
WÓDKA
(VODKA)
40% VOL.
Cuisinart
VISIONS of JAPAN TOMATSU Shomei
SNOW CRYSTALS
FOTOGRAFIE UND DAS UNSICHTBARE
DER KLEINE STOWASSER
MATHEMATICS
MAN AND SPACE
DRUGS
STACK-ON

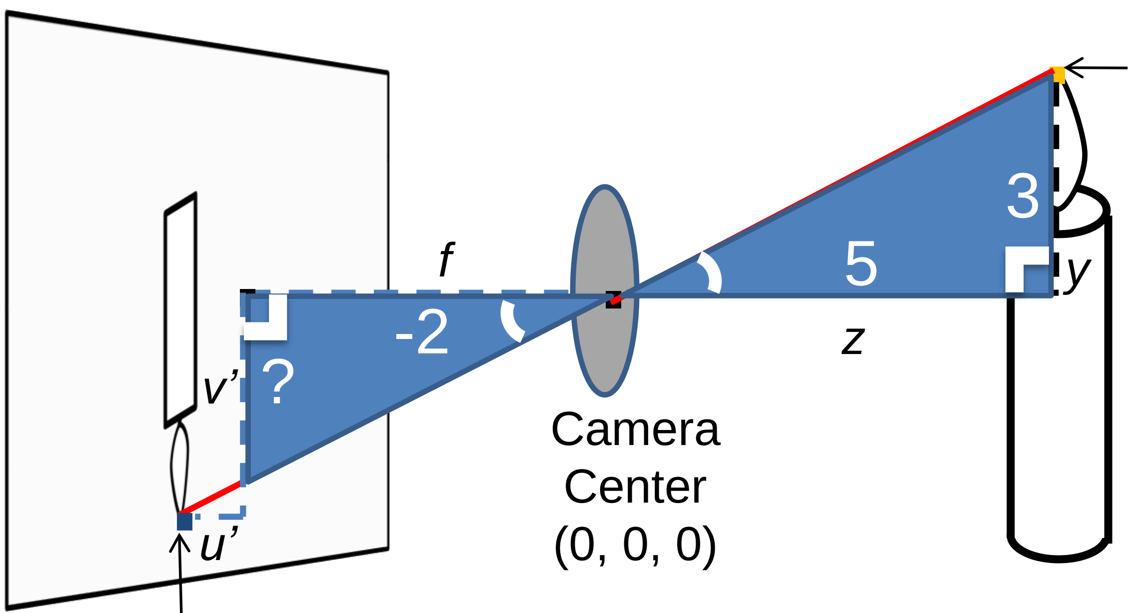

Cuisinart
WÓDKA
SNOW CRYSTALS
MATHEMATICS
MAN AND SPACE
DRUGS
STACK-ON

Light

Computer Vision

James Hays

Projection: world coordinates→image coordinates



$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$x = \begin{bmatrix} u' \\ v' \end{bmatrix}$$

If X = 2, Y = 3,
Z = 5, and f = 2
What are U and V?

$$\frac{v'}{-f} = \frac{y}{z}$$

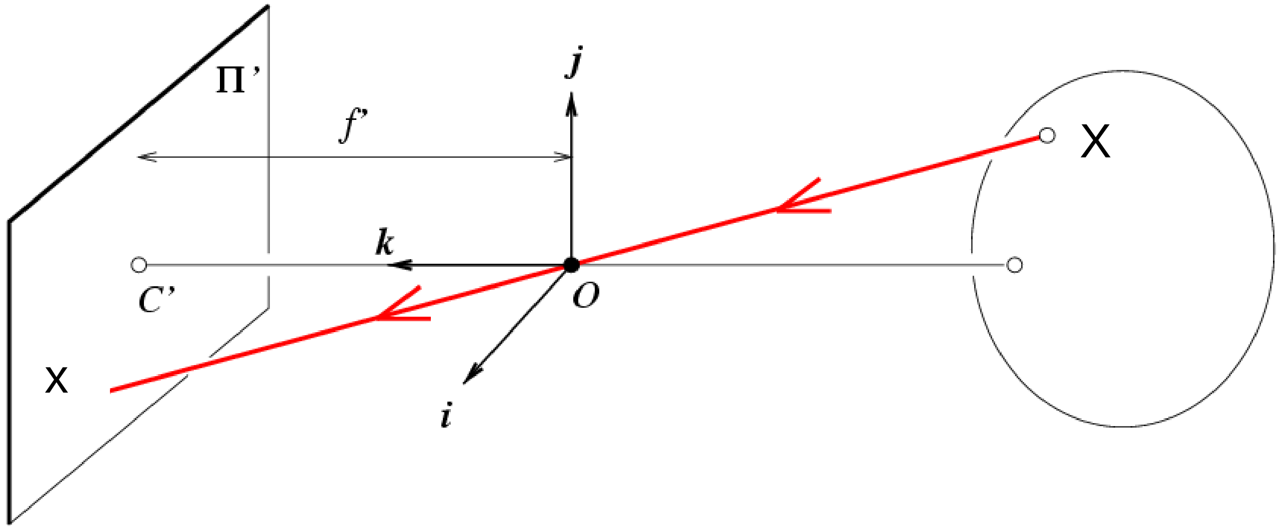
$$u' = -x * \frac{f}{z} \qquad u' = -2 * \frac{2}{5}$$

$$v' = -y * \frac{f}{z} \qquad v' = -3 * \frac{2}{5}$$

Interlude: why does this matter?

Relating multiple views

Projection matrix



Intrinsic Assumptions

- Unit aspect ratio
- Optical center at (0,0)
- No skew

Extrinsic Assumptions

- No rotation
- Camera at (0,0,0)

K

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{X} \Rightarrow w\begin{bmatrix}u \\ v \\ 1\end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix}x \\ y \\ z \\ 1\end{bmatrix}$$

Slide Credit: Savarese

Remove assumption: optical center at origin

Intrinsic Assumptions

- Unit aspect ratio
- No skew

Extrinsic Assumptions

- No rotation
- Camera at (0,0,0)

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{X} \Rightarrow w\begin{bmatrix}u \\ v \\ 1\end{bmatrix} = \begin{bmatrix}f & 0 & u_0 & 0 \\ 0 & f & v_0 & 0 \\ 0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}x \\ y \\ z \\ 1\end{bmatrix}$$

Remove assumption: square pixels

Intrinsic Assumptions
- No skew

Extrinsic Assumptions
- No rotation
- Camera at (0,0,0)

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{X} \Rightarrow w\begin{bmatrix}u\\v\\1\end{bmatrix} = \begin{bmatrix}\alpha & 0 & u_0 & 0\\0 & \beta & v_0 & 0\\0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}x\\y\\z\\1\end{bmatrix}$$

Remove assumption: non-skewed pixels

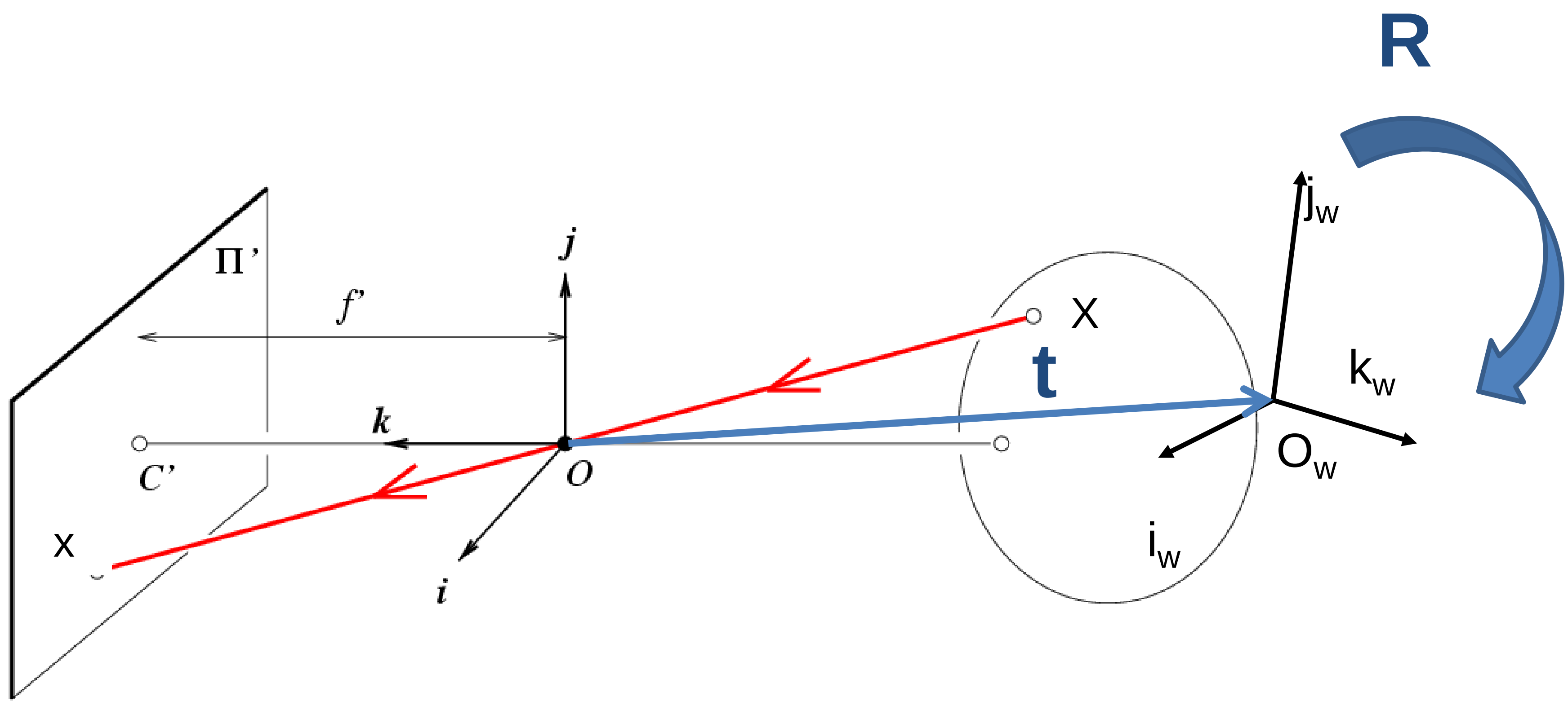
Intrinsic Assumptions

Extrinsic Assumptions
- No rotation
- Camera at (0,0,0)

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{X} \quad \Rightarrow \quad w\begin{bmatrix}u\\v\\1\end{bmatrix} = \begin{bmatrix}\alpha & s & u_0 & 0\\ 0 & \beta & v_0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}x\\y\\z\\1\end{bmatrix}$$

Note: different books use different notation for parameters

Oriented and Translated Camera



Allow camera translation

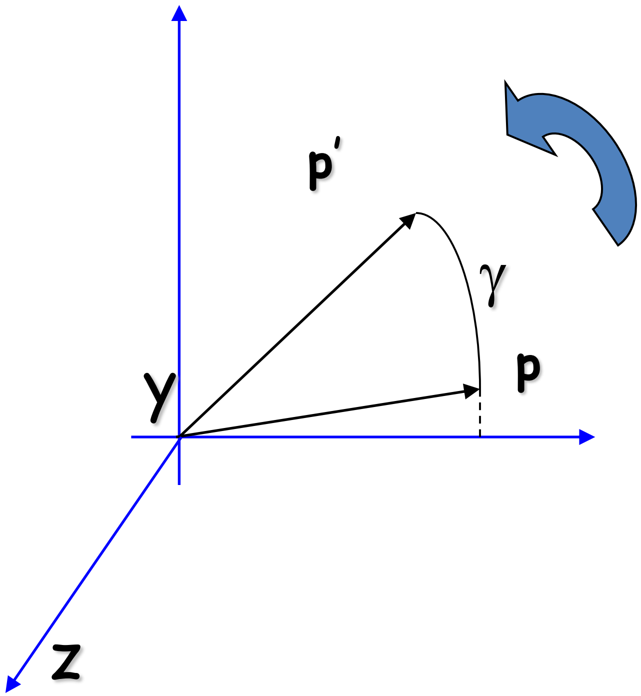
Intrinsic Assumptions

Extrinsic Assumptions
- No rotation

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{I} & \mathbf{t}\end{bmatrix}\mathbf{X} \Rightarrow w\begin{bmatrix}u\\v\\1\end{bmatrix} = \begin{bmatrix}\alpha & 0 & u_0\\0 & \beta & v_0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0 & t_x\\0 & 1 & 0 & t_y\\0 & 0 & 1 & t_z\end{bmatrix}\begin{bmatrix}x\\y\\z\\1\end{bmatrix}$$

3D Rotation of Points

Rotation around the coordinate axes, counter-clockwise:



$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}$$

$$R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}$$

$$R_z(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Allow camera rotation

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

$$w\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Degrees of freedom

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

⇓

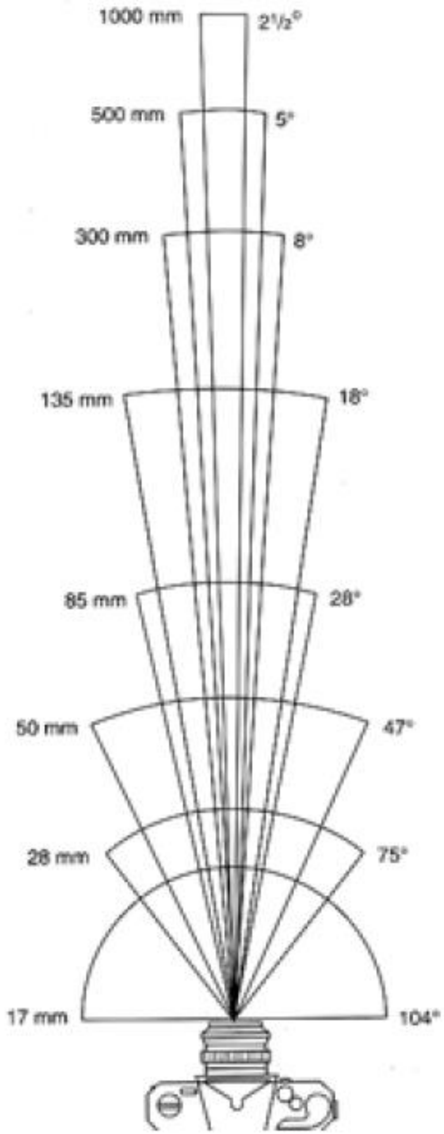
5 6

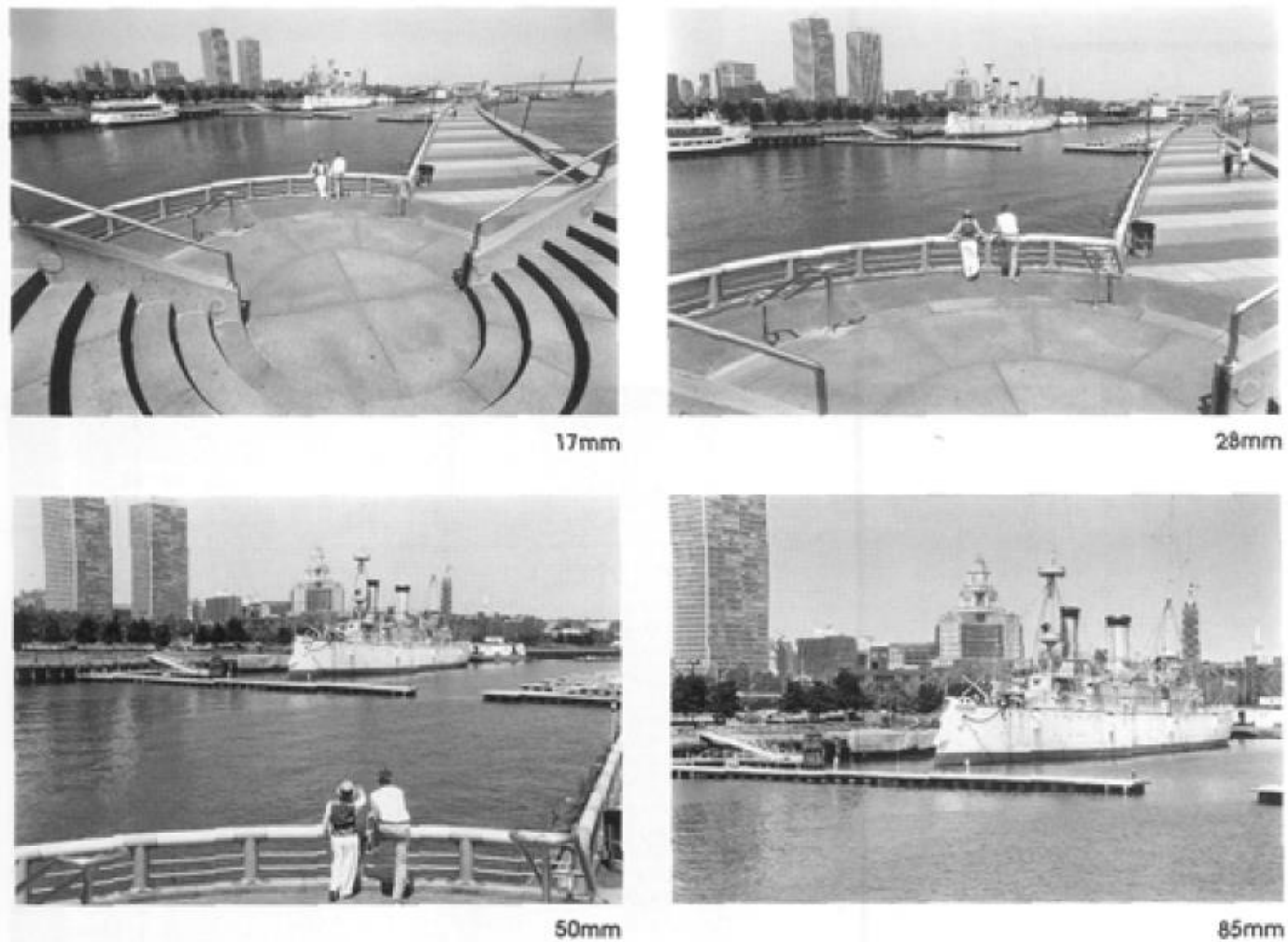
$$w\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Reminder: read your book

- Lectures have assigned readings
- Szeliski 2.1 and especially 2.1.5 cover the geometry of image formation

Field of View (Zoom, focal length)

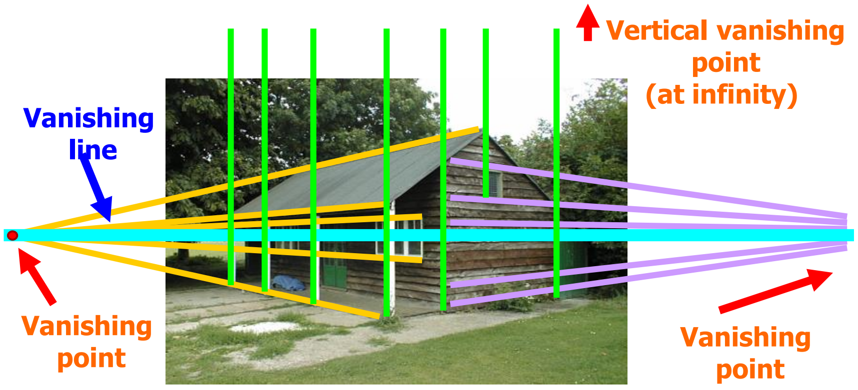




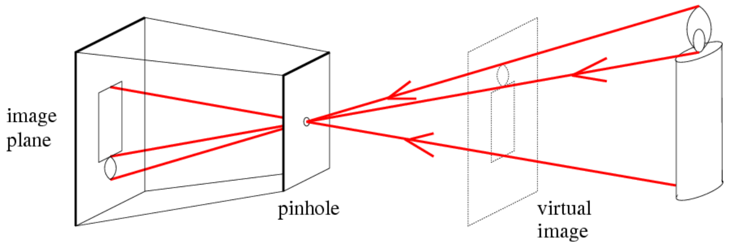
From London and Upton

Things to remember

- Vanishing points and vanishing lines



- Pinhole camera model and camera projection matrix

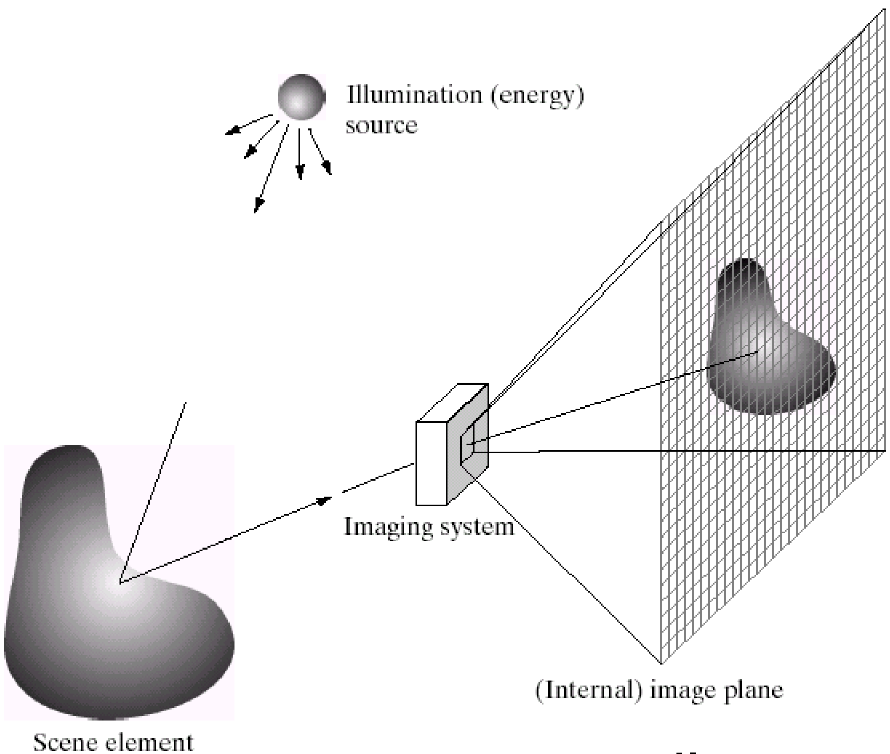


$$\mathbf{x} = \mathbf{K}[\mathbf{R} \quad \mathbf{t}]\mathbf{X}$$

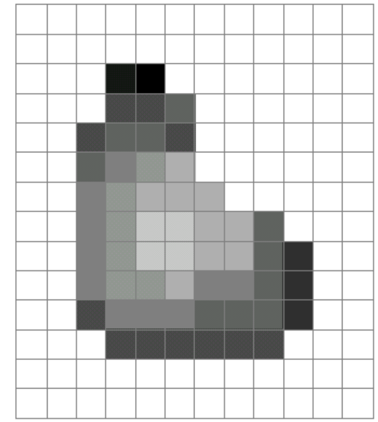
- Homogeneous coordinates

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

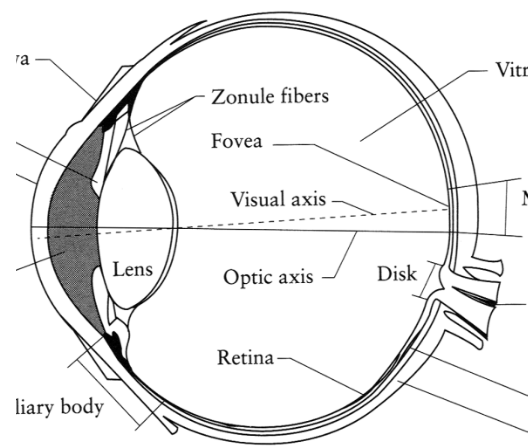
Image Formation



Film



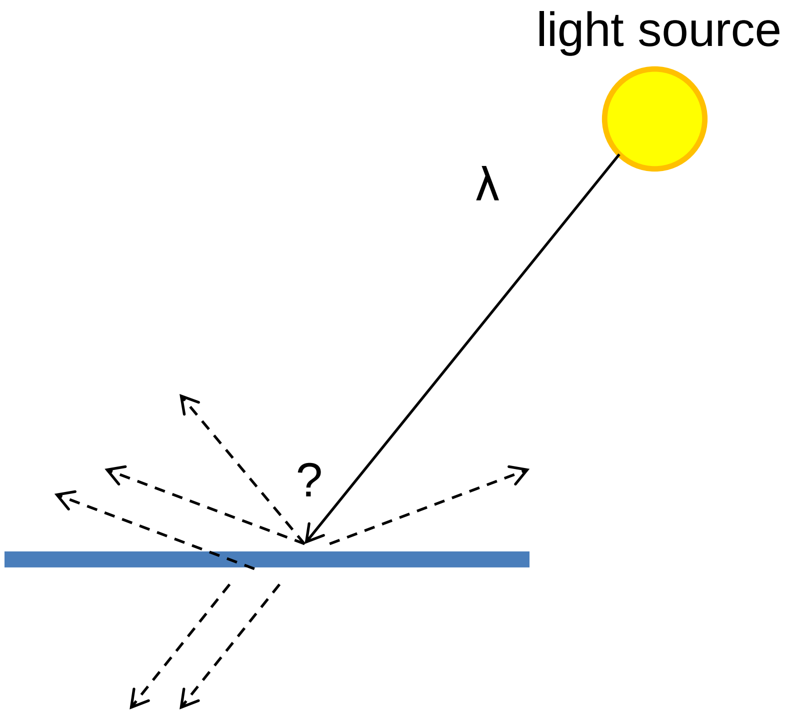
Digital Camera



The Eye

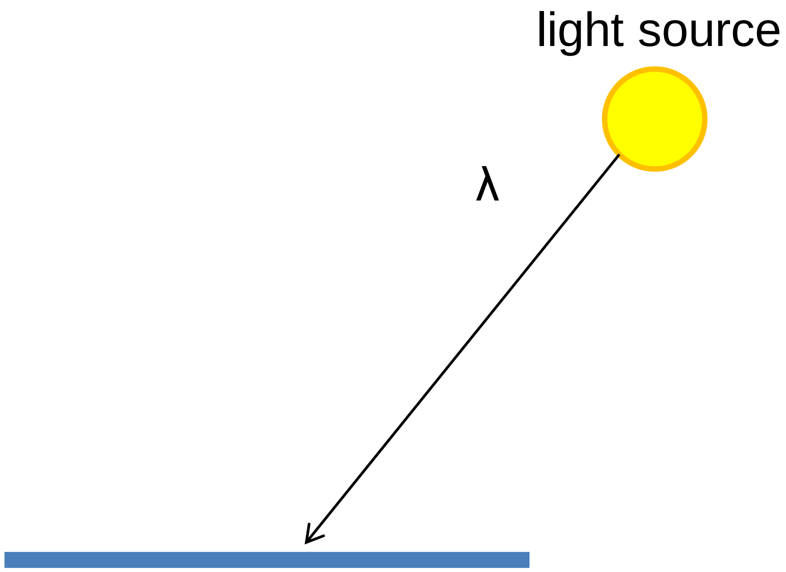
A photon's life choices

- Absorption
- Diffusion
- Reflection
- Transparency
- Refraction
- Fluorescence
- Subsurface scattering
- Phosphorescence
- Interreflection



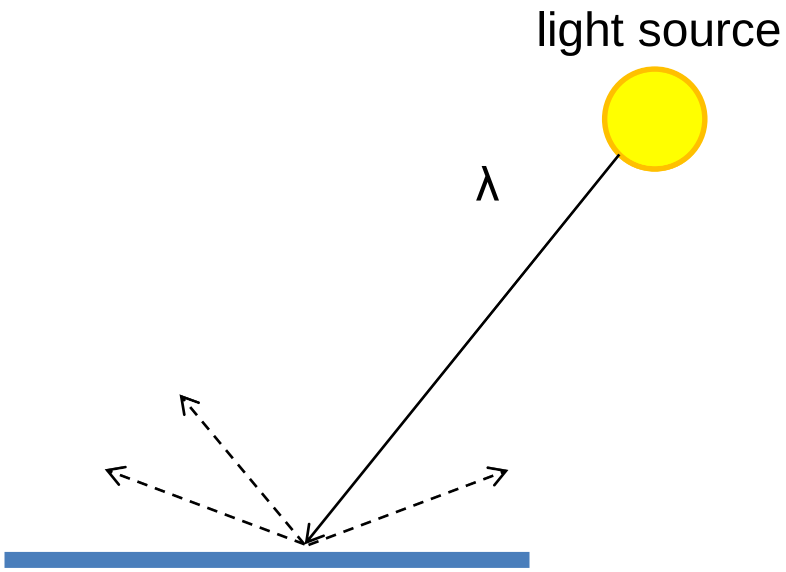
A photon’s life choices

- **Absorption**
- Diffusion
- Reflection
- Transparency
- Refraction
- Fluorescence
- Subsurface scattering
- Phosphorescence
- Interreflection



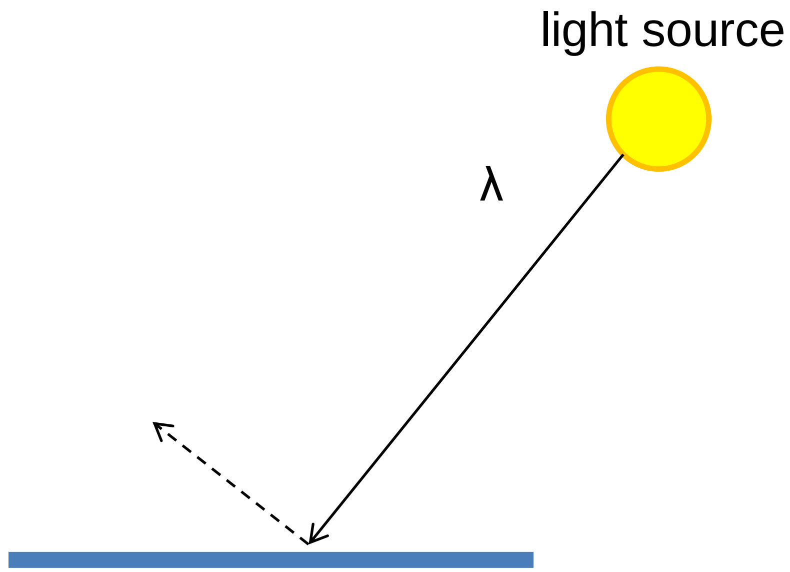
A photon's life choices

- Absorption
- **Diffuse Reflection**
- Reflection
- Transparency
- Refraction
- Fluorescence
- Subsurface scattering
- Phosphorescence
- Interreflection



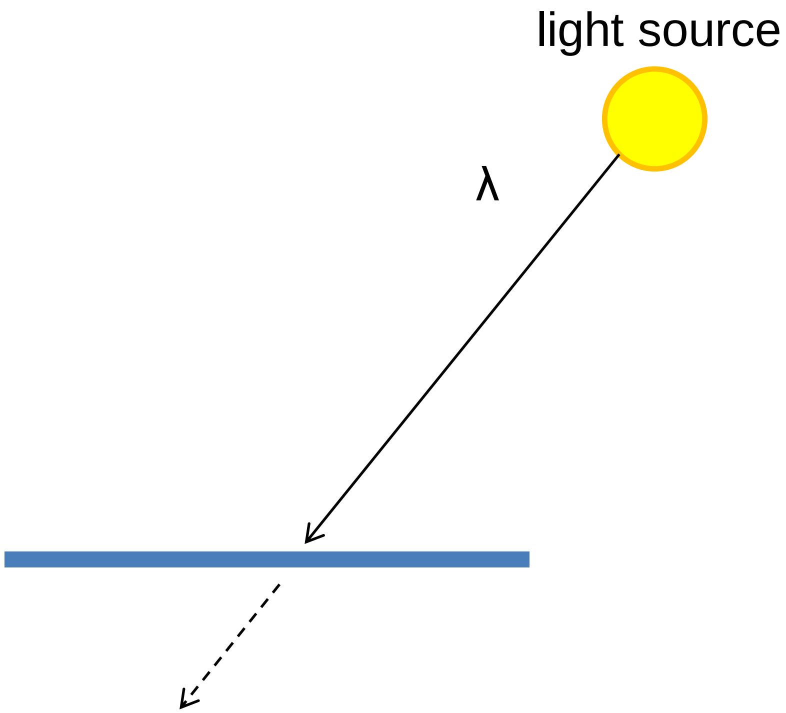
A photon's life choices

- Absorption
- Diffusion
- **Specular Reflection**
- Transparency
- Refraction
- Fluorescence
- Subsurface scattering
- Phosphorescence
- Interreflection



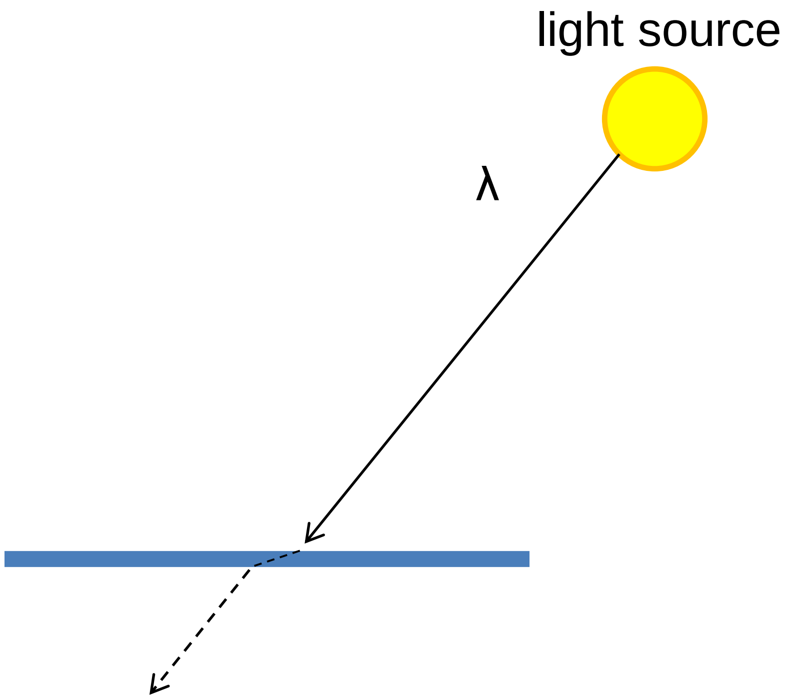
A photon’s life choices

- Absorption
- Diffusion
- Reflection
- **Transparency**
- Refraction
- Fluorescence
- Subsurface scattering
- Phosphorescence
- Interreflection



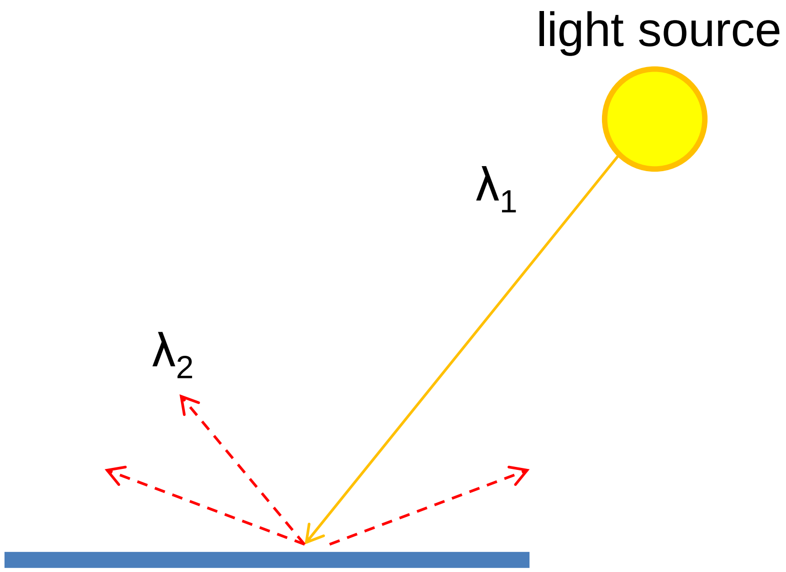
A photon's life choices

- Absorption
- Diffusion
- Reflection
- Transparency
- **Refraction**
- Fluorescence
- Subsurface scattering
- Phosphorescence
- Interreflection



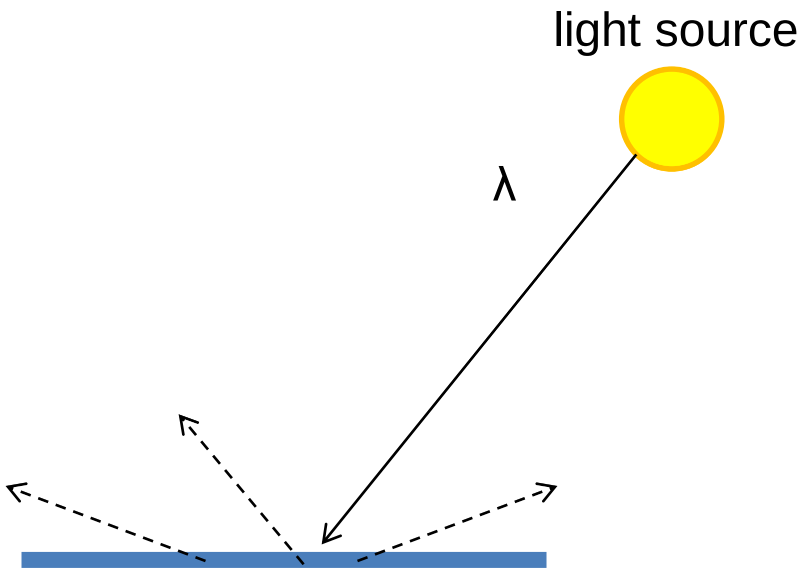
A photon's life choices

- Absorption
- Diffusion
- Reflection
- Transparency
- Refraction
- **Fluorescence**
- Subsurface scattering
- Phosphorescence
- Interreflection



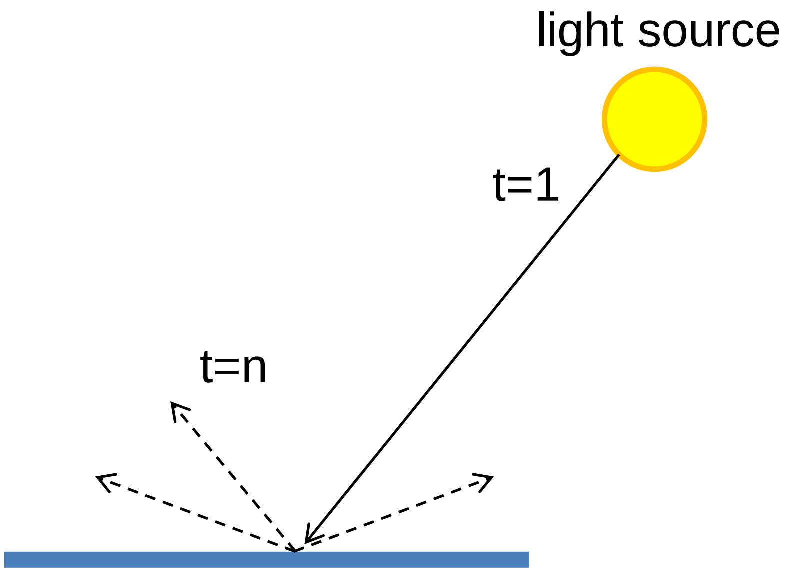
A photon’s life choices

- Absorption
- Diffusion
- Reflection
- Transparency
- Refraction
- Fluorescence
- **Subsurface scattering**
- Phosphorescence
- Interreflection



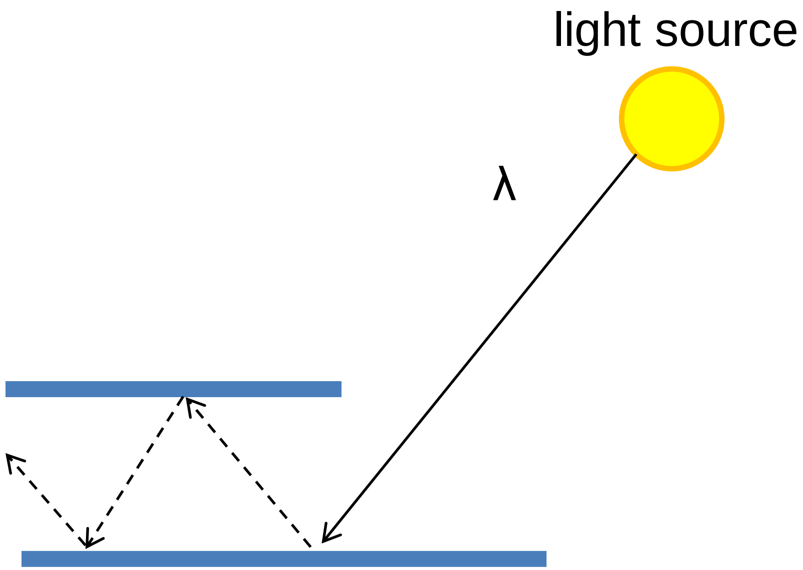
A photon's life choices

- Absorption
- Diffusion
- Reflection
- Transparency
- Refraction
- Fluorescence
- Subsurface scattering
- **Phosphorescence**
- Interreflection



A photon's life choices

- Absorption
- Diffusion
- Reflection
- Transparency
- Refraction
- Fluorescence
- Subsurface scattering
- Phosphorescence
- **Interreflection**



(Specular Interreflection)

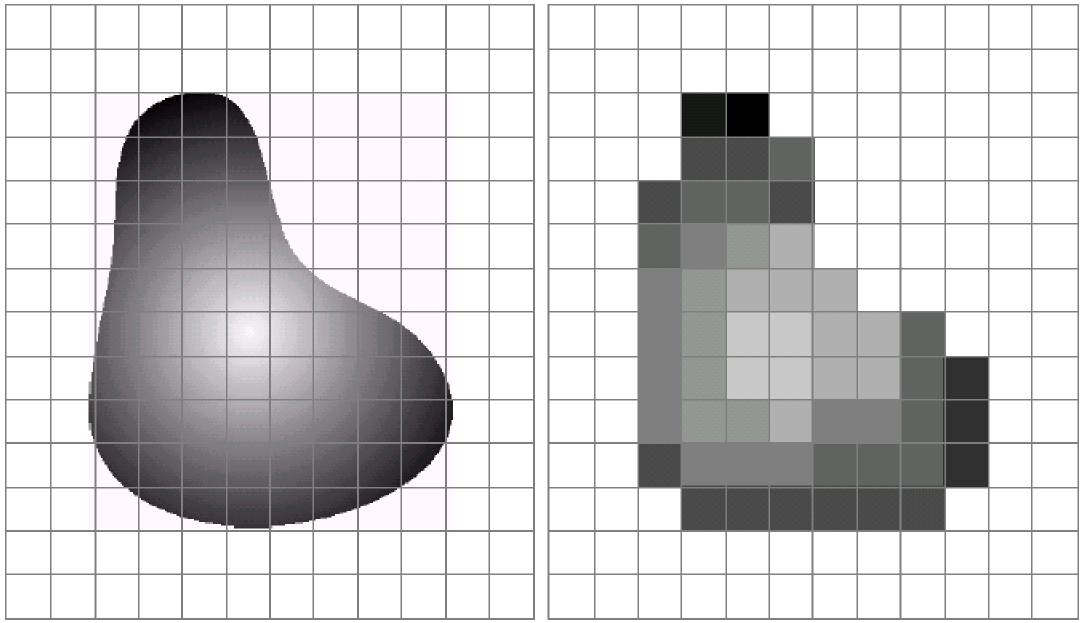
Lambertian Reflectance

- In computer vision, surfaces are often assumed to be ideal diffuse reflectors with no dependence on viewing direction.

Digital camera

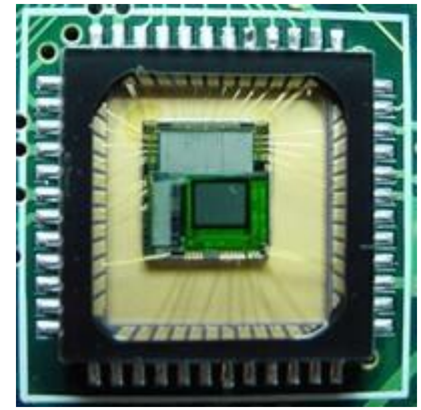
- A digital camera replaces film with a sensor array
  - Each cell in the array is light-sensitive diode that converts photons to electrons
  - Two common types
    - Charge Coupled Device (CCD)
    - CMOS
  - http://electronics.howstuffworks.com/digital-camera.htm

Sensor Array



a b

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



CMOS sensor

Sampling and Quantization

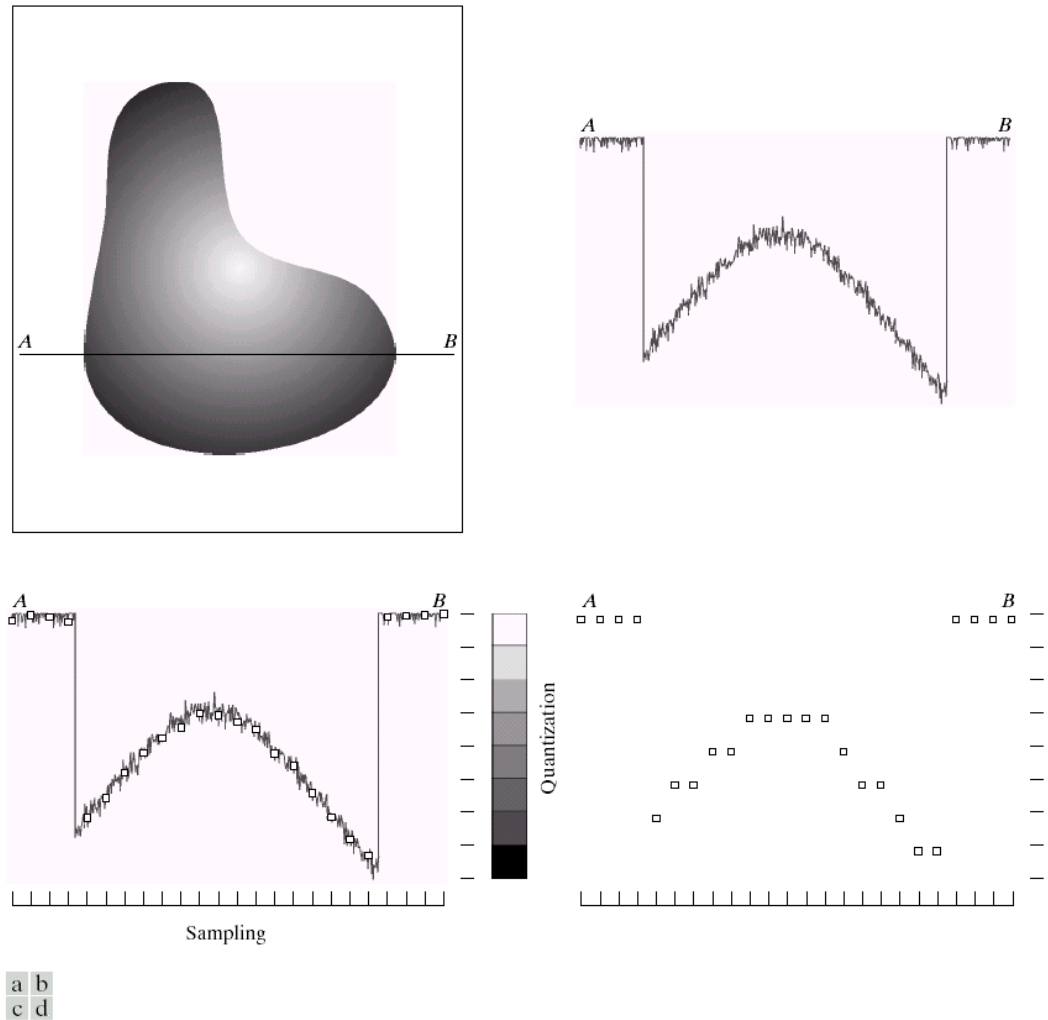
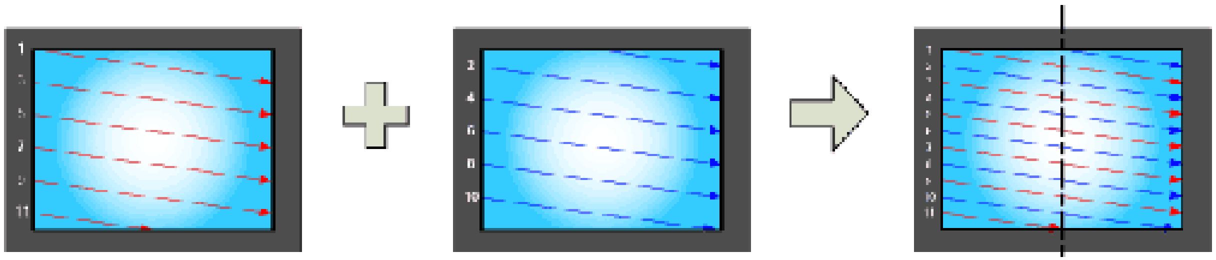


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

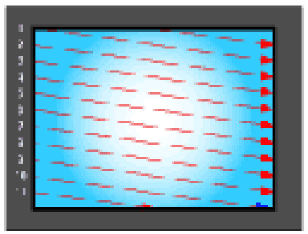
Interlace vs. progressive scan



1st field: Odd field

2nd field: Even field

One complete frame
using interlaced scanning



One complete frame
using progressive scanning

http://www.axis.com/products/video/camera/progressive_scan.htm

Slide by Steve Seitz

Progressive scan

http://www.axis.com/products/video/camera/progressive_scan.htm

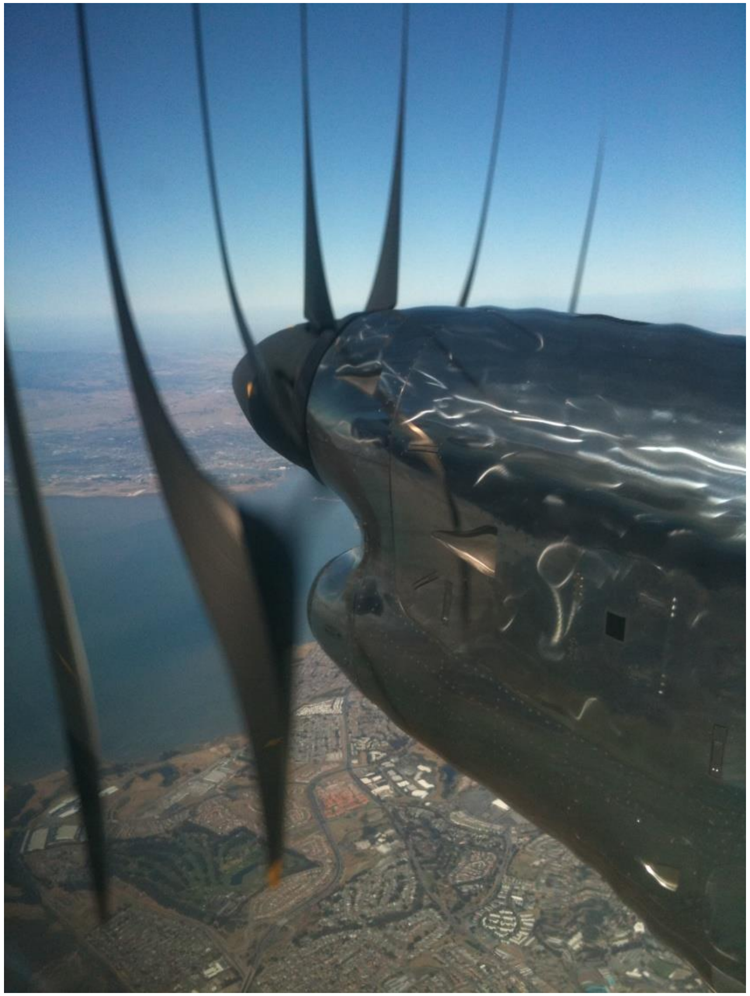
Slide by Steve Seitz

Interlace

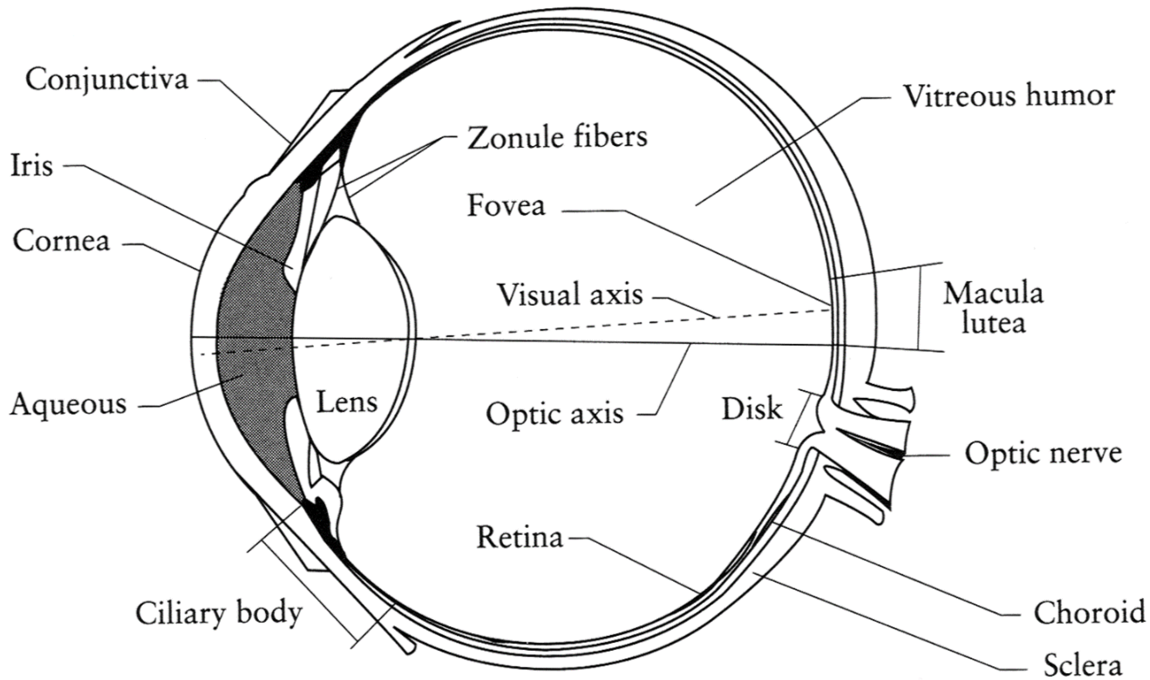
http://www.axis.com/products/video/camera/progressive_scan.htm

Slide by Steve Seitz

Rolling Shutter



The Eye



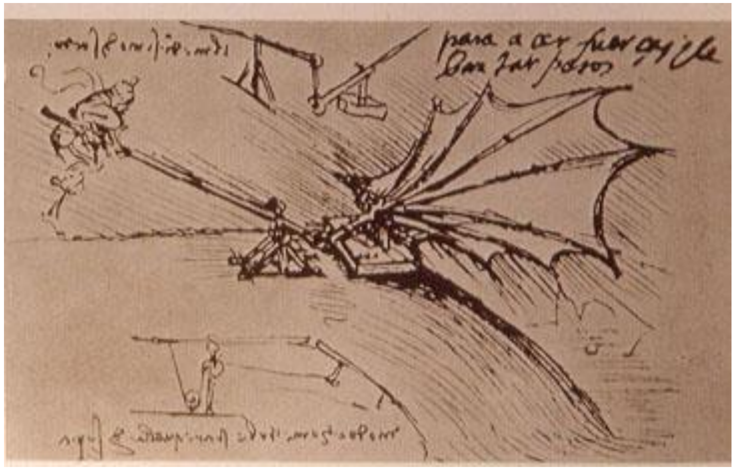
- The human eye is a camera!
  - **Iris** - colored annulus with radial muscles
  - **Pupil** - the hole (aperture) whose size is controlled by the iris
  - What’s the “film”?
    - photoreceptor cells (rods and cones) in the **retina**

Slide by Steve Seitz

Aside: why do we care about human vision in this class?

- We don’t, necessarily.

Ornithopters



Why do we care about human vision?

- We don’t, necessarily.
- But cameras necessarily imitate the frequency response of the human eye, so we should know that much.
- Also, computer vision probably wouldn’t get as much scrutiny if biological vision (especially human vision) hadn’t proved that it was possible to make important judgements from 2d images.

Does computer vision "understand" images?

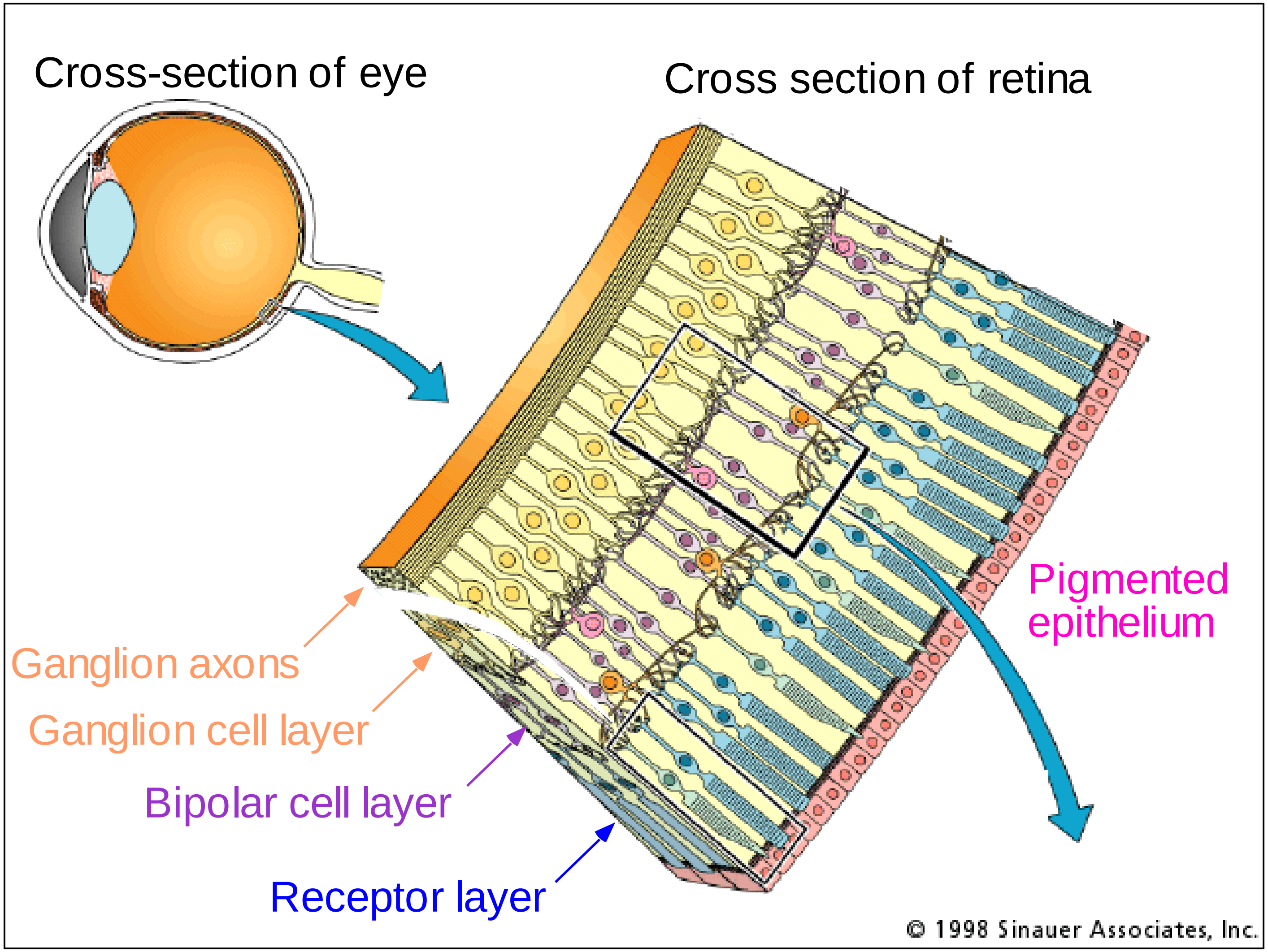
"Can machines fly?" The answer is yes, because airplanes fly.

"Can machines swim?" The answer is no, because submarines don't swim.

"Can machines think?" Is this question like the first, or like the second?

Source: Norvig

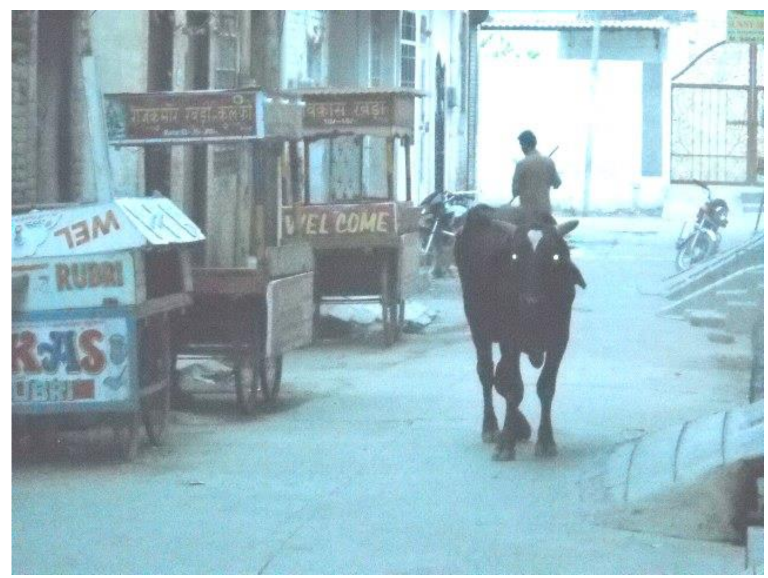
The Retina



Cross-section of eye
Cross section of retina
Pigmented epithelium
Ganglion axons
Ganglion cell layer
Bipolar cell layer
Receptor layer
© 1998 Sinauer Associates, Inc.

What humans don’t have: tapetum lucidum

Human eyes can reflect a tiny bit and blood in the retina makes this reflection red.



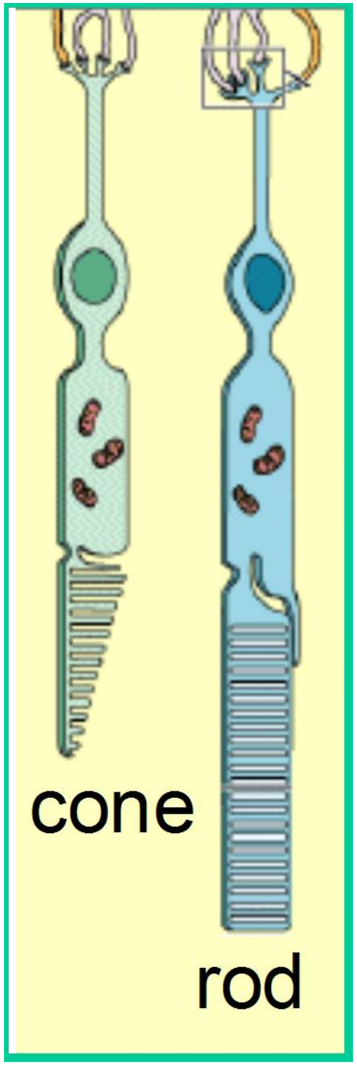
Two types of light-sensitive receptors

Cones

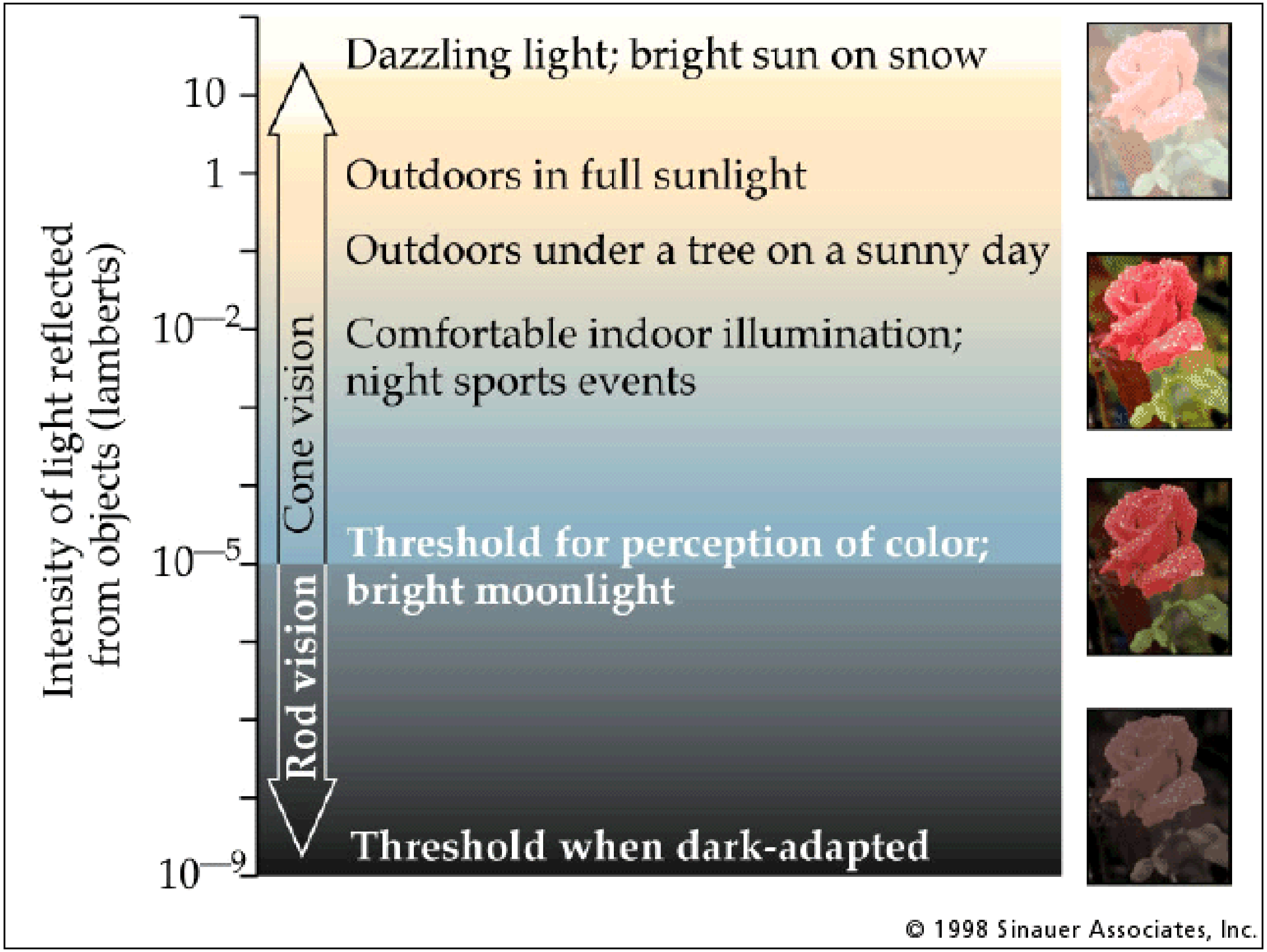
cone-shaped
less sensitive
operate in high light
color vision

Rods

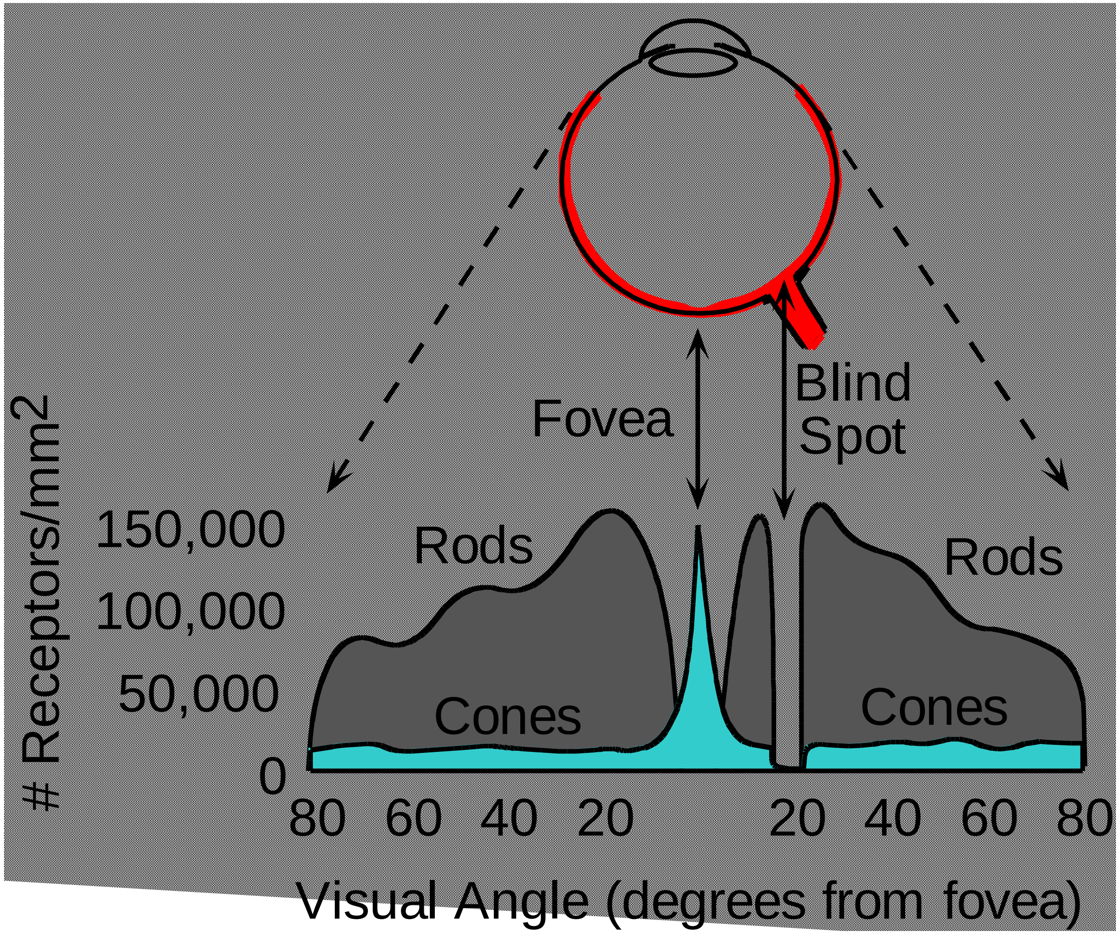
rod-shaped
highly sensitive
operate at night
gray-scale vision



Rod / Cone sensitivity



Distribution of Rods and Cones



Night Sky: why are there more stars off-center?

Averted vision: http://en.wikipedia.org/wiki/Averted_vision

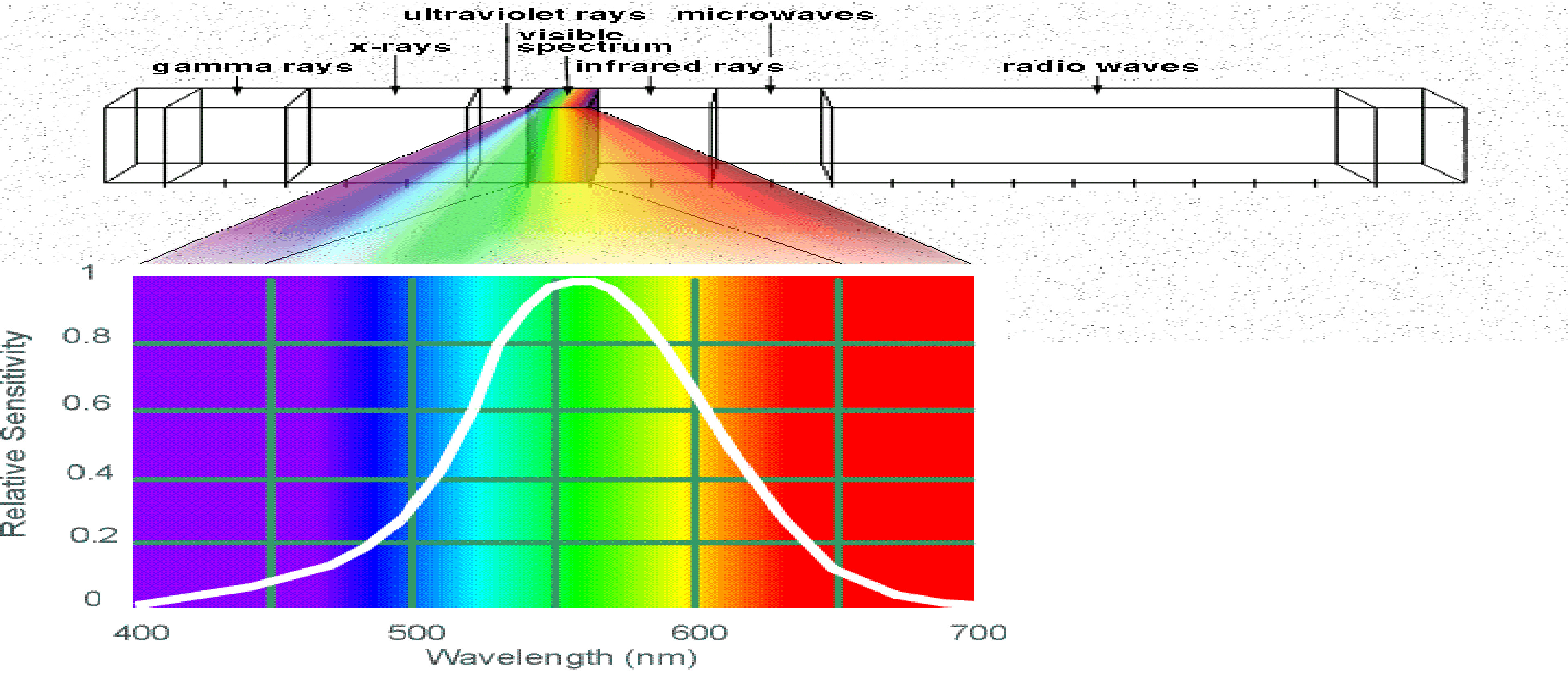
Wait, the blood vessels are in front of the photoreceptors??

https://www.youtube.com/watch?v=L_W-IXqoxHA

Eye Movements

- Saccades
- Can be consciously controlled. Related to perceptual attention.
- 200ms to initiation, 20 to 200ms to carry out. Large amplitude.
- Microsaccades
- Involuntary. Smaller amplitude. Especially evident during prolonged fixation. Function debated.
- Ocular microtremor (OMT)
- involuntary. high frequency (up to 80Hz), small amplitude.
- Smooth pursuit – tracking an object

Electromagnetic Spectrum

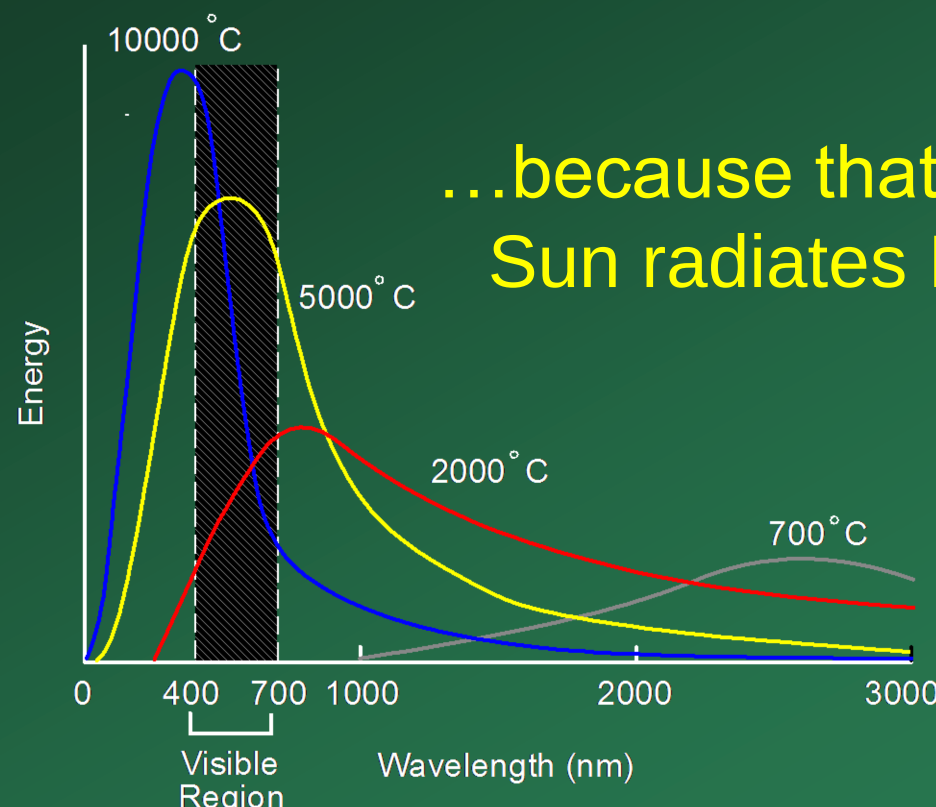


Human Luminance Sensitivity Function

http://www.yorku.ca/eye/photopik.htm

Visible Light

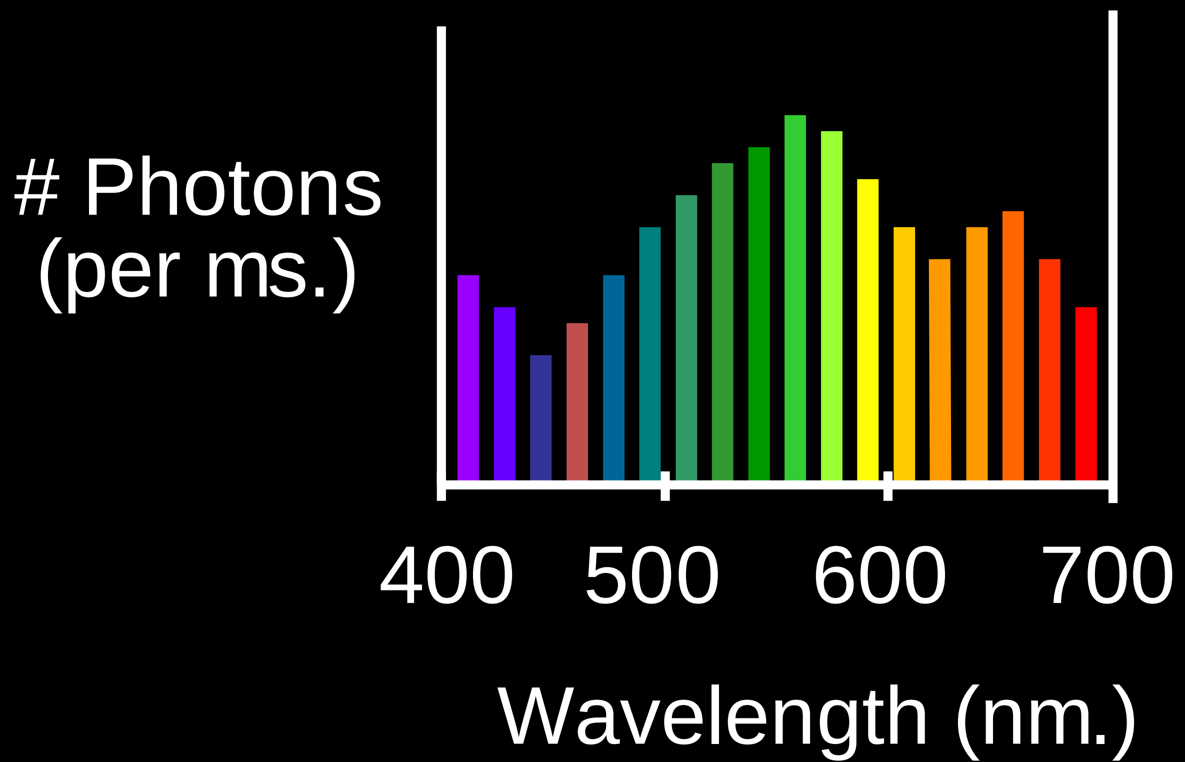
Why do we see light of these wavelengths?



…because that's where the Sun radiates EM energy

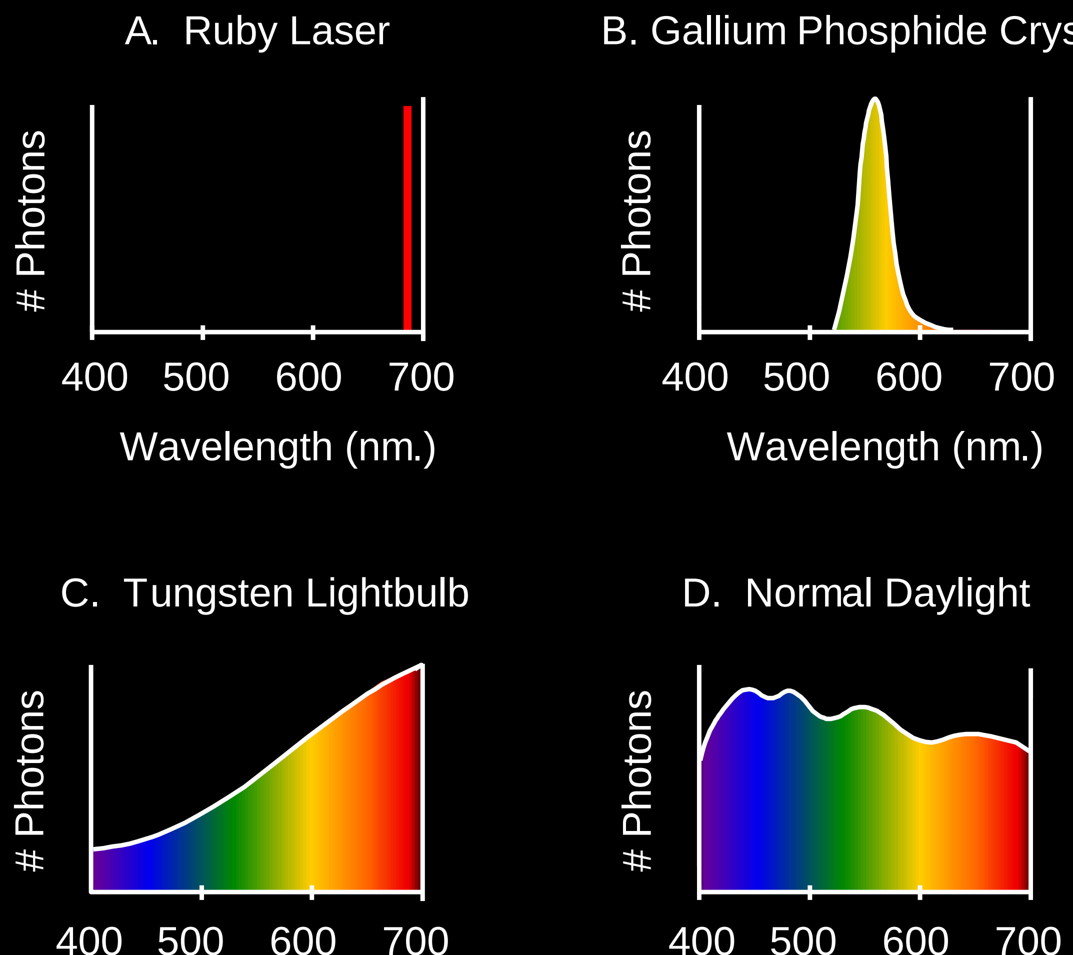
The Physics of Light

Any patch of light can be completely described physically by its spectrum: the number of photons (per time unit) at each wavelength 400 - 700 nm.



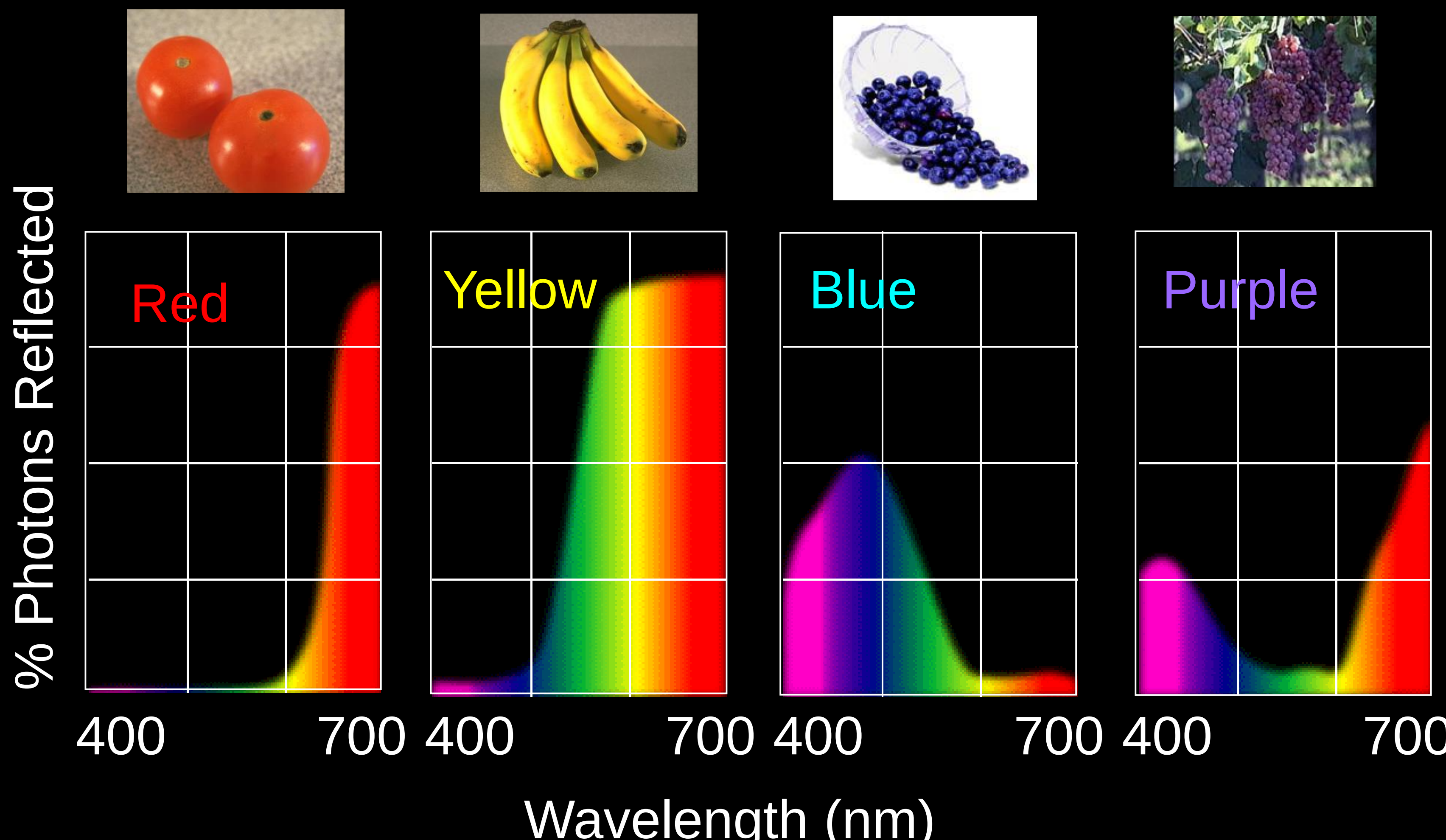
The Physics of Light

Some examples of the spectra of light sources



The Physics of Light

Some examples of the reflectance spectra of surfaces

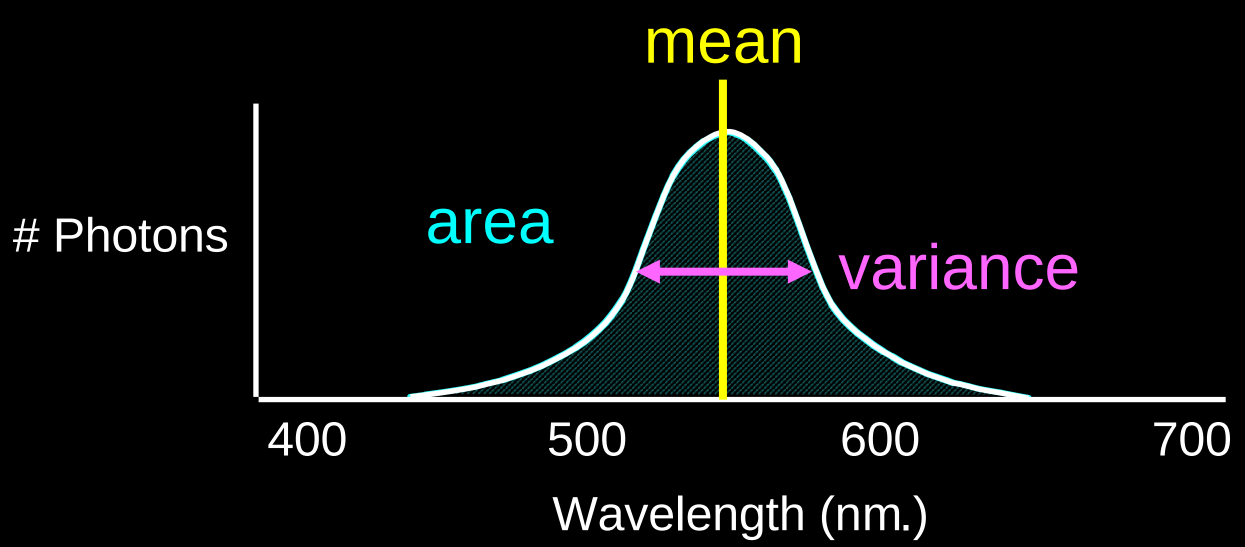


The Psychophysical Correspondence

There is no simple functional description for the perceived color of all lights under all viewing conditions, but …...

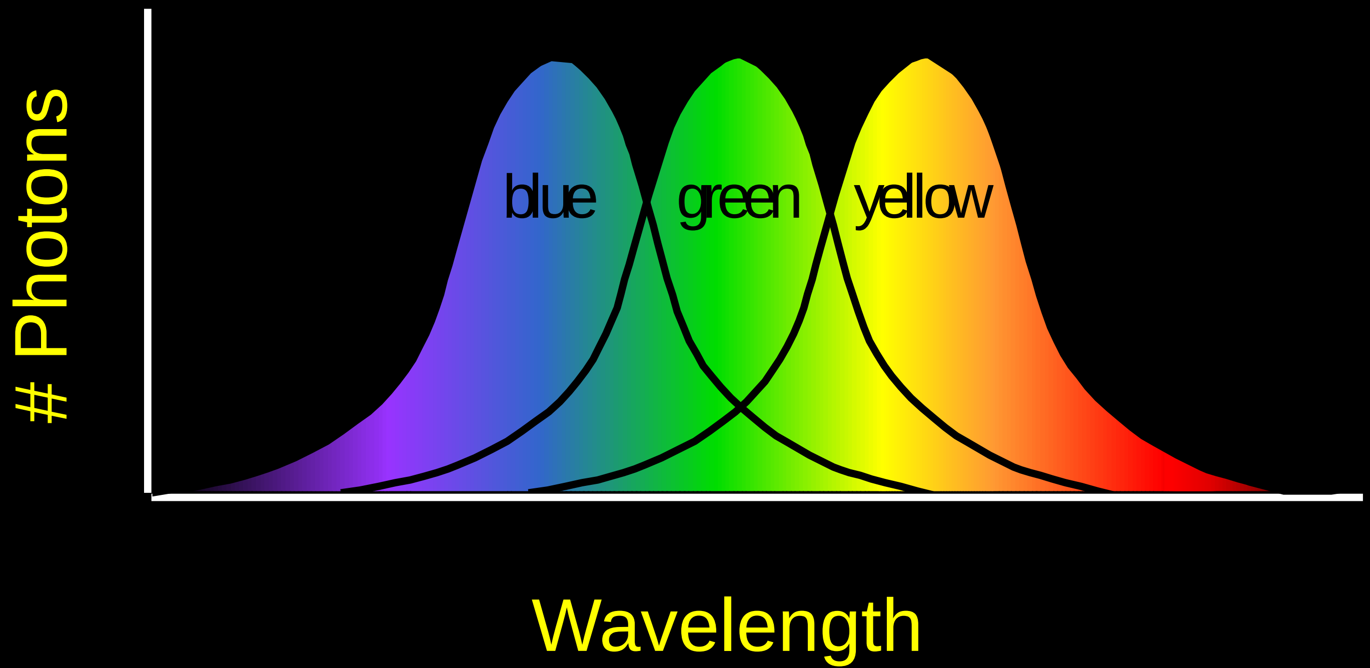
A helpful constraint:

Consider only physical spectra with normal distributions



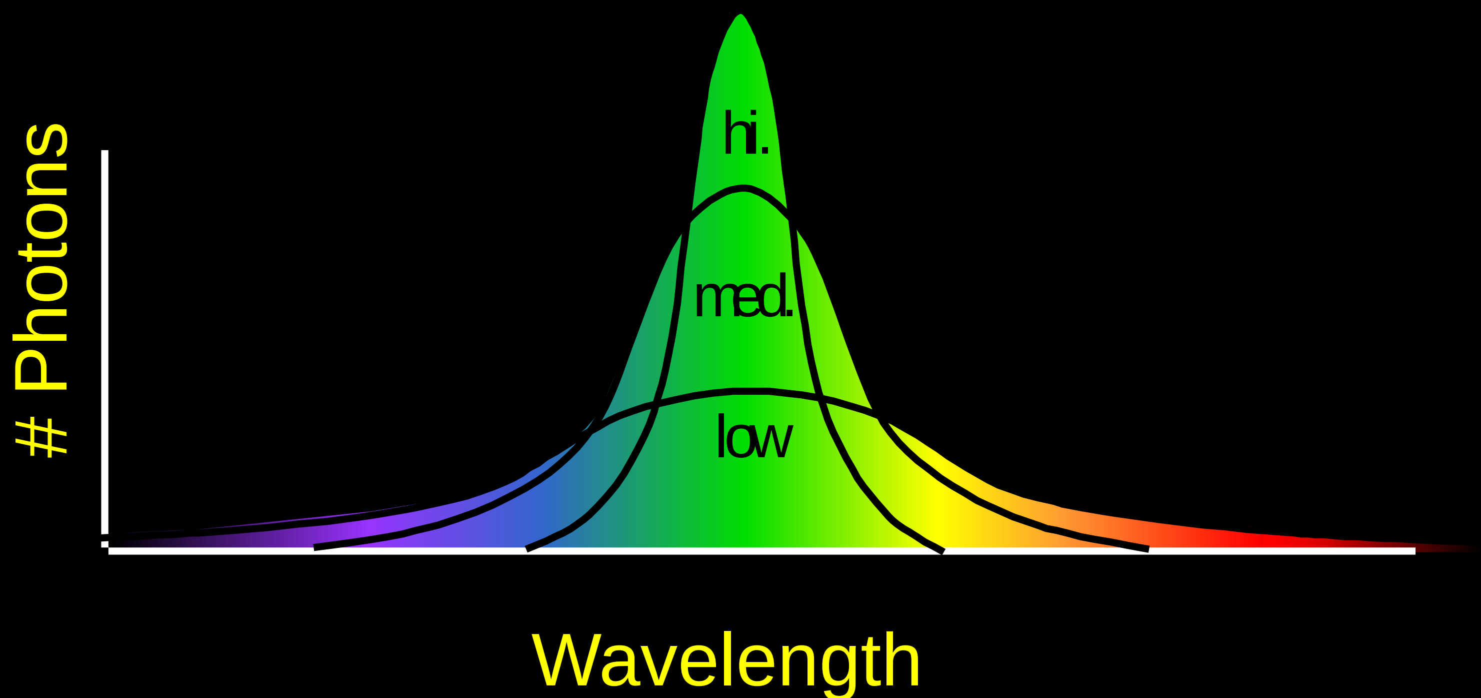
The Psychophysical Correspondence

Mean ⟷ Hue



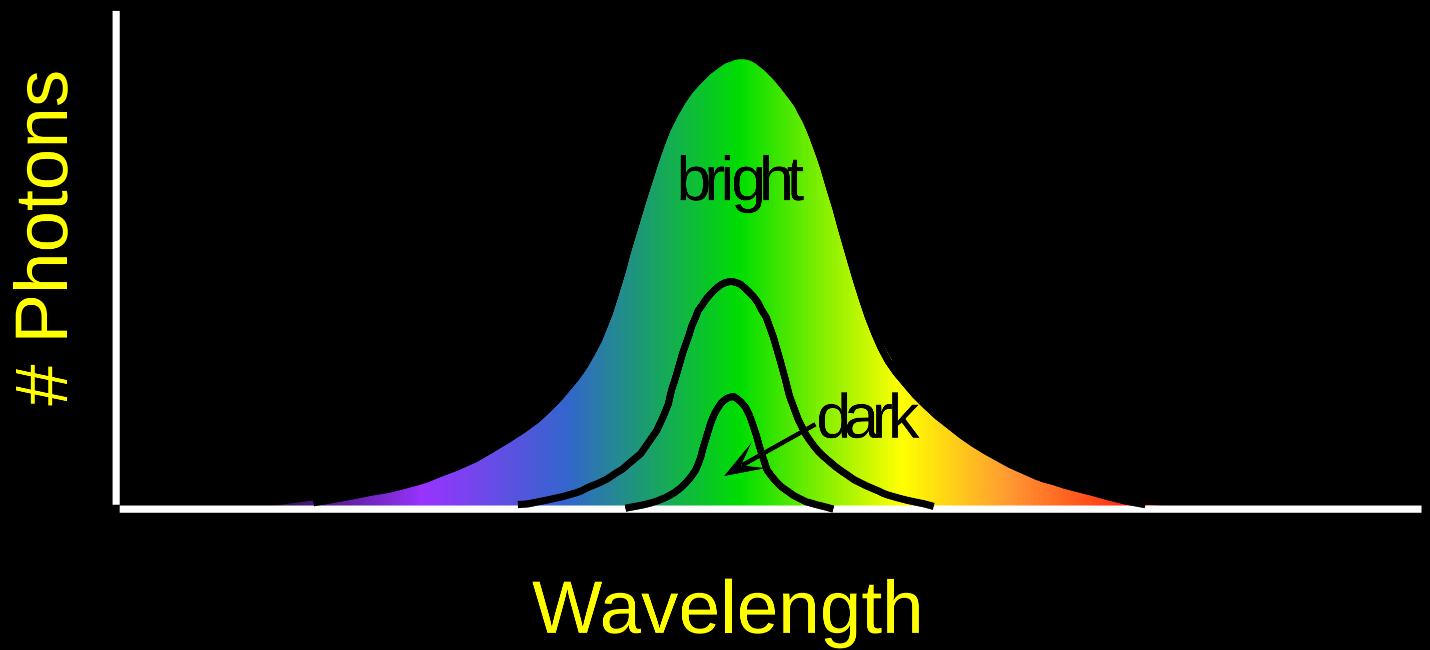
The Psychophysical Correspondence

Variance ⟷ Saturation



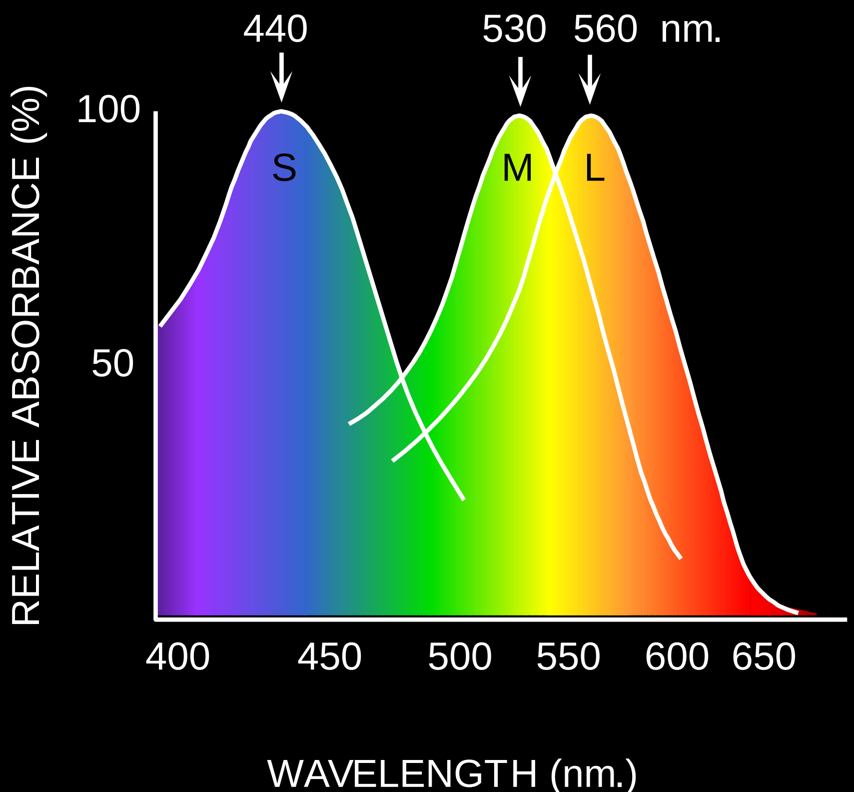
The Psychophysical Correspondence

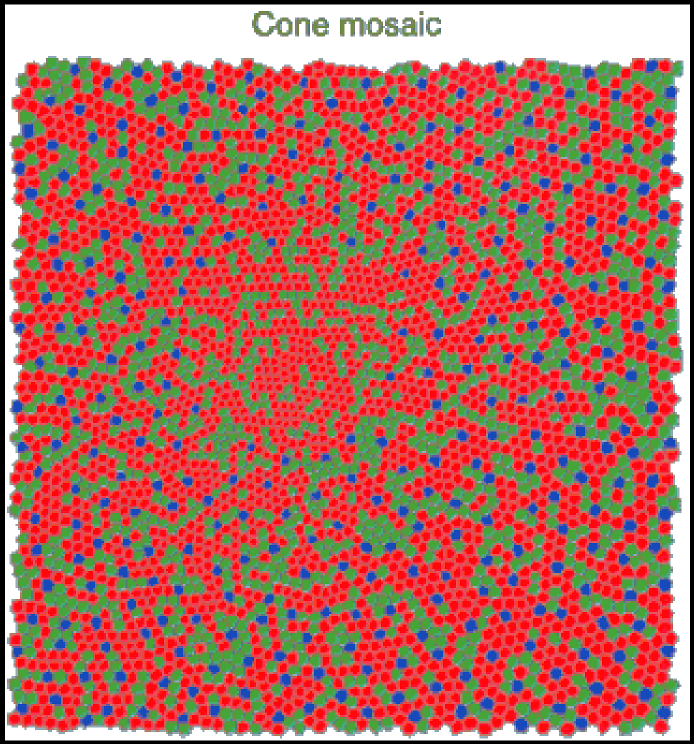
Area ⟷ Brightness



Physiology of Color Vision

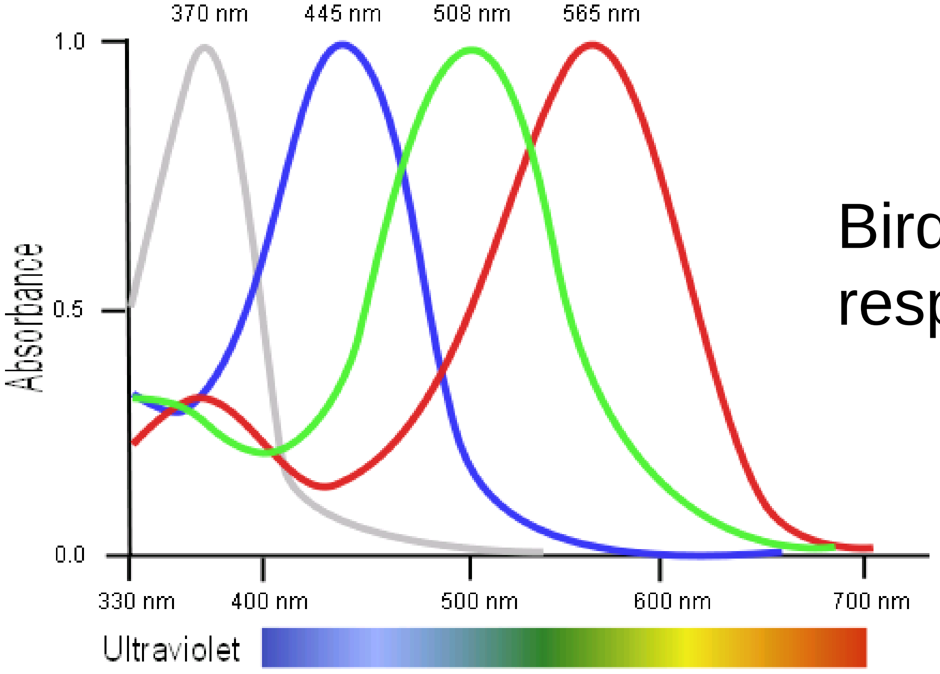
Three kinds of cones:





- Why are M and L cones so close?
- Why are there 3?

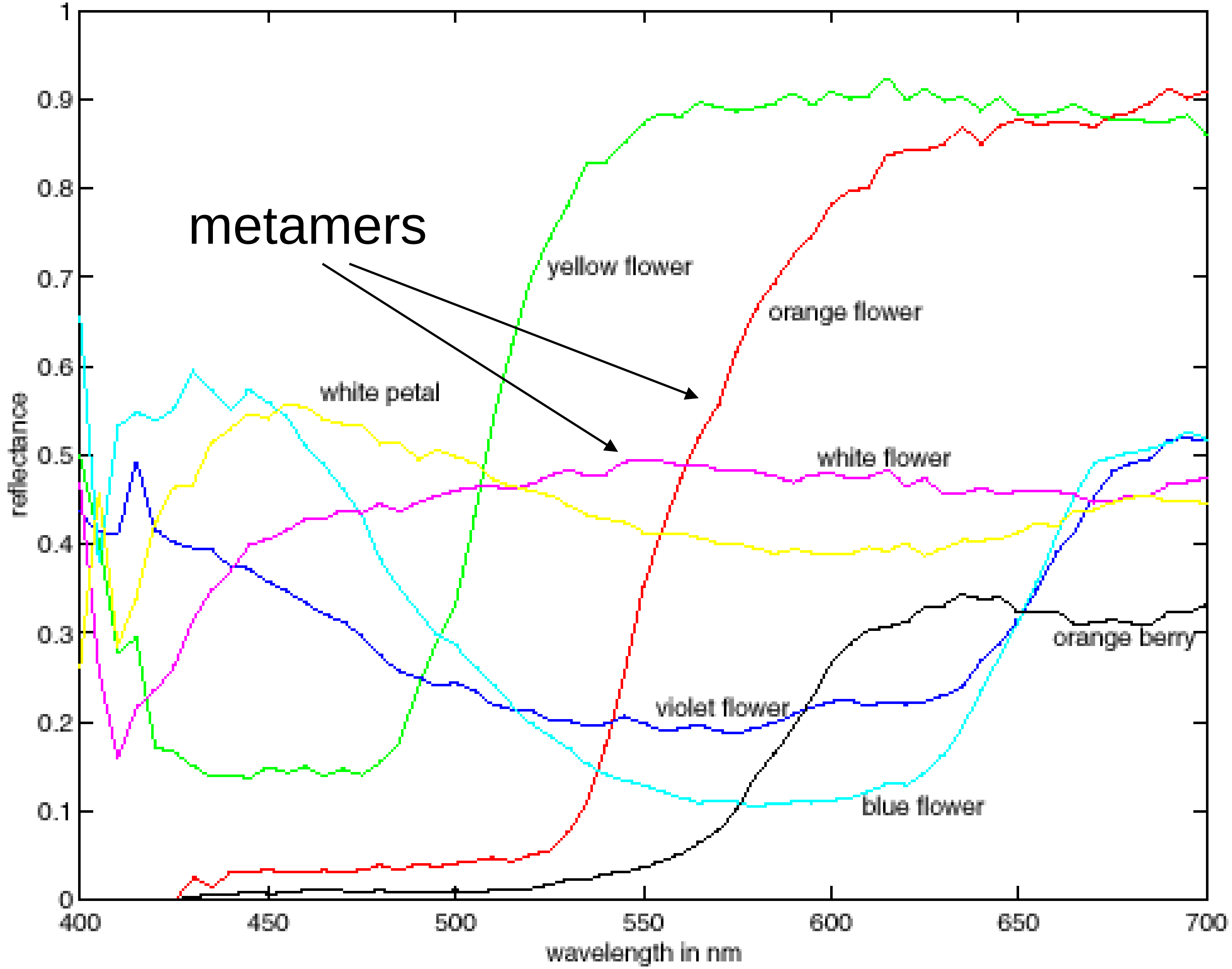
Tetrachromatism



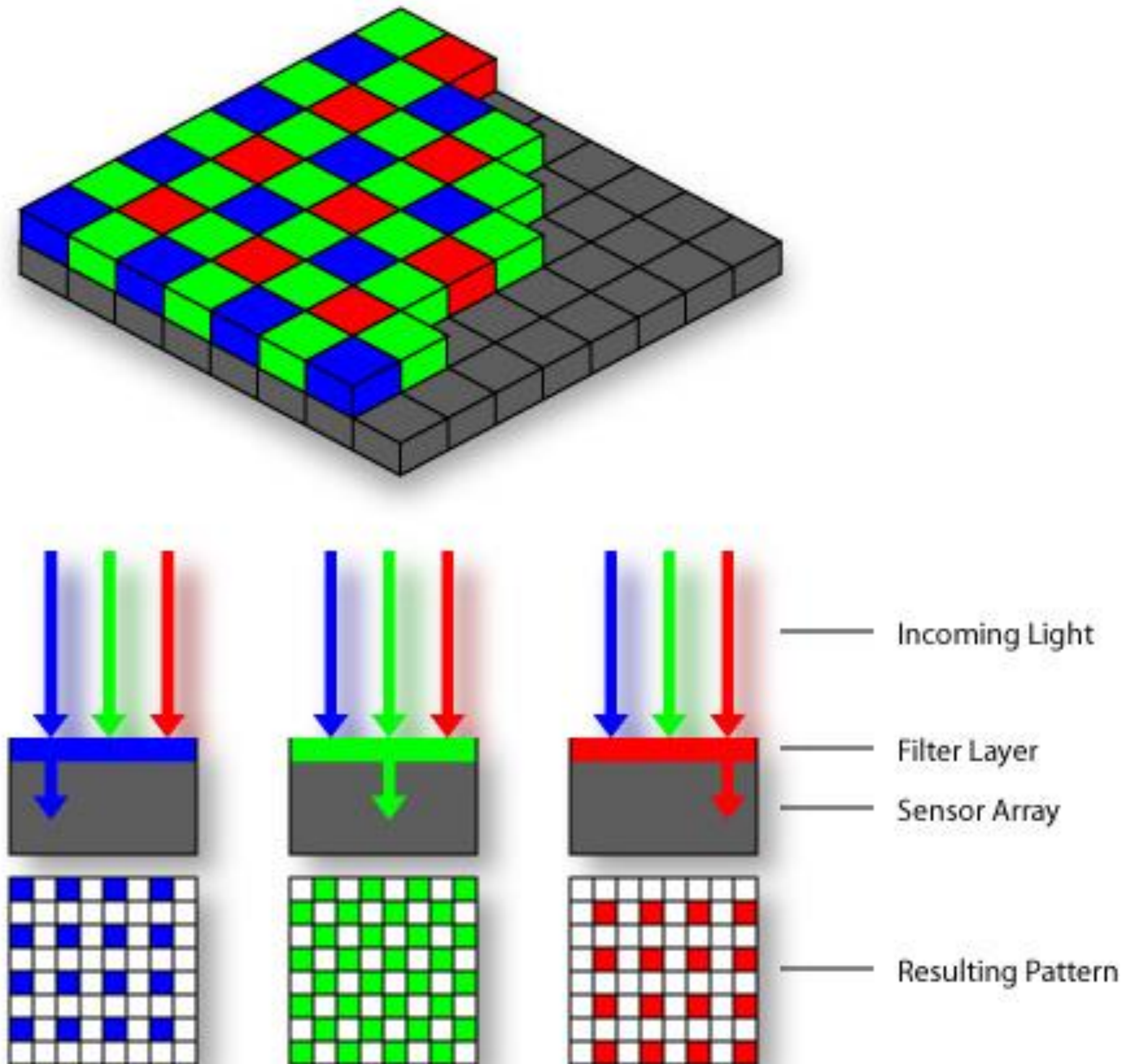
Bird cone responses

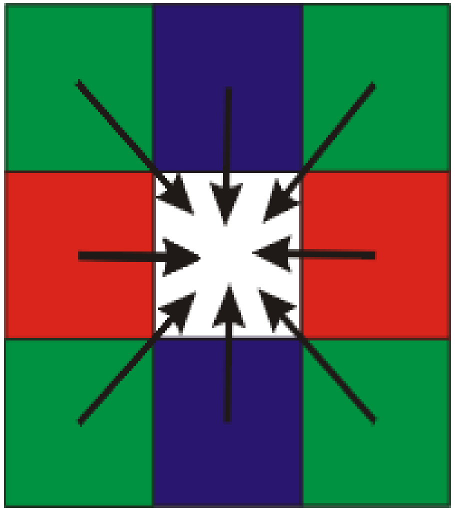
- Most birds, and many other animals, have cones for ultraviolet light.
- Some humans, mostly female, seem to have slight tetrachromatism.

More Spectra



Practical Color Sensing: Bayer Grid





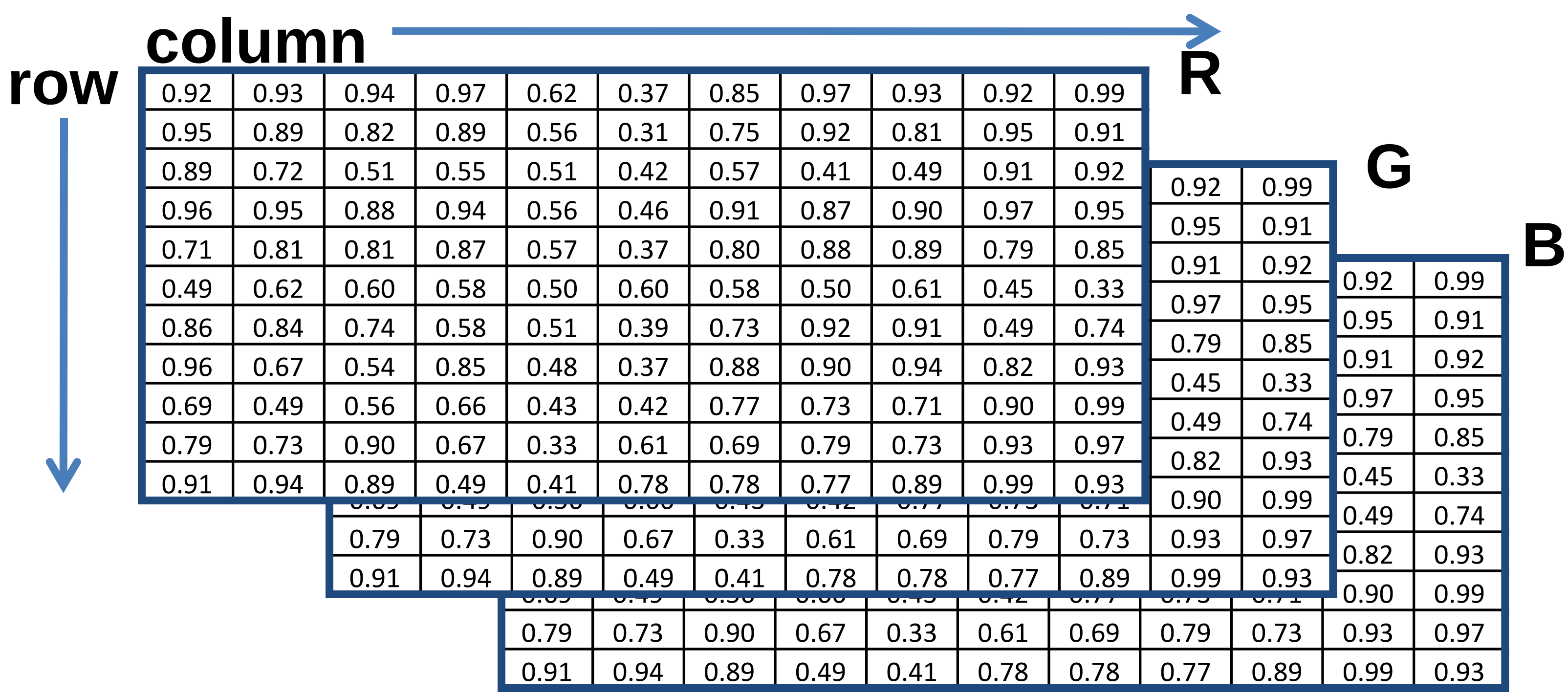
- Estimate RGB at ‘G’ cells from neighboring values

Slide by Steve Seitz

Color Image

Images in Matlab

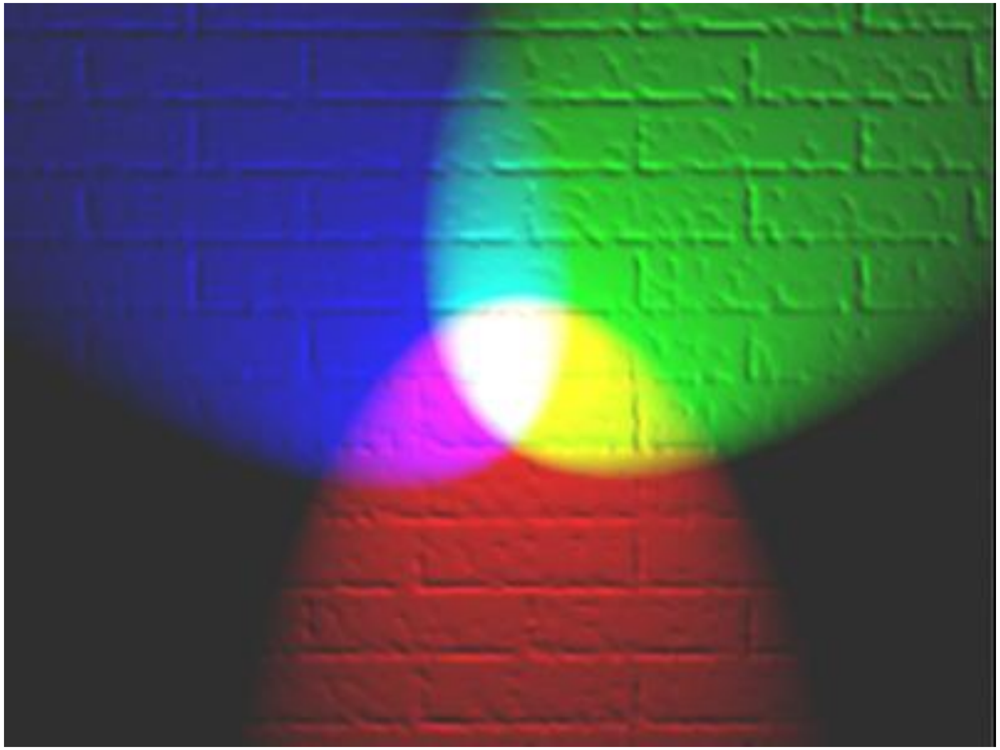
- Images represented as a matrix
- Suppose we have a NxM RGB image called "im"
  - im(1,1,1) = top-left pixel value in R-channel
  - im(y, x, b) = y pixels down, x pixels to right in the b^{th} channel
  - im(N, M, 3) = bottom-right pixel in B-channel
- imread(filename) returns a uint8 image (values 0 to 255)
  - Convert to double format (values 0 to 1) with im2double



0.92	0.93	0.94	0.97	0.62	0.37	0.85	0.97	0.93	0.92	0.99
0.95	0.89	0.82	0.89	0.56	0.31	0.75	0.92	0.81	0.95	0.91
0.89	0.72	0.51	0.55	0.51	0.42	0.57	0.41	0.49	0.91	0.92
0.96	0.95	0.88	0.94	0.56	0.46	0.91	0.87	0.90	0.97	0.95
0.71	0.81	0.81	0.87	0.57	0.37	0.80	0.88	0.89	0.79	0.85
0.49	0.62	0.60	0.58	0.50	0.60	0.58	0.50	0.61	0.45	0.33
0.86	0.84	0.74	0.58	0.51	0.39	0.73	0.92	0.91	0.49	0.74
0.96	0.67	0.54	0.85	0.48	0.37	0.88	0.90	0.94	0.82	0.93
0.69	0.49	0.56	0.66	0.43	0.42	0.77	0.73	0.71	0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

Color spaces

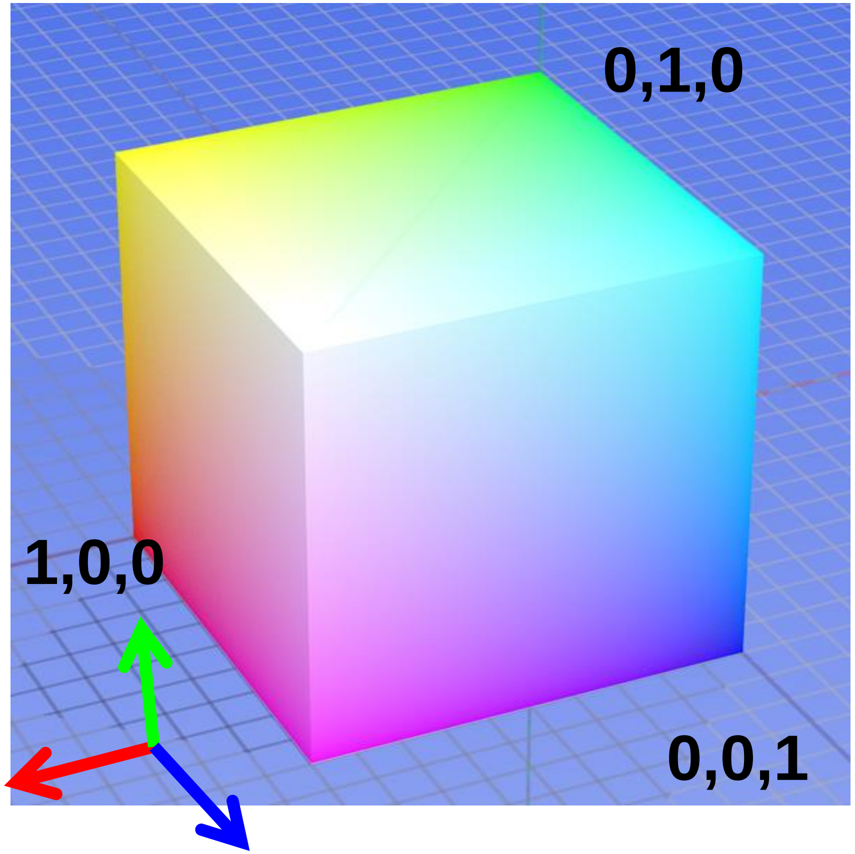
- How can we represent color?



http://en.wikipedia.org/wiki/File:RGB_illumination.jpg

Color spaces: RGB

Default color space



R
(G=0,B=0)

G
(R=0,B=0)

B
(R=0,G=0)

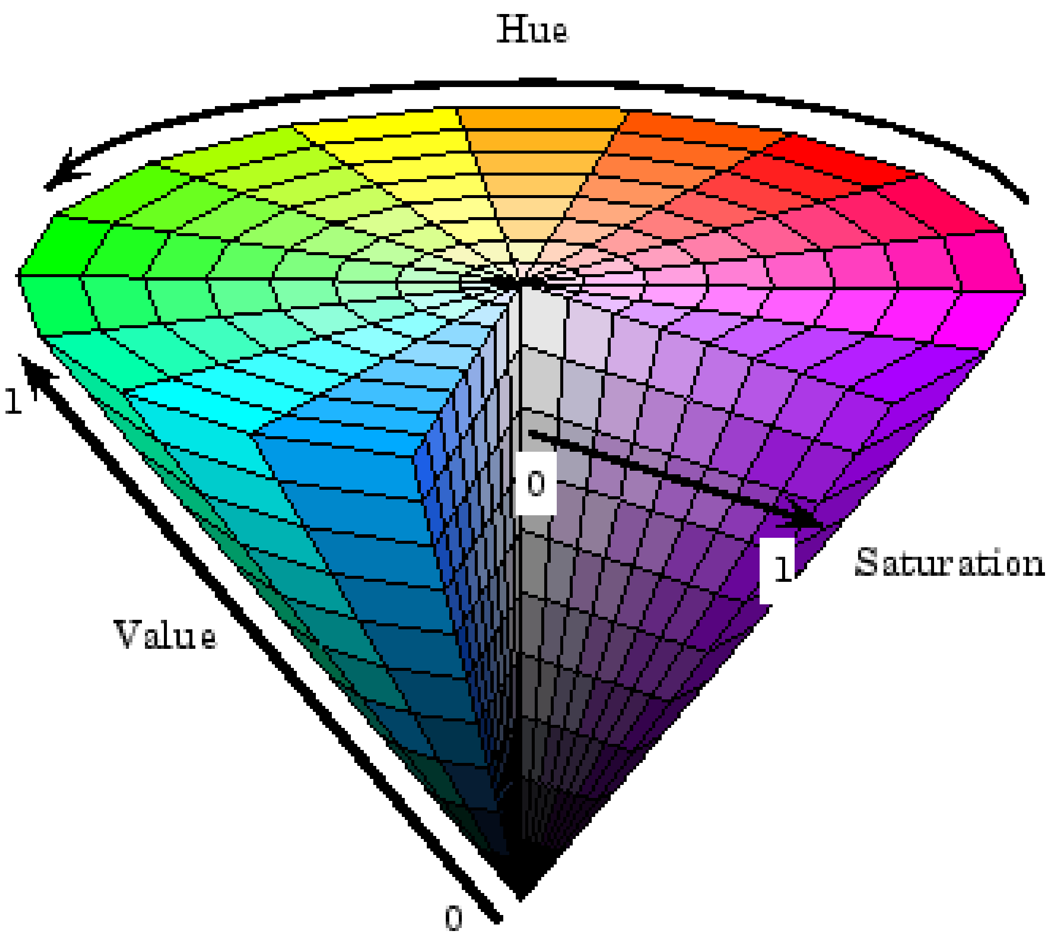
Some drawbacks

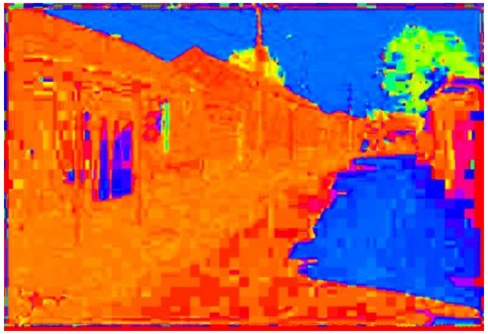
- Strongly correlated channels
- Non-perceptual

Image from: http://en.wikipedia.org/wiki/File:RGB_color_solid_cube.png

Color spaces: HSV

Intuitive color space





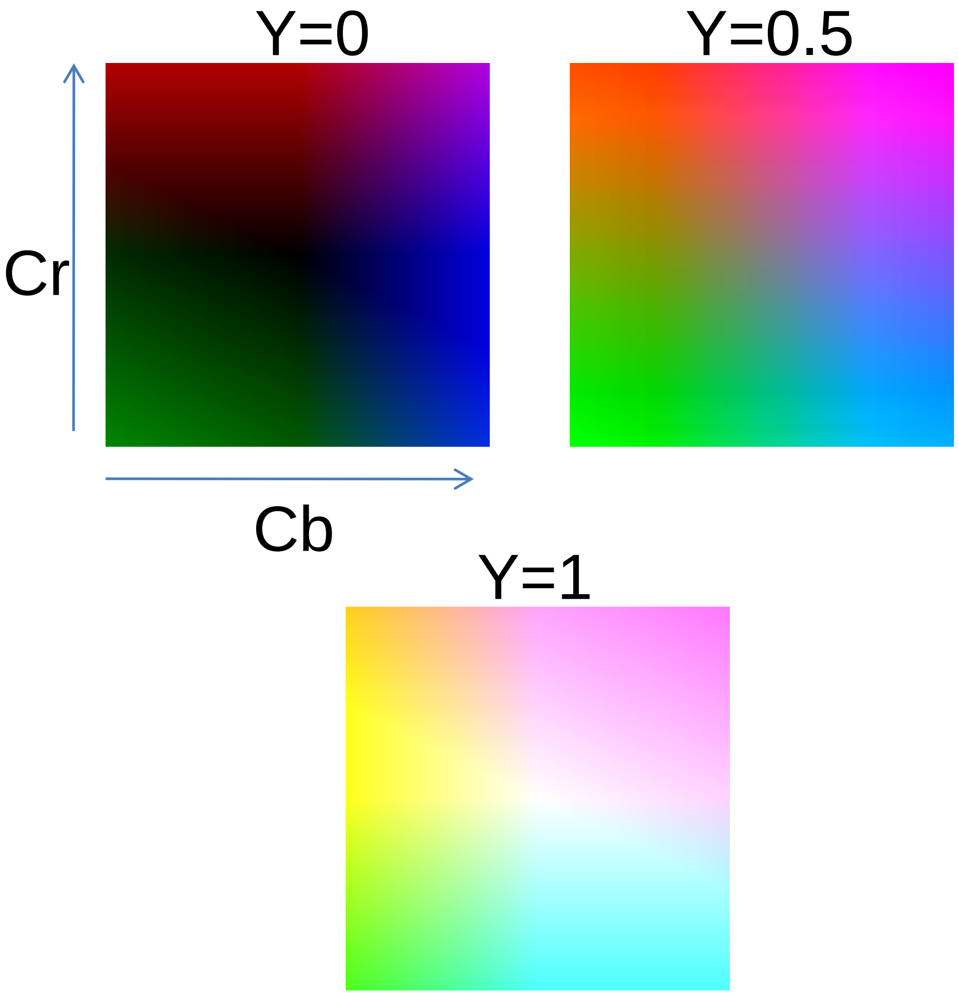
H
(S=1,V=1)

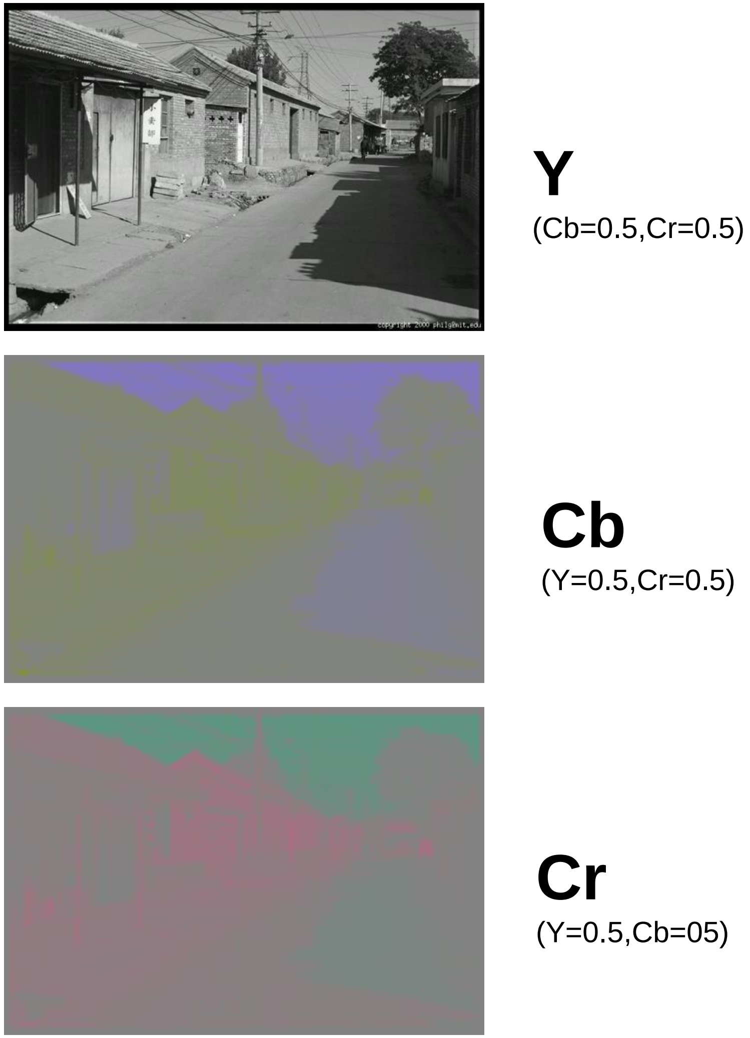
S
(H=1,V=1)

V
(H=1,S=0)

Color spaces: YCbCr

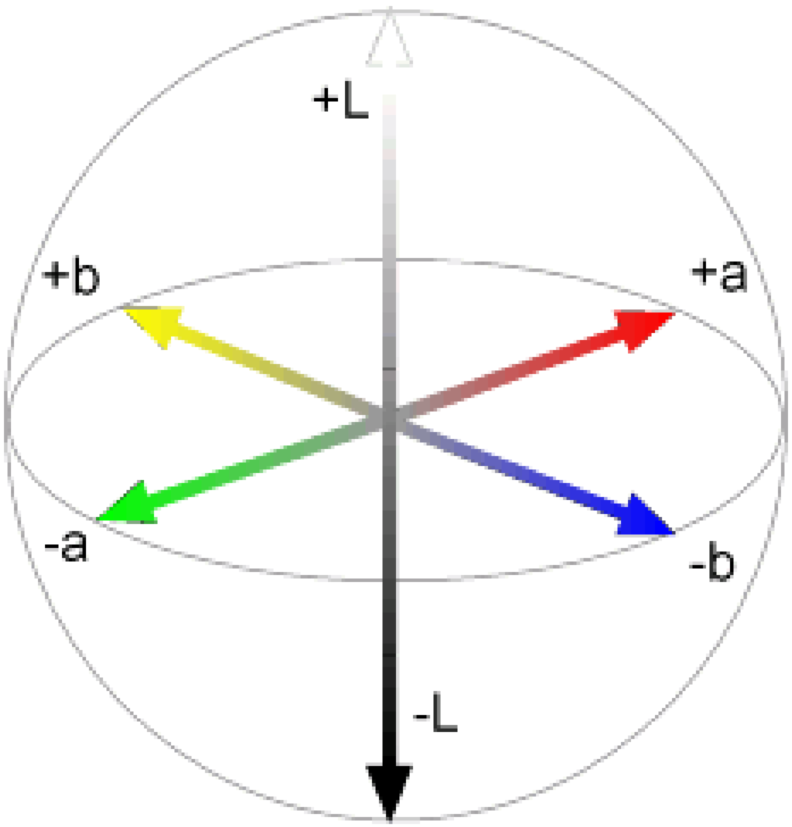
Fast to compute, good for compression, used by TV





Color spaces: L*a*b*

“Perceptually uniform”* color space



L
(a=0,b=0)

a
(L=65,b=0)

b
(L=65,a=0)

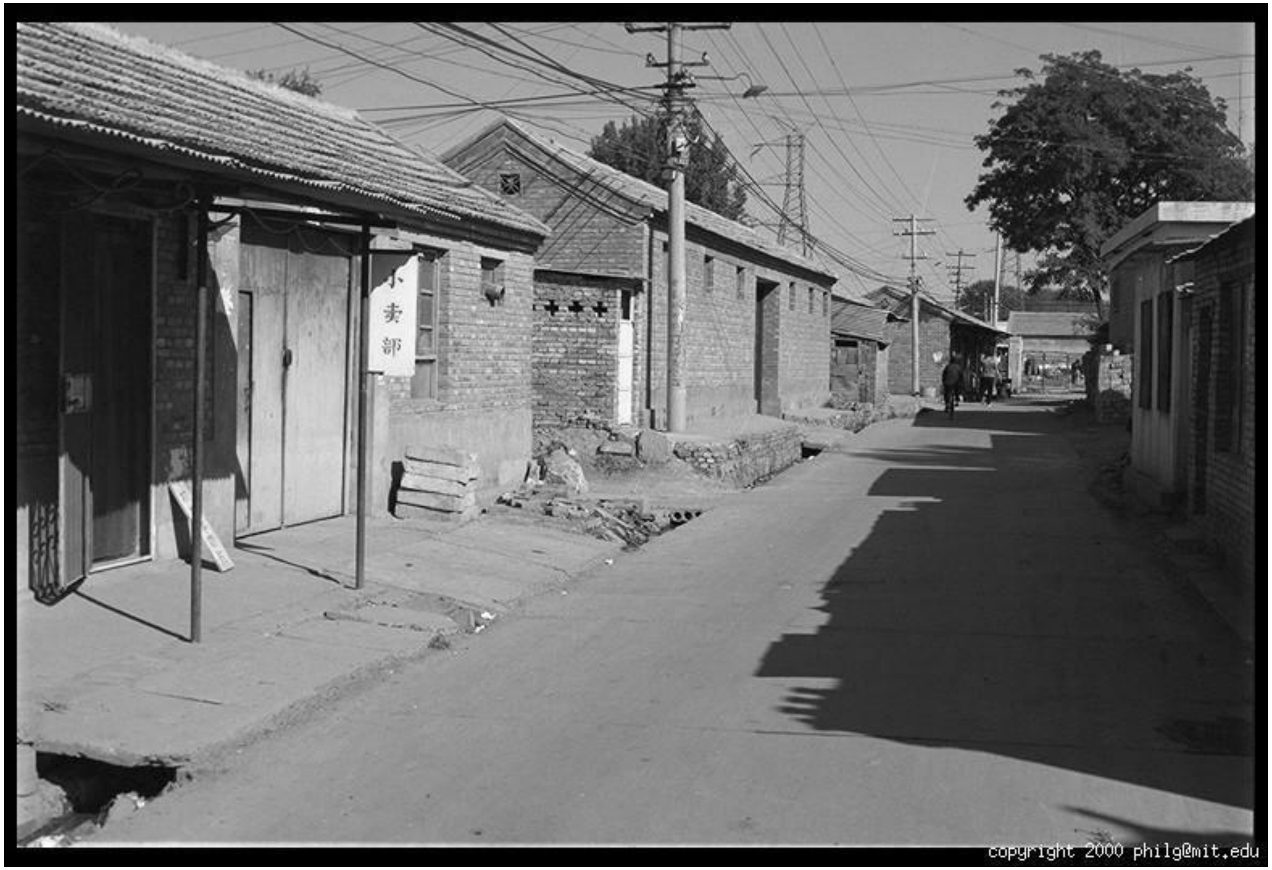
If you had to choose, would you rather go without luminance or chrominance?

If you had to choose, would you rather go without luminance or chrominance?

Most information in intensity

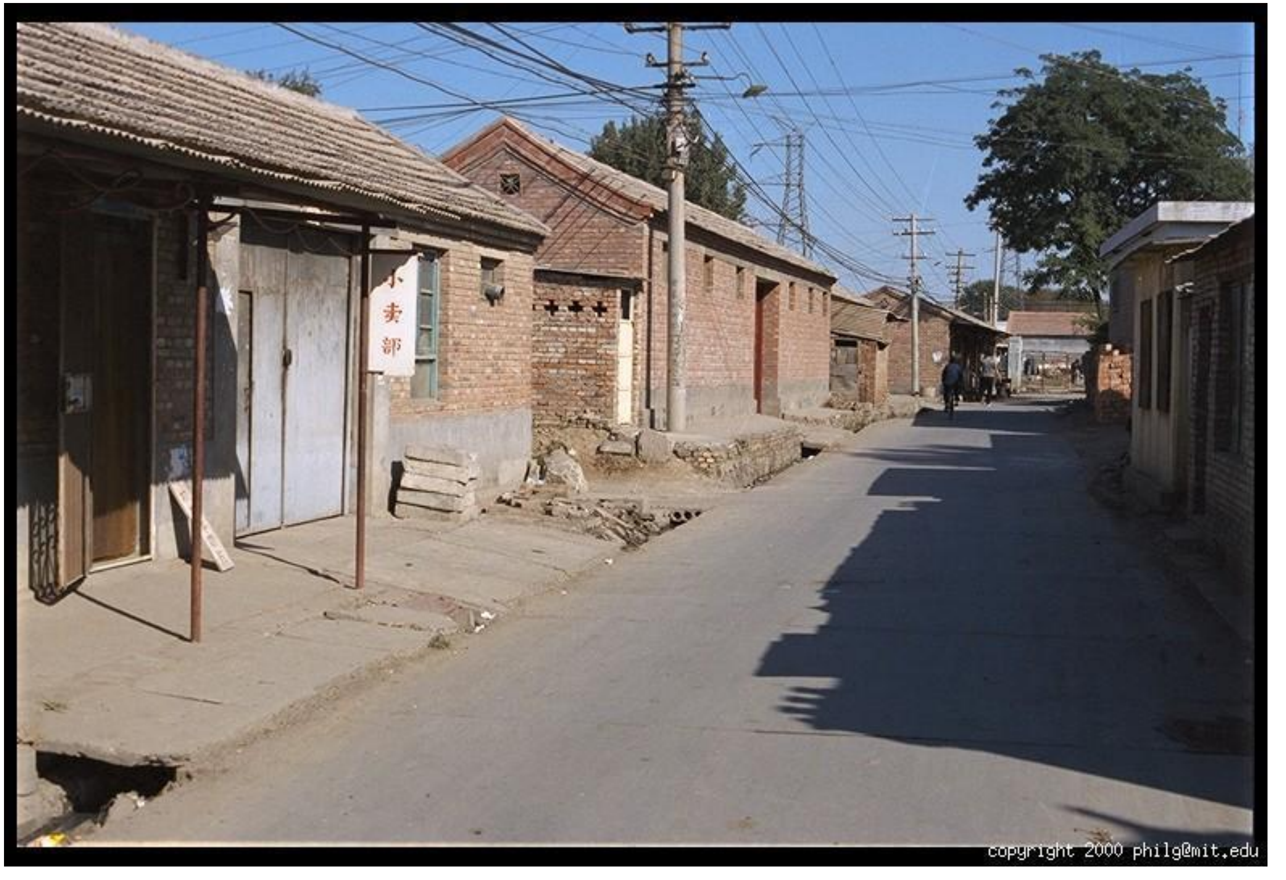
Only color shown – constant intensity

Most information in intensity



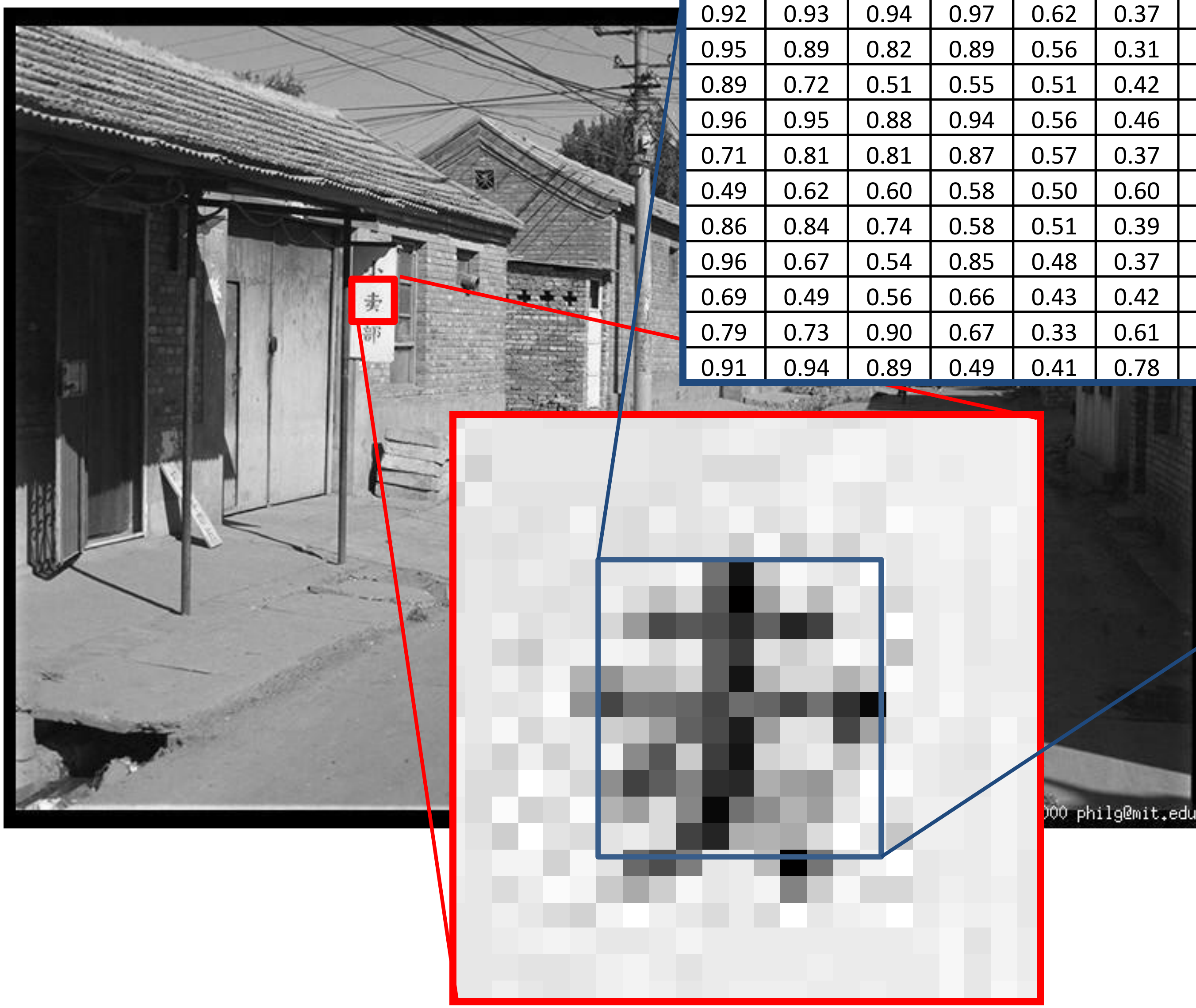
Only intensity shown – constant color

Most information in intensity



Original image

Back to grayscale intensity



0.92	0.93	0.94	0.97	0.62	0.37	0.85	0.97	0.93	0.92	0.99
0.95	0.89	0.82	0.89	0.56	0.31	0.75	0.92	0.81	0.95	0.91
0.89	0.72	0.51	0.55	0.51	0.42	0.57	0.41	0.49	0.91	0.92
0.96	0.95	0.88	0.94	0.56	0.46	0.91	0.87	0.90	0.97	0.95
0.71	0.81	0.81	0.87	0.57	0.37	0.80	0.88	0.89	0.79	0.85
0.49	0.62	0.60	0.58	0.50	0.60	0.58	0.50	0.61	0.45	0.33
0.86	0.84	0.74	0.58	0.51	0.39	0.73	0.92	0.91	0.49	0.74
0.96	0.67	0.54	0.85	0.48	0.37	0.88	0.90	0.94	0.82	0.93
0.69	0.49	0.56	0.66	0.43	0.42	0.77	0.73	0.71	0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

Next week

- Convolution, Filtering, Image Pyramids, Frequencies, Project 1