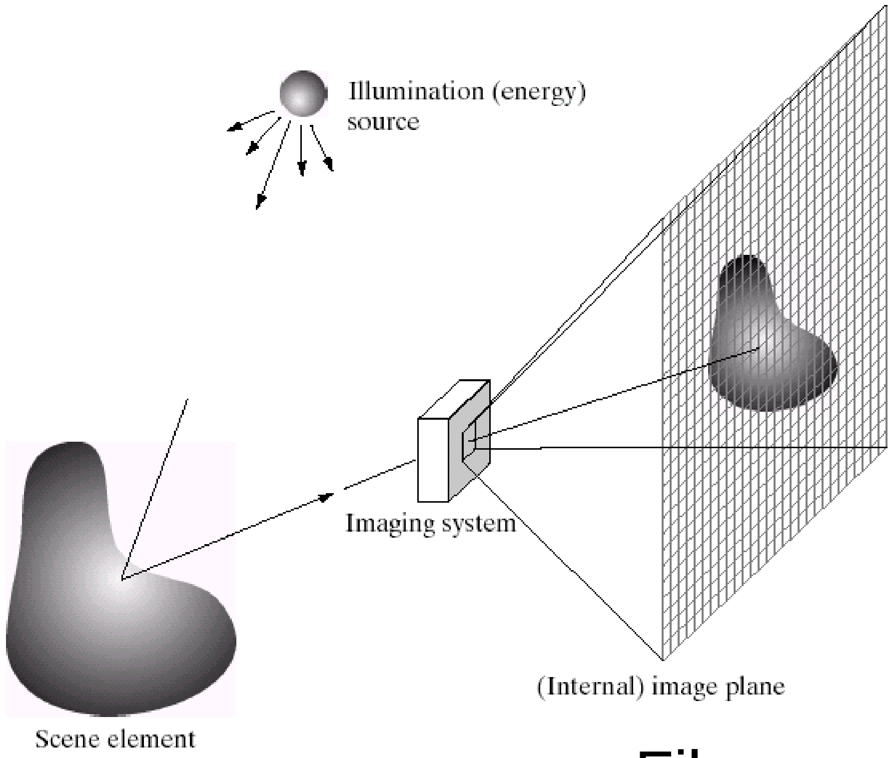
19CSE435: Computer Vision

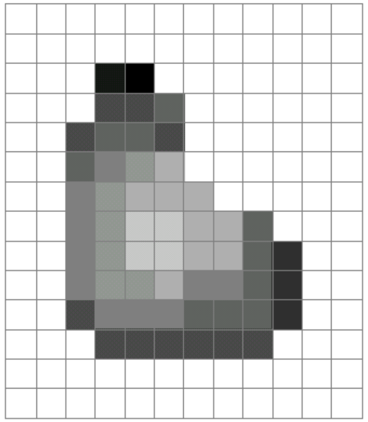
Image formation: Photometry

Adopted from Computer Vision Textbook and course materials R_Szeliski

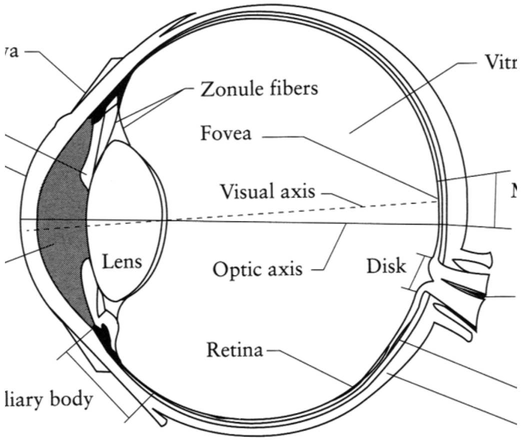
Image Formation



Film



Digital Camera

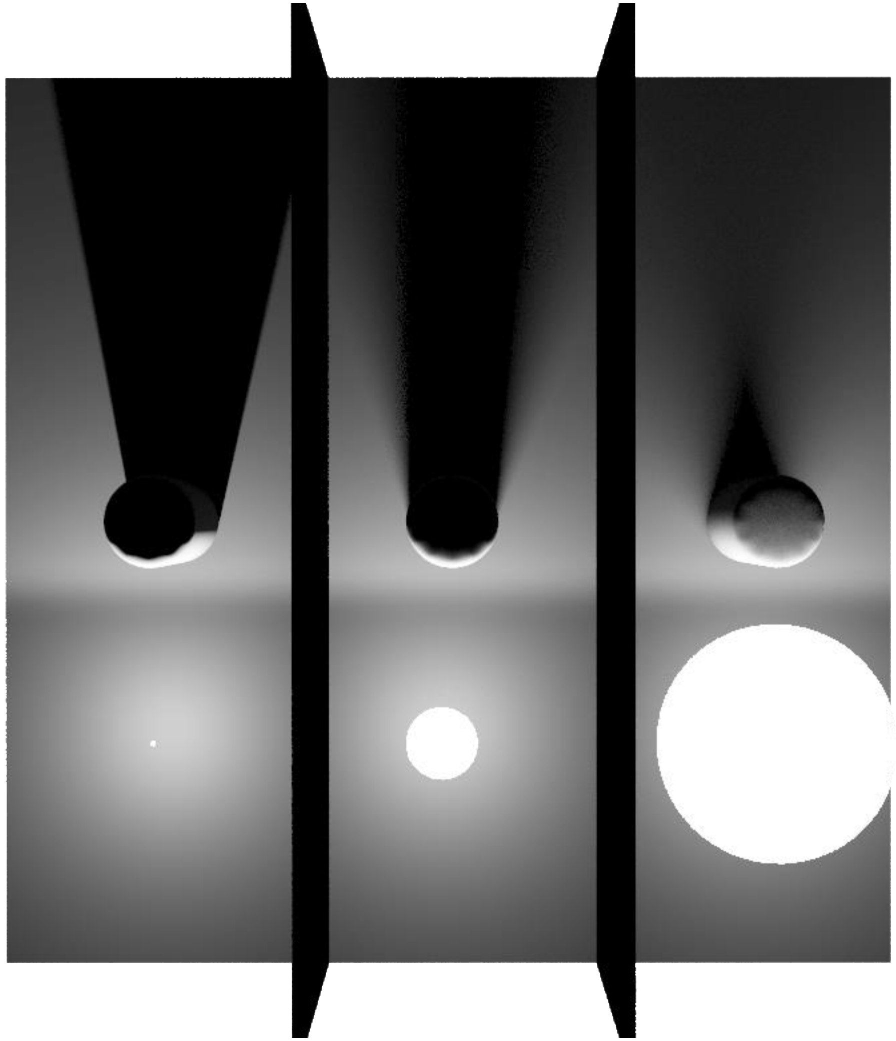


The Eye

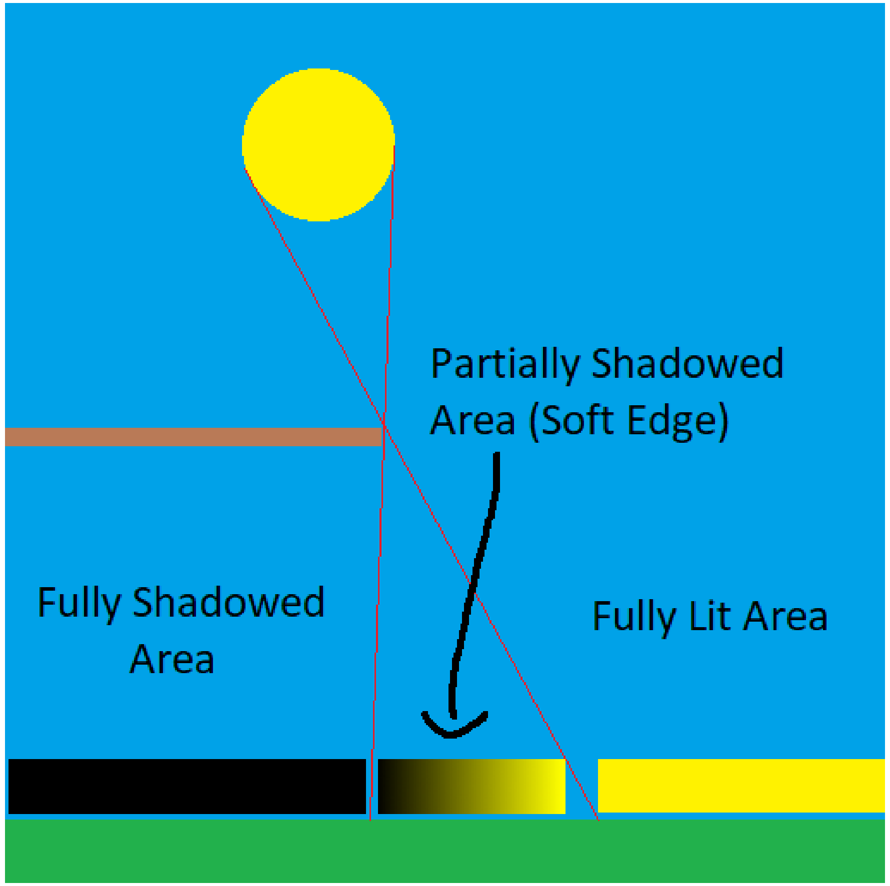
2.2.1 Lighting

- Point light source
- Area light source
- Environment map

Point and area light sources



sun & moon both subtend half a degree



https://blog.demofox.org/2017/07/01/why-are-some-shadows-soft-and-other-shadows-hard/

https://cg-masters.com/nicks-rants-and-raves/contact-shadows-cast-shadows-myth/

Environment map

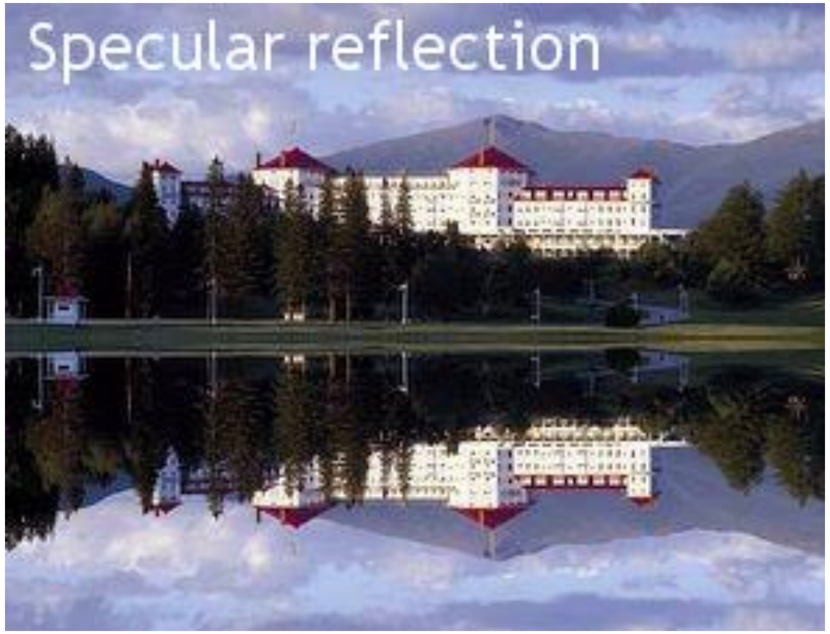
$$L(\hat{\boldsymbol{v}}; \lambda),$$

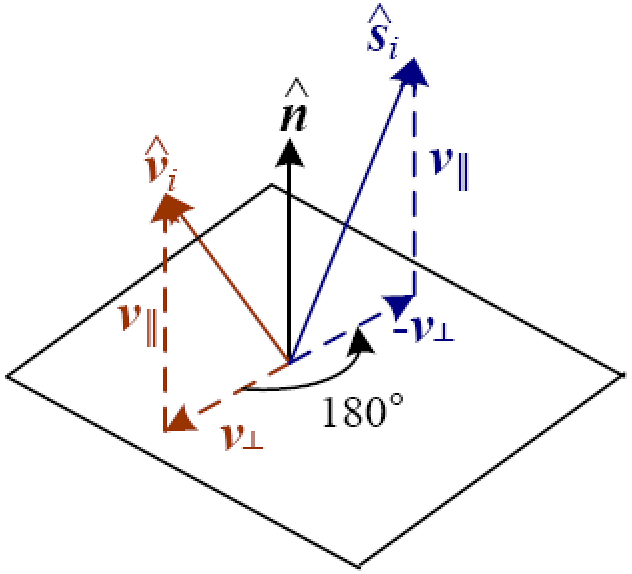
http://www.sparse.org/3d.html

2.2.2 Reflectance and Shading

- Specular reflection
- Diffuse reflection
- Oren-Nayar
- Phong shading
    - ambient illumination
    - Phong formula
- BRDF
- Isotropic vs anisotropic
- Global illumination
- Photon’s life choices

Specular reflection



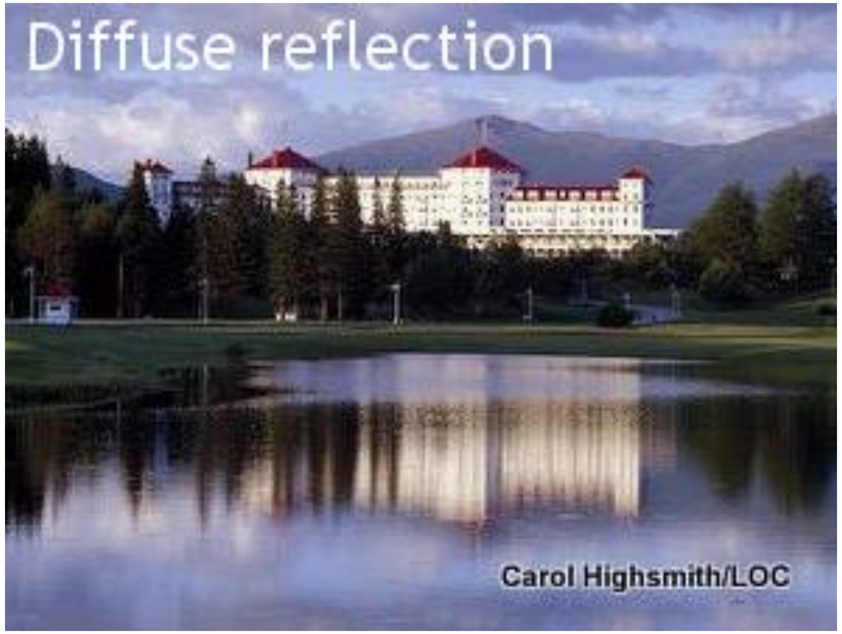


Specular reflection direction for light source i = deterministic function of incoming light direction v_i and normal n:

$$\hat{\boldsymbol{s}}_i = \boldsymbol{v}_\parallel - \boldsymbol{v}_\perp = (2\hat{\boldsymbol{n}}\hat{\boldsymbol{n}}^T - \boldsymbol{I})\boldsymbol{v}_i.$$

Photo Carol Highsmith/LOC

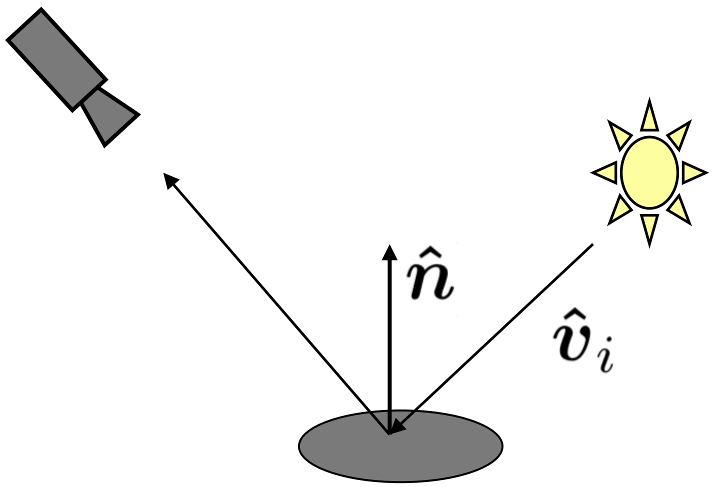
Diffuse Reflection



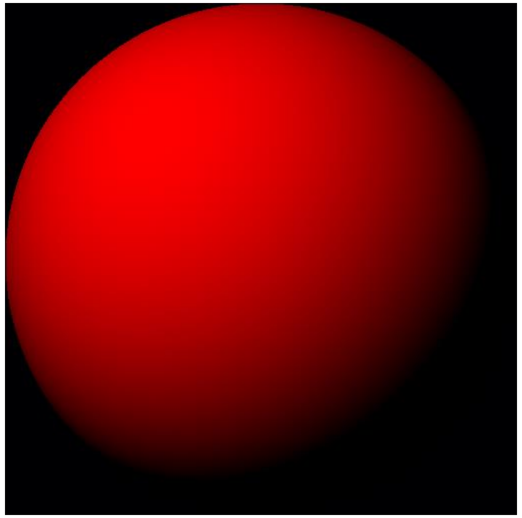
- Lambertian
- Oren-Nayar
- Fully general: BRDF

Trevor Darrell slide

Lambertian Reflectance Model



Surface normal n
Direction of illumination s



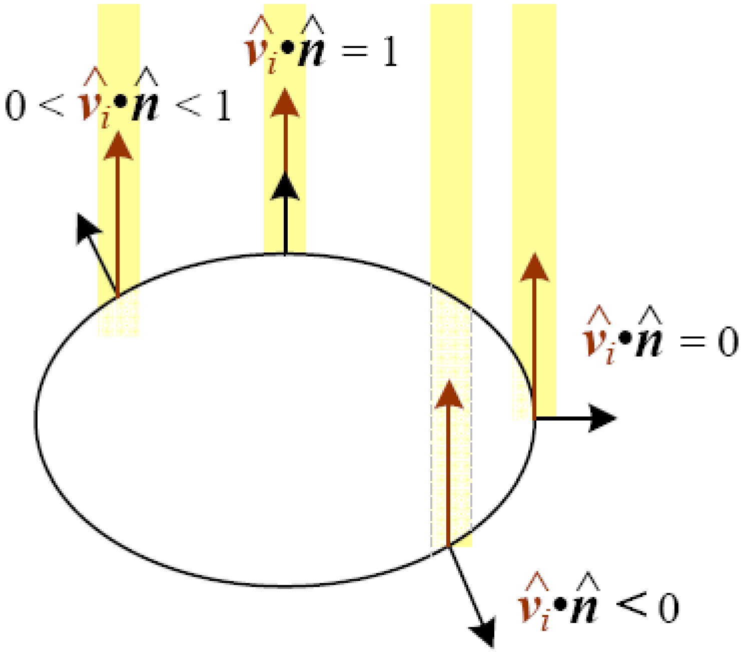
A Lambertian sphere

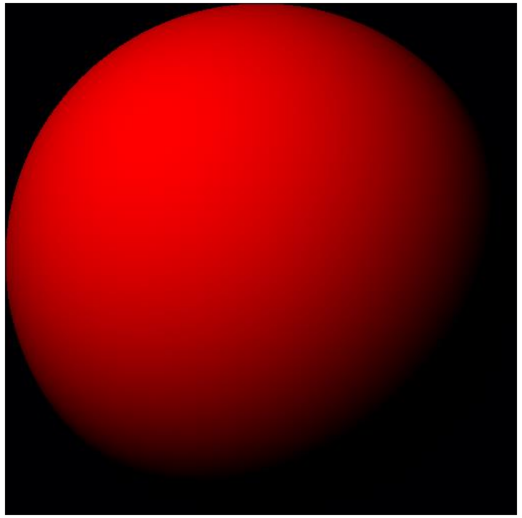
$$k_d(\lambda) \sum_i L_i(\lambda) [\hat{v}_i \cdot \hat{n}]^+$$

Commonly used in computer vision and computer graphics

Based on slide by Ioannis Stamos

Foreshortening





The diminution of returned light caused by foreshortening *depends on* $\hat{\boldsymbol{v}}_i \cdot \hat{\boldsymbol{n}}$, *the cosine of the angle between the incident light direction* $\hat{\boldsymbol{v}}_i$ *and the surface normal* $\hat{\boldsymbol{n}}$.

Trevor Darrell slide

Confusion around Lambert's cosine law

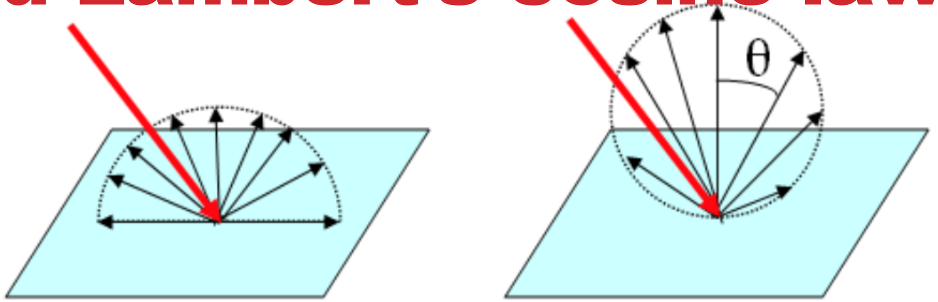


Figure: 1. The mental images corresponding to the two descriptions of Lambertian surfaces.

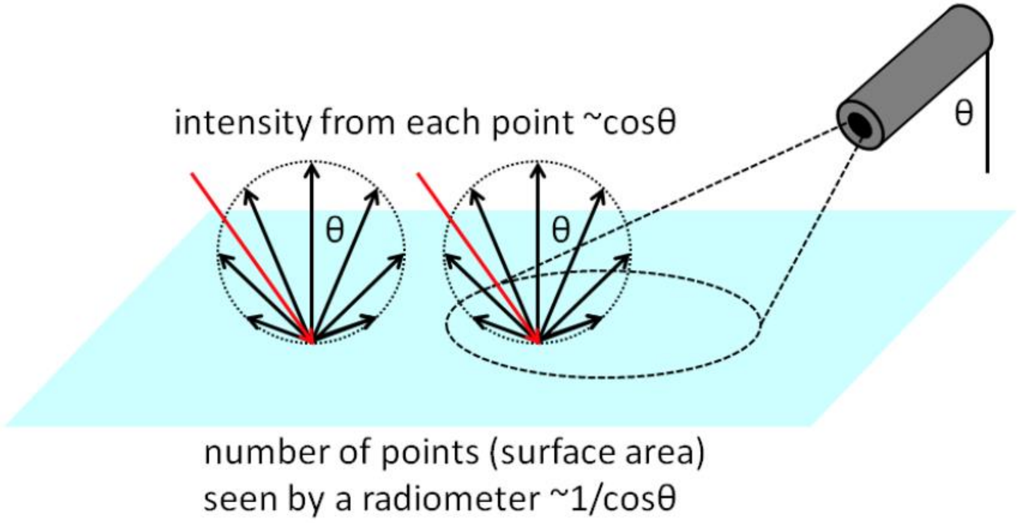
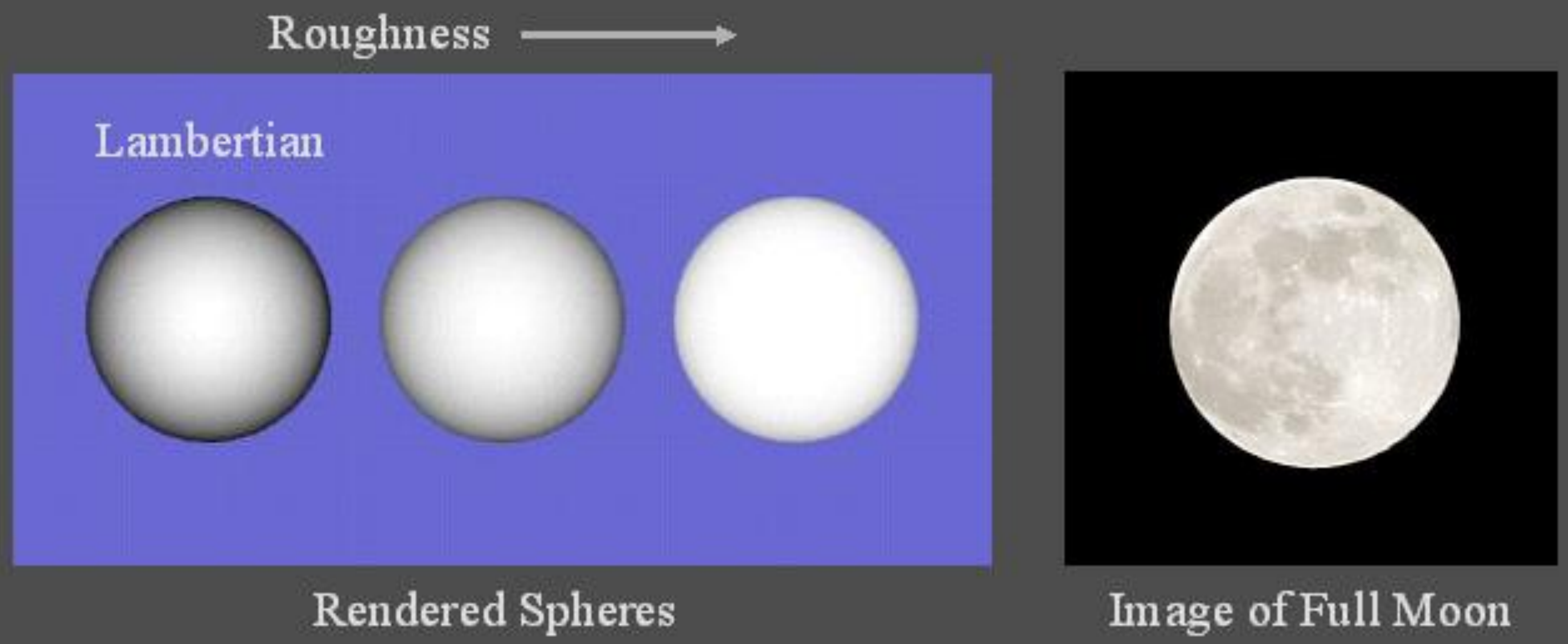


Figure: 2. Resolution of the paradoxical statements about how Lambertian surfaces reflect light. png

http://www.oceanopticsbook.info/view/surfaces/lambertian_brdfs

n-Nayar Reflectance Model



Moon is a counter-example! Roughness of surface makes that more light is reflected back to the viewer than expected under a Lambertian model.

http://www1.cs.columbia.edu/CAVE/projects/oren/oren.php

Phong Reflectance Model

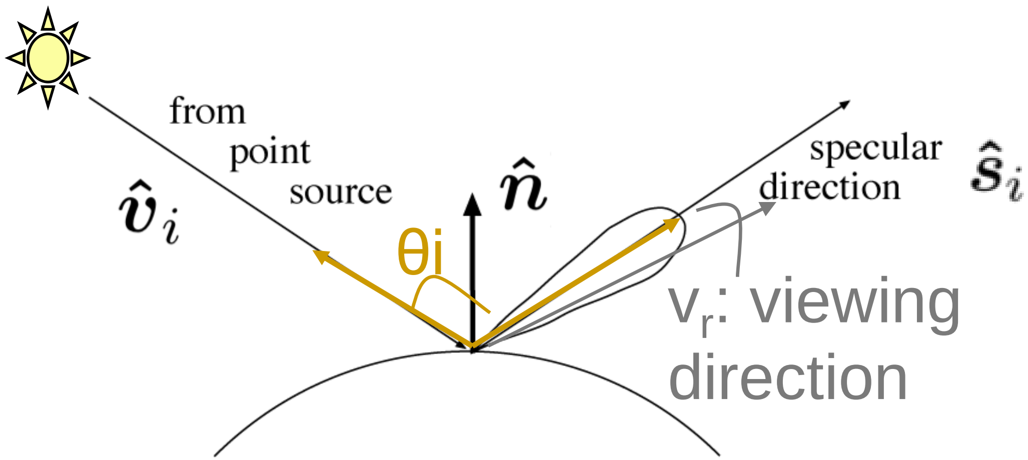
Combines diffuse (Lambertian) and specular lobe

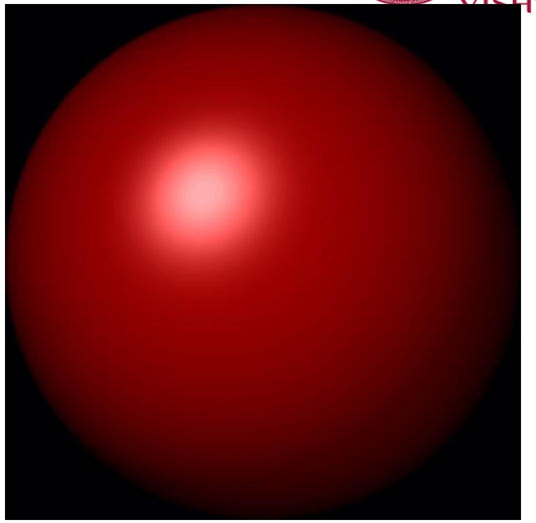
$$k_a(\lambda)L_a(\lambda) +$$ frequently with "ambient" component

$$k_d(\lambda)\sum_i L_i(\lambda)[\hat{\boldsymbol{v}}_i \cdot \hat{\boldsymbol{n}}]^+ + k_s(\lambda)\sum_i L_i(\lambda)(\hat{\boldsymbol{v}}_r \cdot \hat{\boldsymbol{s}}_i)^{k_e}$$

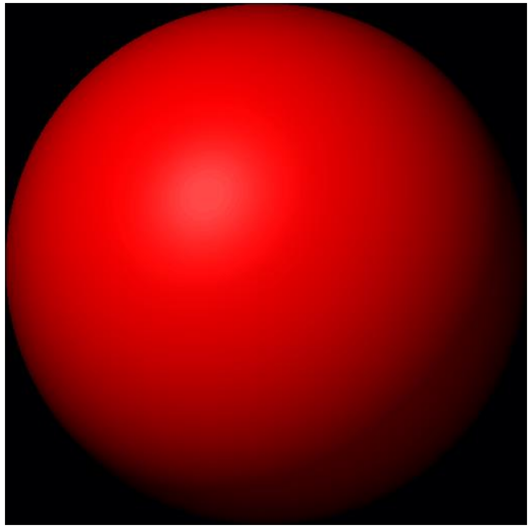
diffuse

specular





k_d=0.3, k_s=0.7, k_e=2



k_d=0.7, k_s=0.3, k_e=0.5

Based on slide by Ioannis Stamos

BDRF

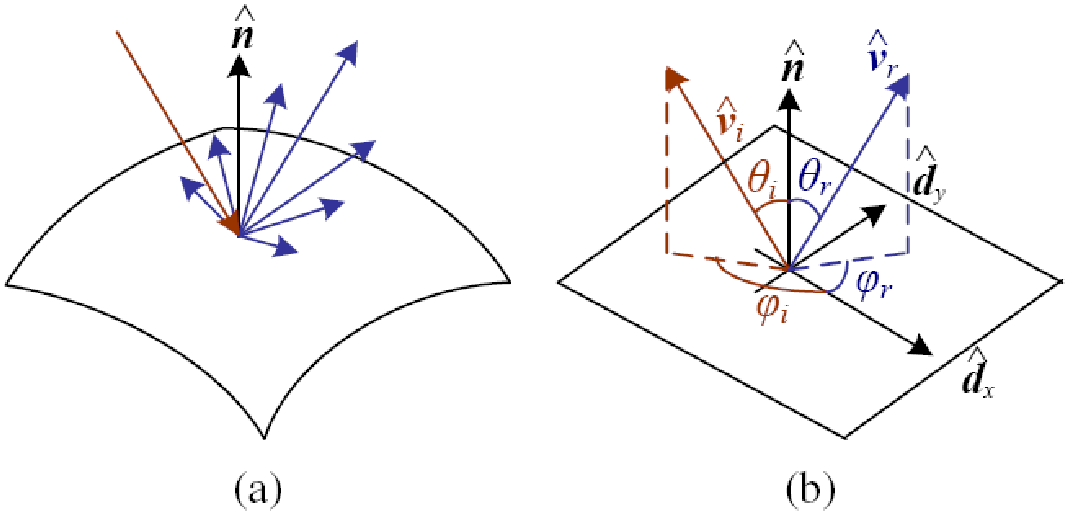


Figure 2.15: *(a) Light scattering when hitting a surface. (b) The bidirectional reflectance distribution function (BRDF) $f(\theta_i, \phi_i, \theta_r, \phi_r)$ is parameterized by the angles the incident $\hat{\boldsymbol{v}}_i$ and reflected $\hat{\boldsymbol{v}}_r$ light ray directions make with the local surface coordinate frame $(\hat{\boldsymbol{d}}_x, \hat{\boldsymbol{d}}_y, \hat{\boldsymbol{n}})$.*

For an isotropic material, we can simplify the BRDF to

$$f_r(\theta_i, \theta_r, |\phi_r - \phi_i|; \lambda) \quad \text{or} \quad f_r(\hat{\boldsymbol{v}}_i, \hat{\boldsymbol{v}}_r, \hat{\boldsymbol{n}}; \lambda),$$

While light is scattered uniformly in all directions, i.e., the BRDF is constant,

$$f_d(\hat{\boldsymbol{v}}_i, \hat{\boldsymbol{v}}_r, \hat{\boldsymbol{n}}; \lambda) = f_d(\lambda),$$

Trevor Darrell slide

2.2.3 Optics

Pinhole size / aperture

How does the size of the aperture affect the image we'd get?

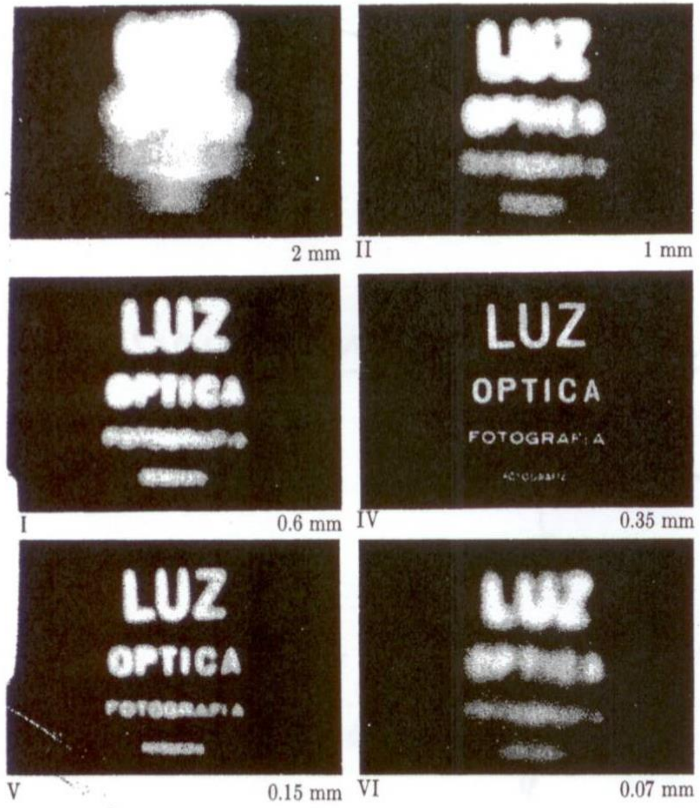
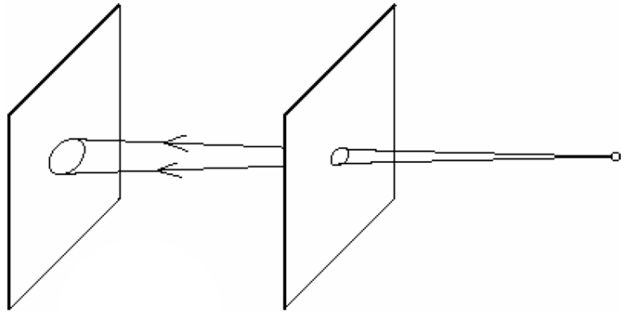


Fig. 5.96 The pinhole camera. Note the variation in image clarity as the hole diameter decreases. [Photos courtesy Dr. N. Joel, UNESCO.]

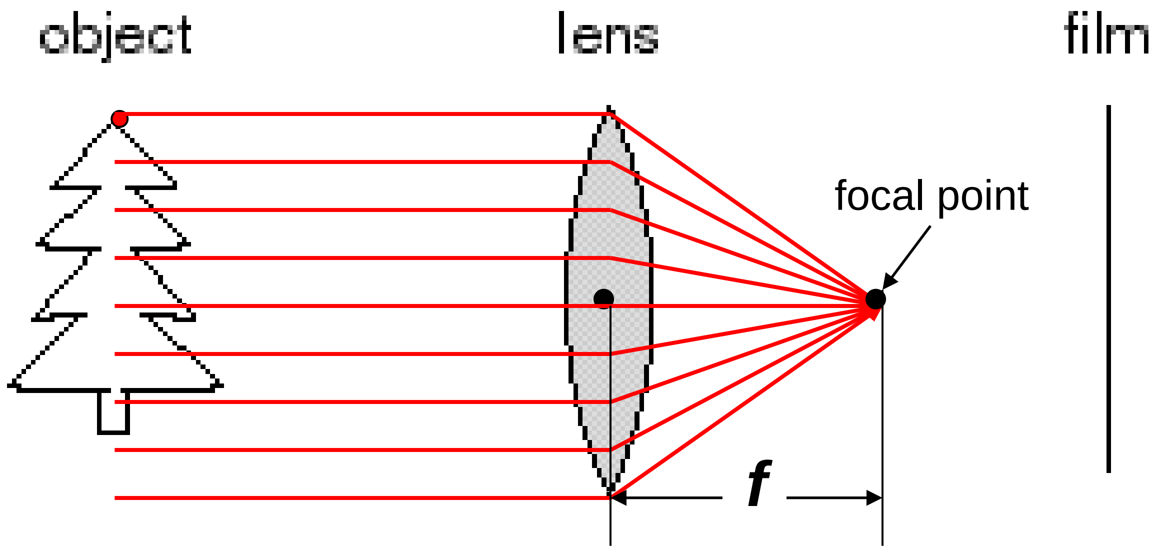
Larger

Smaller



K. Grauman

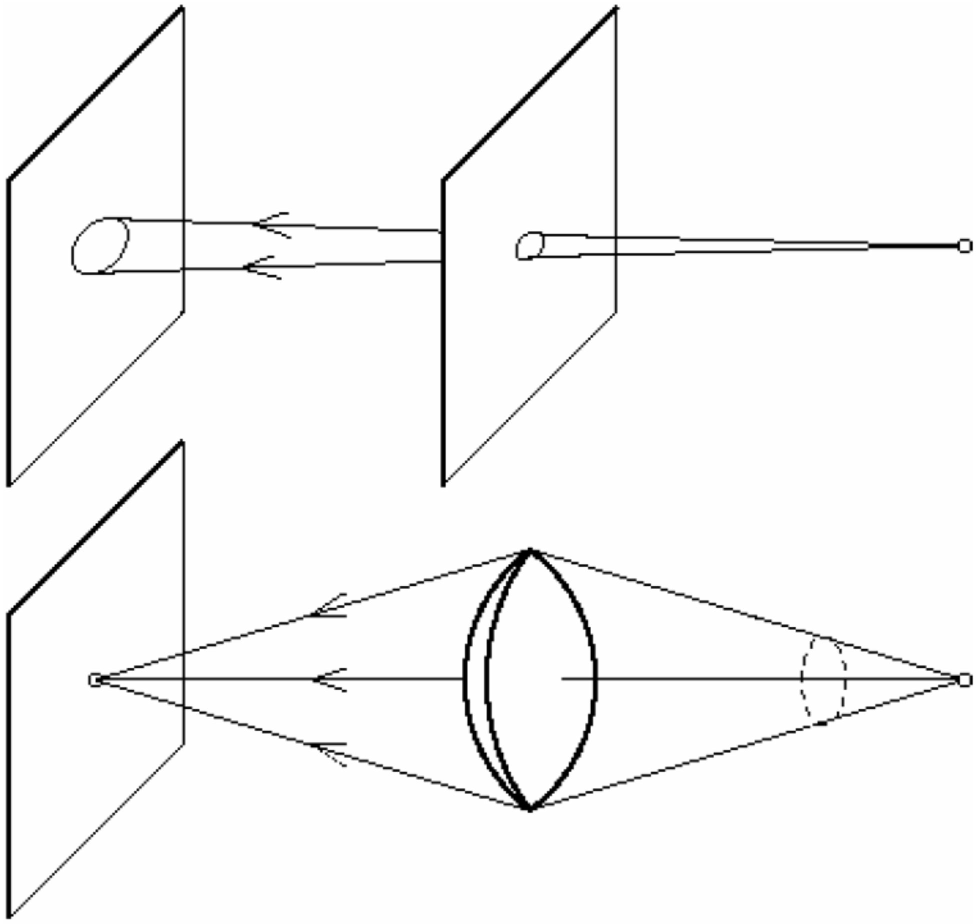
Adding a lens



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

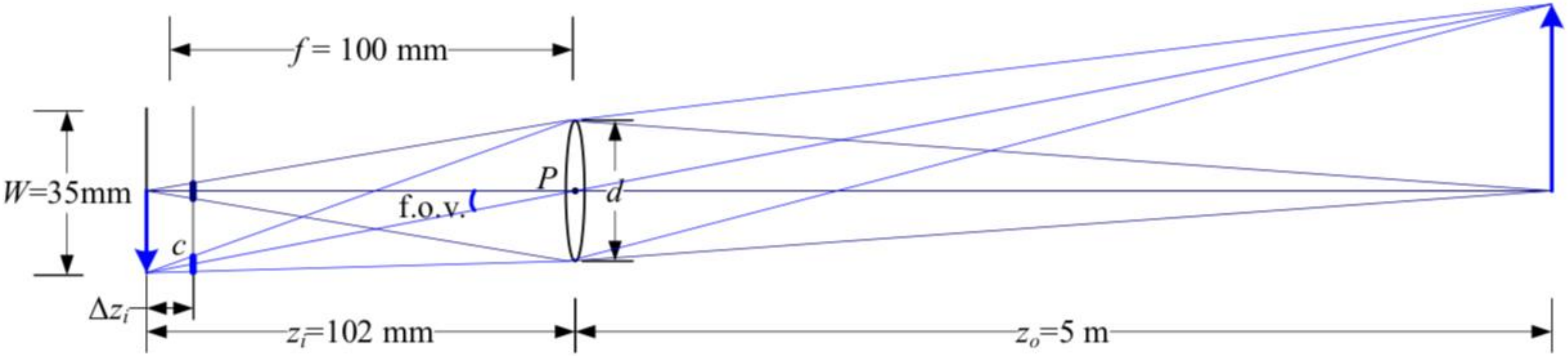
Slide by Steve Seitz

Pinhole vs. lens



K. Grauman

Thin lens model



$$\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}$$

- In a camera, we can adjust image plane to be at z_i
- If z_i == f, focus is at infinity
- If we increase zi > f, we bring the focal plane back from infinity.
  - E.g.: z_i == 102mm, f == 100mm => z_o == 5m

Image from Szeliski book

Focus and depth of field

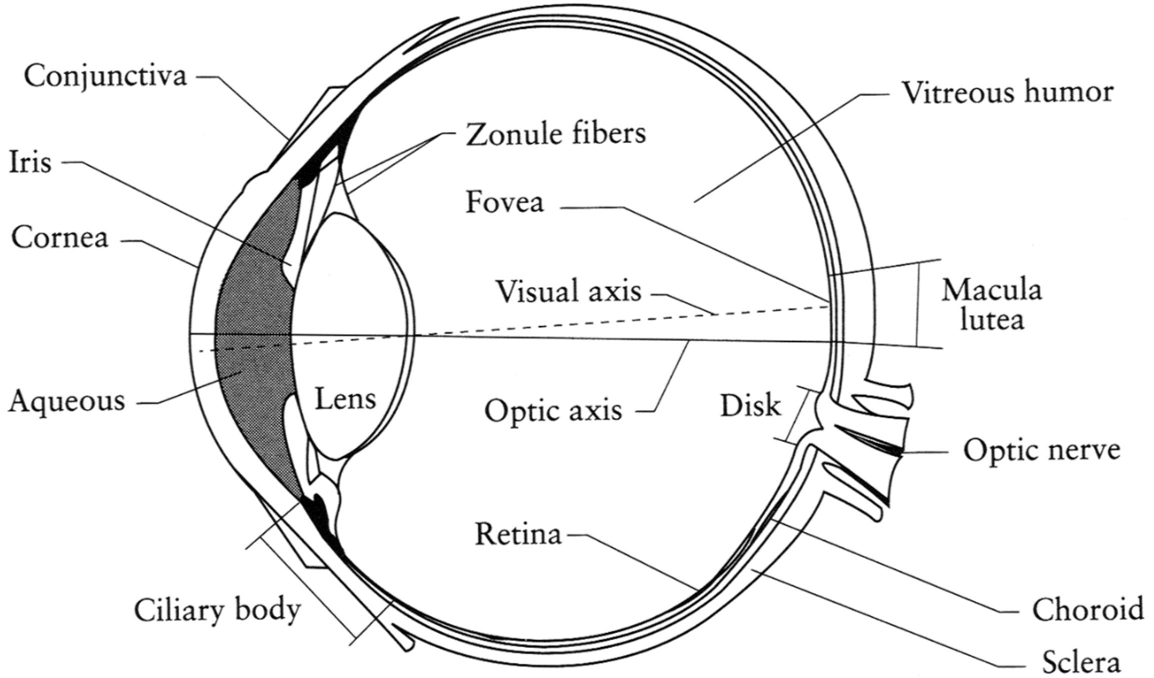
The smaller the aperture (area that lets light through), the more a lens behaves as a pinhole, the more everything is in focus.

Phones: fake "depth of field" with deep learning and stereo ☺

Image credit: cambridgeincolour.com

Image Formation

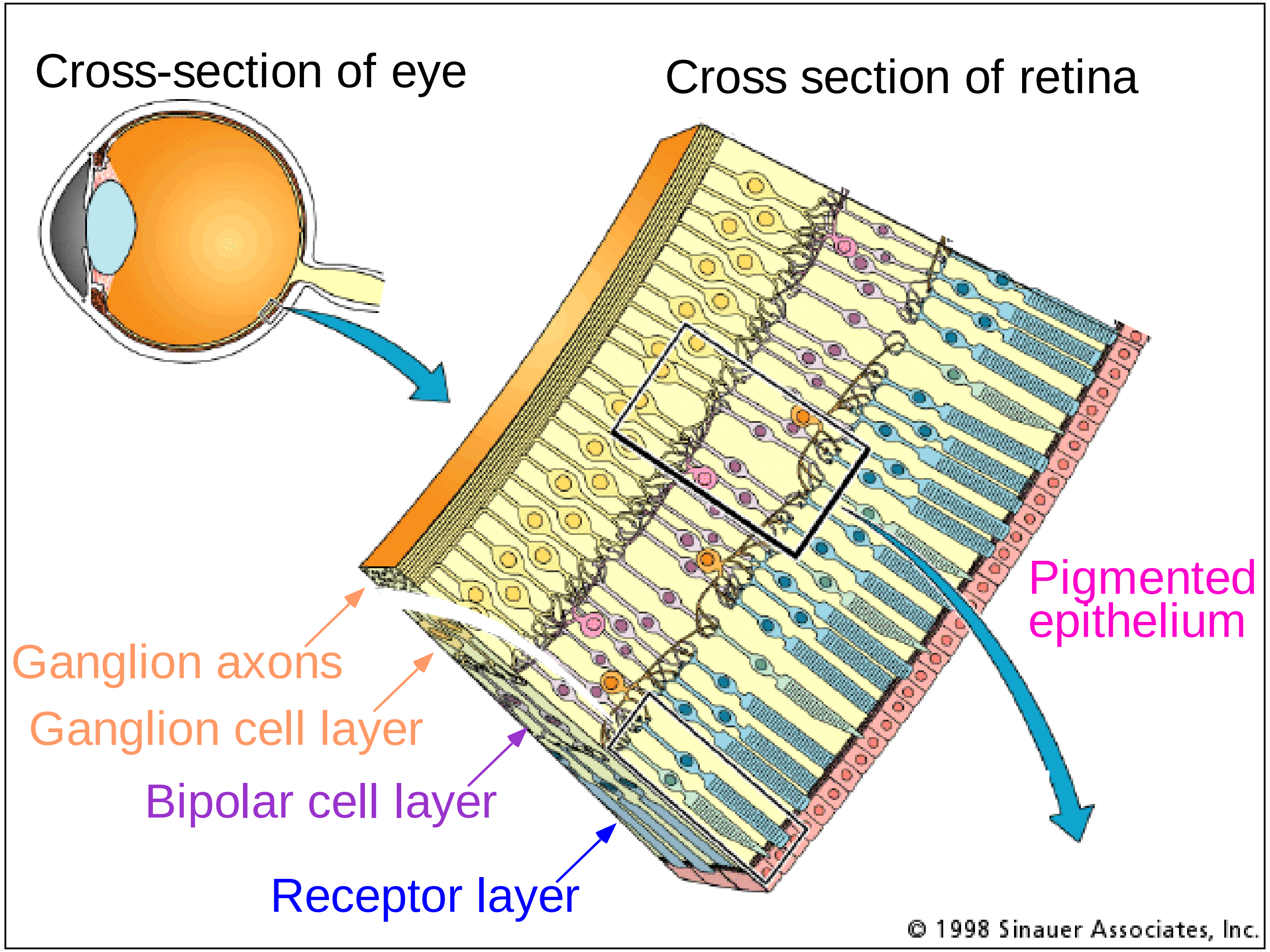
2.3.0 Human Vision (not in book)



- **The human eye is a pinhole 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

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.

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

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

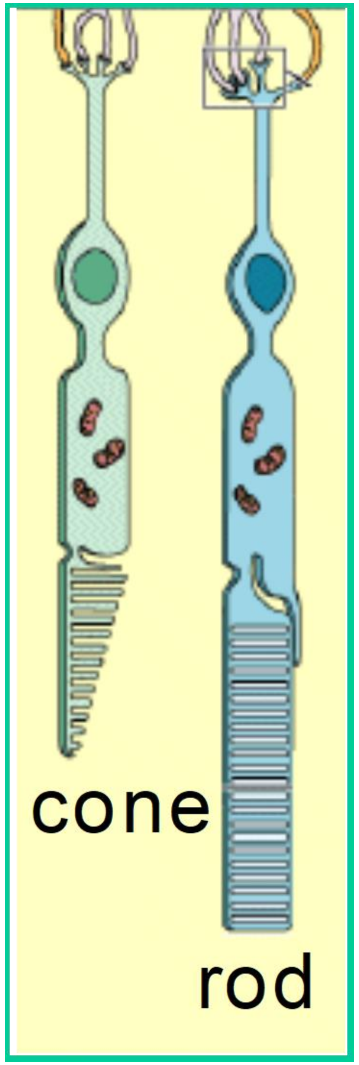
Two types of light-sensitive receptors

Cones

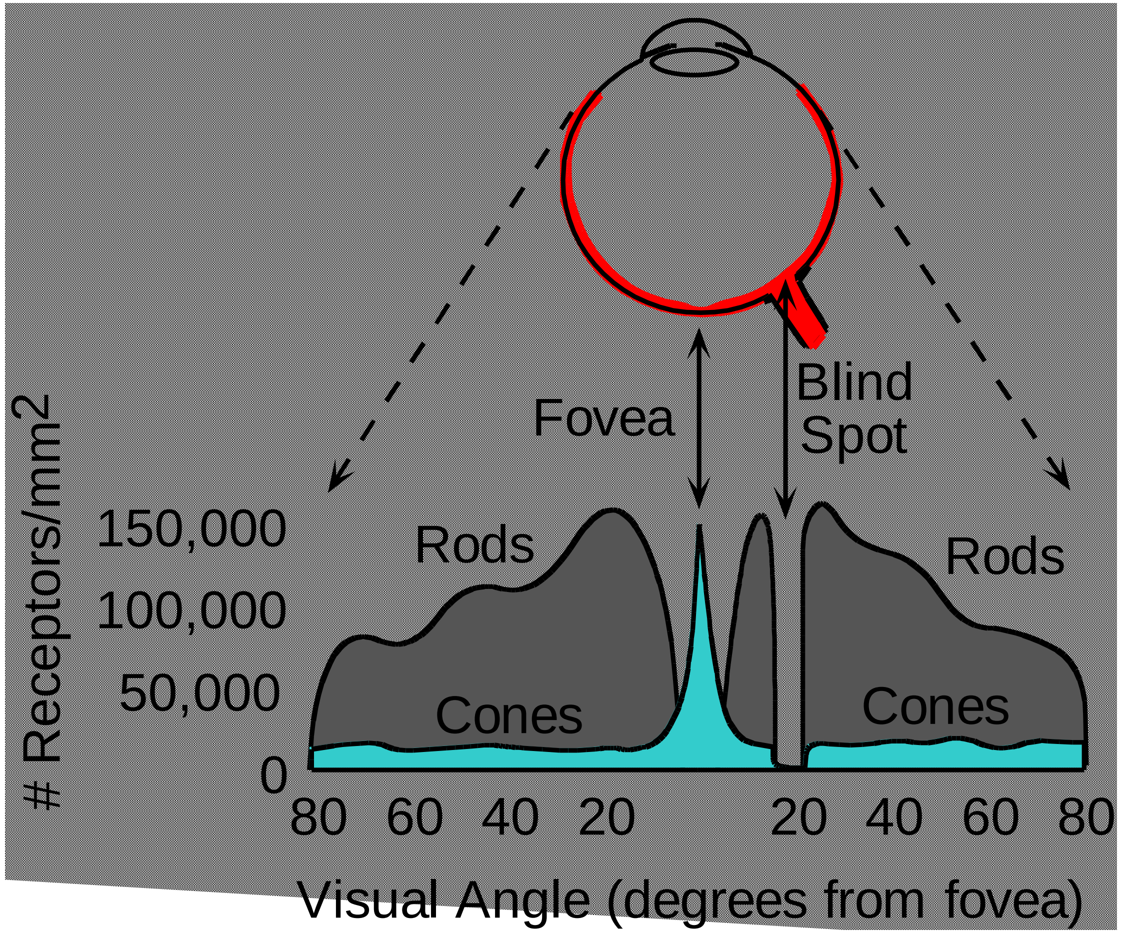
cone-shaped
less sensitive
operate in high light
color vision

Rods

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



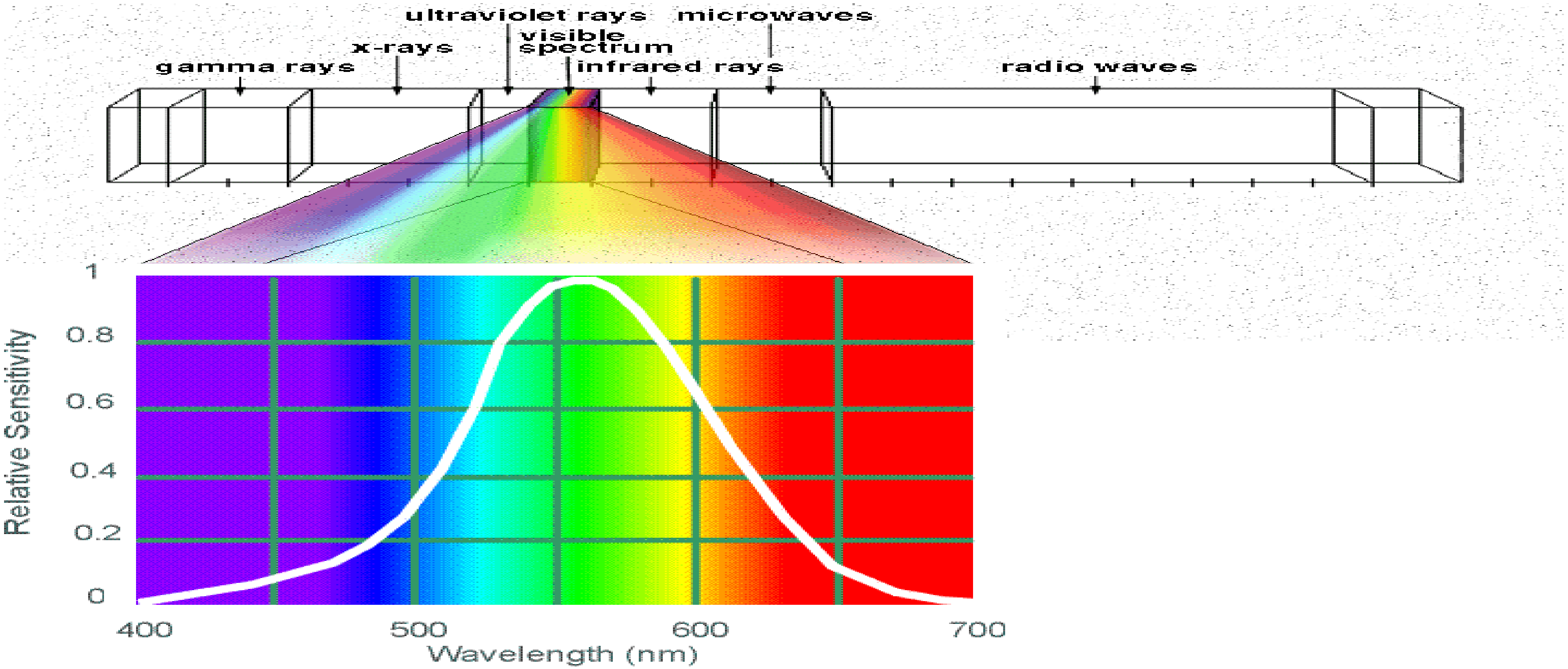
Distribution of Rods and Cones



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

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

Electromagnetic Spectrum

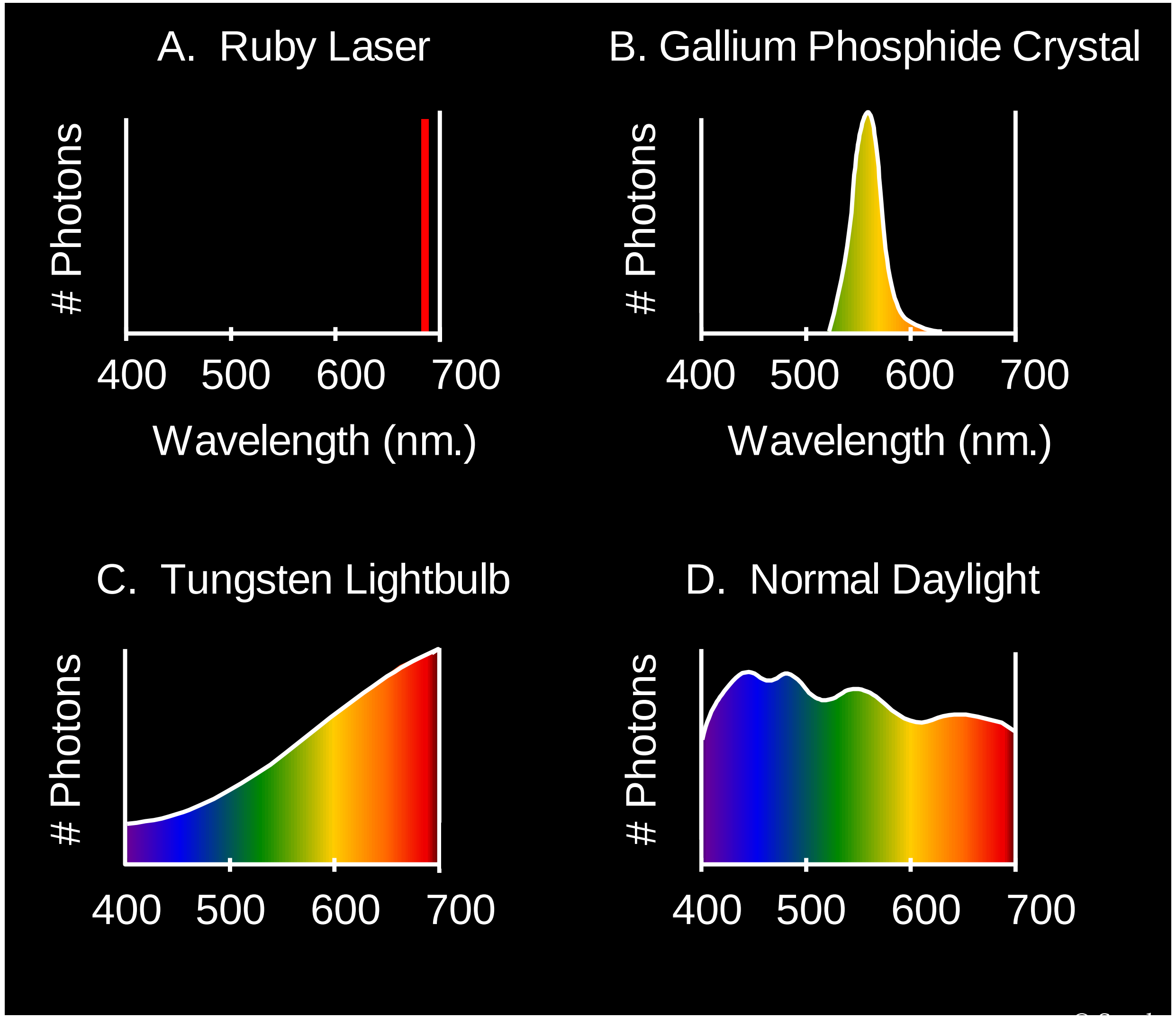


Human Luminance Sensitivity Function

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

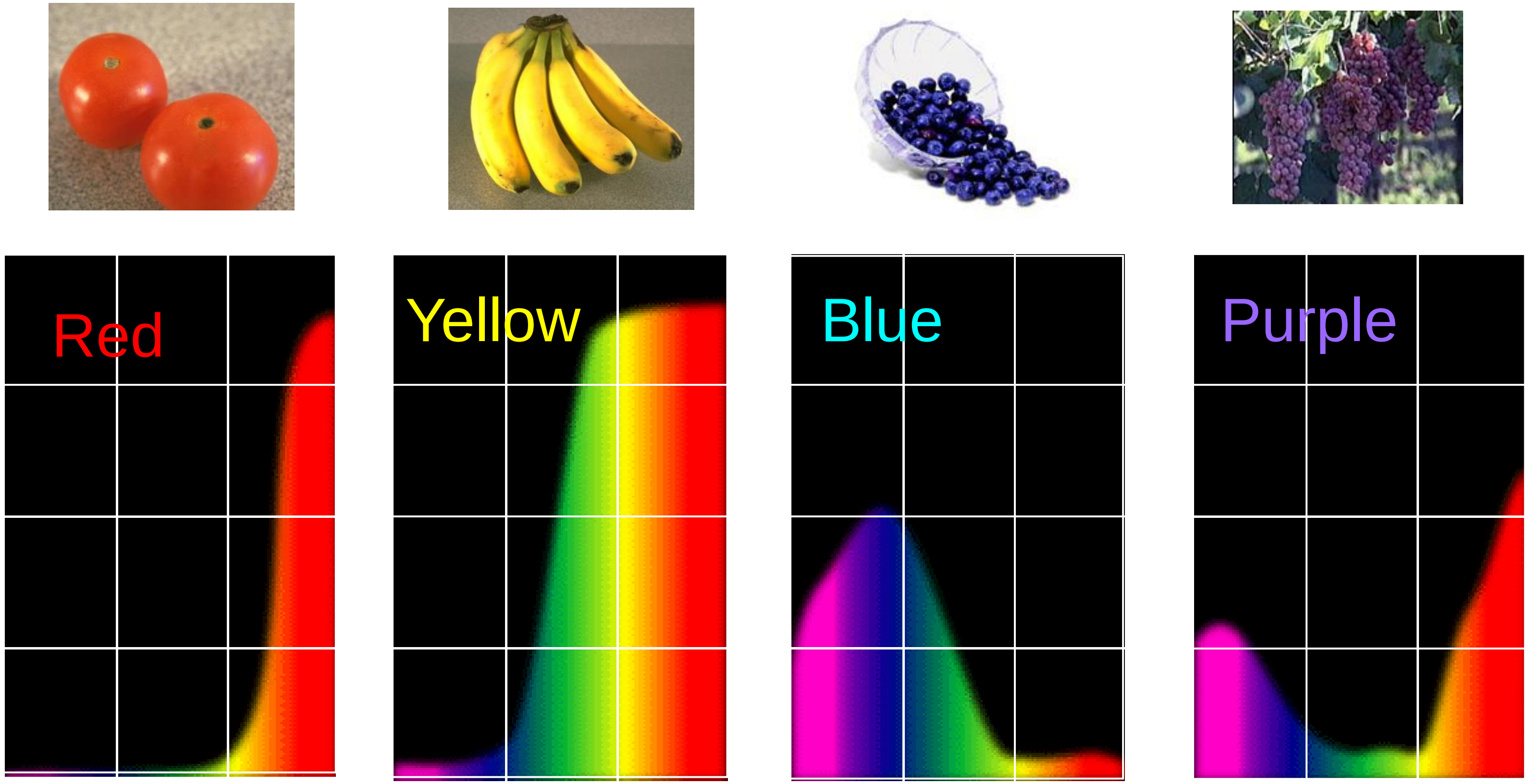
The Physics of Light

Some examples of the spectra of light sources



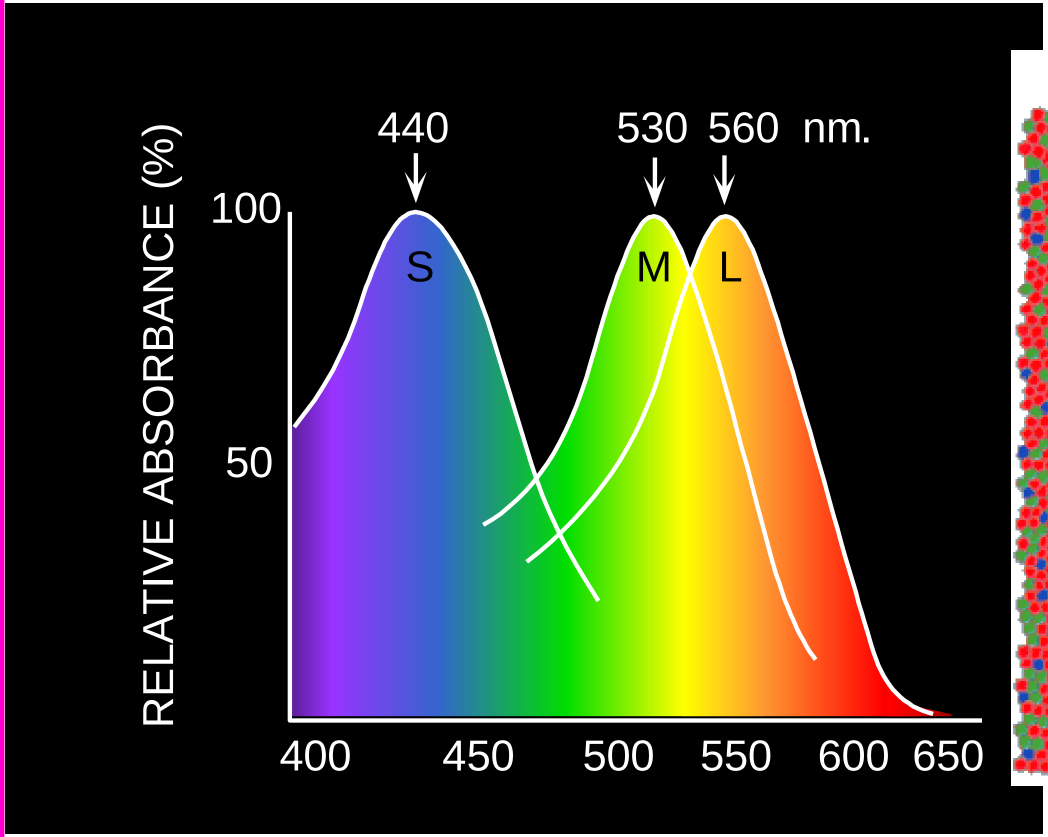
The Physics of Light

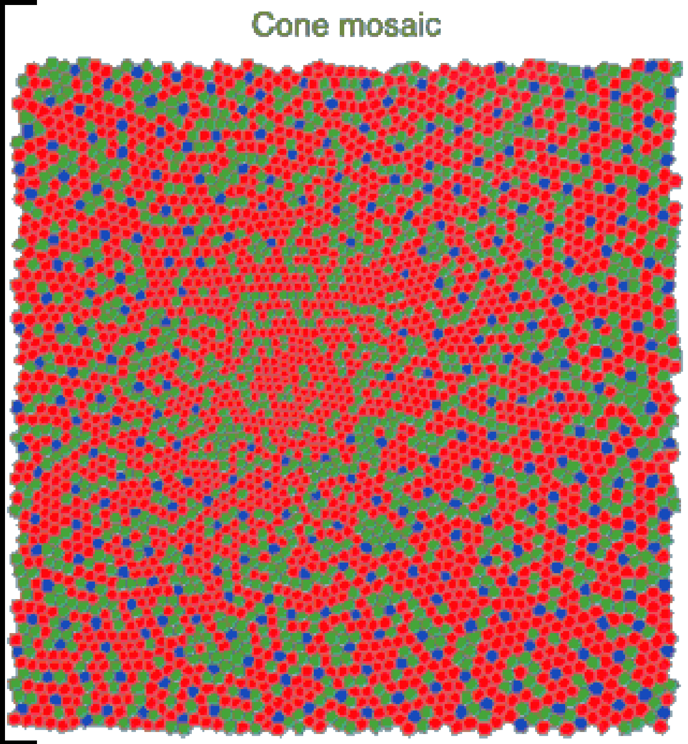
Some examples of the reflectance spectra of surfaces



Physiology of Color Vision

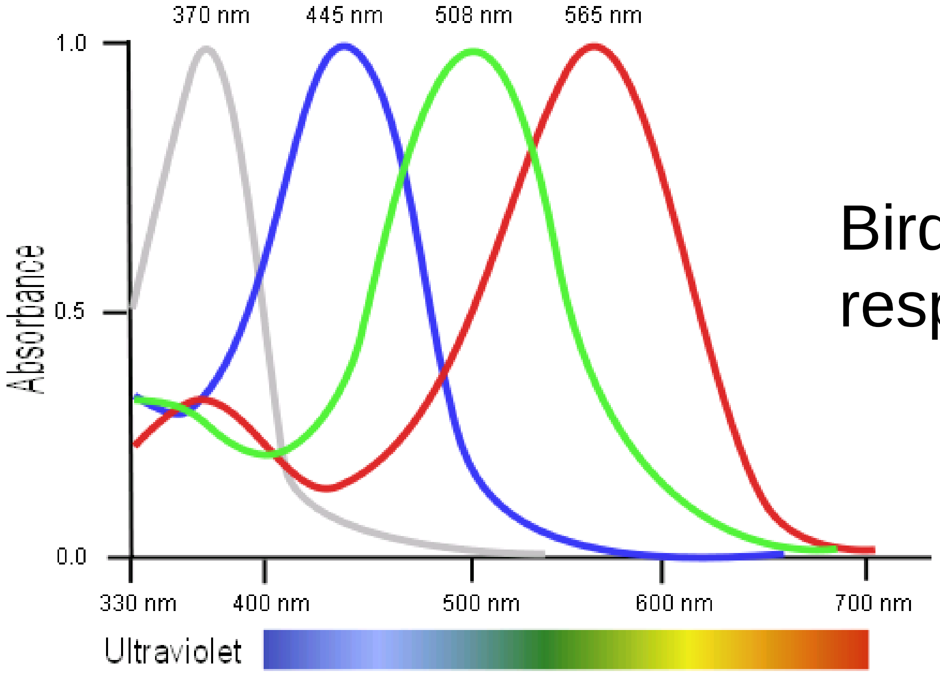
Three kinds of cones:





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

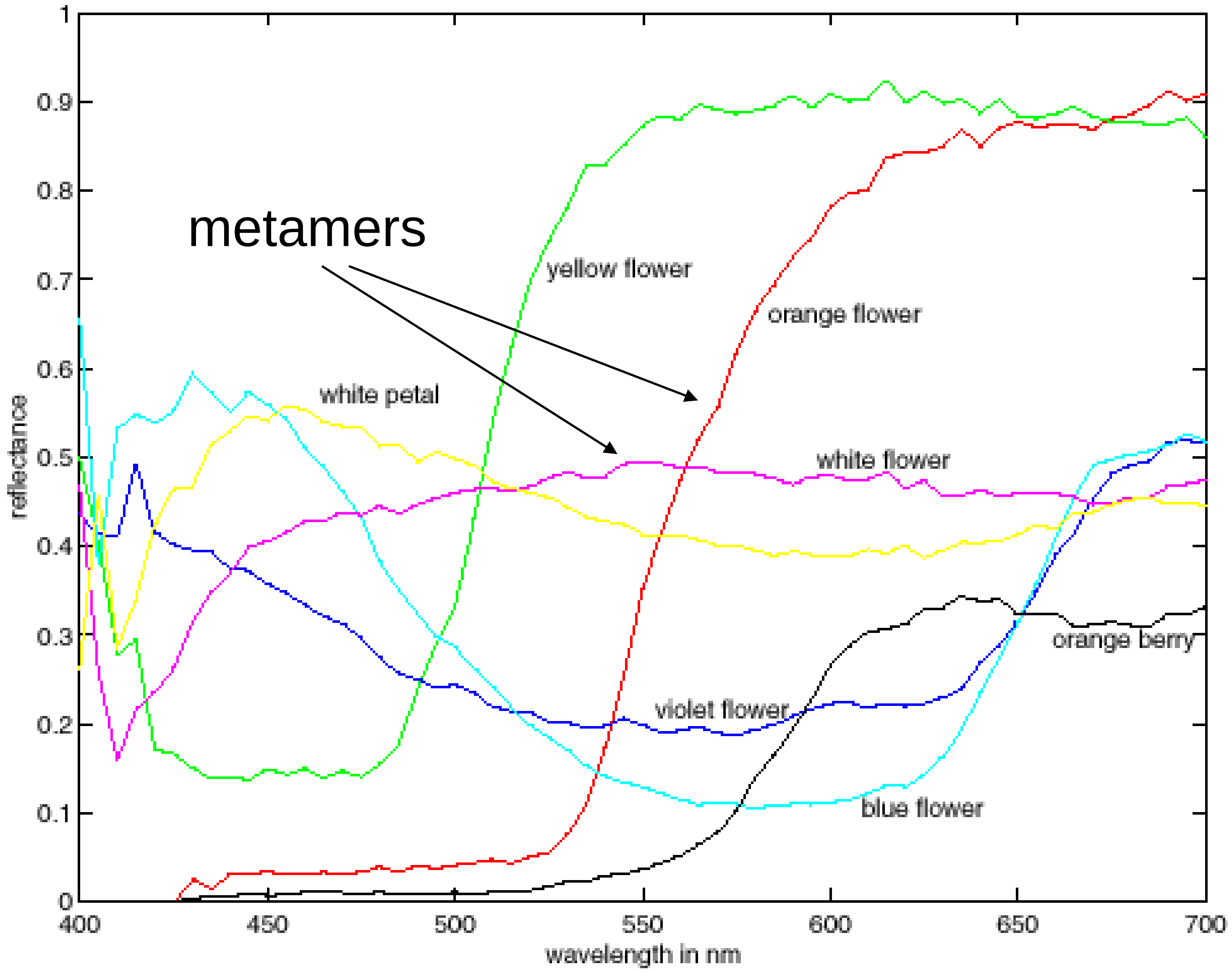
Tetrachromatism



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

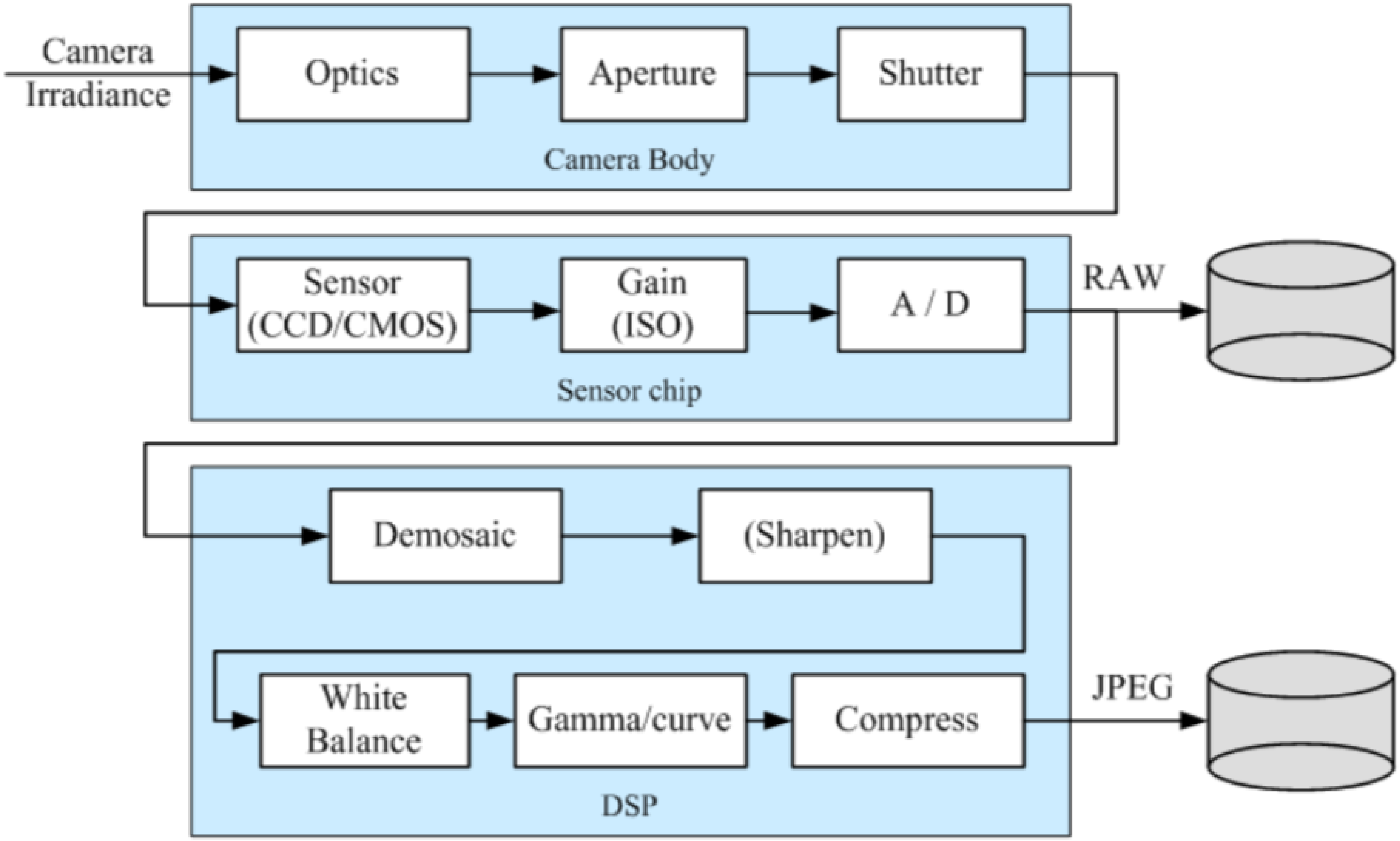


2.3 The Digital camera

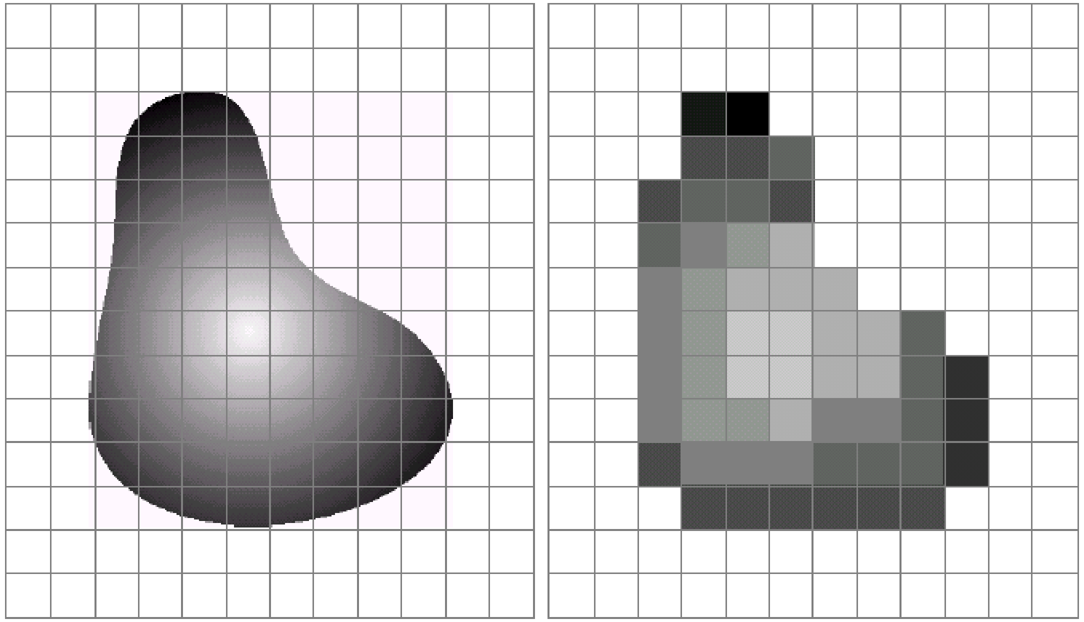
- **A digital camera replaces film with a sensor array**
  - Each cell in the array is light-sensitive diode that converts photons to electrons
  - Two common types:
    - Charge Coupled Device (CCD)
    - CMOS
  - http://electronics.howstuffworks.com/digital-camera.htm

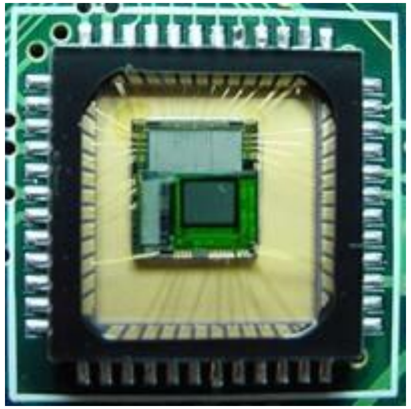
Slide by Steve Seitz

The sensing pipeline



Sensor Array





CMOS sensor

a b

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

Sampling and Quantization

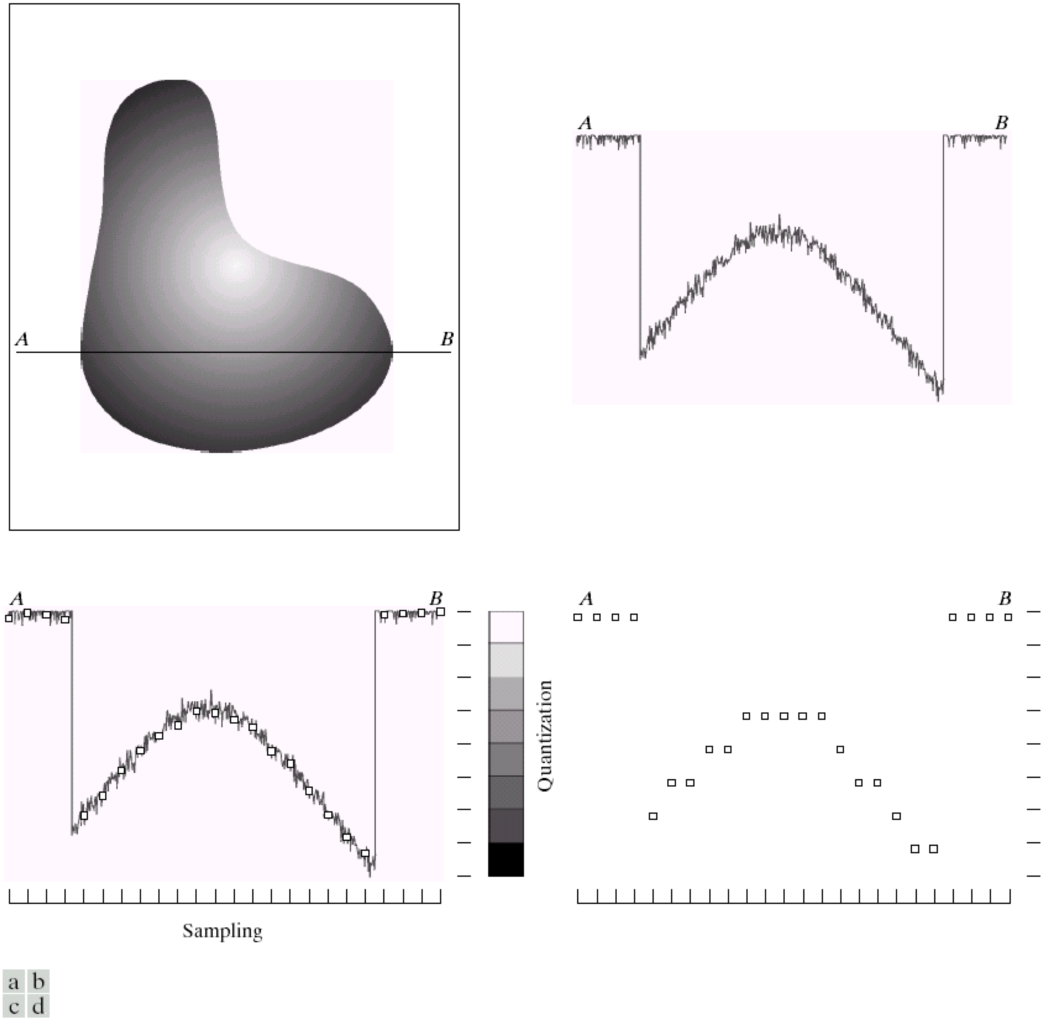
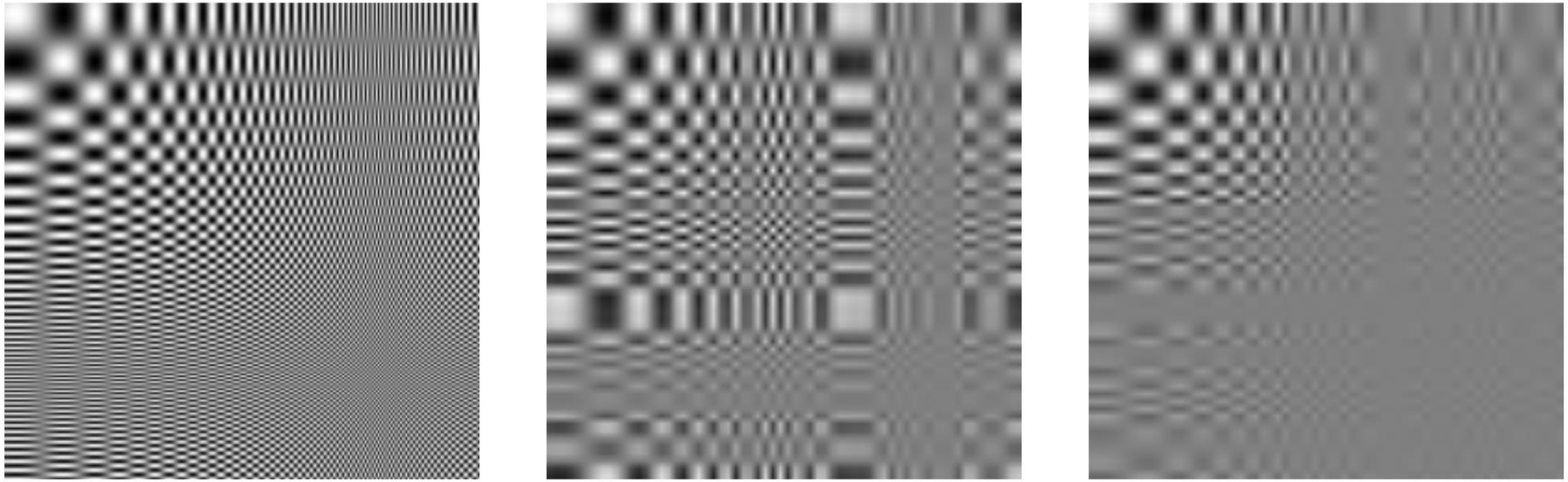
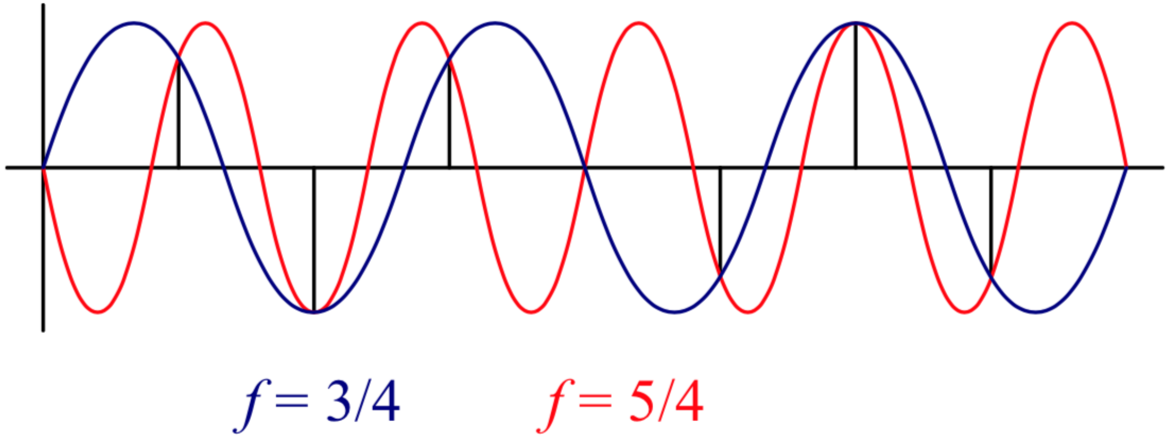


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

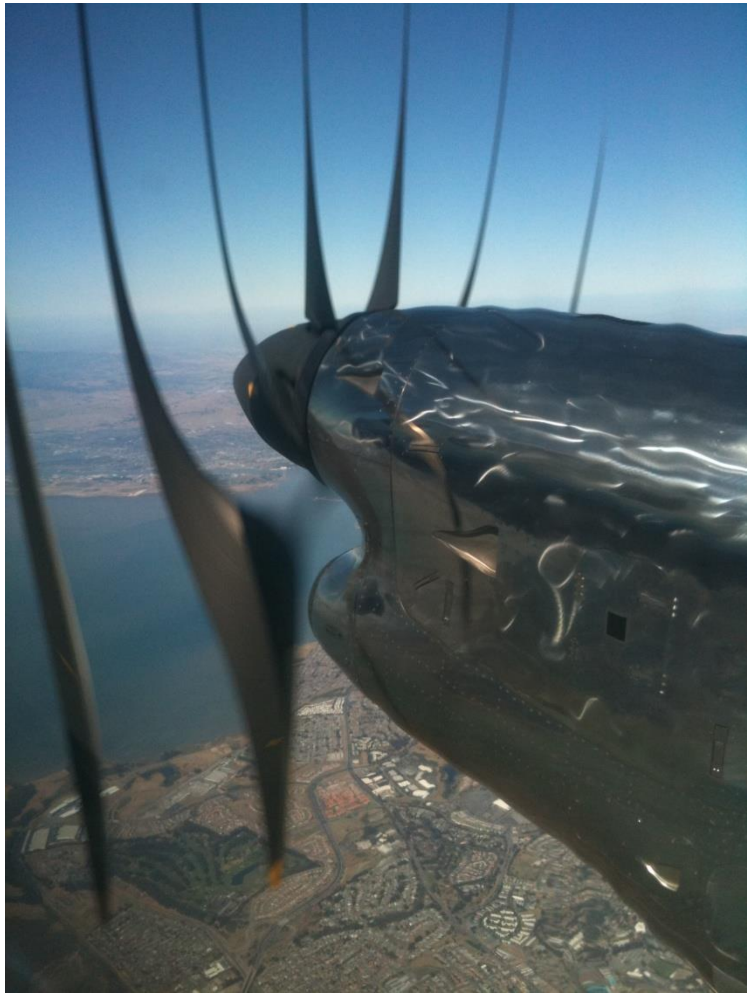
2.3.1 Sampling and Aliasing



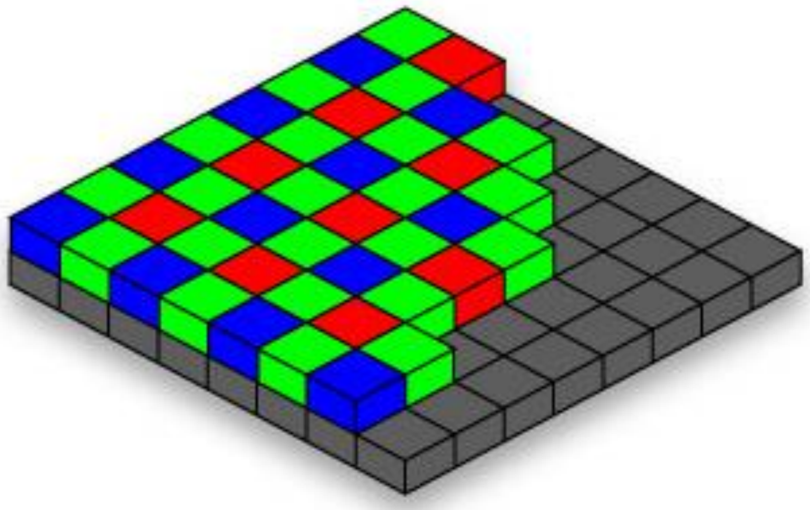
Violation of Shannon's sampling theorem: $f_s >= 2 f_{max}$

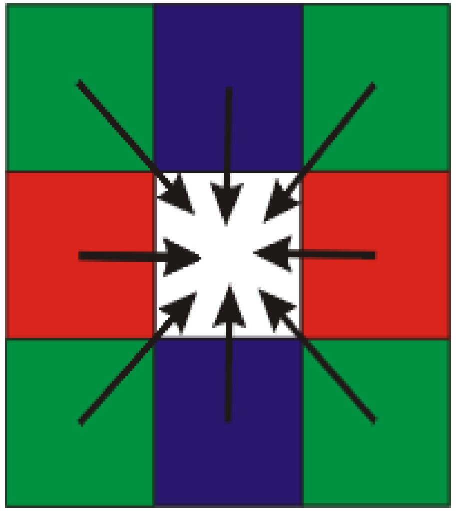


Rolling Shutter

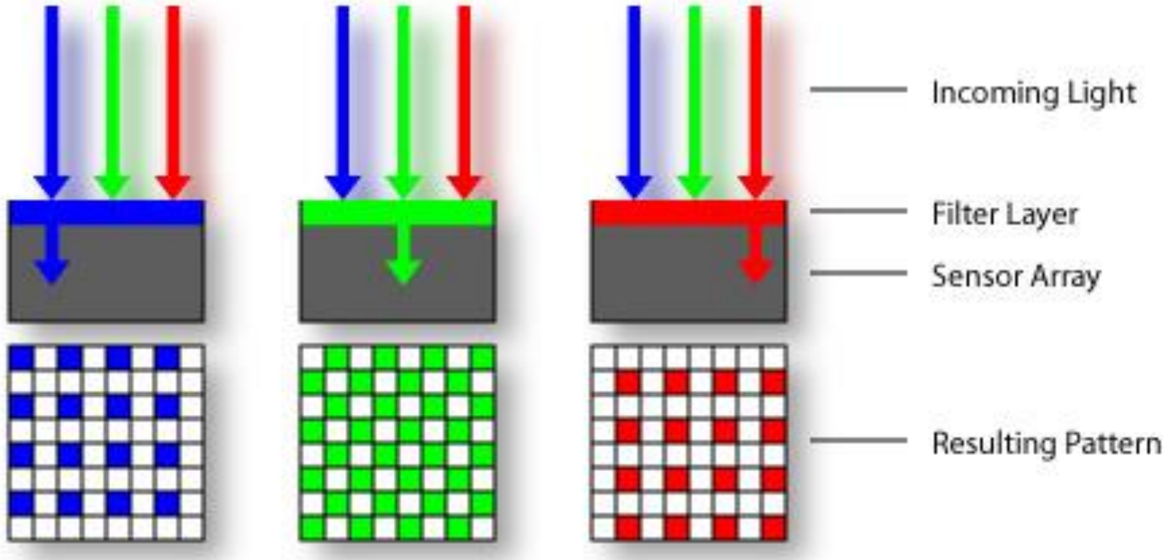


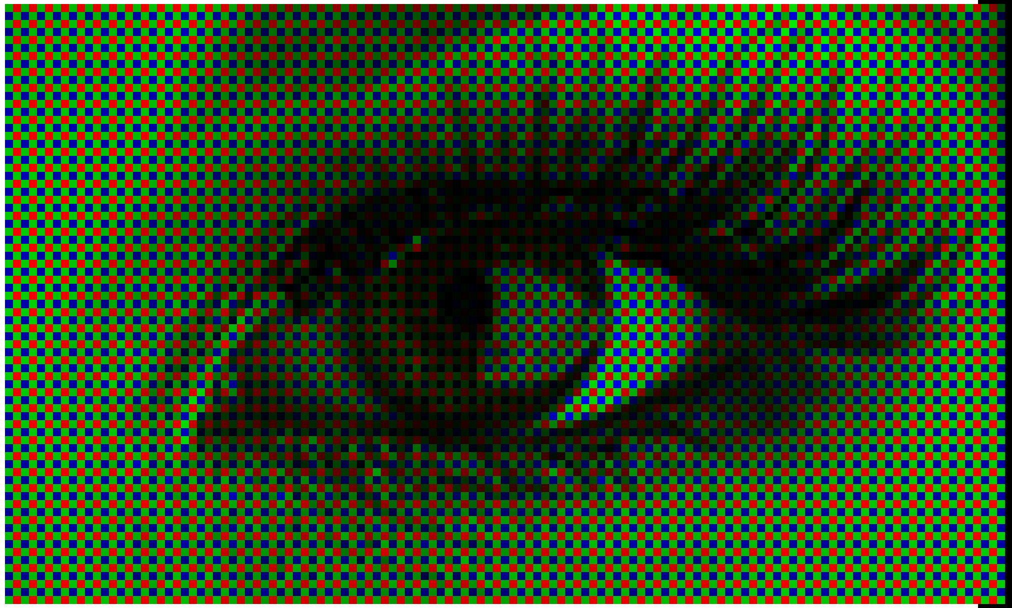
2.3.2 Color: the Bayer grid





Estimate RGB at ‘G’ cells from neighboring values





Slide by Steve Seitz

Color Image

Images in ~~Matlab~~ Python

- **Images represented as a matrix**
- **Suppose we have a NxM RGB image called “im”**
  - im(0,0,0) = 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-1, M-1, 2) = bottom-right pixel in B-channel

column →

row ↓

R

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

G

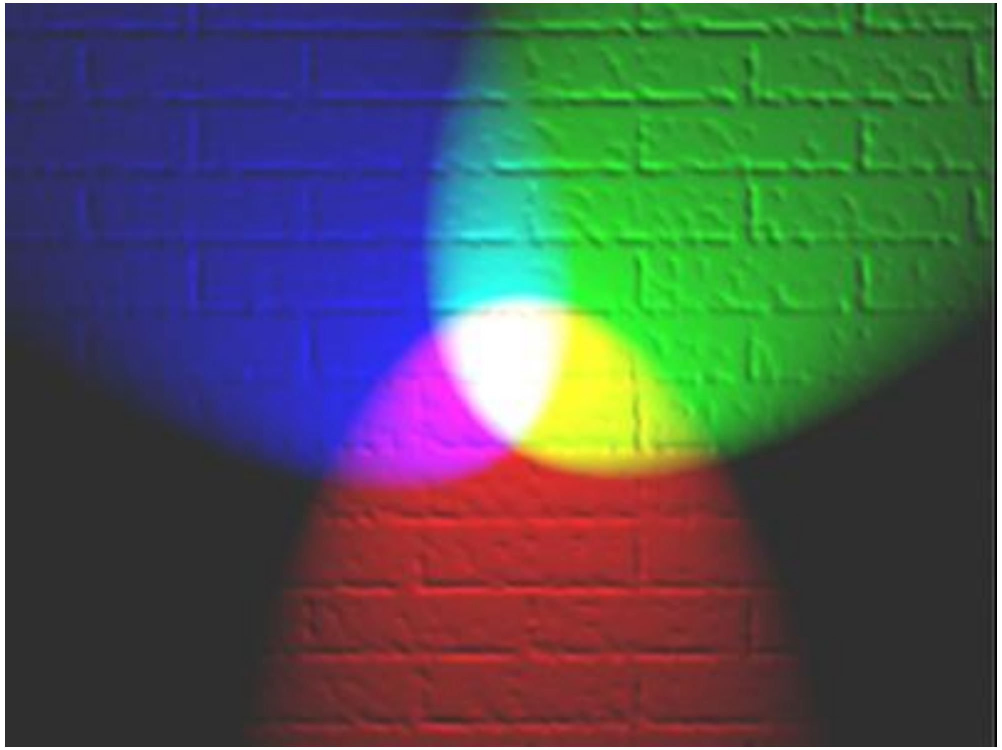
									0.92	0.99
									0.95	0.91
									0.91	0.92
									0.97	0.95
									0.79	0.85
									0.45	0.33
									0.49	0.74
									0.82	0.93
									0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

B

									0.92	0.99
									0.95	0.91
									0.91	0.92
									0.97	0.95
									0.79	0.85
									0.45	0.33
									0.49	0.74
									0.82	0.93
									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

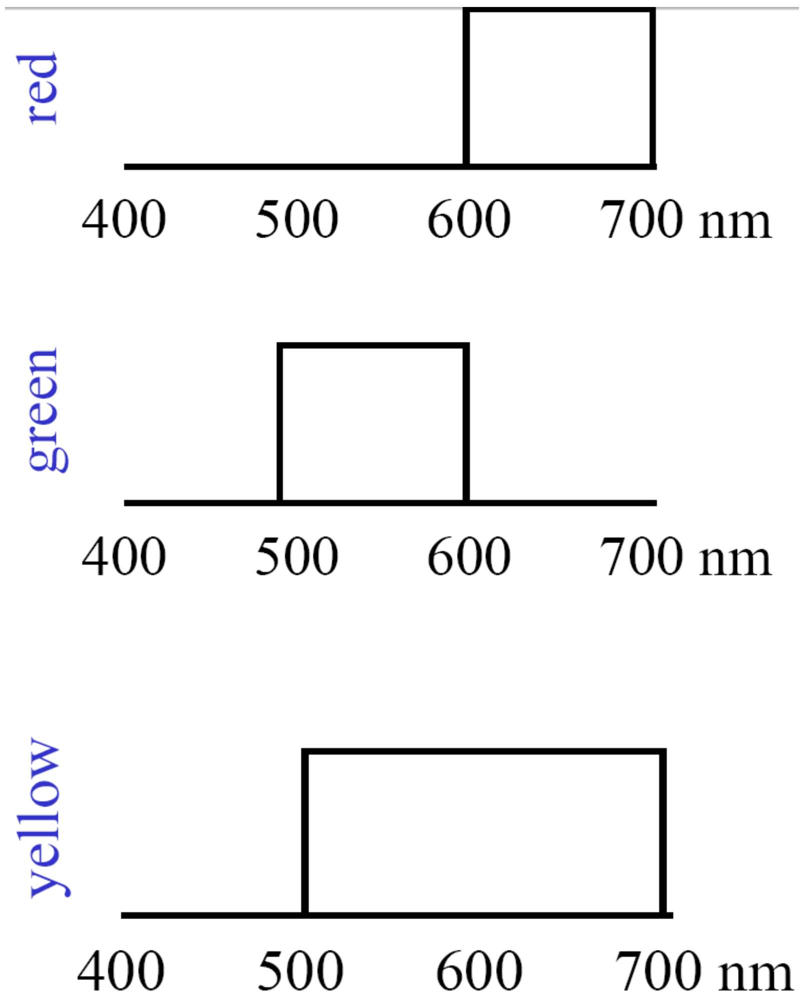
Color spaces

- **How can we represent color?**

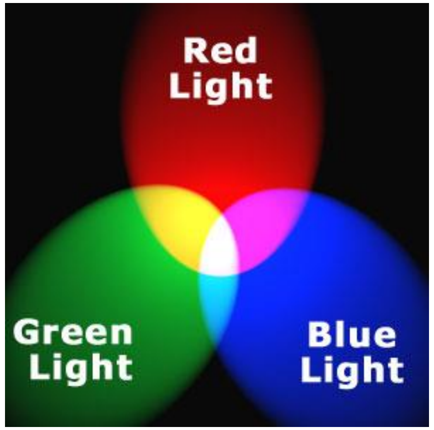


http://en.wikipedia.org/wiki/File:RGB_illumination.jpg

Additive color mixing



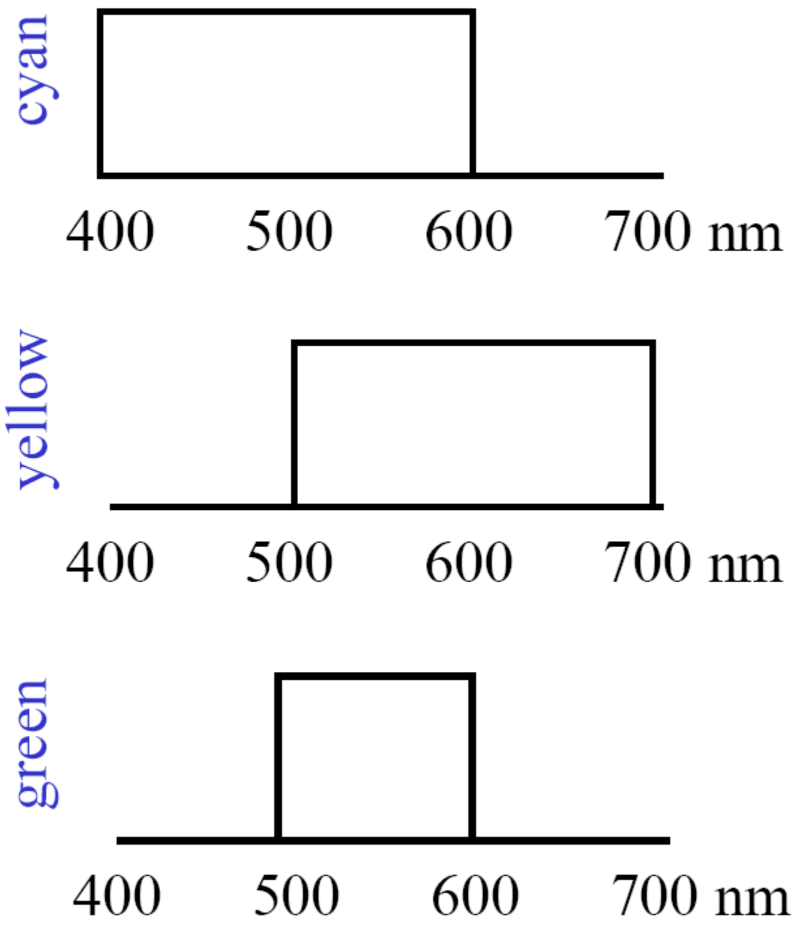
Colors combine by *adding* color spectra



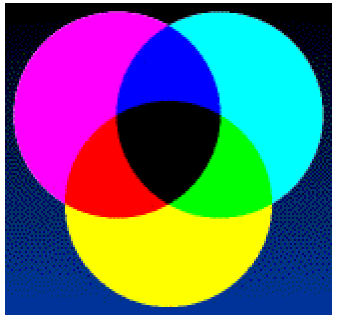
Light *adds* to existing black.

Source: W. Freeman

Subtractive color mixing



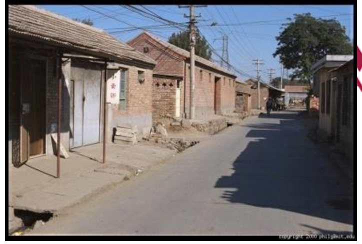
Colors combine by *multiplying* color spectra.



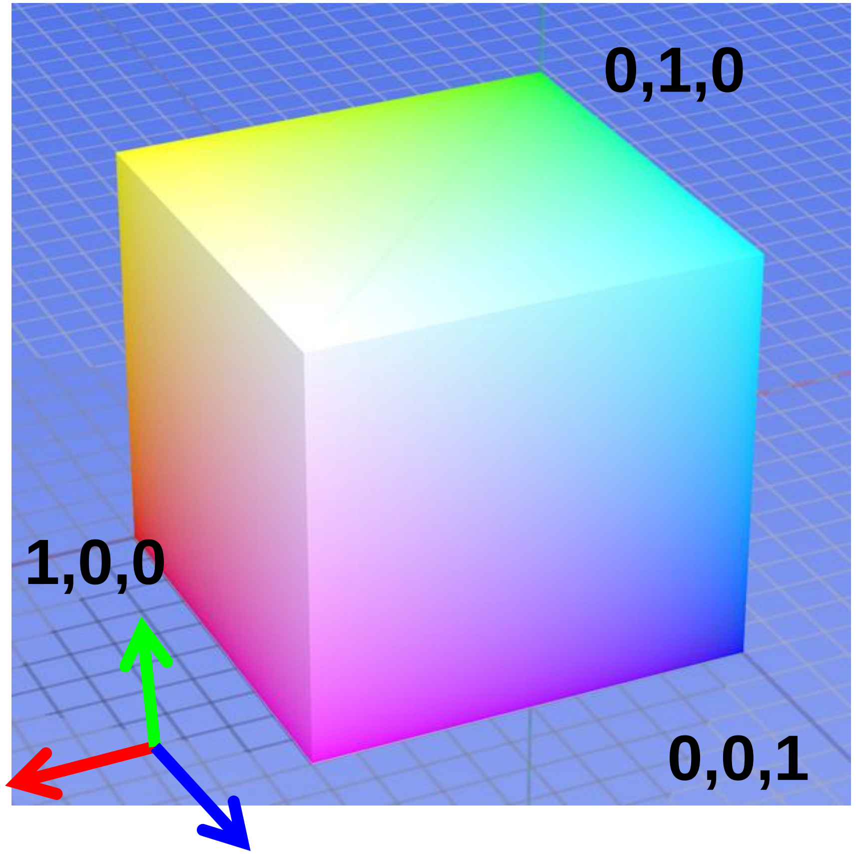
Pigments *remove* color from incident light (white).

Source: W. Freeman

Color spaces: RGB



Default color space



R
(G=0,B=0)

G
(R=0,B=0)

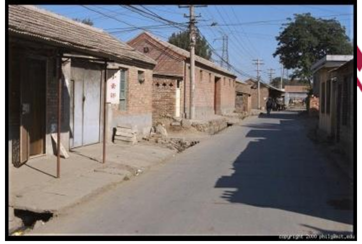
B
(R=0,G=0)

Some drawbacks

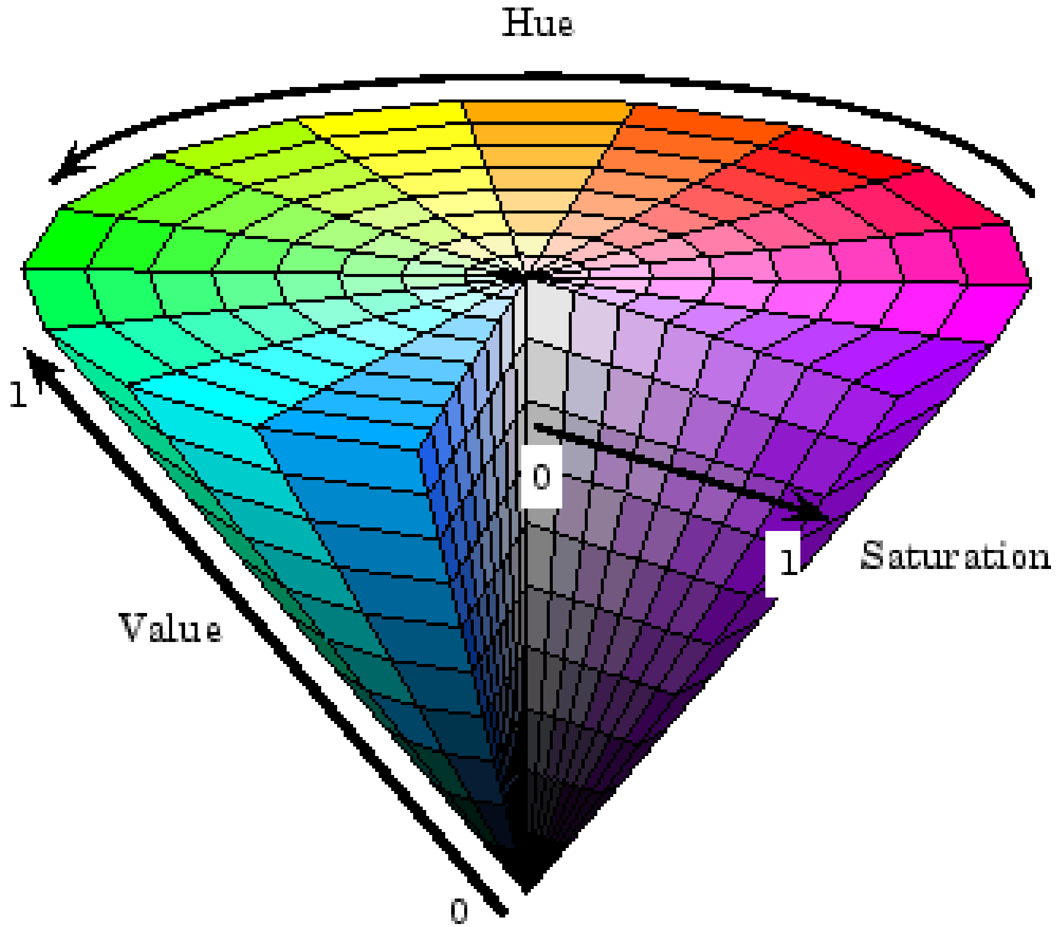
- Strongly correlated channels
- Non-perceptual

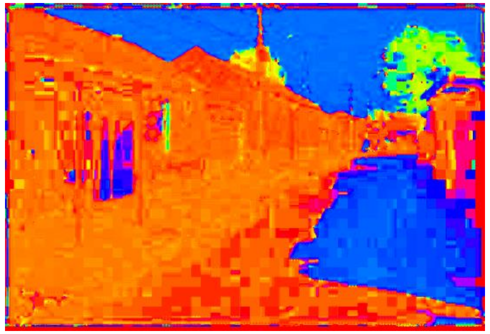
Image from: http://en.wikipedia.org/wiki/File:RGB_color_solid_cube.png

Color spaces: HSV



Intuitive color space



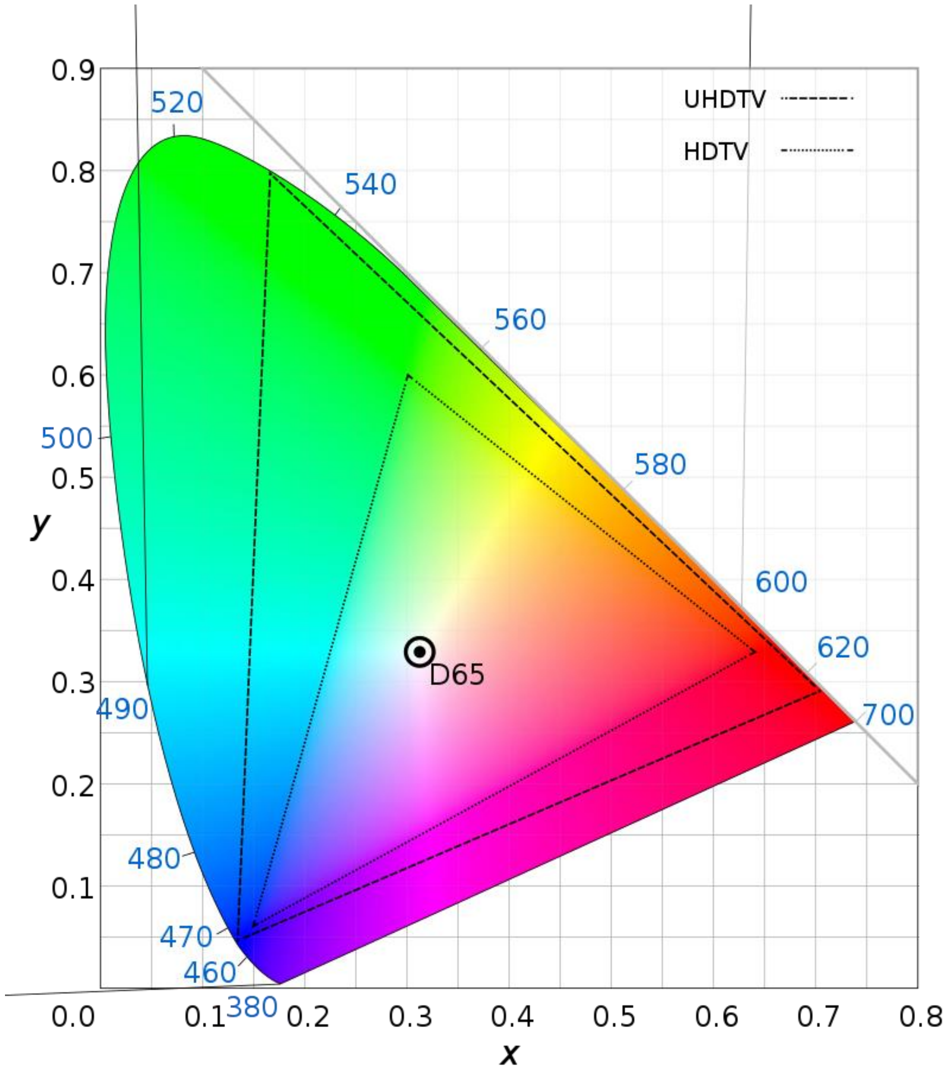


H
(S=1,V=1)

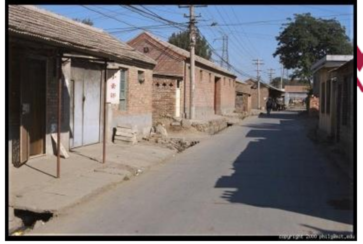
S
(H=1,V=1)

V
(H=1,S=0)

Color gamut

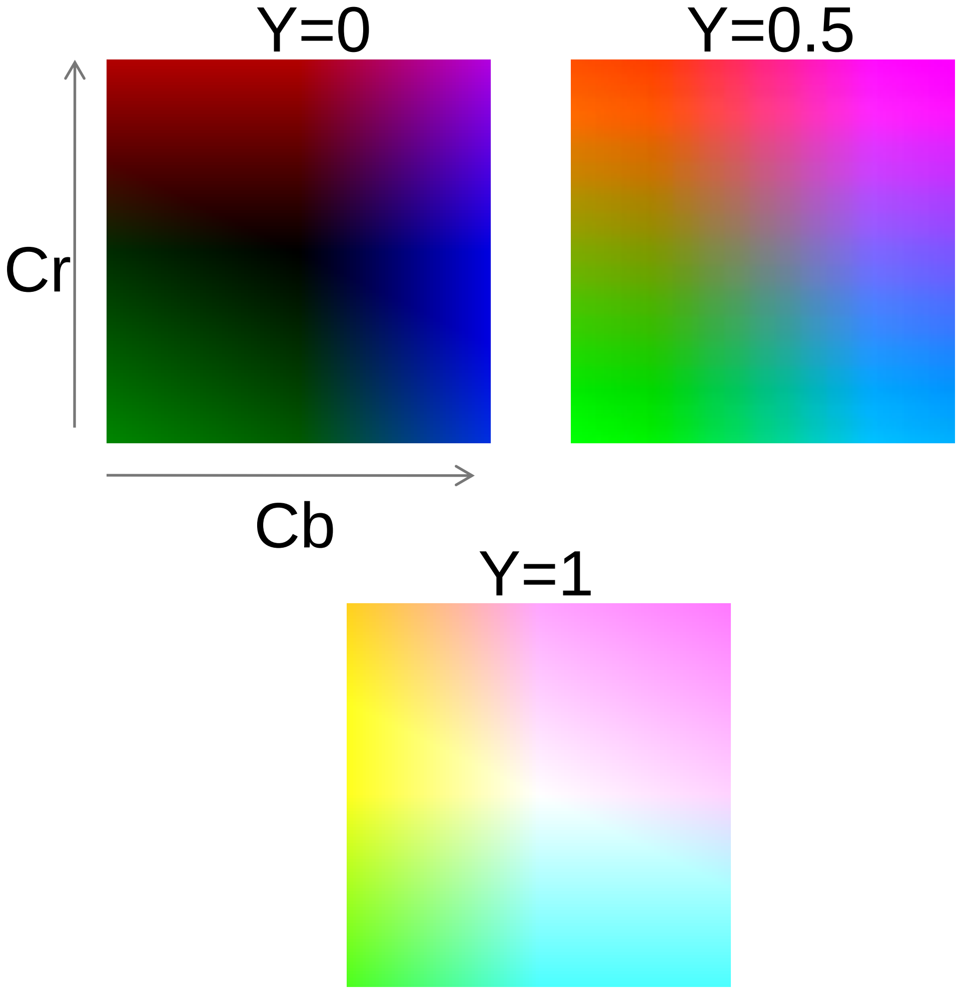


UHDTV
HDTV
520
540
560
580
600
620
700
500
490
480
470
460
380
D65
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
y
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
x



Color spaces: YCbCr

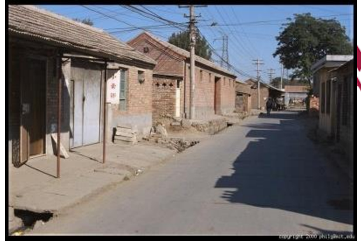
Fast to compute, good for compression, used by TV



Y
(Cb=0.5,Cr=0.5)

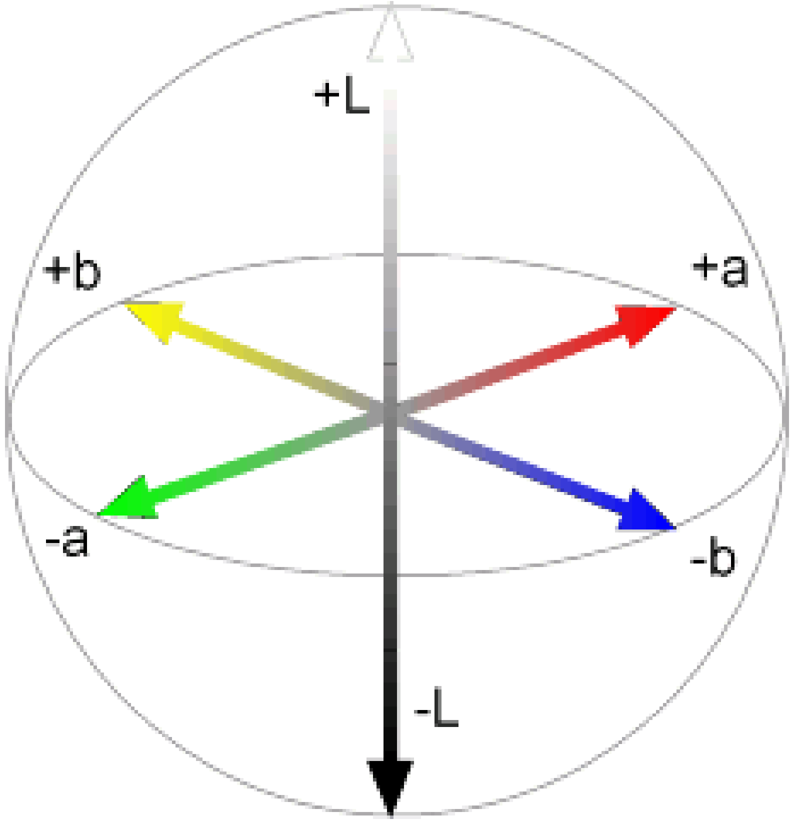
Cb
(Y=0.5,Cr=0.5)

Cr
(Y=0.5,Cb=05)



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?

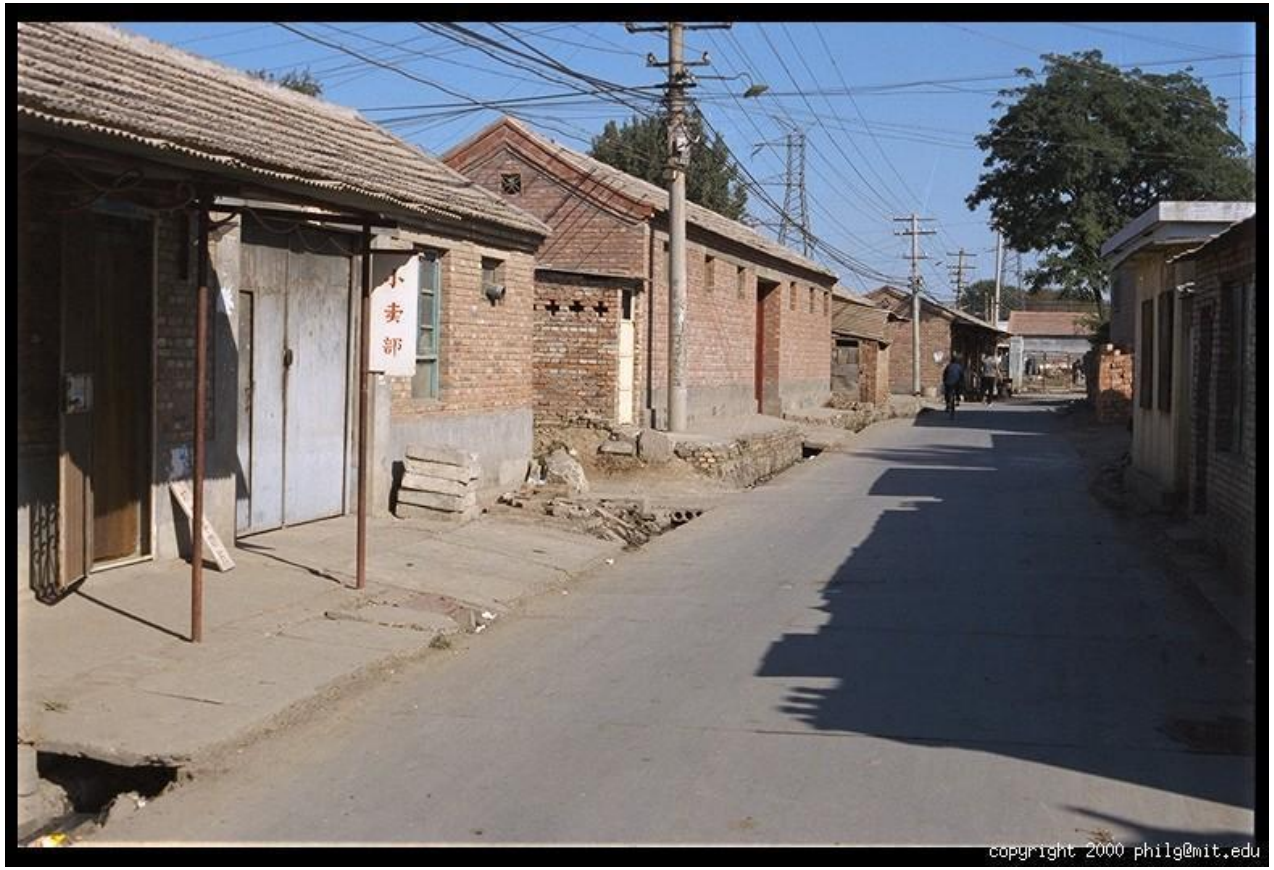
Most information in intensity

Only color shown – constant intensity

Most information in intensity

Only intensity shown – constant color

Most information in intensity



Original image

2.3.3 Compression

Figure 2.33 Image compressed with JPEG at three quality settings. Note how the amount of block artifact and high-frequency aliasing ("mosquito noise") increases from left to right.

Gamma correction

- **Power-law transformations**
- $s = c(r+\varepsilon)^{\gamma}$ **or** $s = cr^{\gamma}$
  - $\gamma < 1$ maps a narrow range of dark input values into a wider range of output values, while $\gamma > 1$ maps a narrow range of bright input values into a wider range of output values
  - γ : gamma, gamma correction

Perceived (linear) brightness =	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Physical (linear) brightness =	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0

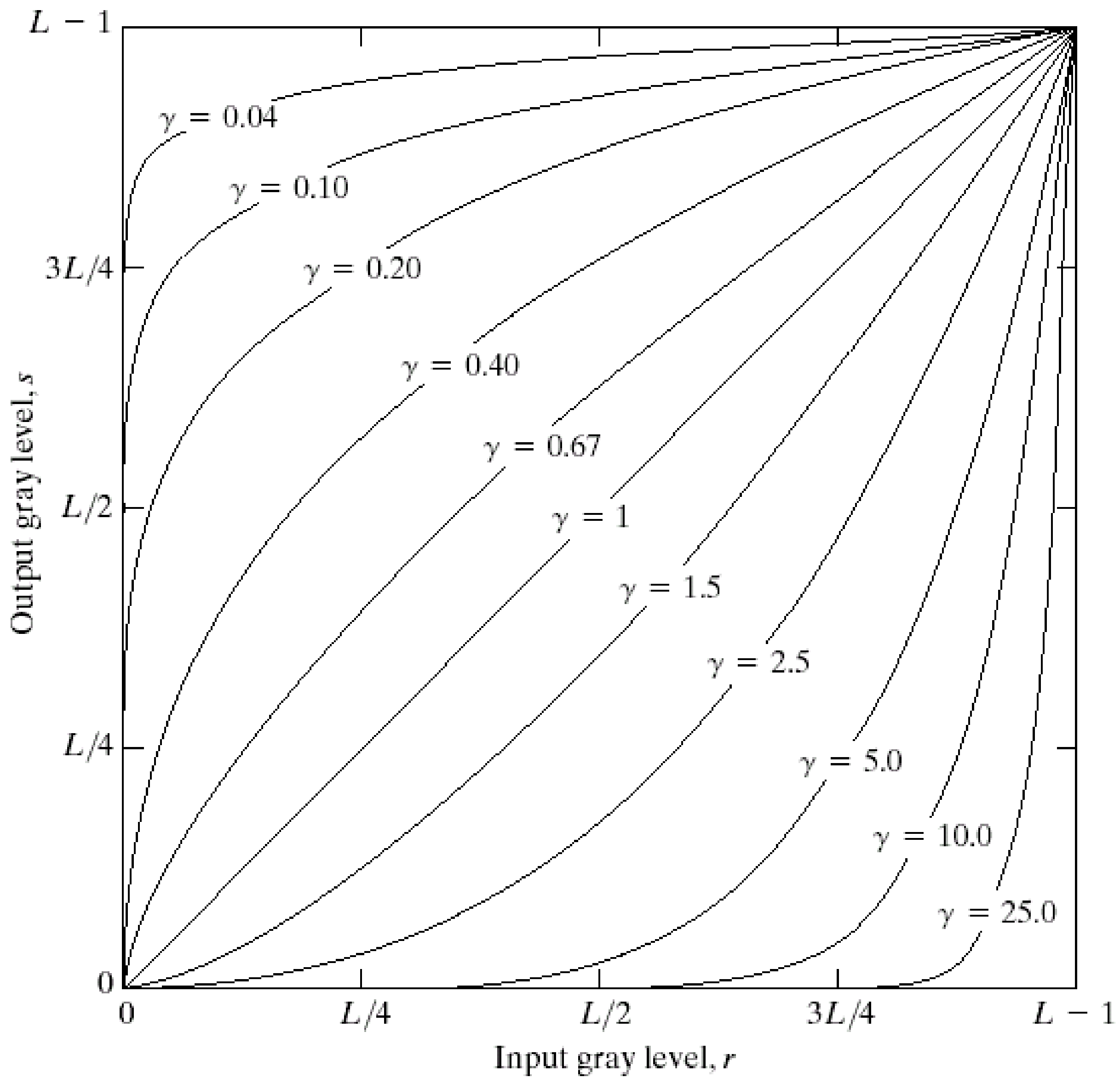
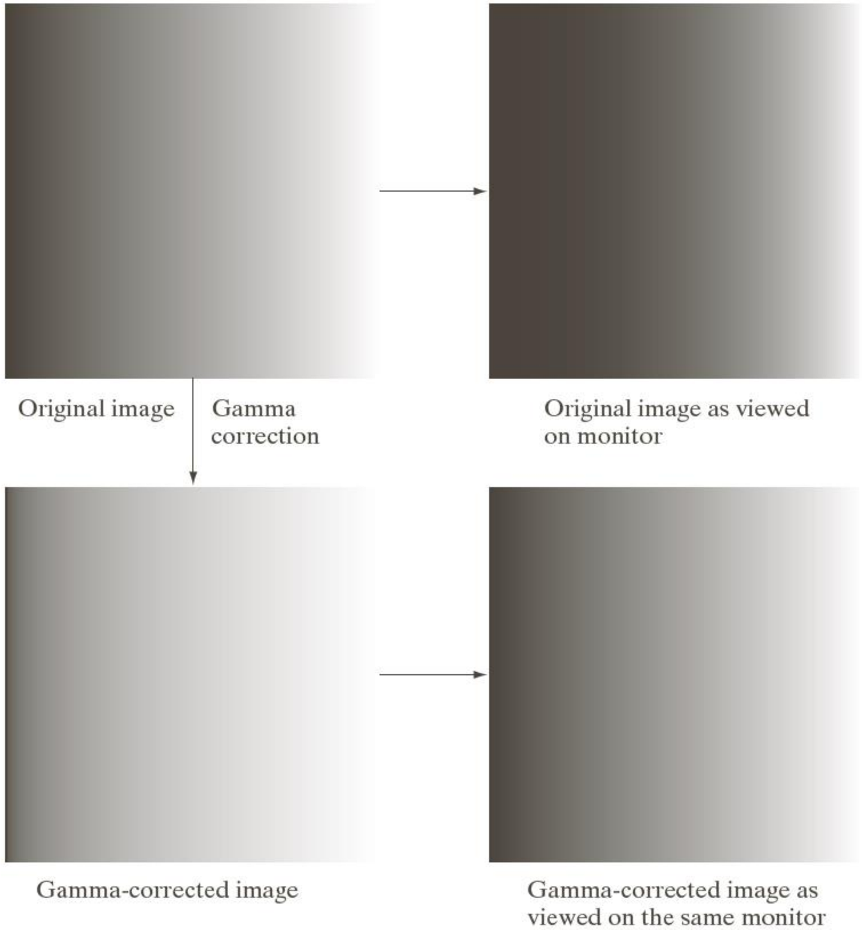


FIGURE 3.6 Plots of the equation $s = cr^{\gamma}$ for various values of γ ($c = 1$ in all cases).

- **Monitor,**

$$\gamma = 2.5$$



a b
c d

FIGURE 3.7
(a) Intensity ramp image. (b) Image as viewed on a simulated monitor with a gamma of 2.5. (c) Gamma-corrected image. (d) Corrected image as viewed on the same monitor. Compare (d) and (a).

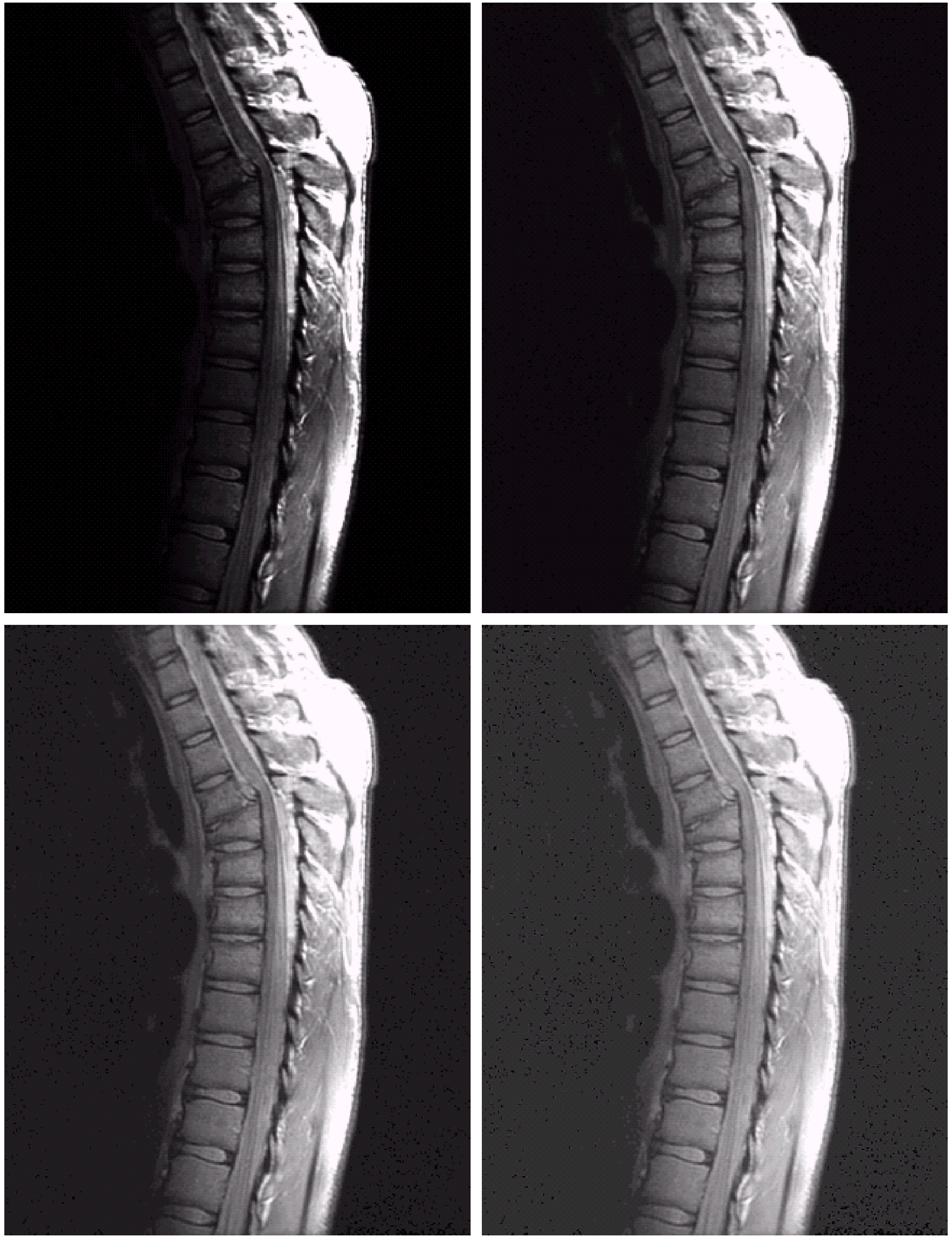


FIGURE 3.8 (a) Magnetic resonance (MR) image of a fractured human spine. (b)–(d) Results of applying the transformation in Eq. (3.2-3) with $c = 1$ and $\gamma = 0.6, 0.4,$ and 0.3, respectively. (Original image for this example courtesy of Dr. David R. Pickens, Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center.)

a b
c d

FIGURE 3.9
(a) Aerial image. (b)–(d) Results of applying the transformation in Eq. (3.2-3) with $c = 1$ and $\gamma = 3.0, 4.0,$ and 5.0, respectively. (Original image for this example courtesy of NASA.)