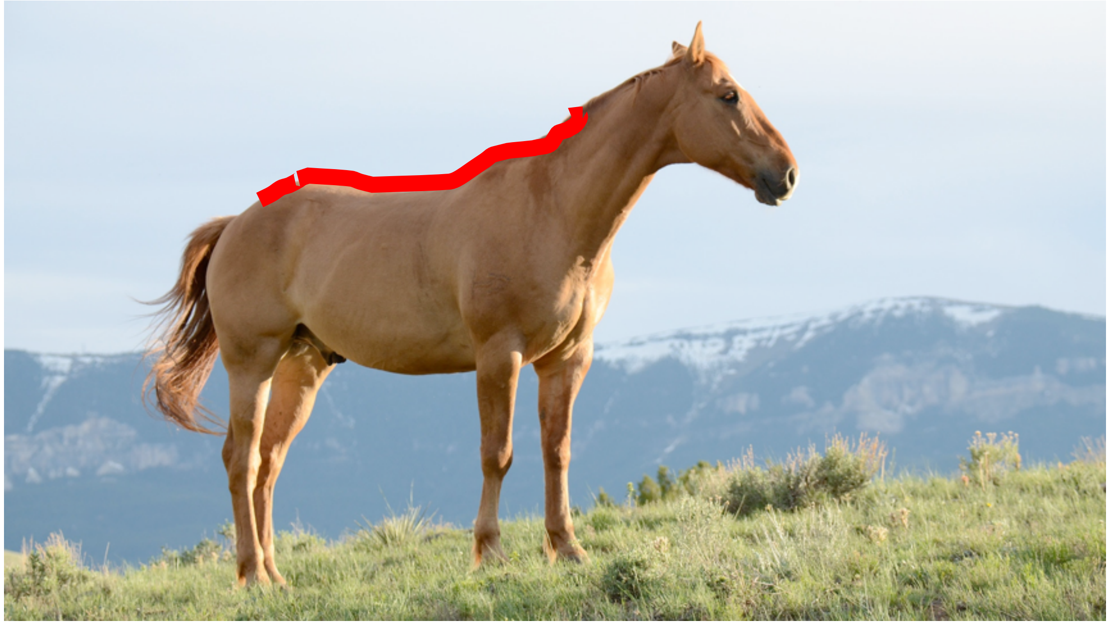
Consequence 1: nearby pixels are similar

- Nearby pixels that are *not* similar tend to lie on *different* objects
- Idea: To find where one object ends and another begins, look for *abrupt changes in color*



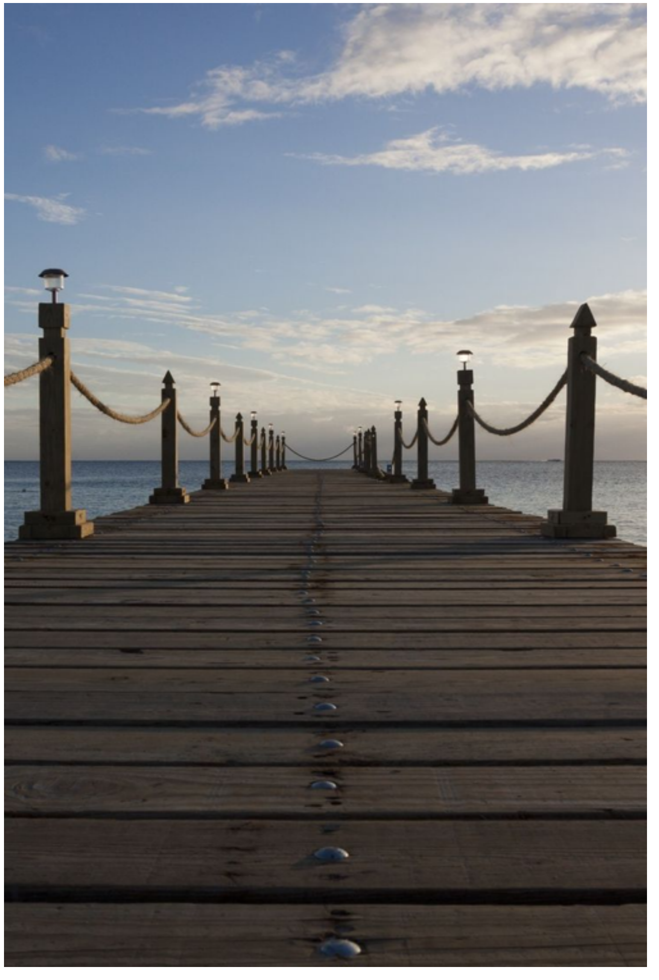
Consequence 1: nearby pixels are similar

- Places of color change might correspond to object boundaries
- Object boundaries are a clue to *object shape*
- Idea: *Use rough boundaries to recognize object(s)*

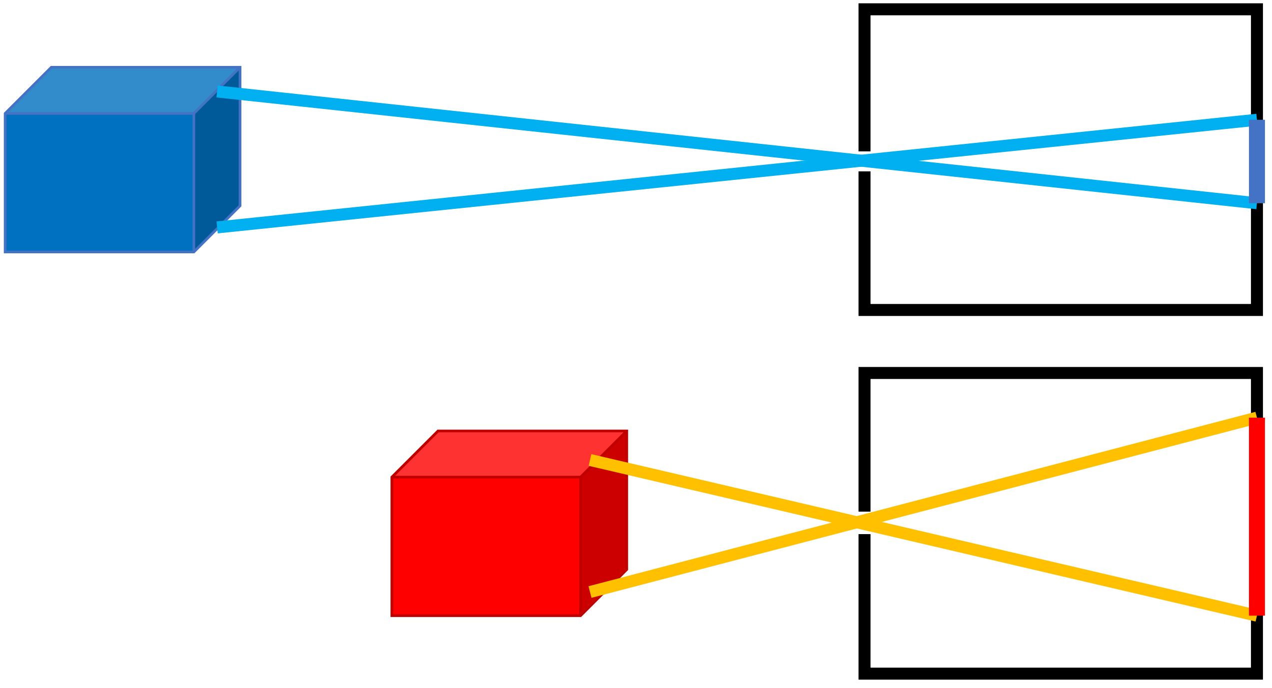


Counterexample: camouflage

Consequence 2: Farther away objects appear smaller

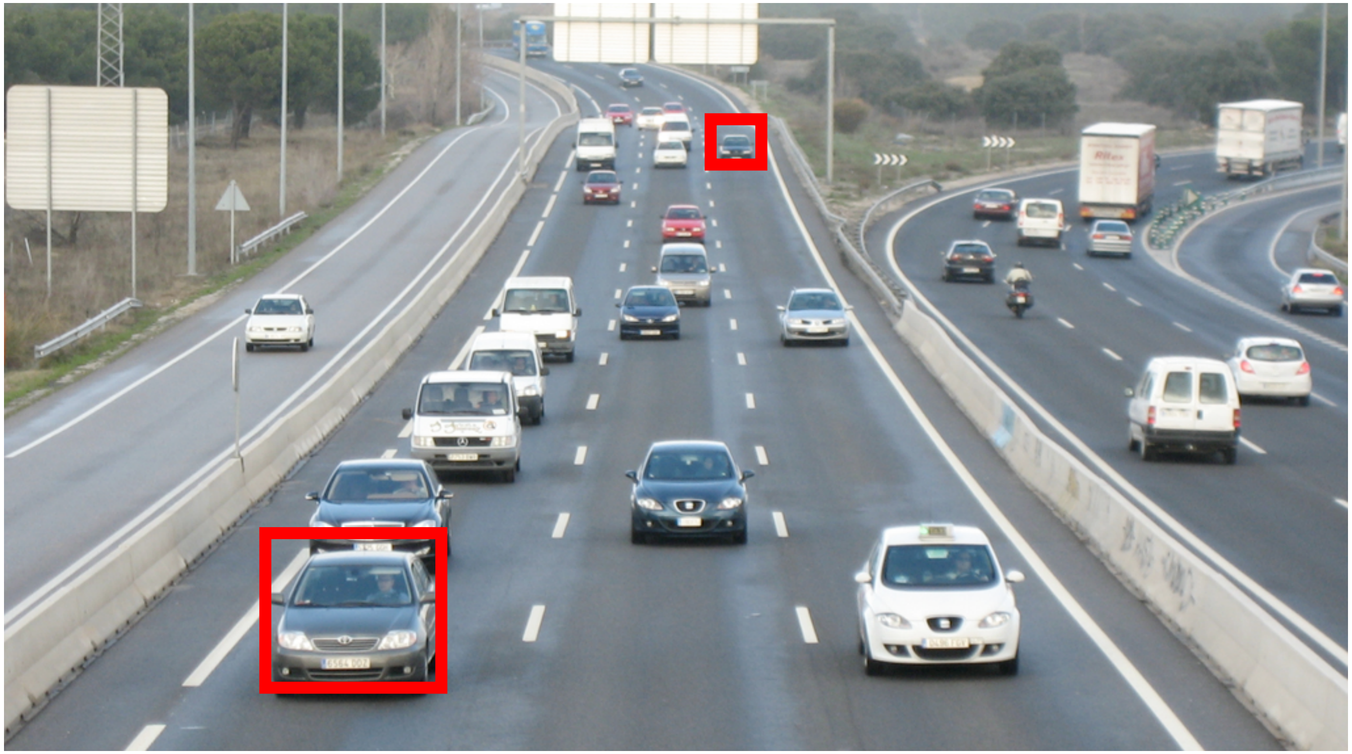


Consequence 2: Farther away objects appear smaller



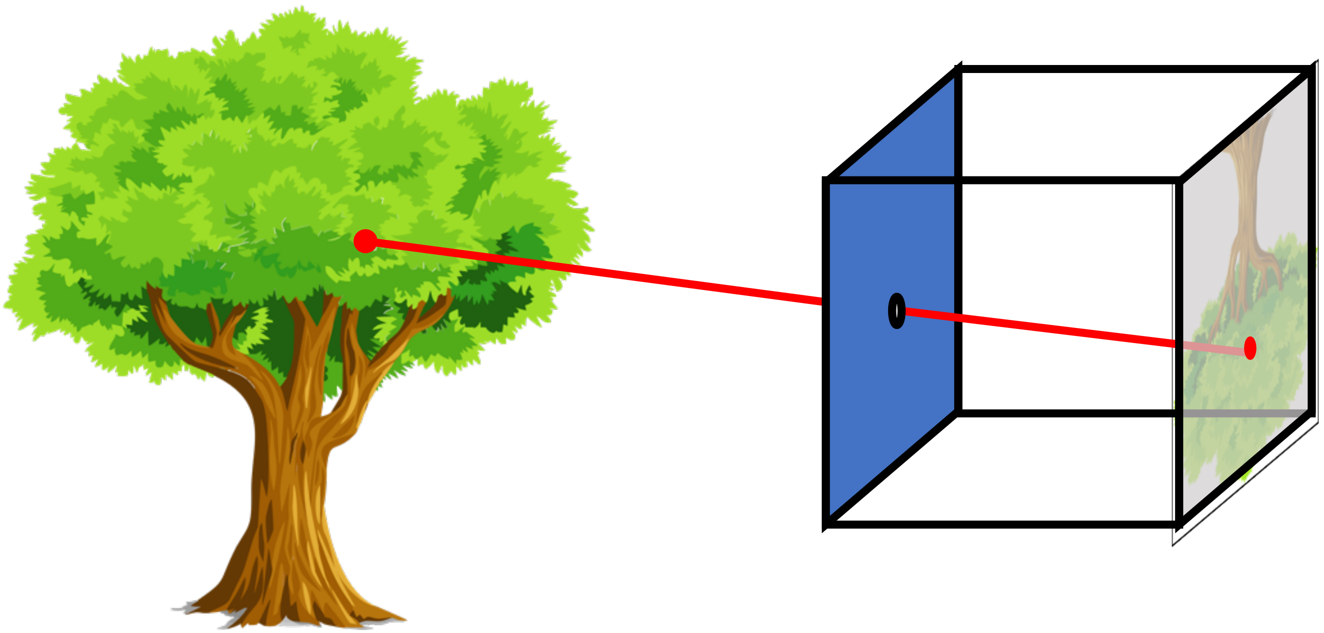
Consequence 2: Farther away objects appear smaller

- Key modules: search over scales, zoom-out/zoom-in



Consequence 3: Image formation is lossy

- We lose depth information



Consequence 3: Image formation is lossy

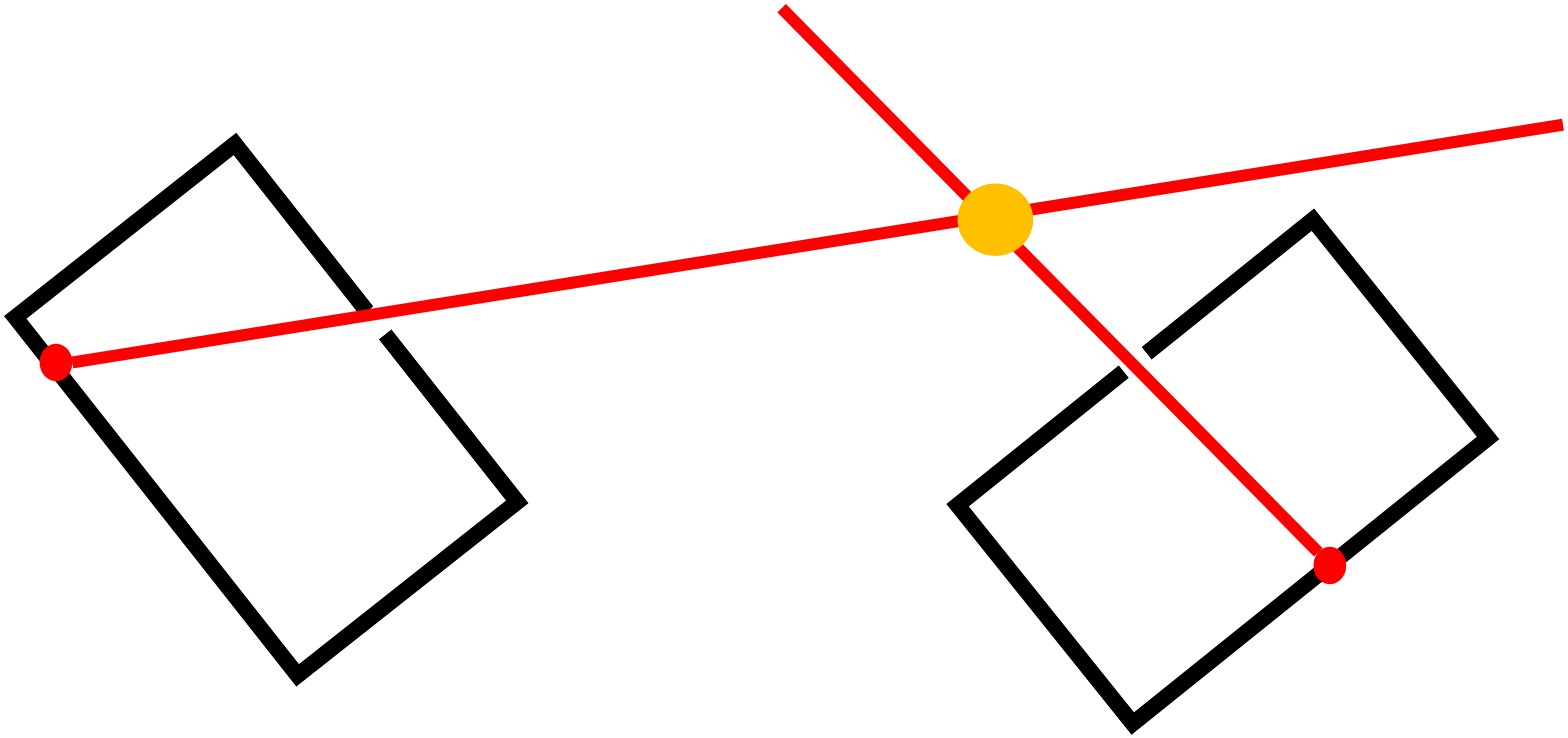
- Idea: use multiple images

Consequence 3: Image formation is lossy

- Idea: use multiple images

Consequence 3: Image formation is lossy

- Idea: use multiple images

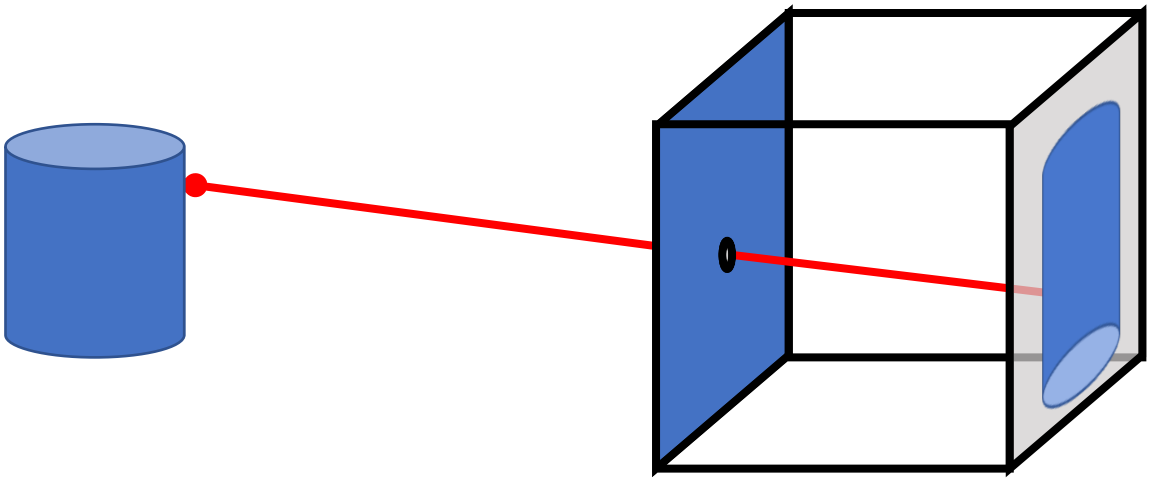


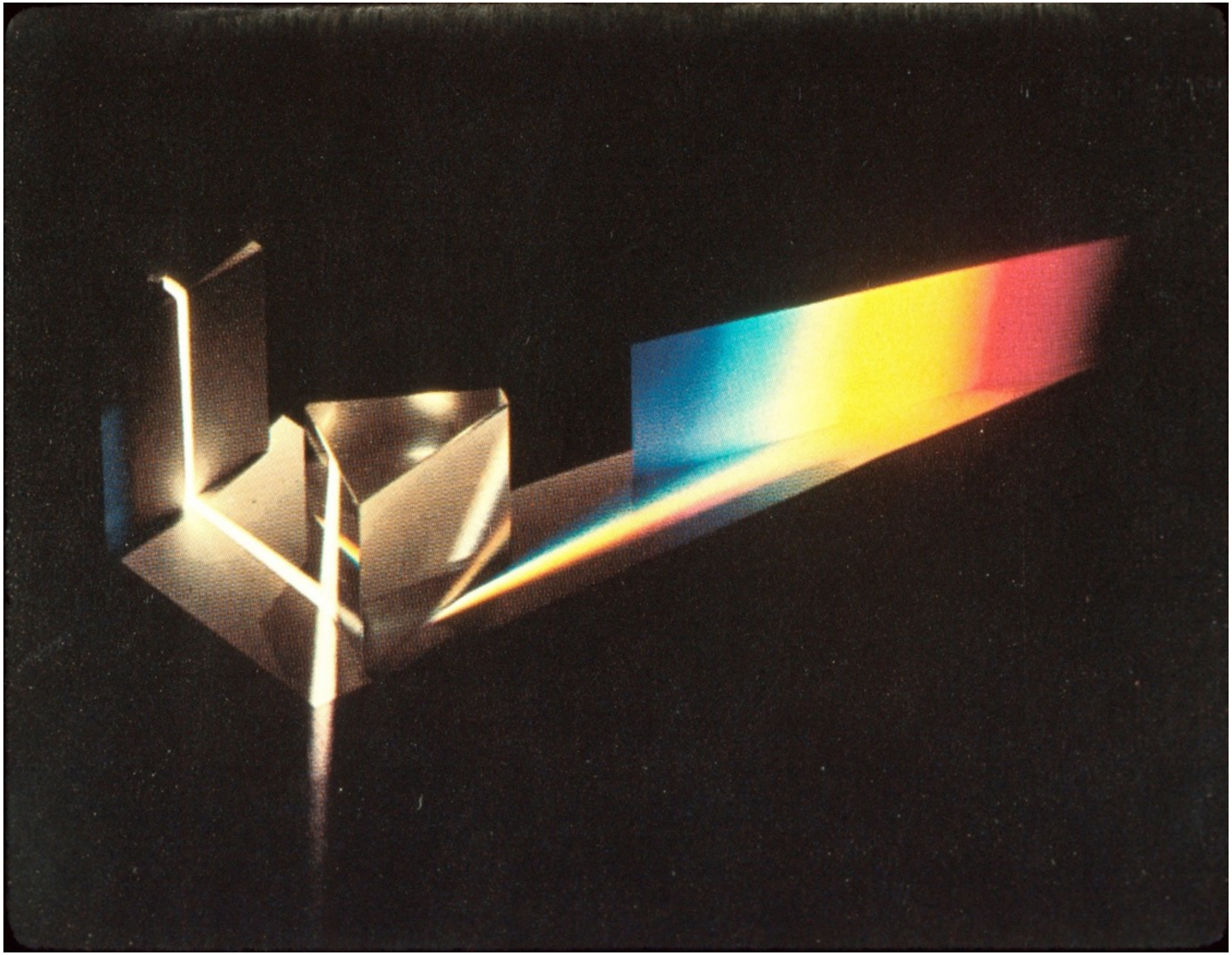
Consequence 4: Image formation is lossy

- Idea: use multiple images
- Need to find which pixel in image 2 matches which in image 1 - the *correspondence* problem

Color

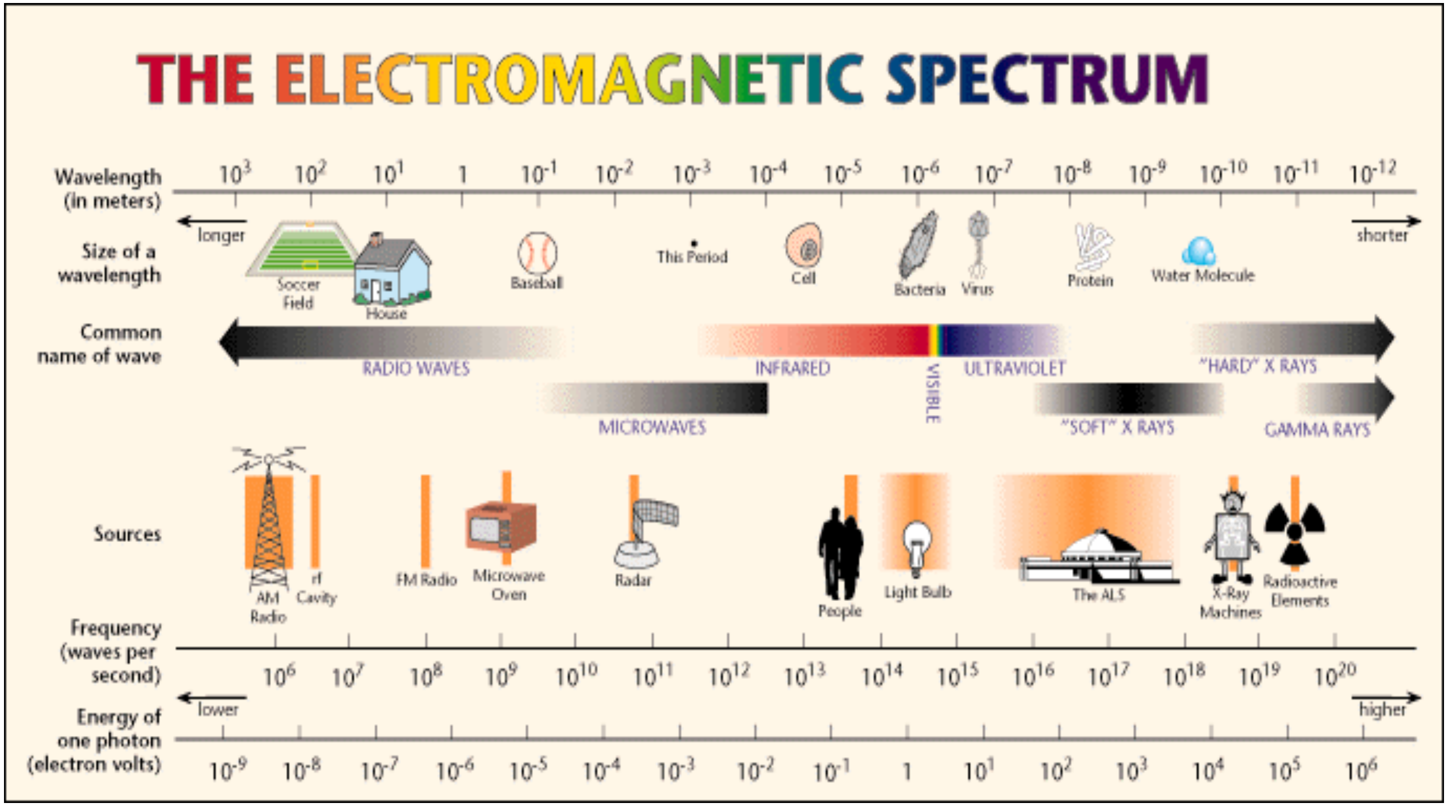
- Each pixel records "color" of a ray
- But *what is color?*





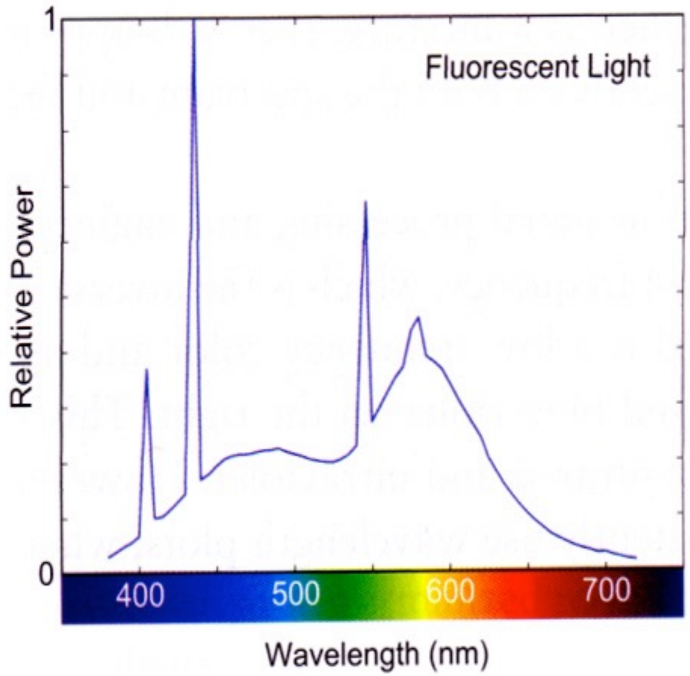
What is light?

- Light is electromagnetic radiation



[Lawrence Berkeley Lab / MicroWorlds]

Physics to Brain



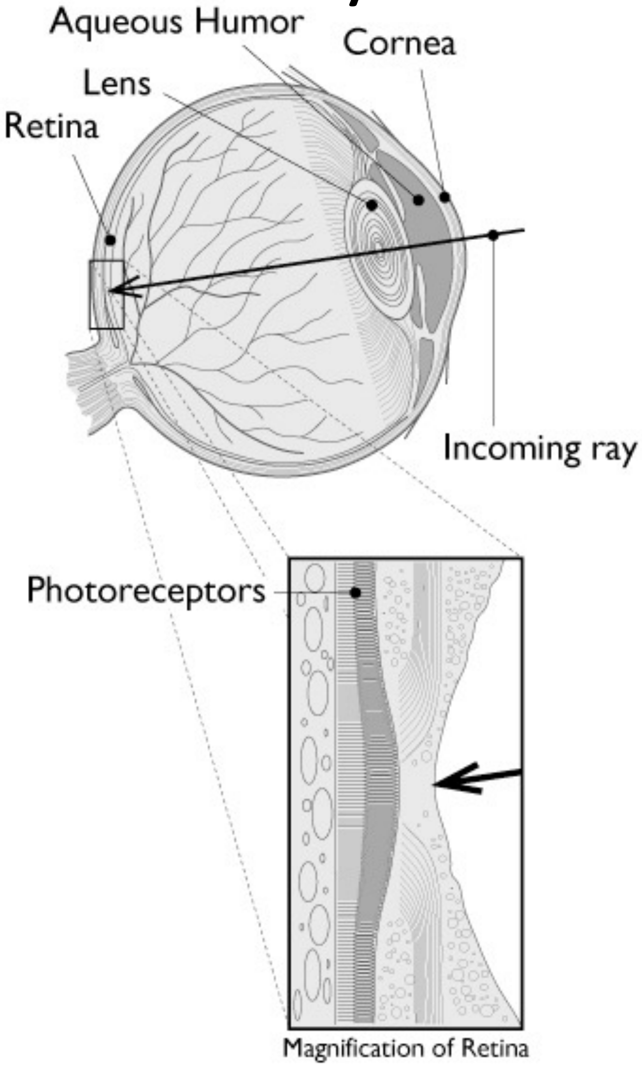
[Stone 2003]

→ ?

Physical

Perceptual

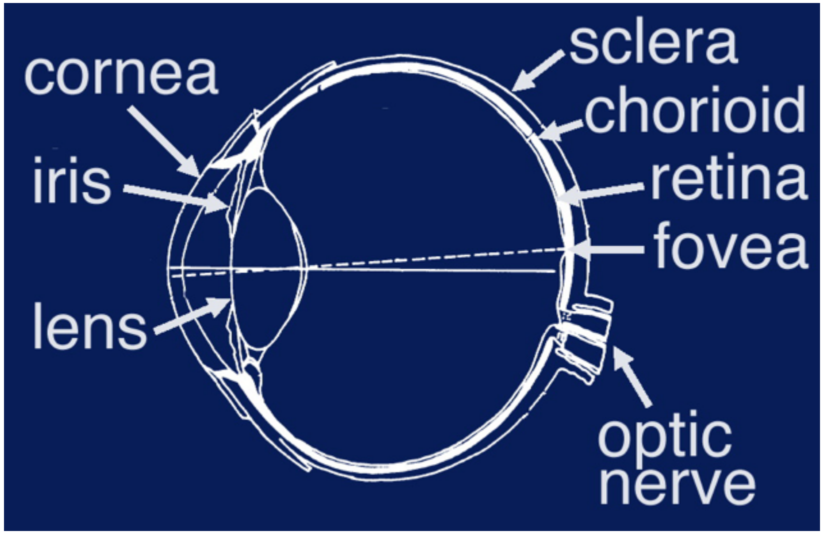
The eye as a measurement device

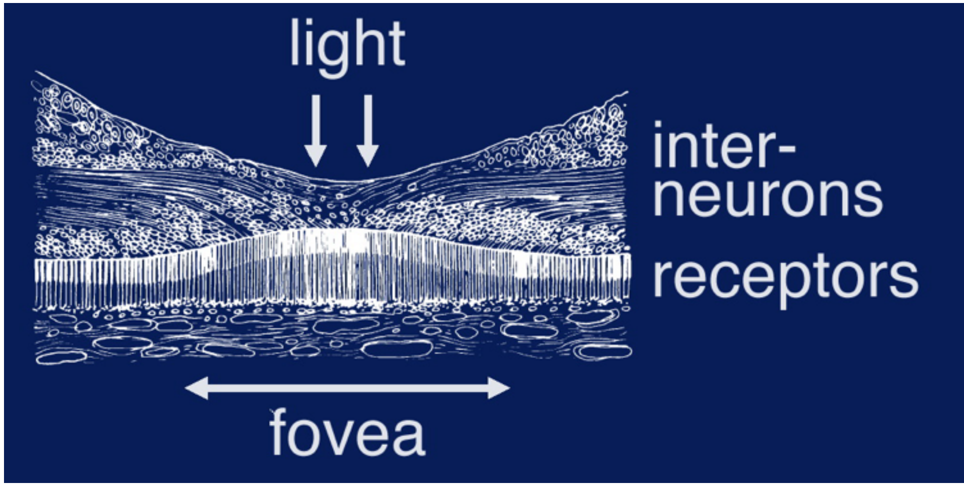


[Greger et al. 1995]

- Light is measured by the *photoreceptors* in the retina
- Photoreceptor cells *absorb* photons and convert to electrical signals
- Different photoreceptor types respond to different wavelengths

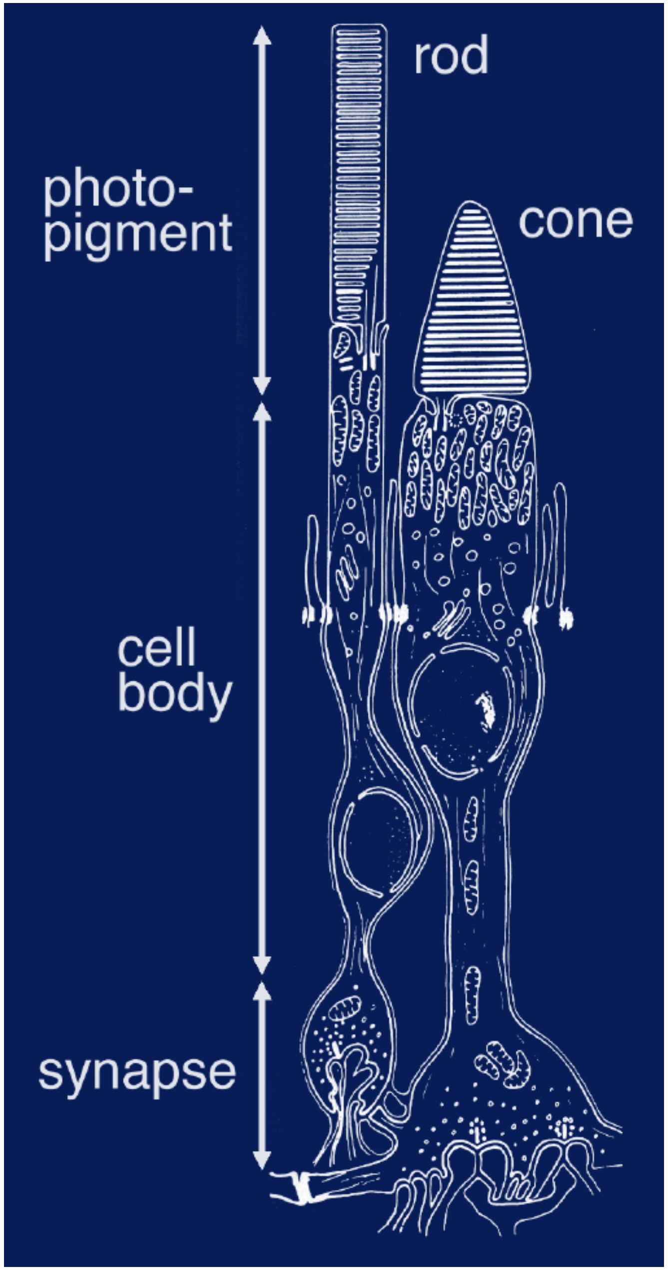
The eye





Left figure from Atkinson, Steven's Handbook of Experimental Psychology © Wiley, 1986.

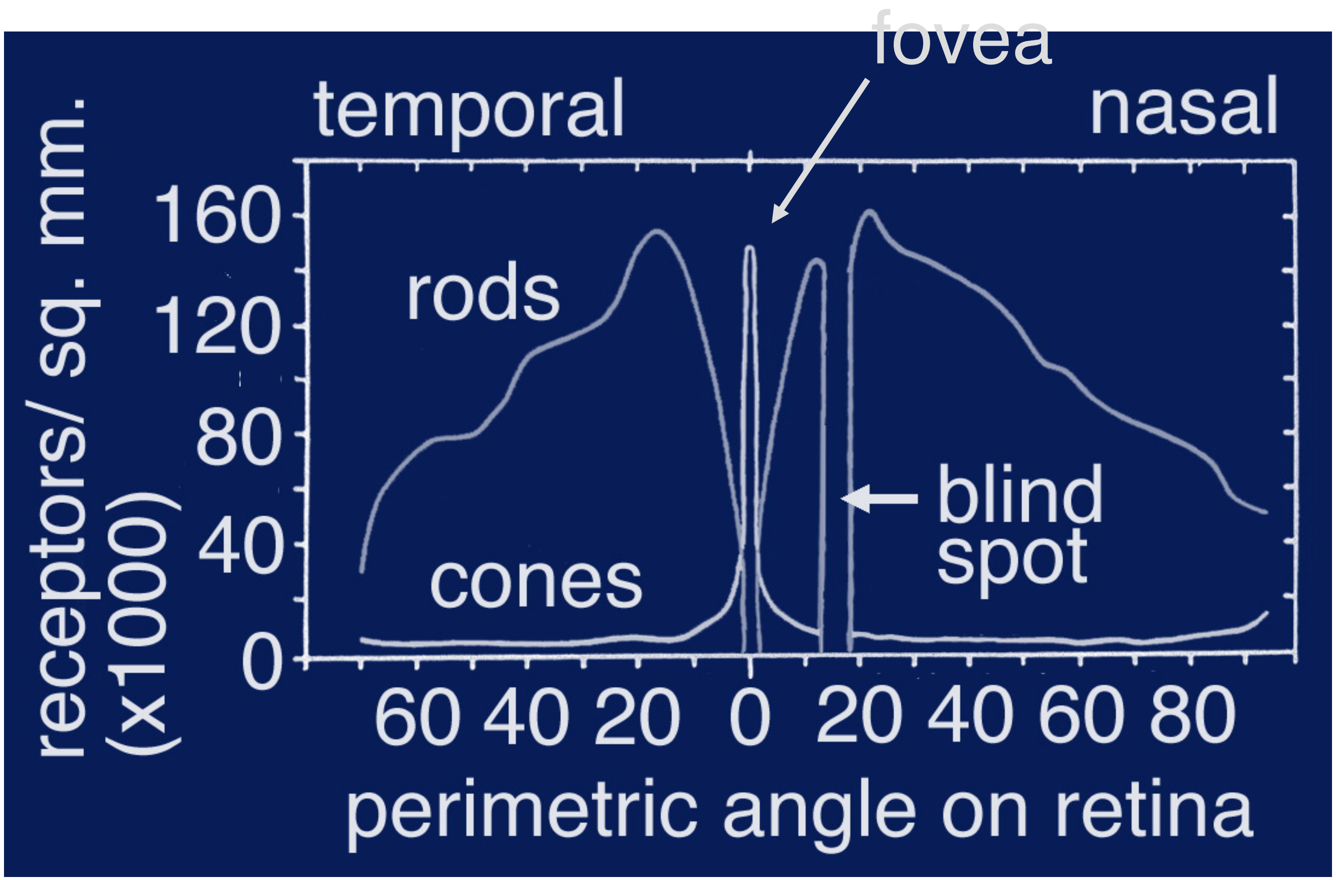
Right figure from Sekuler, Perception © Alfed A. Knopf, 1985.



Photoreceptors

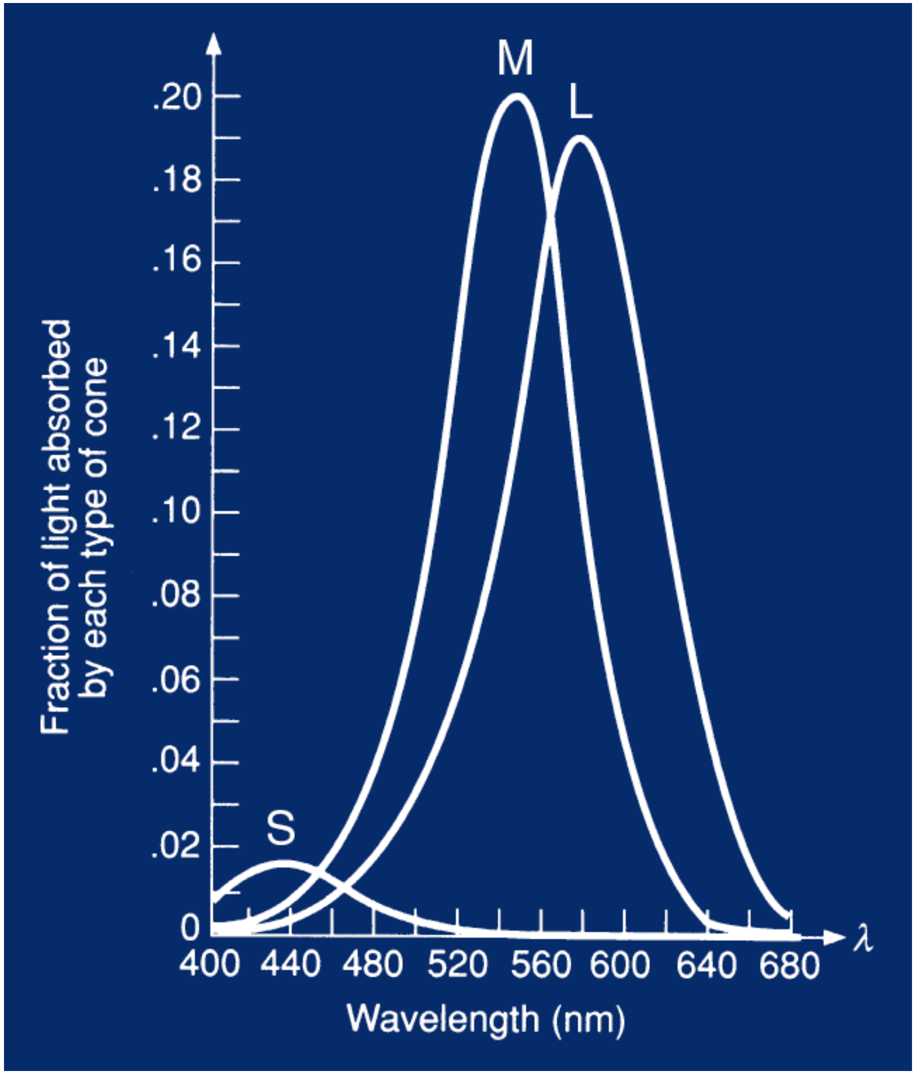
- 120 million rods
- 7-8 million cones
  in each eye

Receptor distribution



Adapted from Levine, Vision in Man and Machine © McGraw-Hill, 1985.

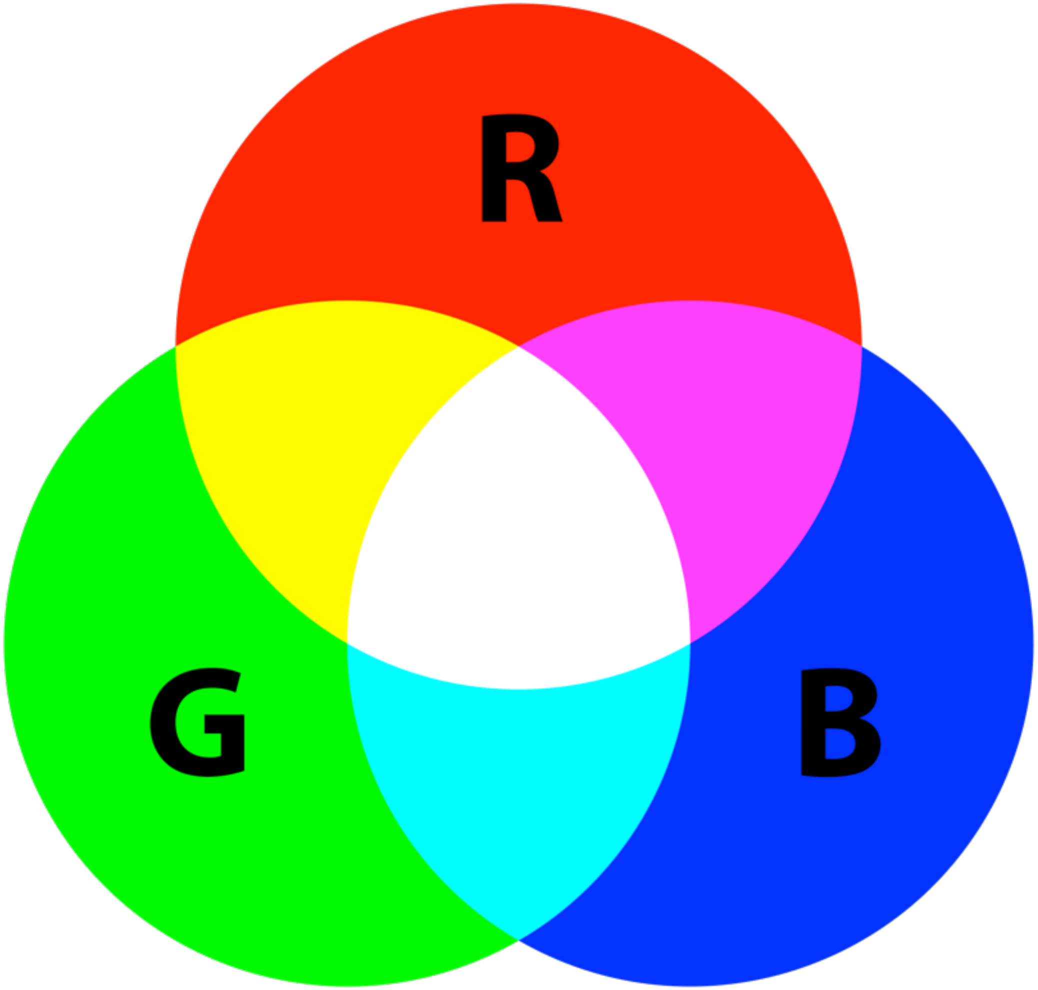
Cone Responses



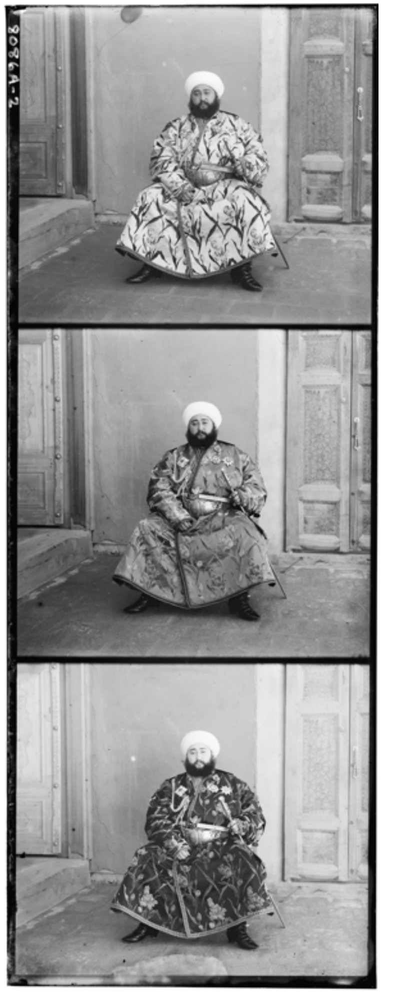
- S,M,L cones have broadband spectral sensitivity
- Converts a *distribution* over wavelength into 3 values
- *Hence 3 colors: blue (S), green (M), red (L)*

[source unknown]

Color

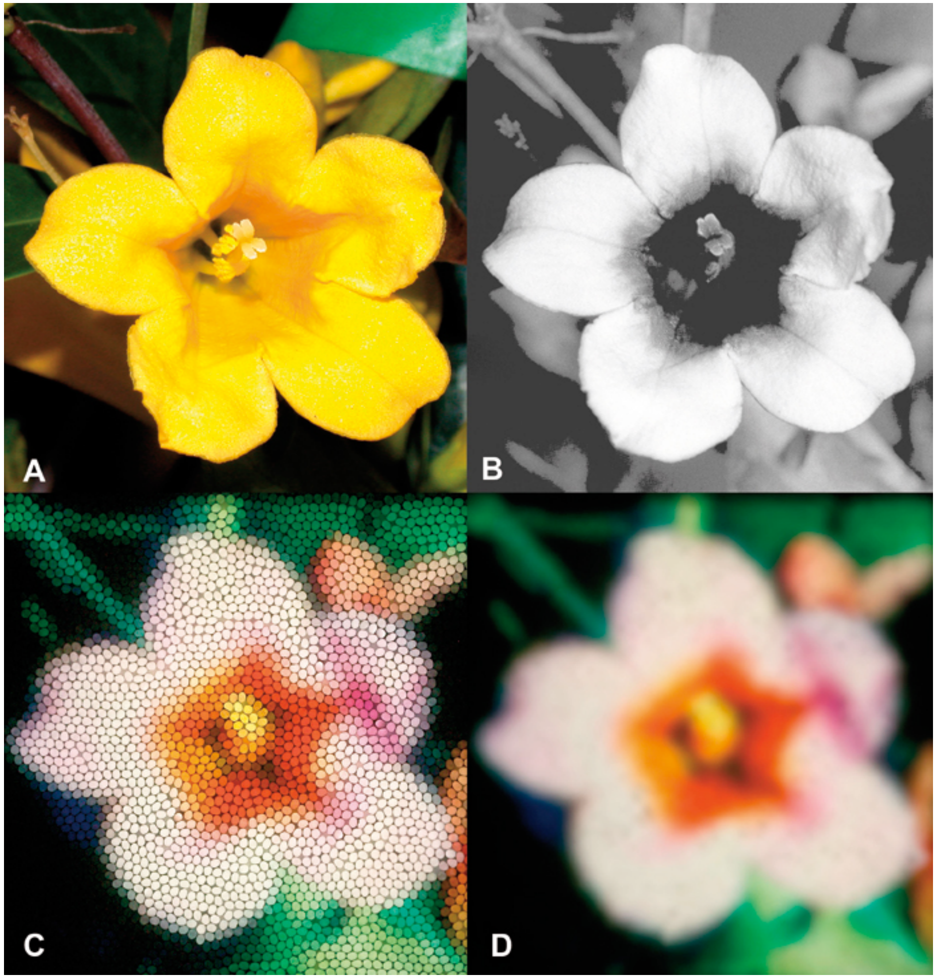


Color photography



The Emir of Bukhara, Alim Khan, in a 1911 color photograph by Sergey Prokudin-Gorsky. At right is the triple color-filtered black-and-white glass plate negative, shown here as a positive. [wikipedia article on color photography]

Color and light

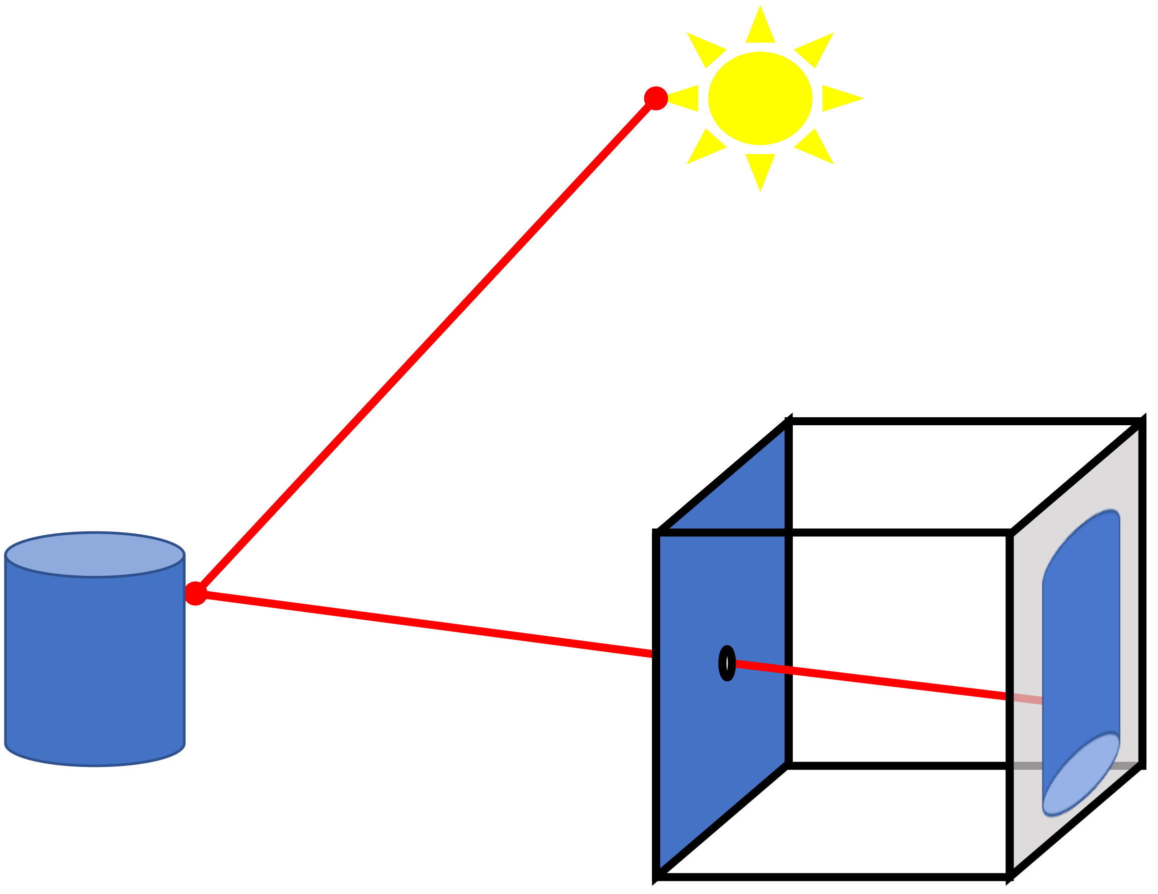


How bees see the world

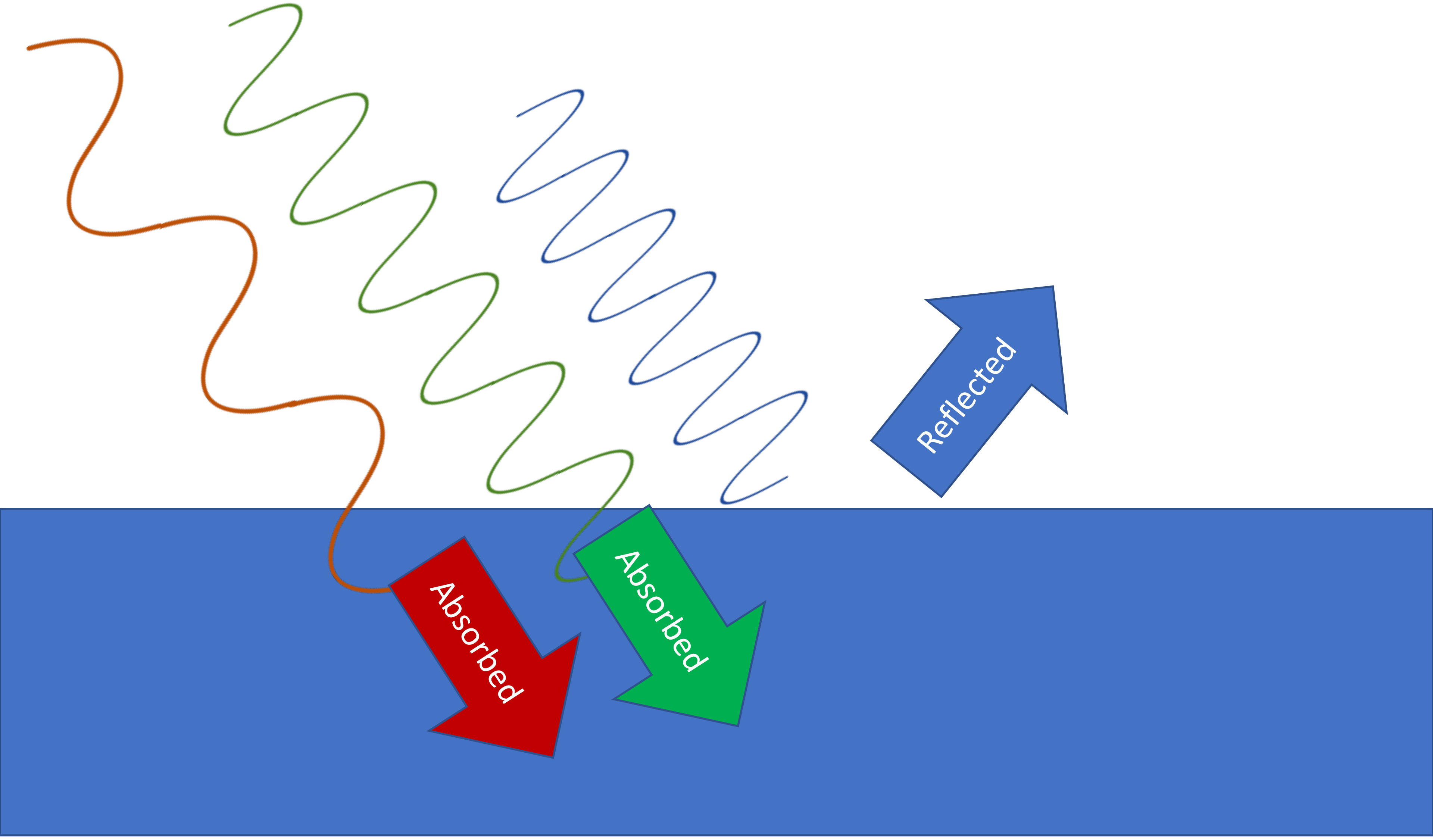
https://beecare.bayer.com/media-center/beenow/detail/vision-science-how-bees-perceive-the-world

Color and light

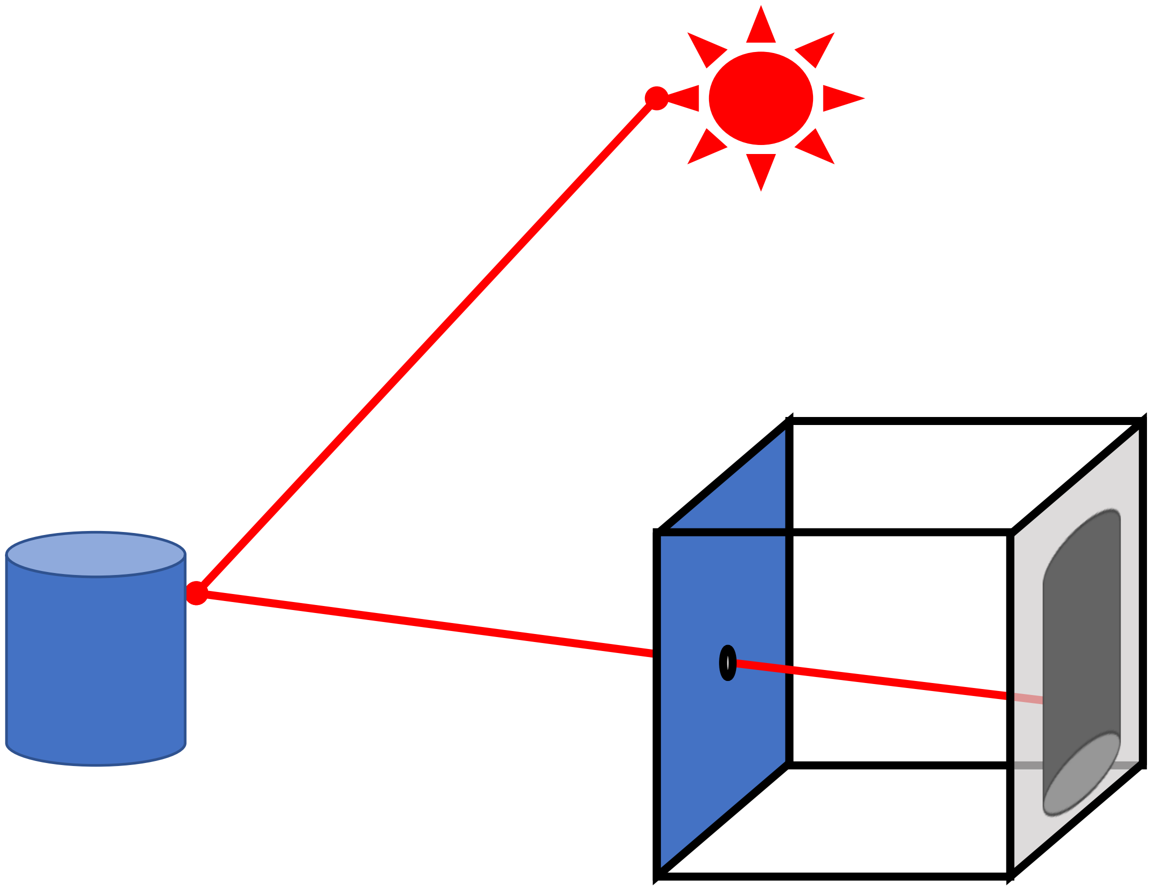
- Each pixel records amount of energy in red light, blue light green light
- But where does light energy come from?



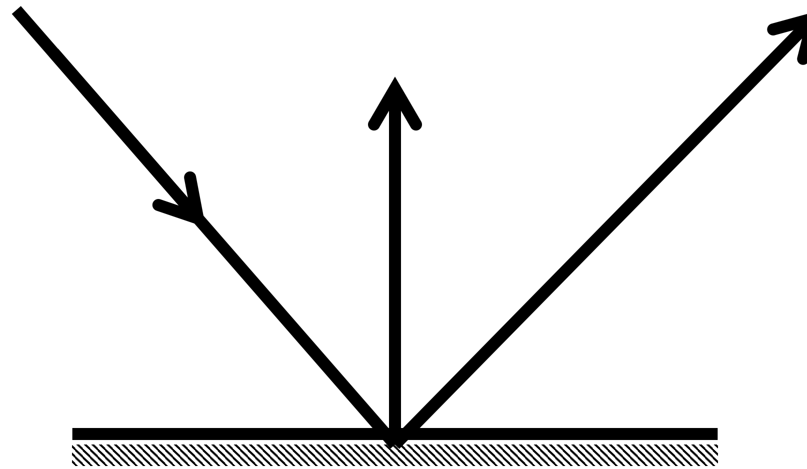
Color and light



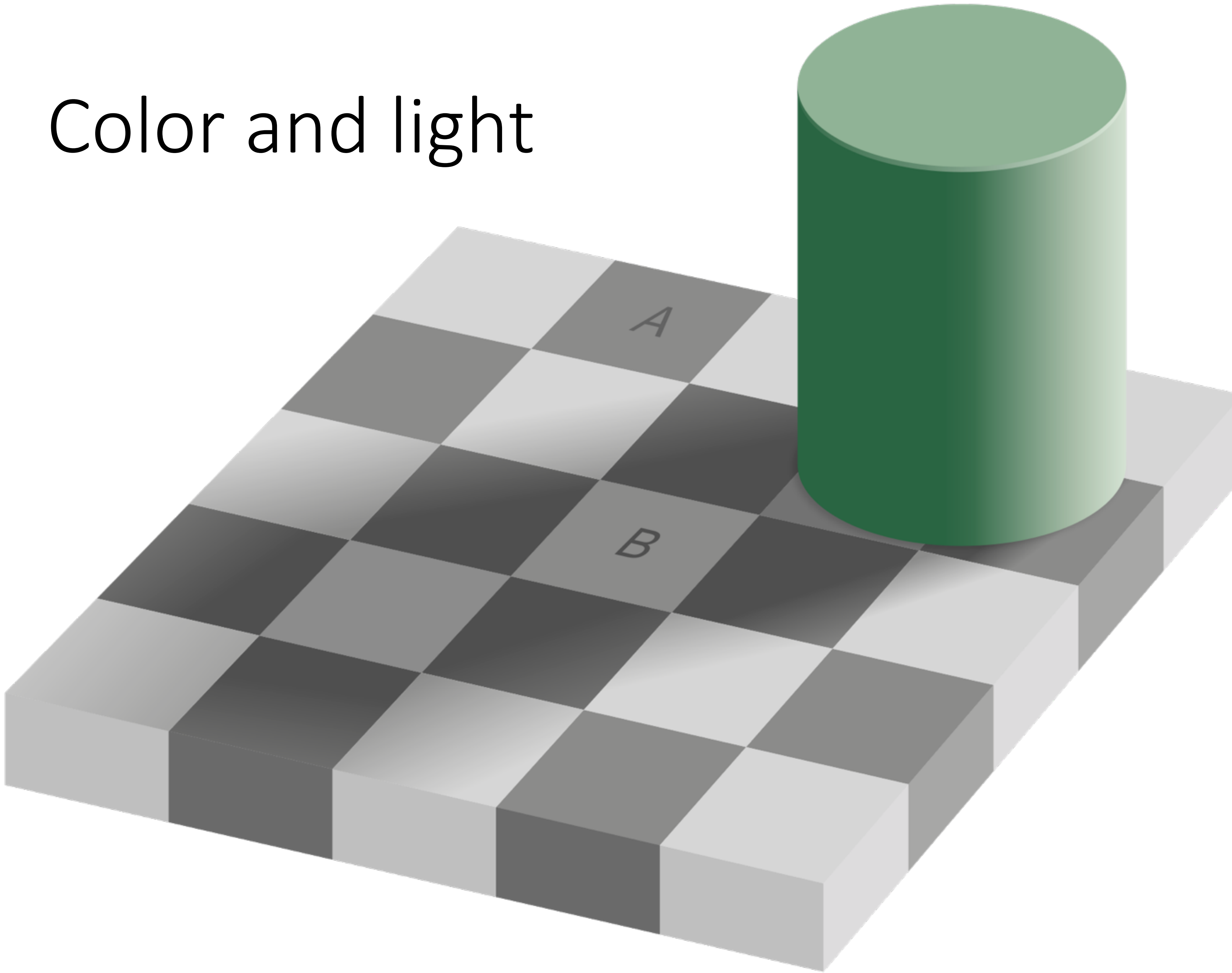
Color and light



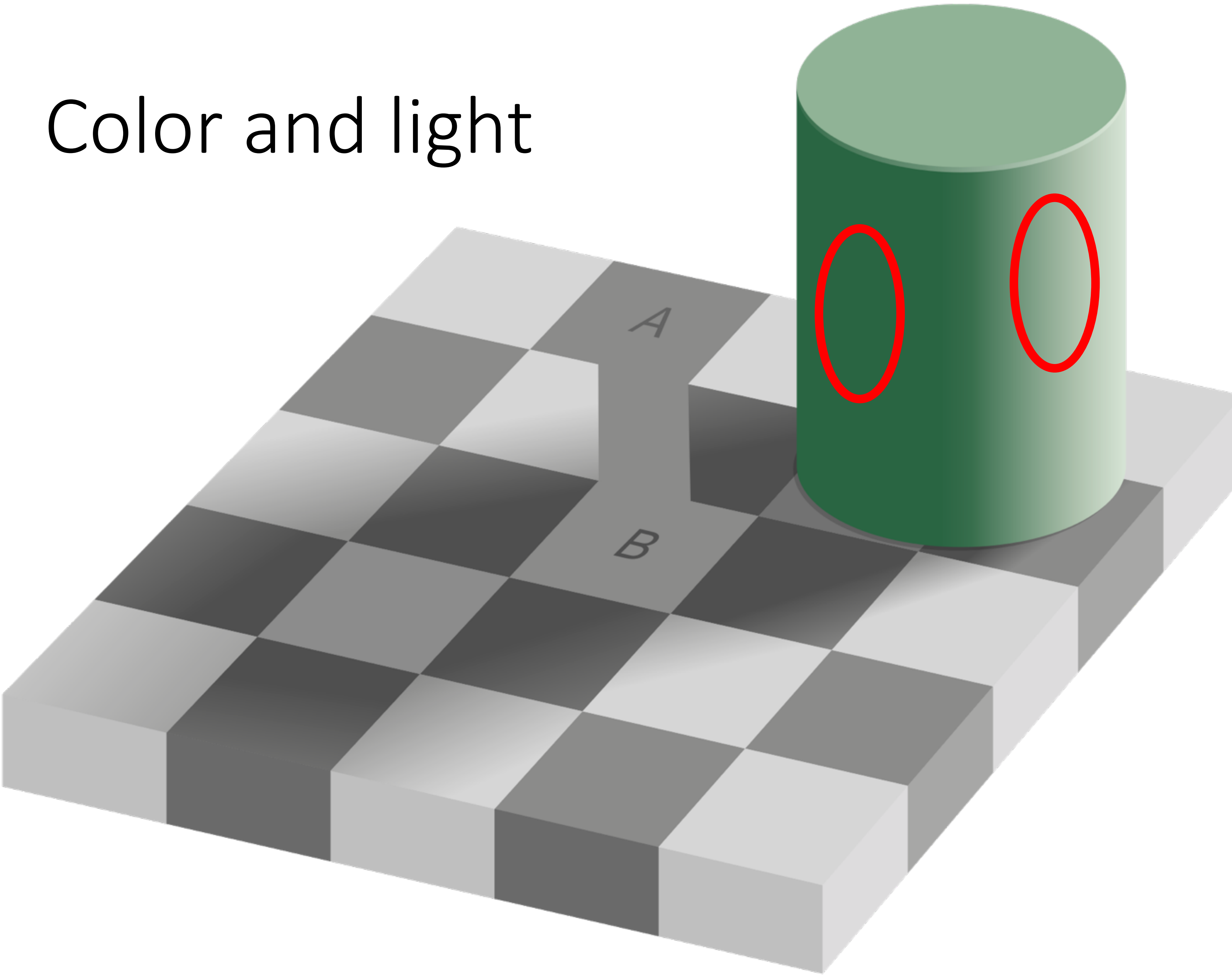
Color and light



Color and light



Color and light



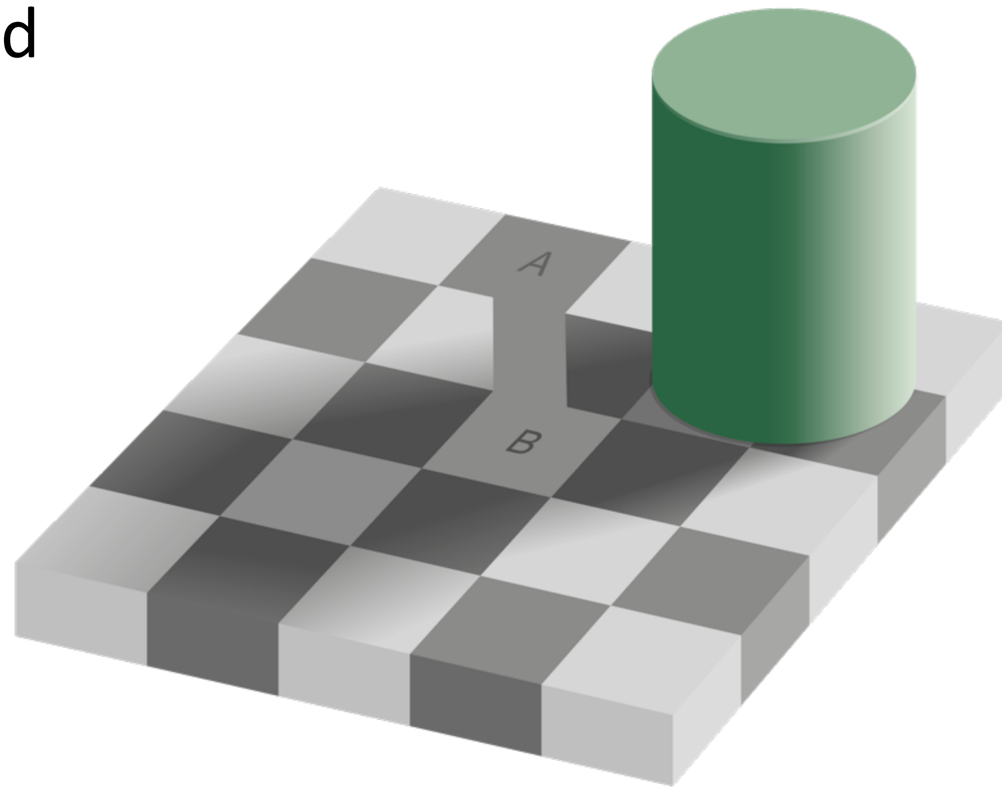
Color and light

- Color of a pixel depends on:
  - Color of light
  - “Paint” on surface
  - Direction of light w.r.t surface
  - Viewing direction
  - Presence/absence of cast shadows

We will get into the math and physics later

Consequence 4: Pixel color is complicated

- Idea: rely less on absolute color. Look at *changes in color* (may be object boundaries or change in paint) instead



Consequence 4: Pixel color is complicated

- Understanding light can give us clues to shape

Take-away

- Natural images are *not* arbitrary 2D arrays
- They have properties resulting from physics / math of image formation
- Solving computer vision requires using these properties

Some primitives

- Edge detection: identifying where pixels change color
  - Cue to object boundary
  - Cue to shape
  - More resilient to lighting than pixel color
- Zooming into or out of images
  - Searching for both nearby and far-off objects
- Matching patches from two different images
  - First step in identifying 3D location

Other related problems

- Image Restoration
  - denoising
  - deblurring
- Image Compression
  - JPEG, JPEG2000, MPEG..
- Again, use the same ``priors”

Next up: Image processing

Let’s enhance

Let's Enhance (HD)