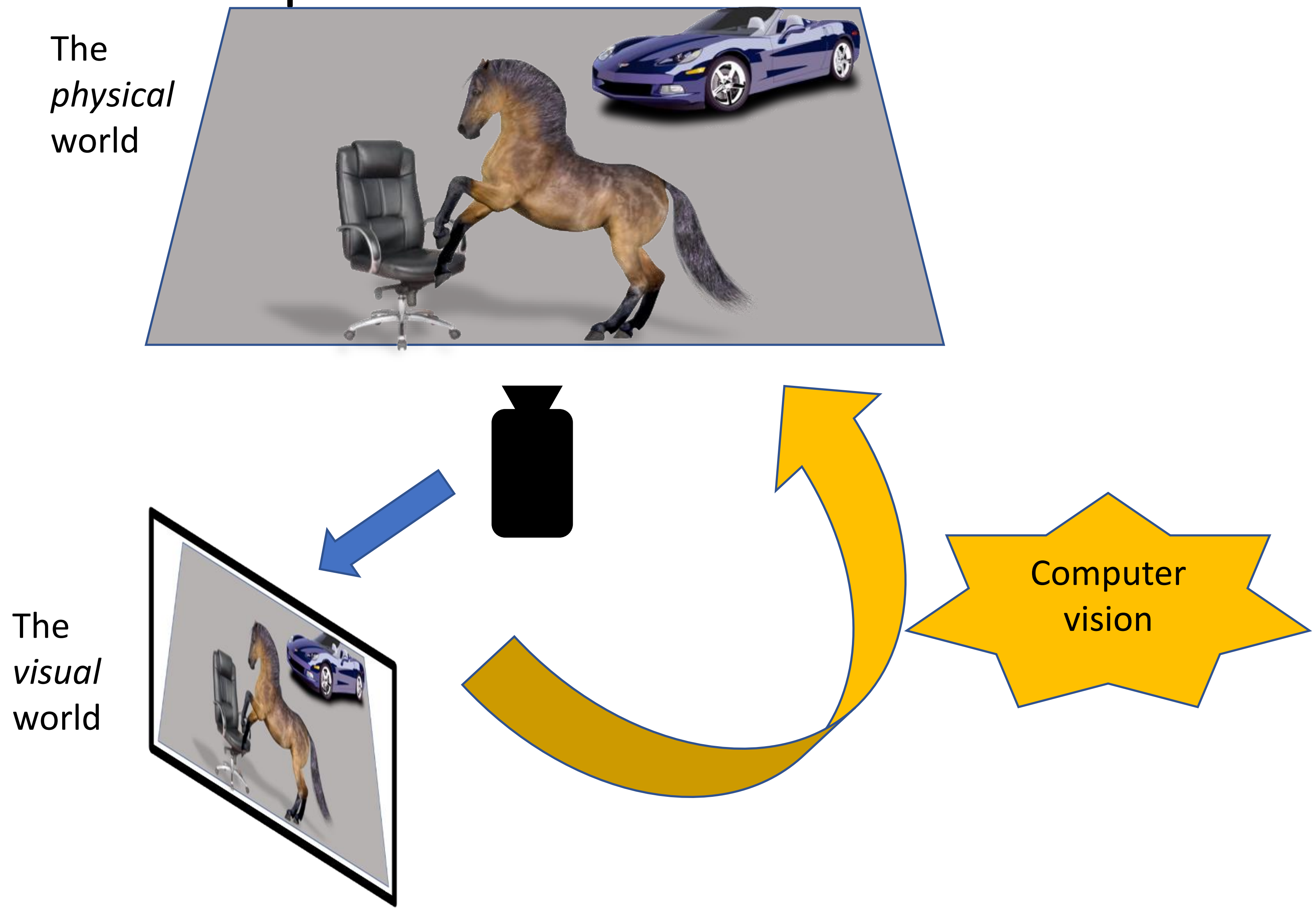
Computer Vision (CS 419/619)
Basics of Computer Vision

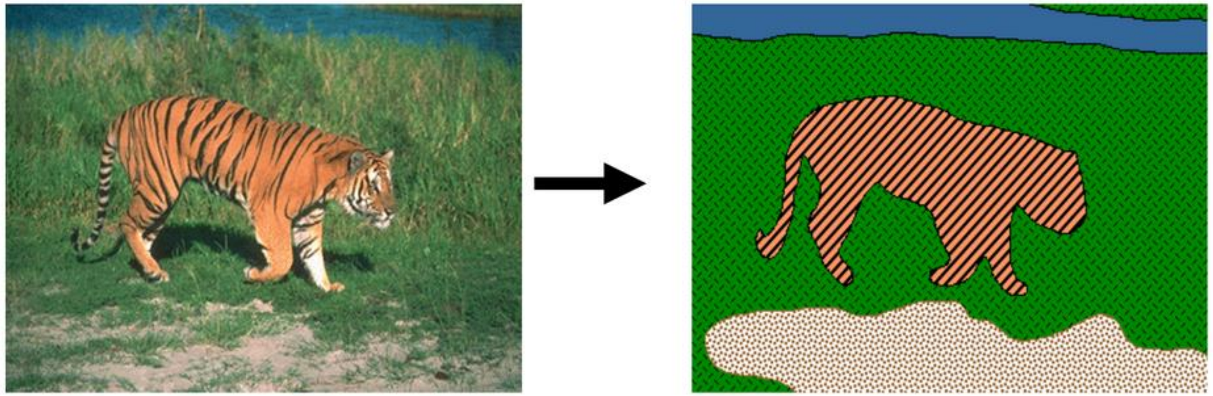
Dr. Puneet Gupta

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?



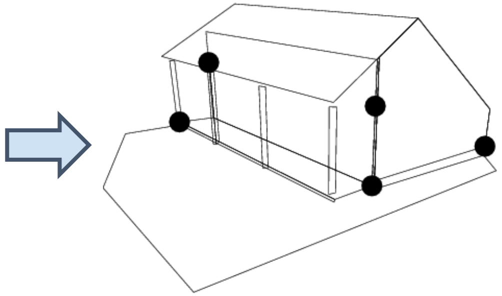
[Blobworld]

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]

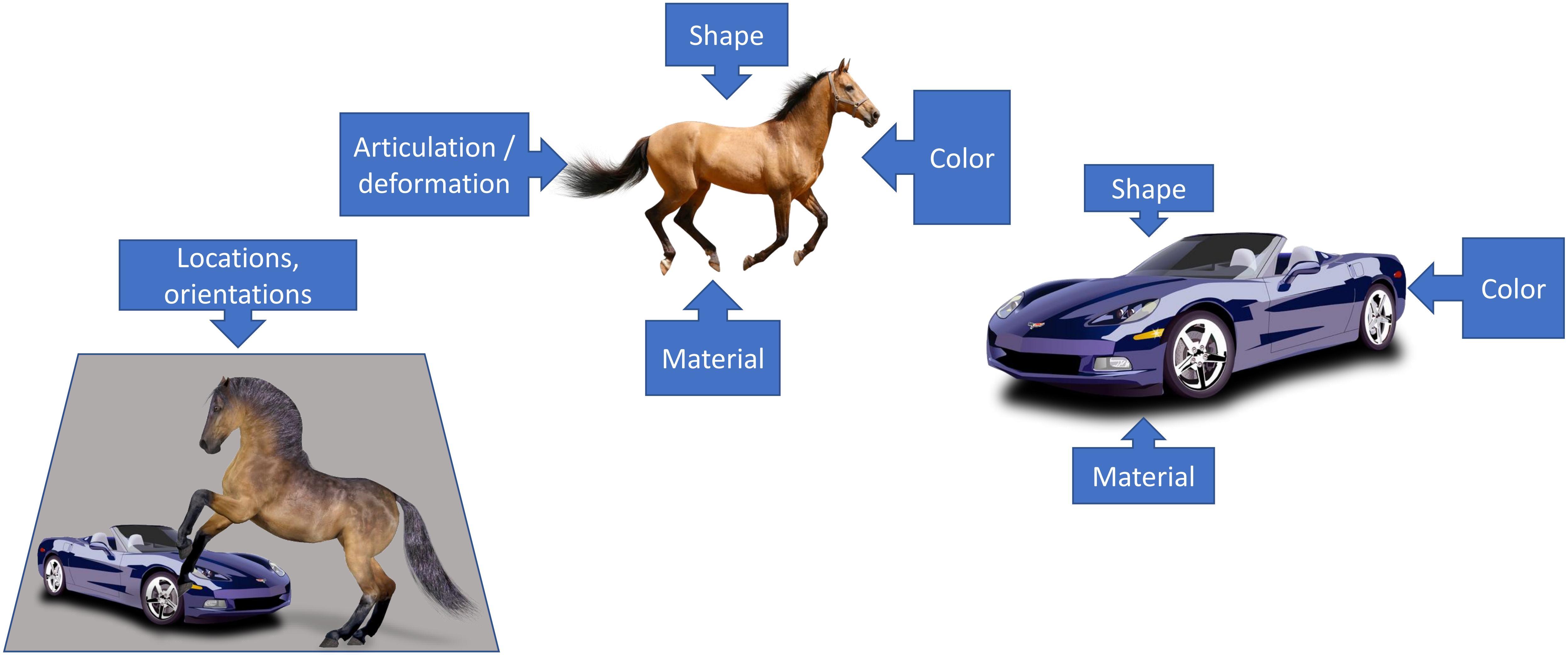
The goal(s) of computer vision

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

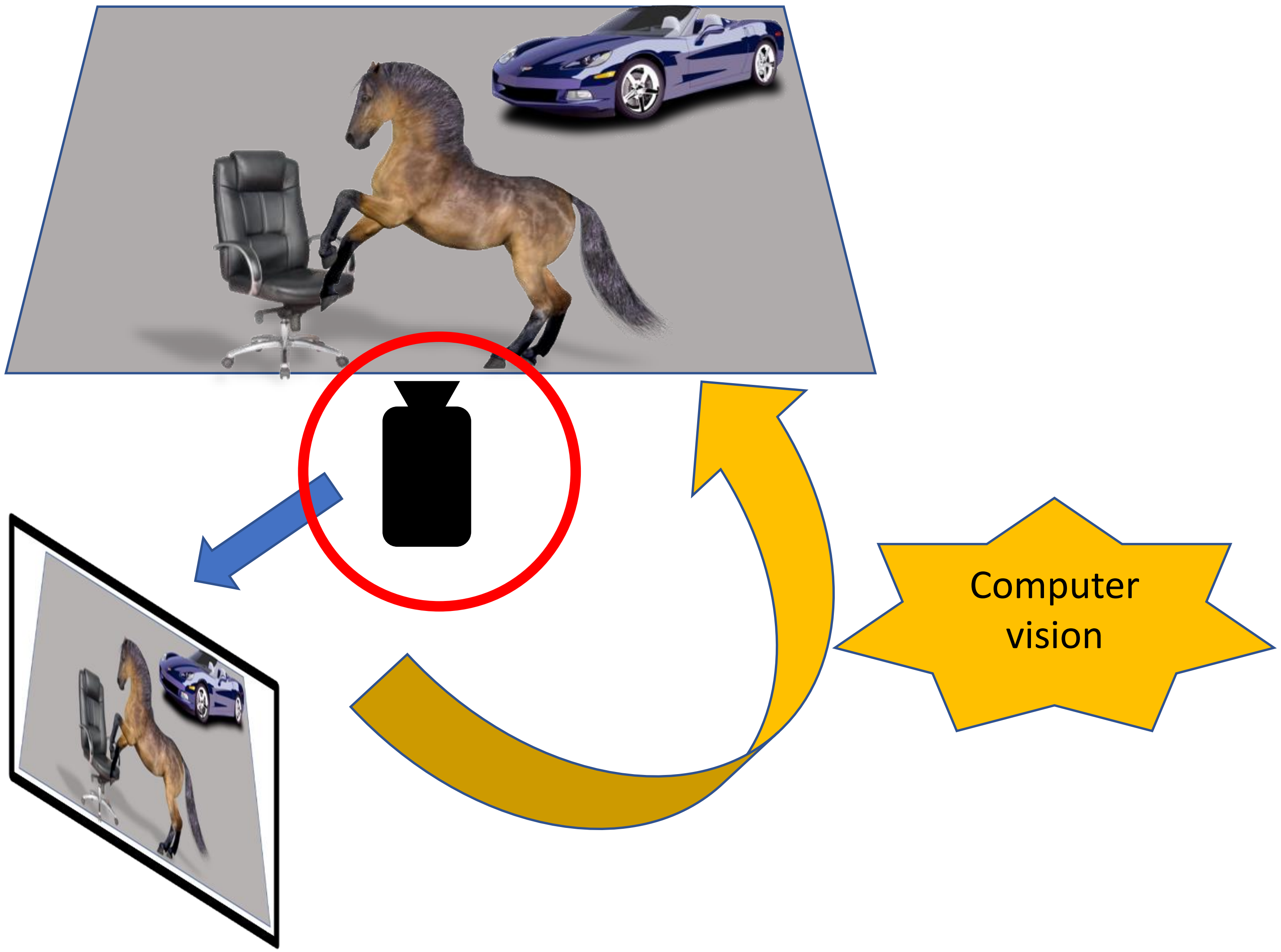
Picture credit: Magritte , Jon Barron

The *physical* world

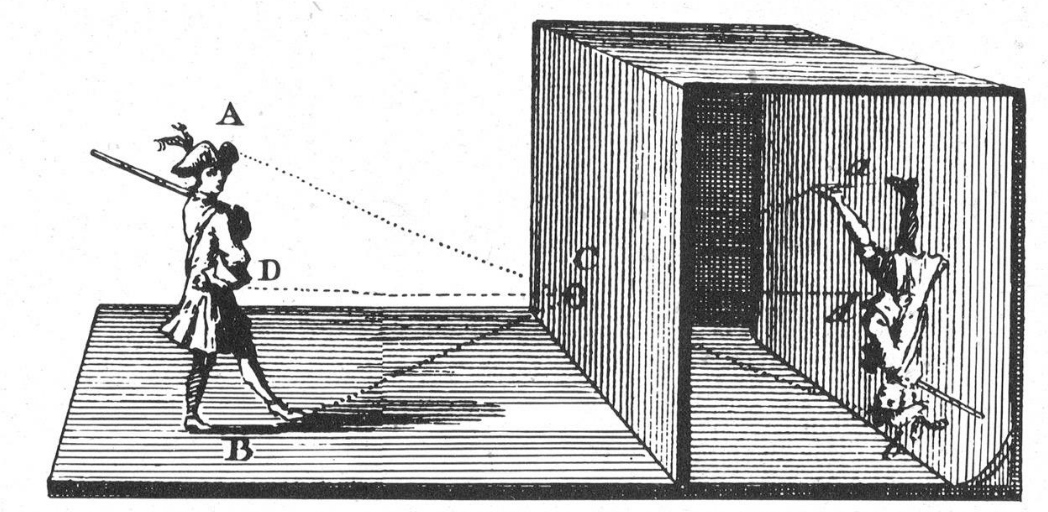
Not pixels!

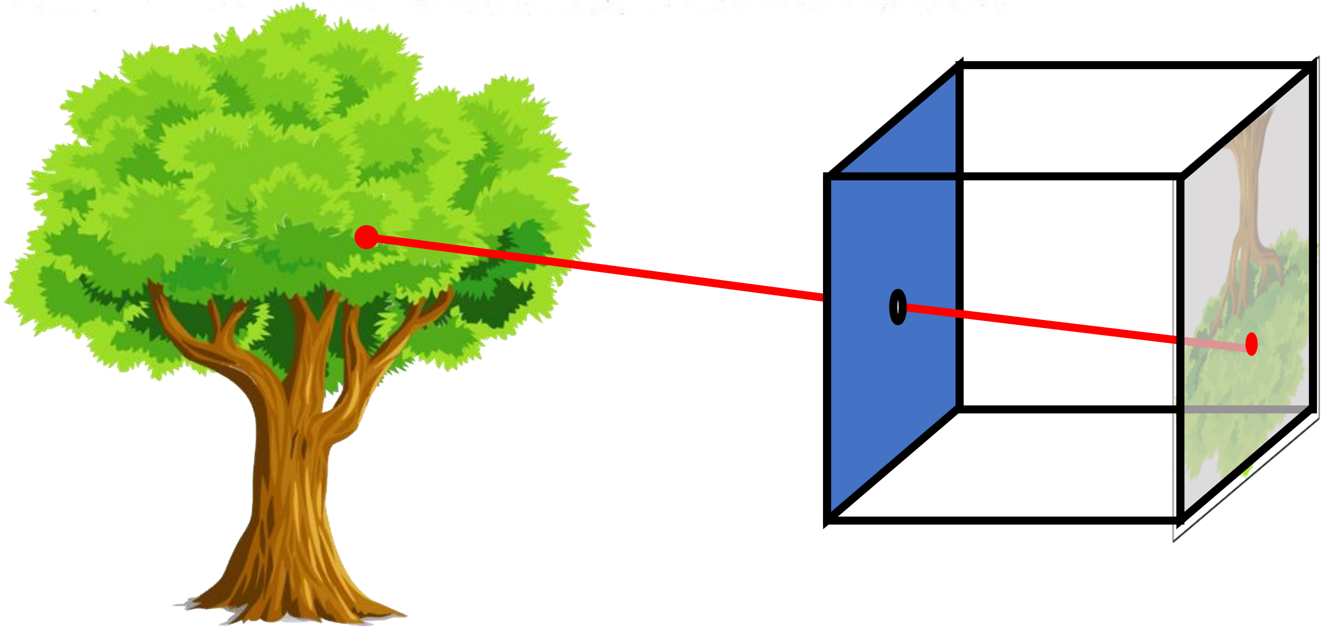


How do we do this?



The pinhole camera - *Camera Obscura*

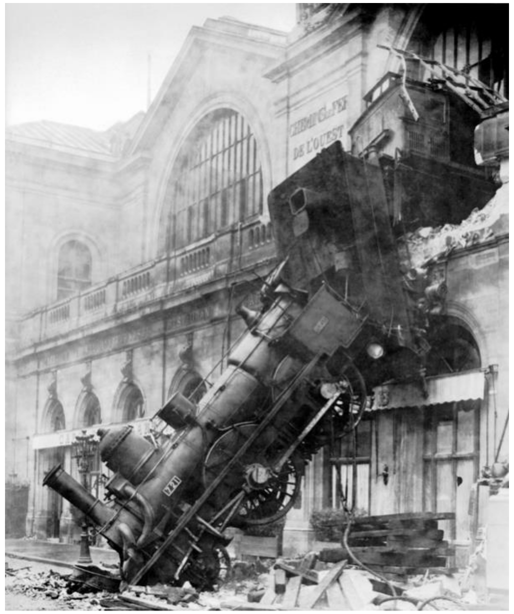


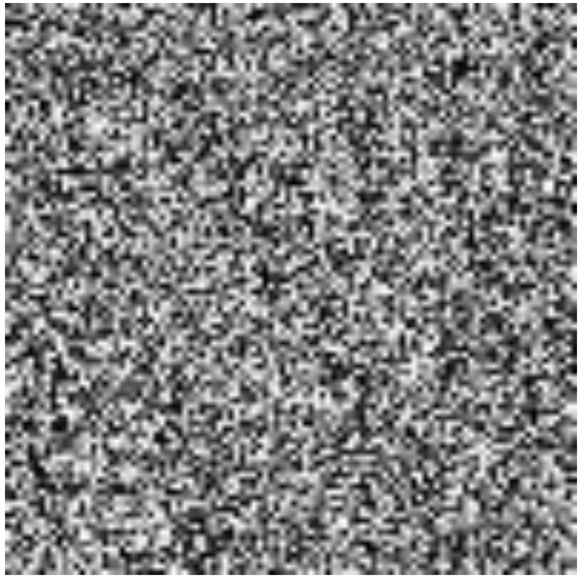


Observation 1: nearby pixels are similar

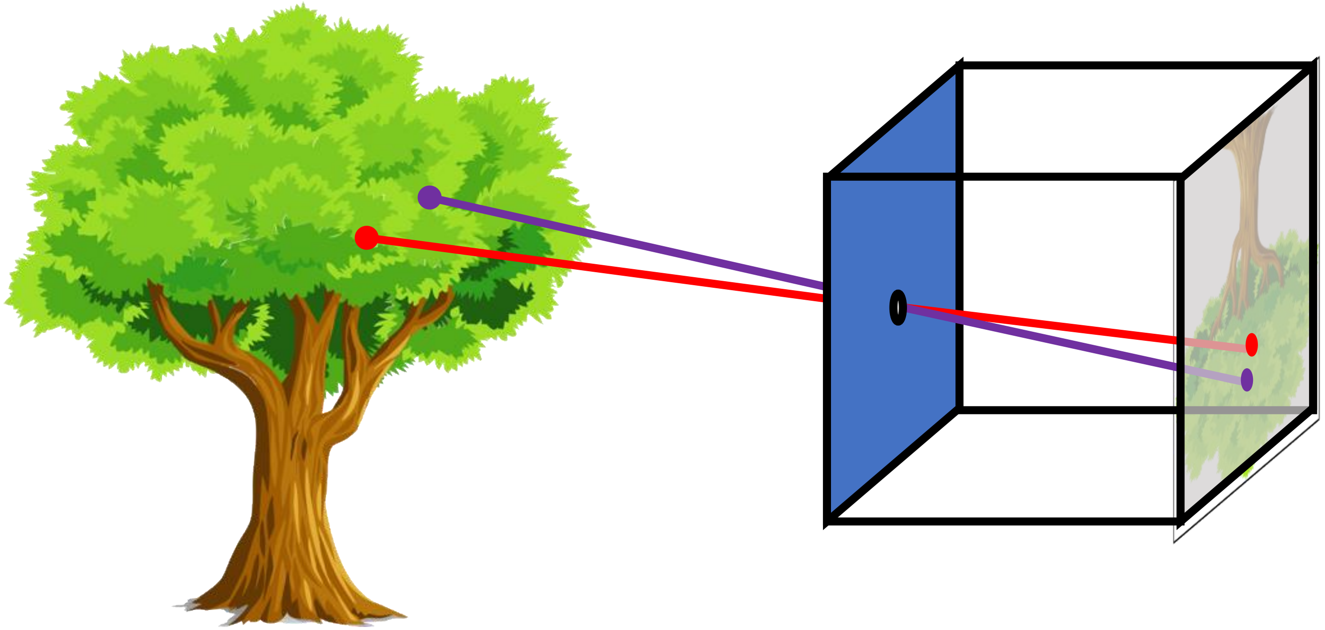
Why?

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



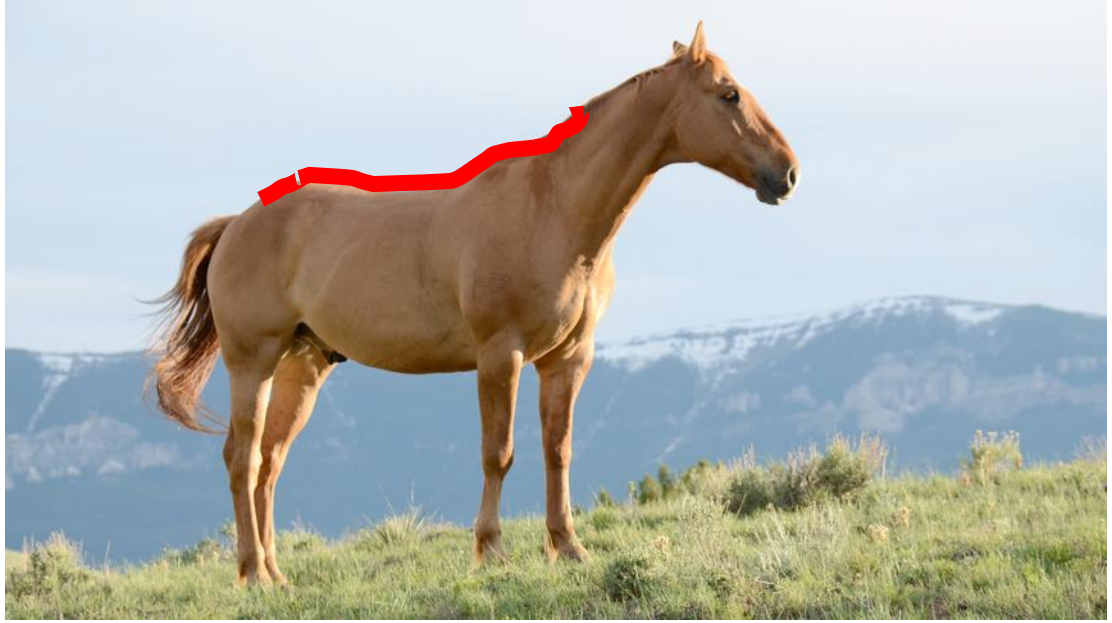


Not all 2D arrays are images



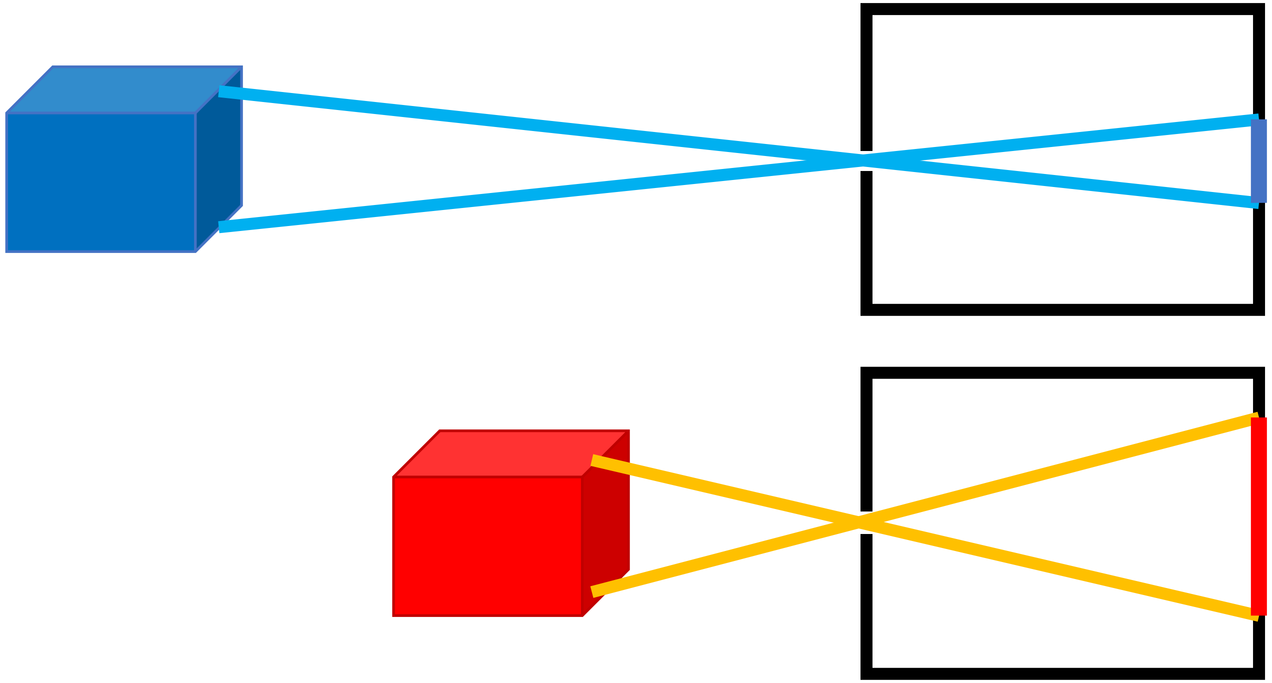
Observation 2: dissimilar pixels are useful cues

- 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)*



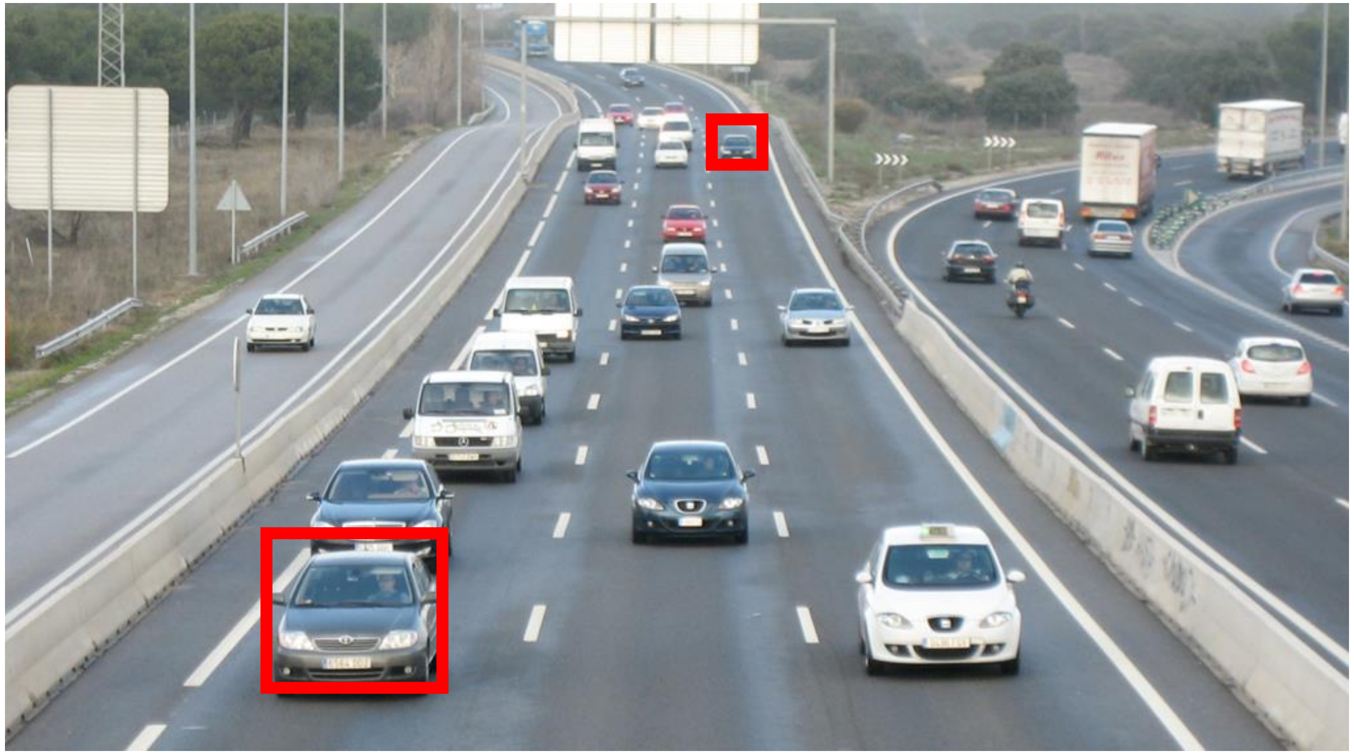
Counterexample: camouflage

Observation 3: Farther away objects appear smaller



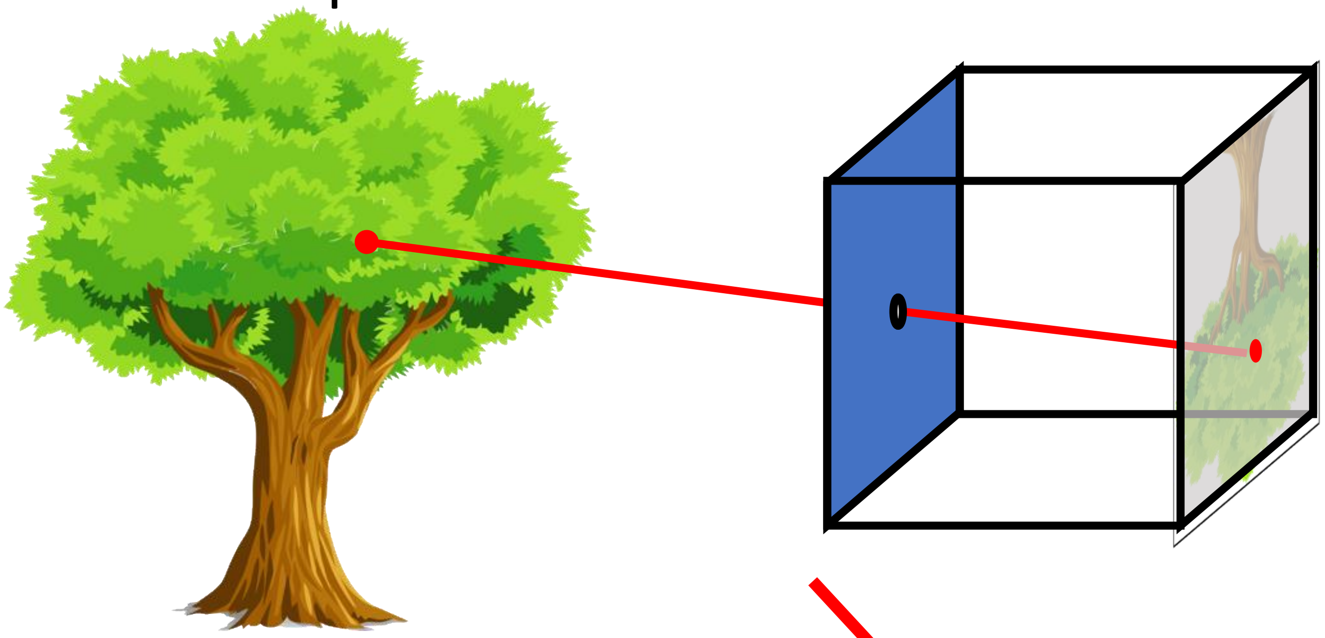
Observation 3: Farther away objects appear smaller

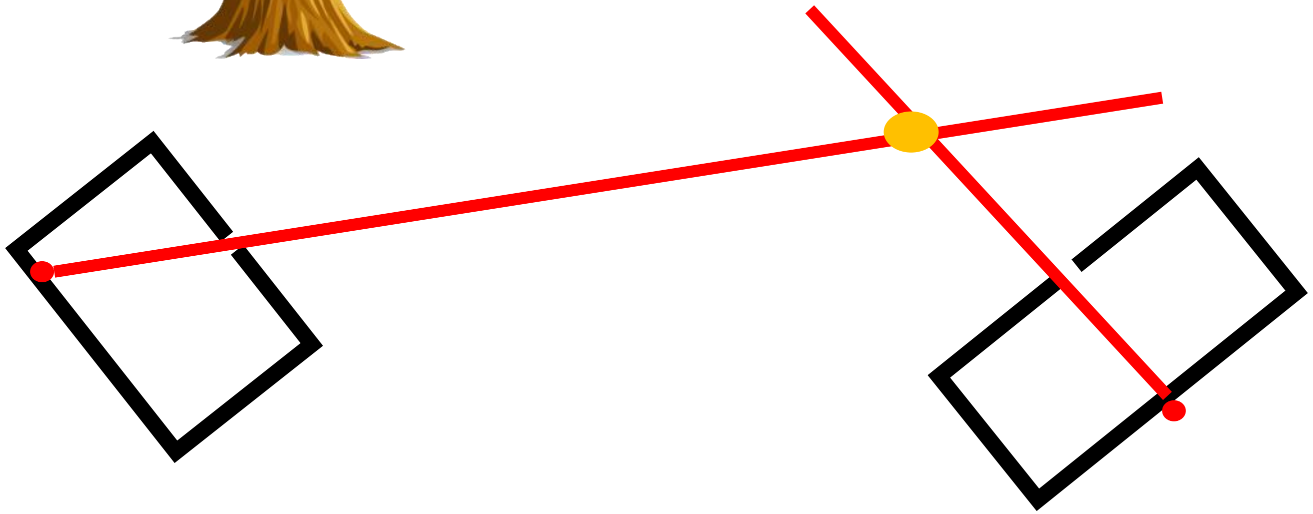
Key modules: search over scales, zoom-out/zoom-in



Observation 4: Image formation is lossy

We lose depth information

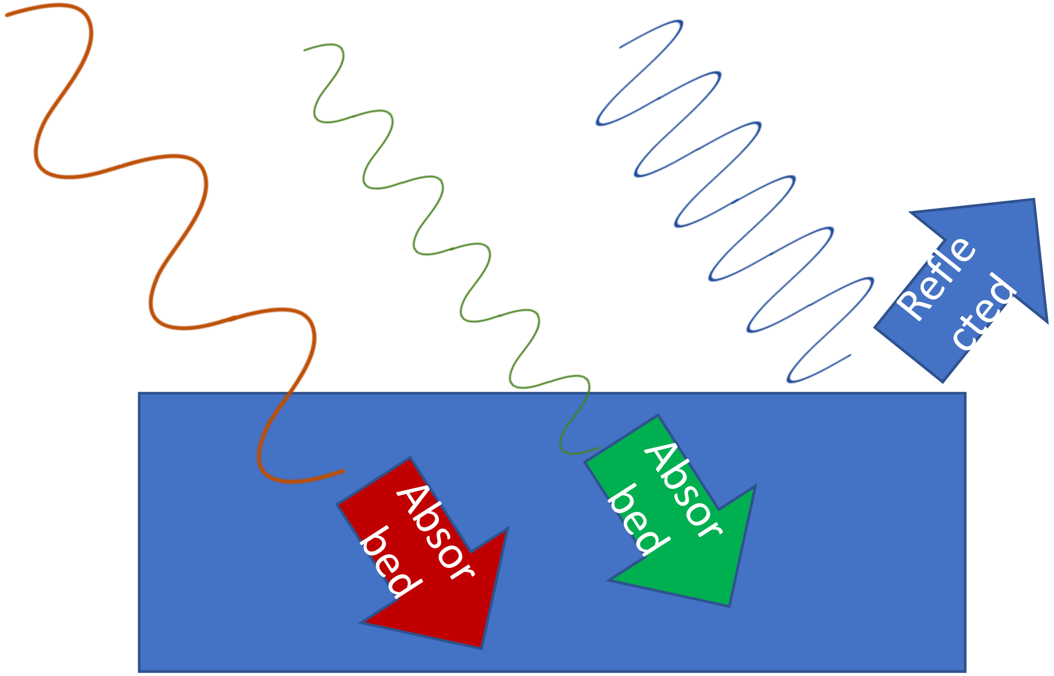


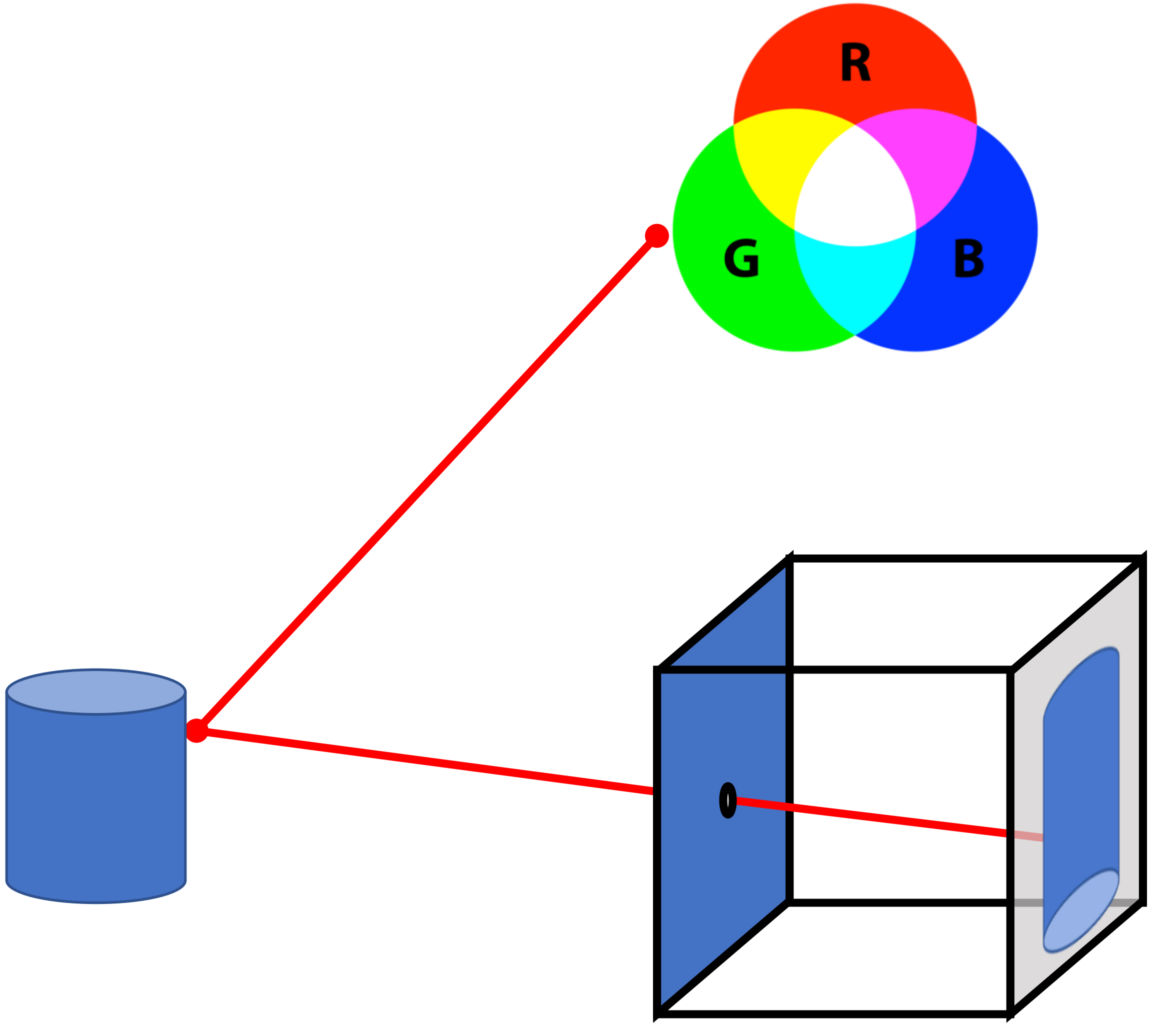


Idea: use multiple images
Need to find which pixel in image 2 matches which in image 1 - the *correspondence* problem

Color and light

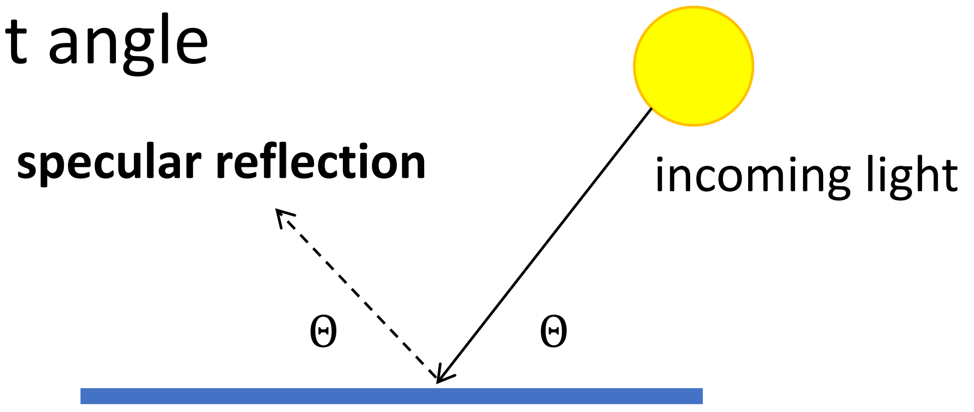
- Each pixel records amount of energy in red light, blue light green light
- But where does light energy come from?



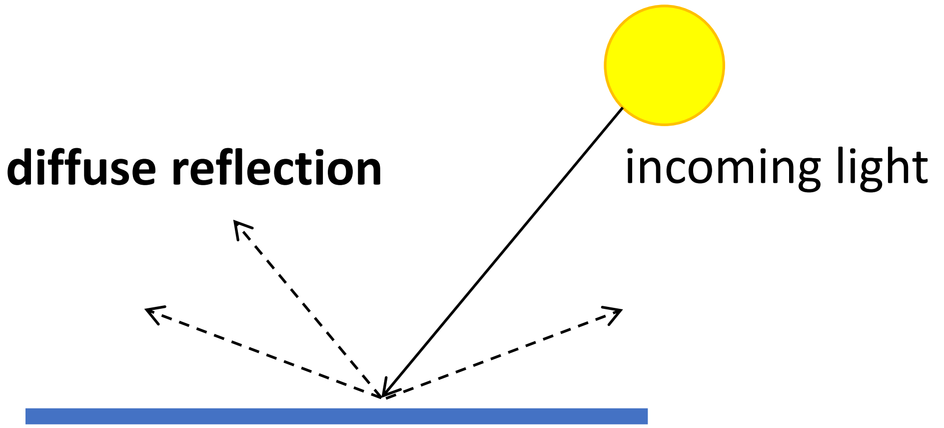


Basic models of reflection

- **Specular**: light bounces off at the incident angle
  - E.g., mirror



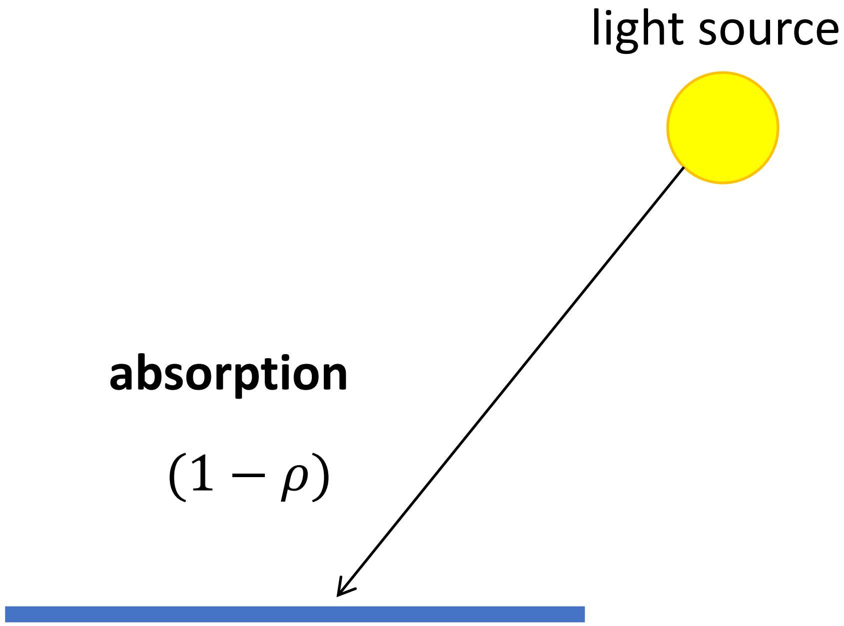
- **Diffuse**: light scatters in all directions
  - E.g., brick, cloth, rough wood

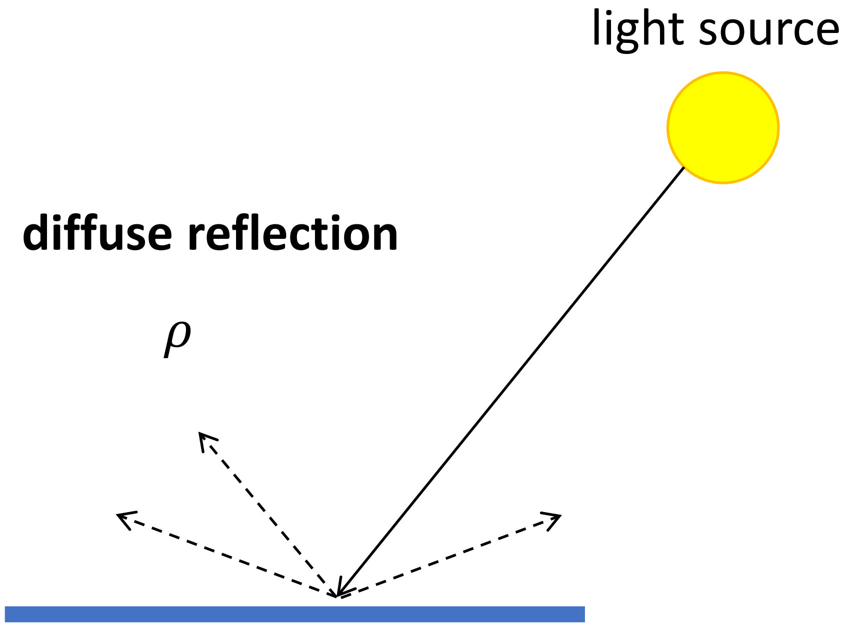


Slide credit: Derek Hoiem

Lambertian reflectance model

- Some light is absorbed (function of albedo ρ)
- Remaining light is scattered (diffuse reflection)
- Examples: soft cloth, concrete, matte paints

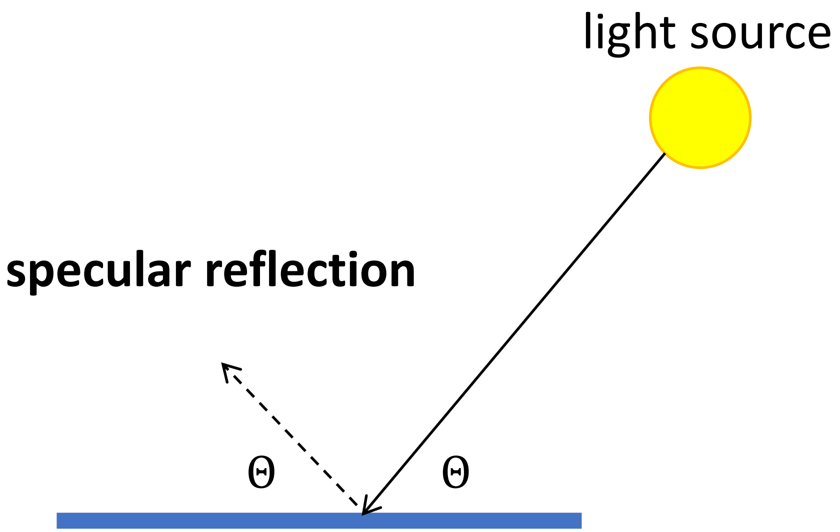




Slide credit: Derek Hoiem

Specular Reflection

- Reflected direction depends on light orientation and surface normal
  - E.g., mirrors are fully specular
  - Most surfaces can be modeled with a mixture of diffuse and specular components



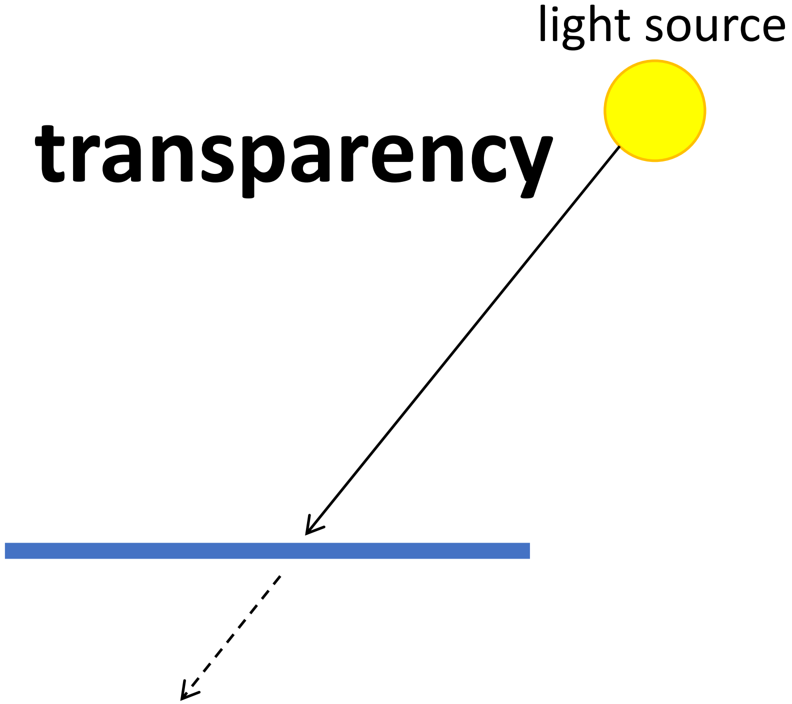
Flickr, by suzysputnik

Flickr, by piratejohnny

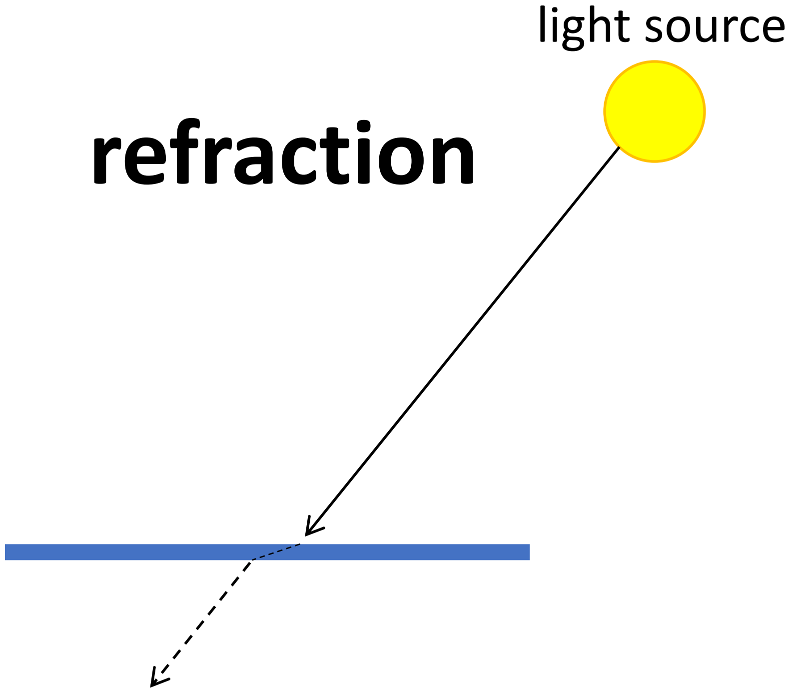
Slide credit: Derek Hoiem

Other possible effects

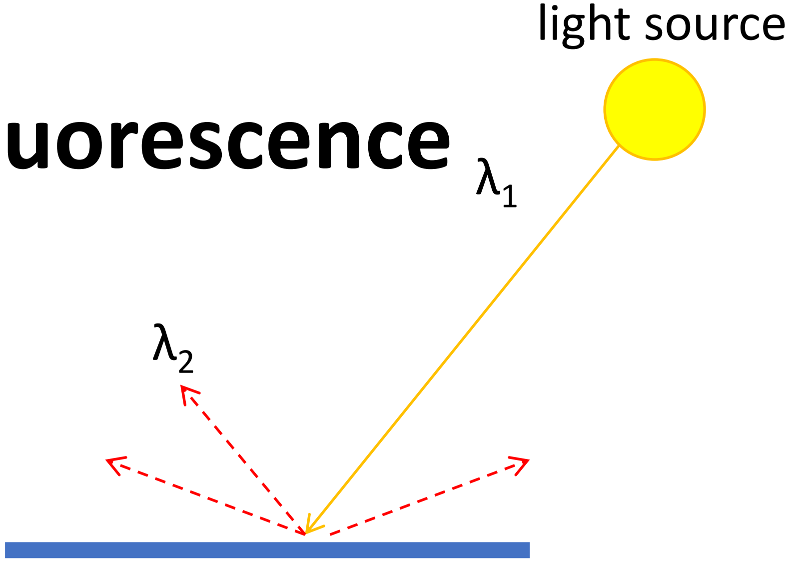
transparency



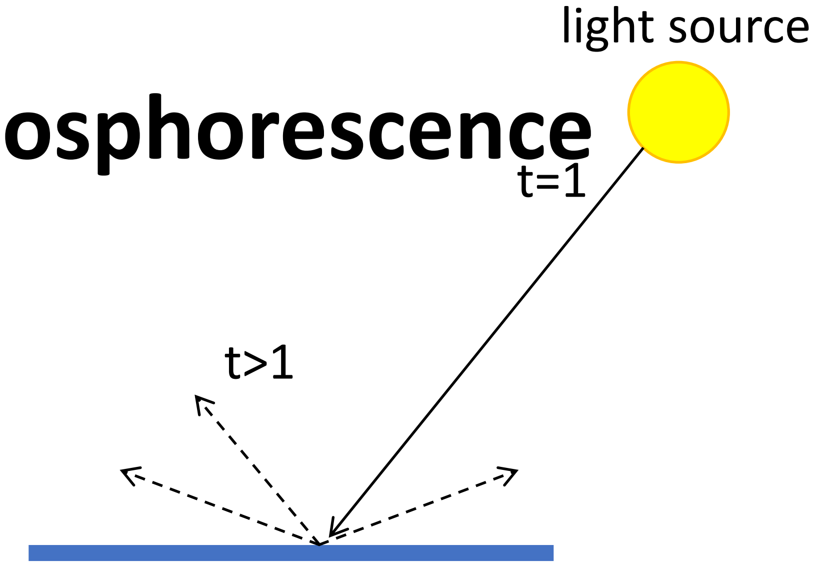
refraction



fluorescence

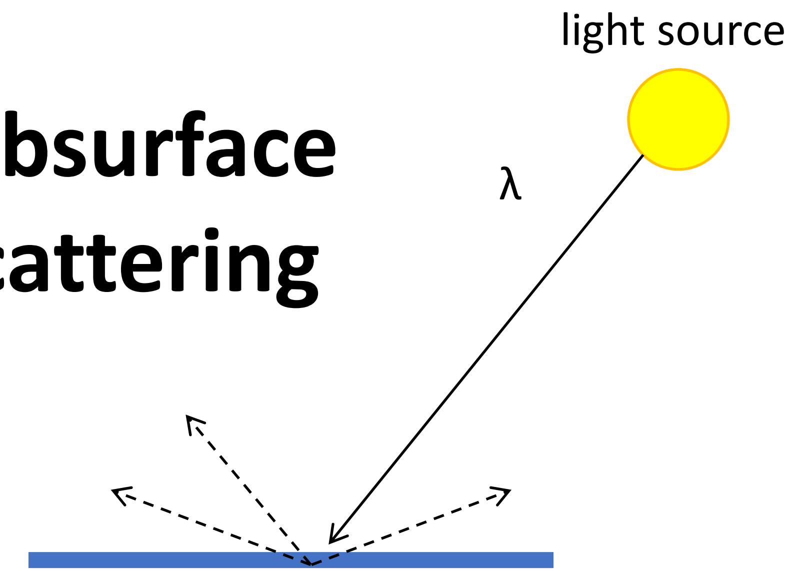


phosphorescence

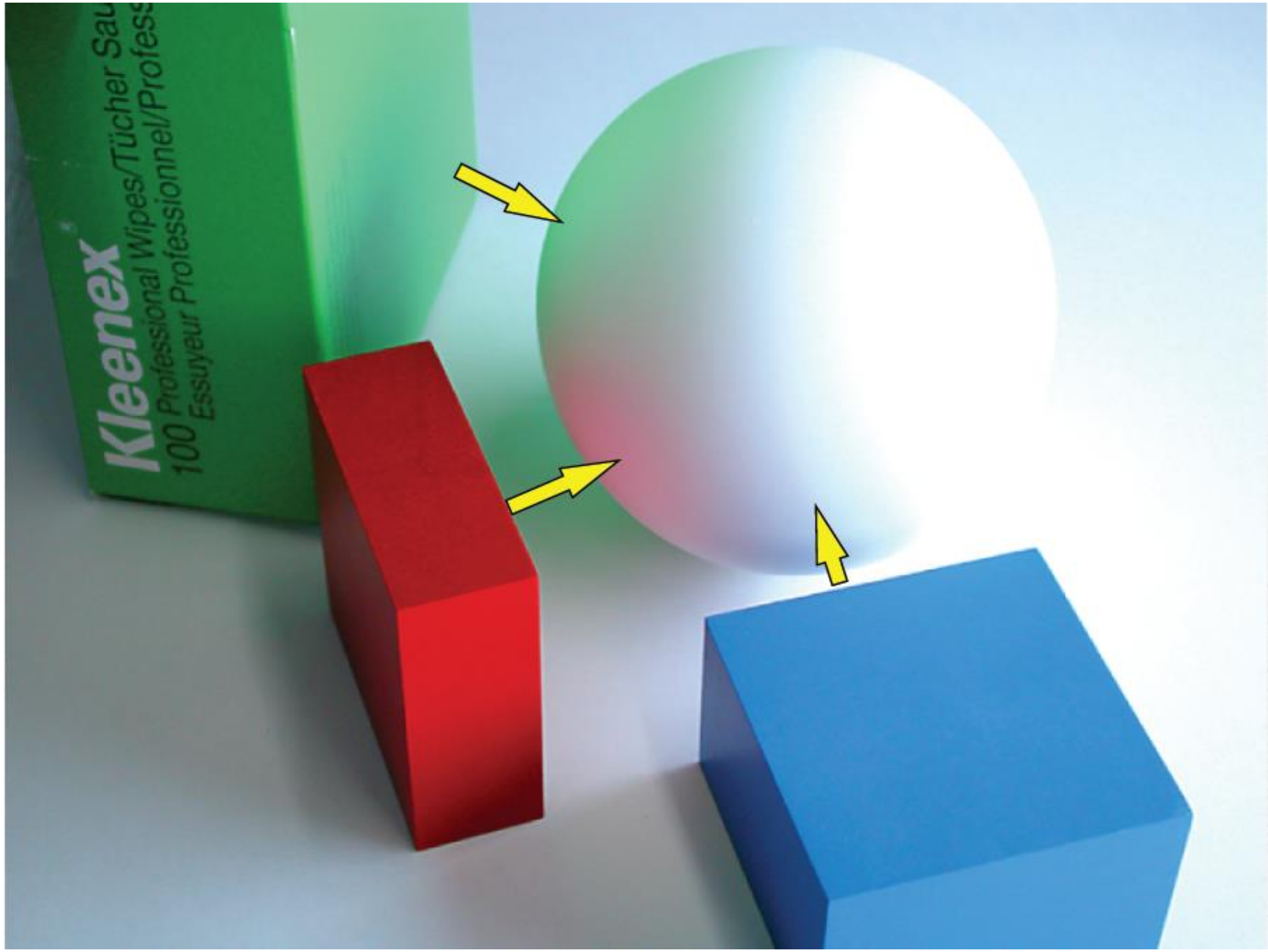


Slide credit: Derek Hoiem

subsurface scattering



Inter-reflection affects the apparent color of objects



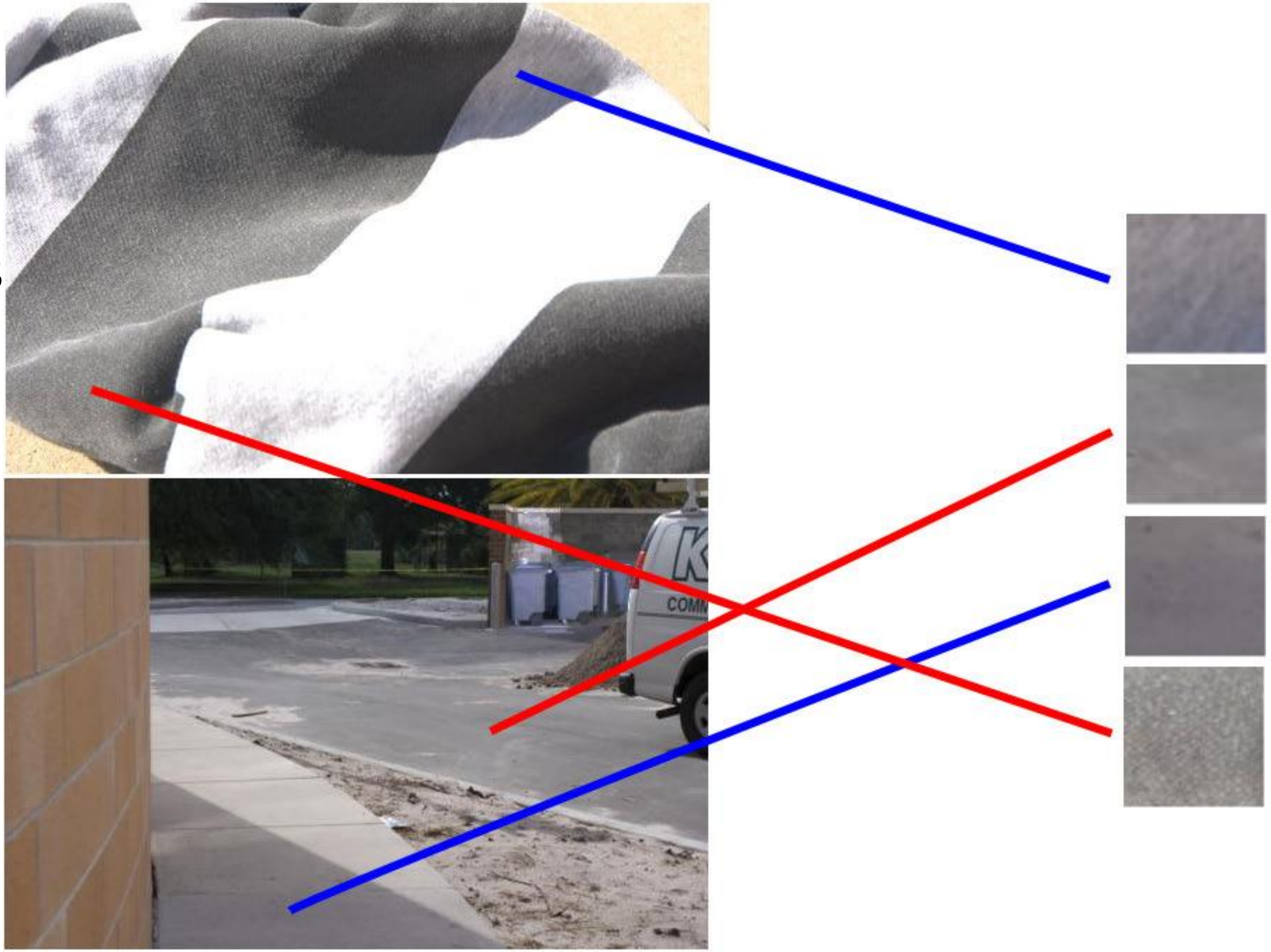
From Koenderink slides on image texture and the flow of light

Color and light

Color of a pixel depends on:

- Color of light
- Surface property
- Direction of light w.r.t surface
- Viewing direction
- Presence/absence of cast shadows

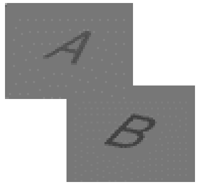
A pixel's brightness tells us nothing by itself

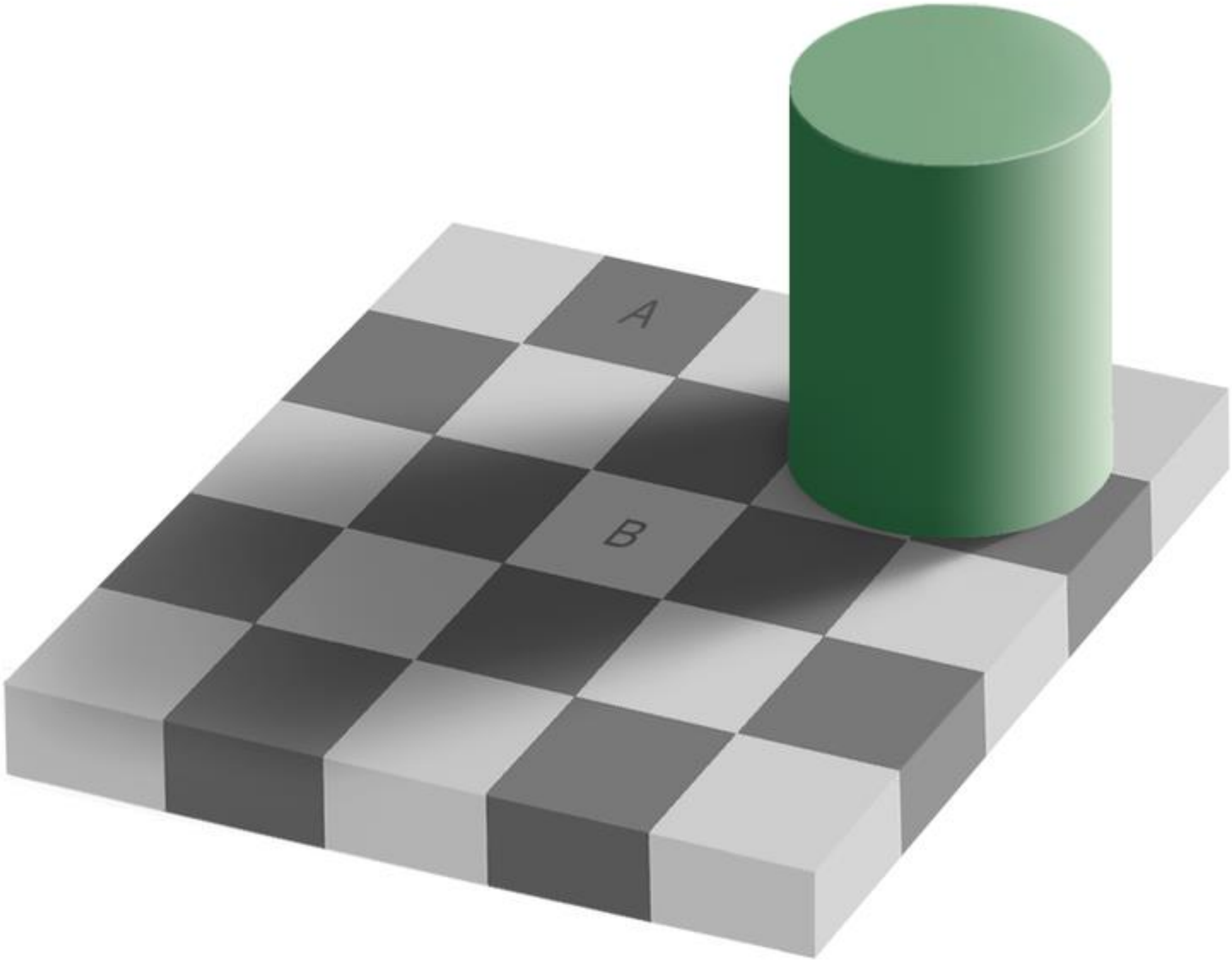


And yet we can interpret images…

- Key idea: for nearby scene points, most factors do not change much
- The information is mainly contained in ***local differences*** of brightness

Observation: Pixel color is complicated

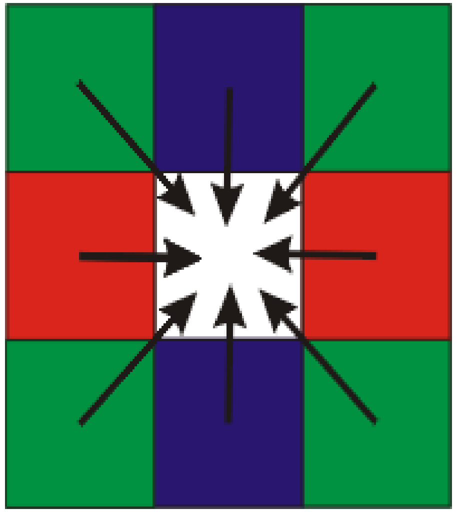


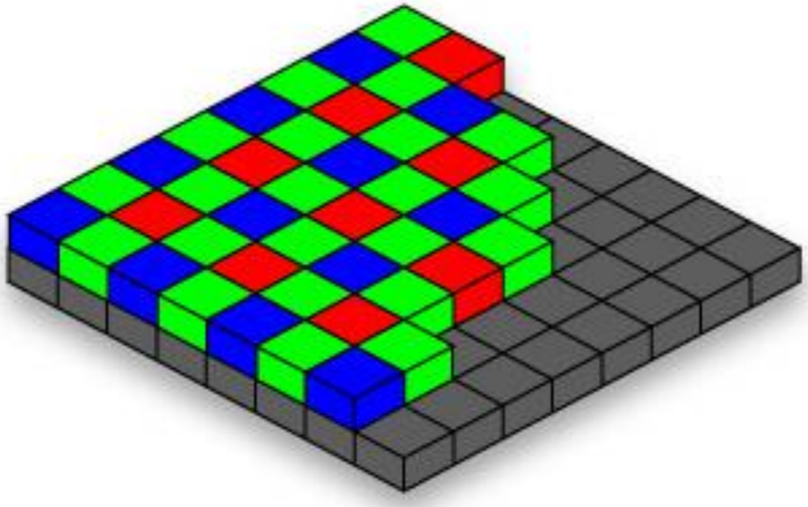


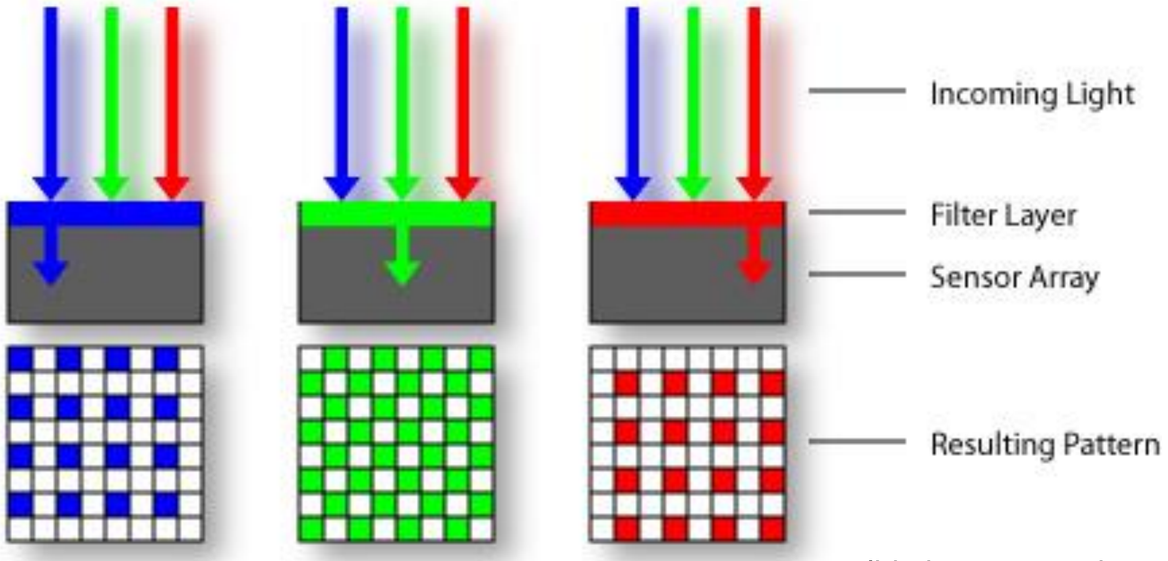
Idea: rely less on absolute color. Look at *changes in color* (may be object boundaries) instead

Adapted from Ted Adelson

Practical Color Sensing: Bayer Grid





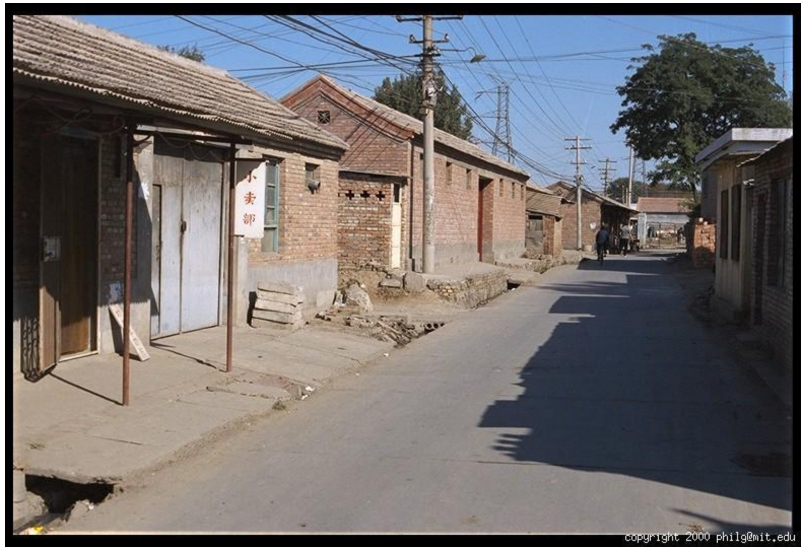


- Estimate RGB at ‘G’ cels from neighboring values

http://www.cooldictionary.com/words/Bayer-filter.wikipedia

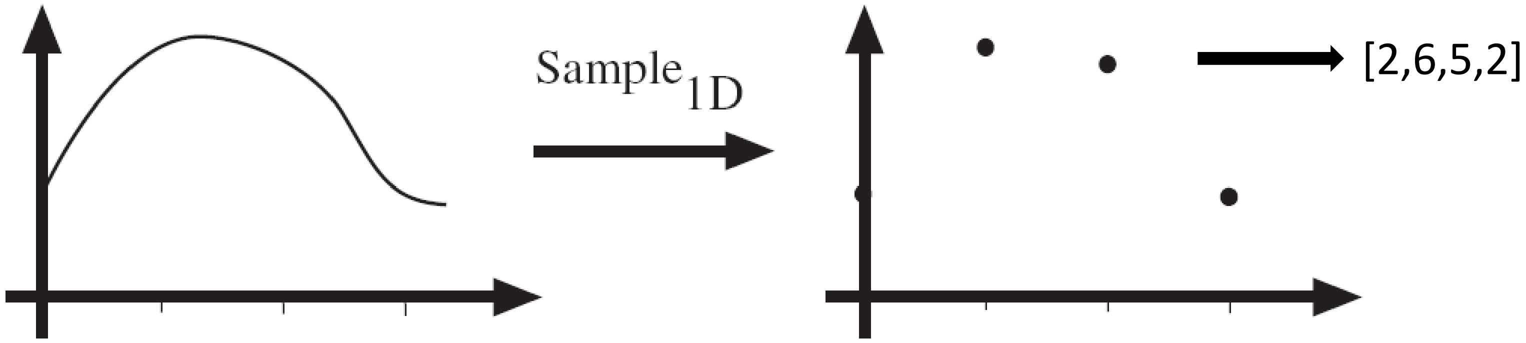
Slide by Steve Seitz

Color Image



What is Image?

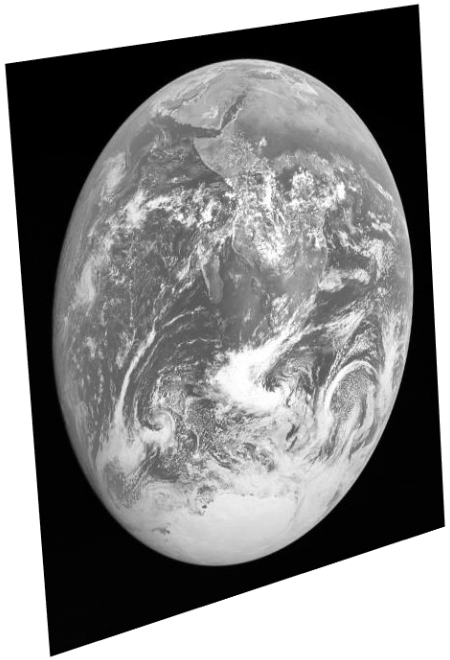
Signal: A (multi-dimensional) function that contains information about a phenomenon.

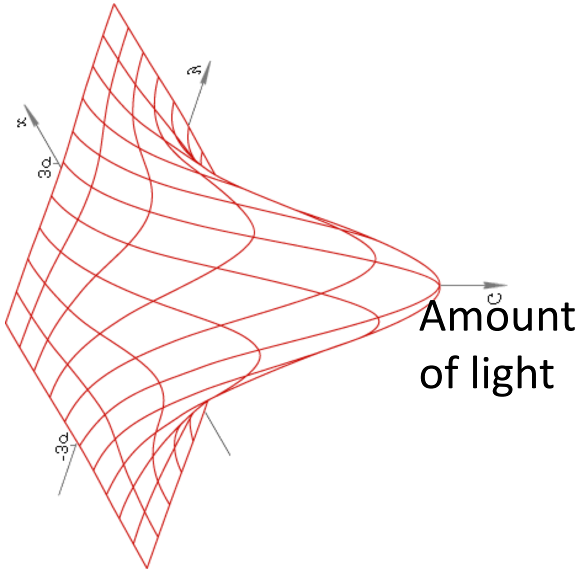


Signals can be

- Continuous: light
- Discrete: measurement of a light
- Sampling: reduction of continuous signal to a discrete signal

Any phenomenon: Light, Heat, Gravity and so on…

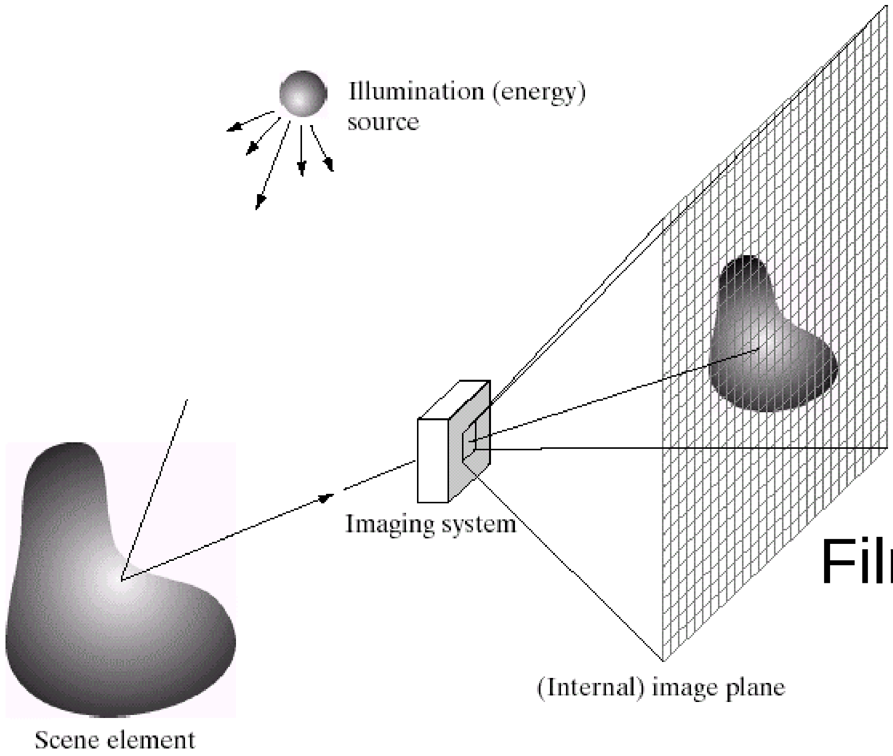




Light received by camera

- A 2D image is a sampling of a 2D signal. Note that the 2D signal can also be a projection (or slice) of a higher-dimensional signal like in MRI or CT scans.
- Image stores intensity or ‘brightness’ along x and y dimensions
- Video: xy-coordinates + time (Time-varying 2D signal)

Image Formation



Film

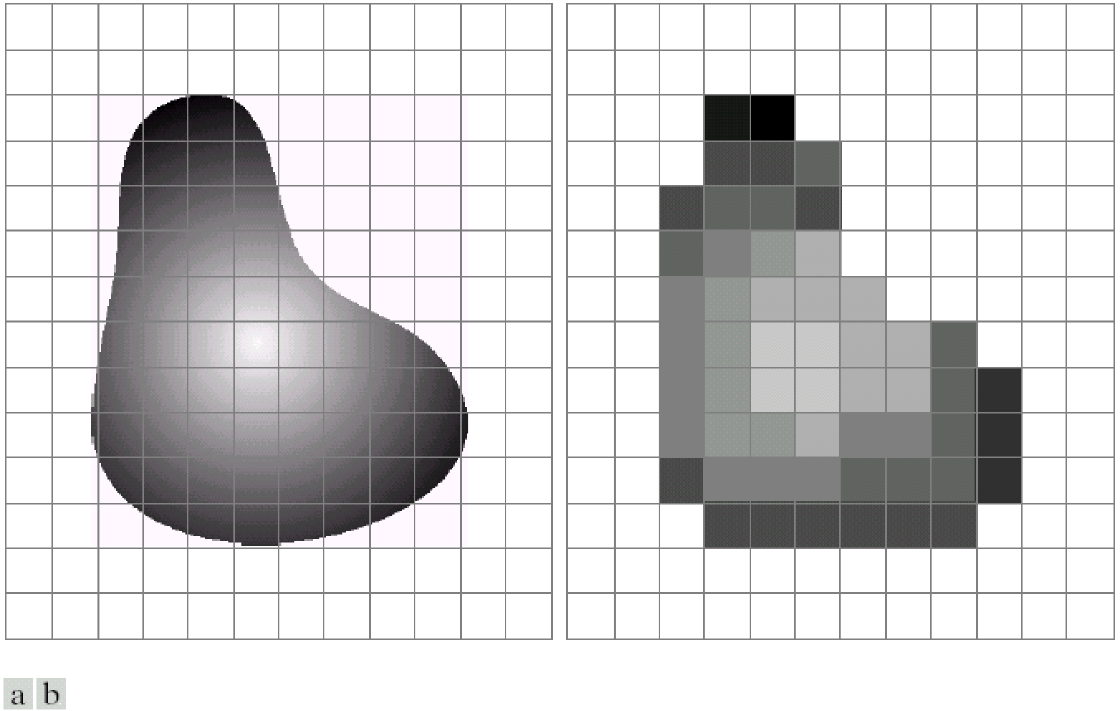
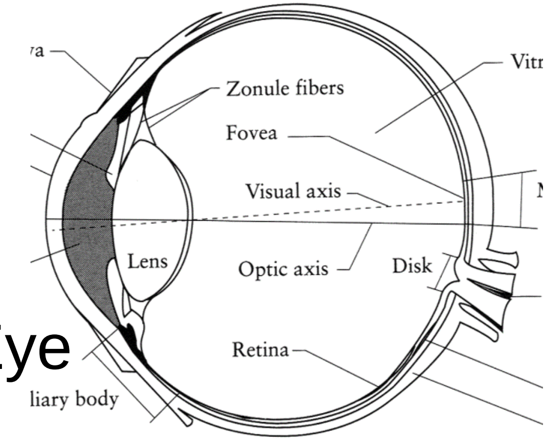
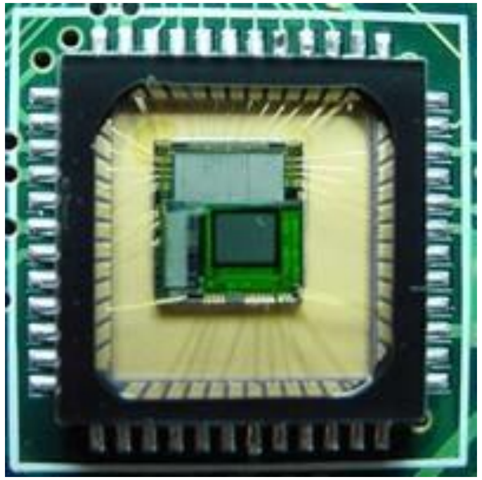


FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.



The Eye

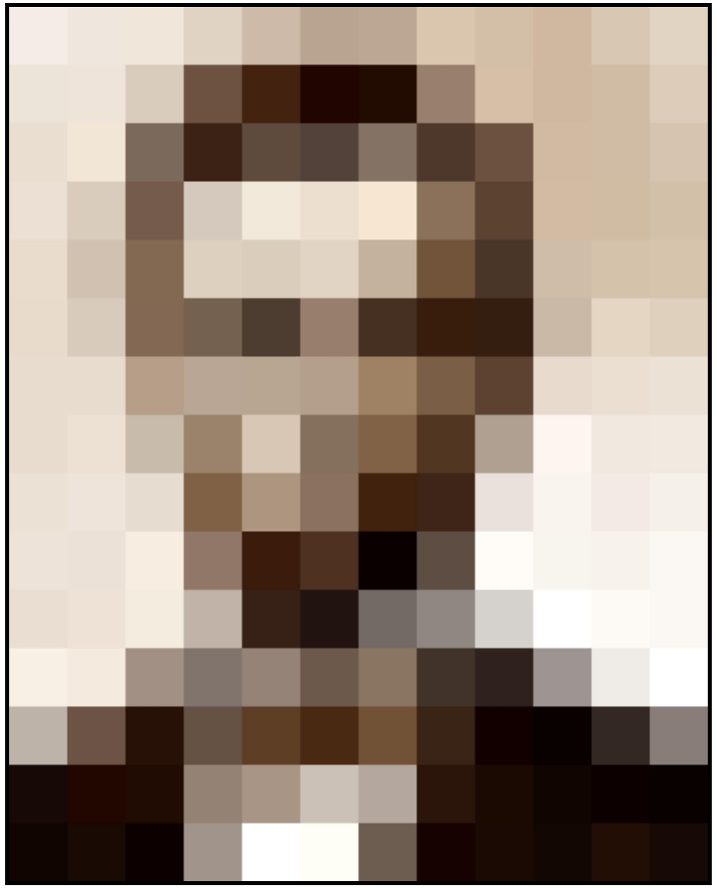


CMOS
sensor

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

What humans see



Slide credit: Larry Zitnick

What computers see

243	239	240	225	206	185	188	218	211	206	216	225
242	239	218	110	67	31	34	152	213	206	208	221
243	242	123	58	94	82	132	77	108	208	208	215
235	217	115	212	243	236	247	139	91	209	208	211
233	208	131	222	219	226	196	114	74	208	213	214
232	217	131	116	77	150	69	56	52	201	228	223
232	232	182	186	184	179	159	123	93	232	235	235
232	236	201	154	216	133	129	81	175	252	241	240
235	238	230	128	172	138	65	63	234	249	241	245
237	236	247	143	59	78	10	94	255	248	247	251
234	237	245	193	55	33	115	144	213	255	253	251
248	245	161	128	149	109	138	65	47	156	239	255
190	107	39	102	94	73	114	58	17	7	51	137
23	32	33	148	168	203	179	43	27	17	12	8
17	26	12	160	255	255	109	22	26	19	35	24

Slide credit: Larry Zitnick

Epitome

- Natural images are *not* arbitrary 2D arrays
- 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
- Solving computer vision requires using these properties