

3. Image Processing



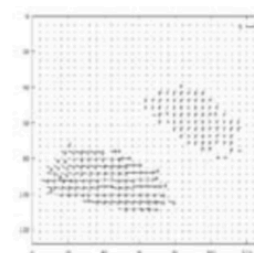
4. Features



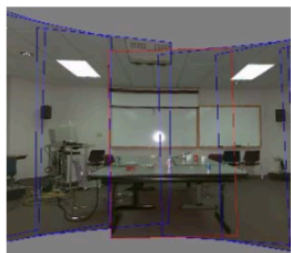
5. Segmentation



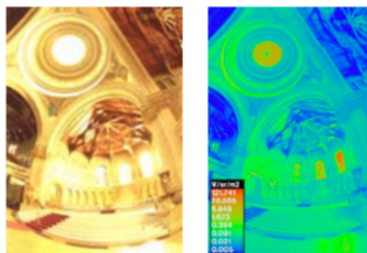
6-7. Structure from Motion



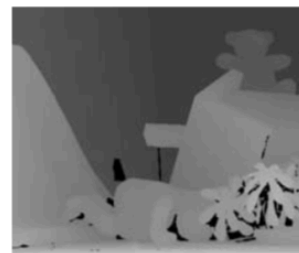
8. Motion



9. Stitching



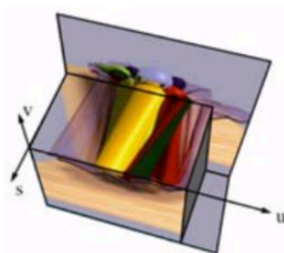
10. Computational Photography



11. Stereo



12. 3D Shape



13. Image-based Rendering

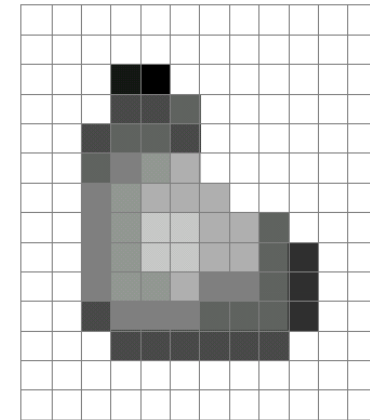
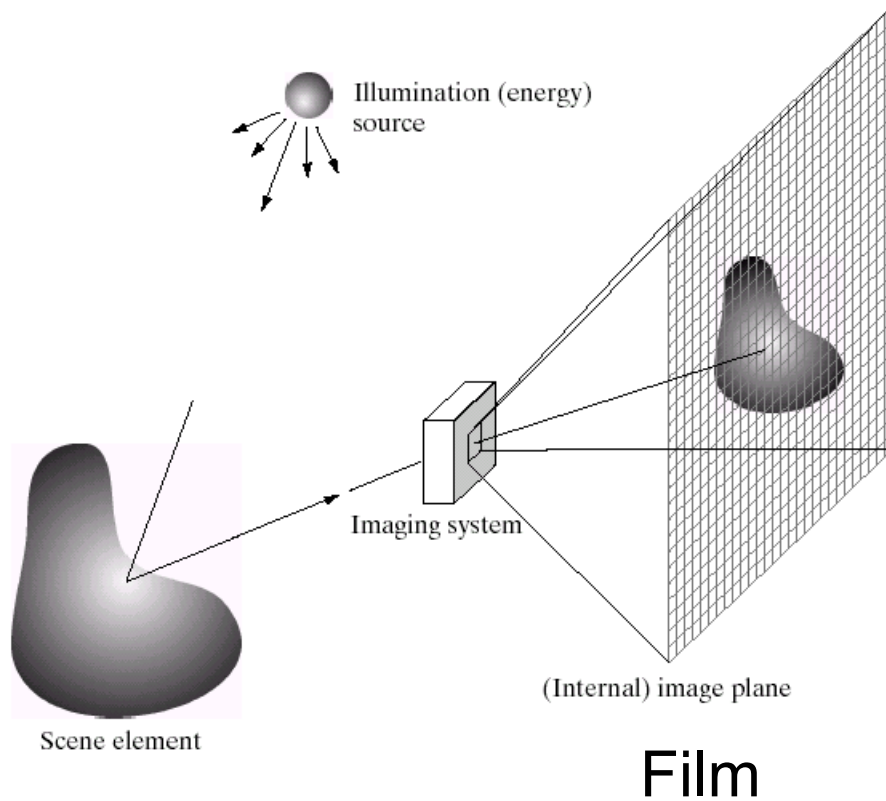


14. Recognition

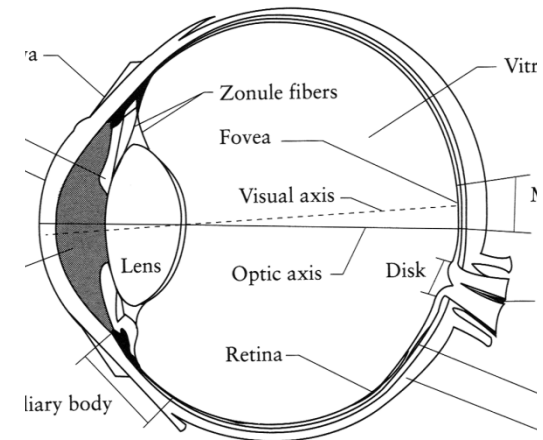
Image Formation

2.1	Geometric primitives and transformations	31
2.1.1	Geometric primitives	32
2.1.2	2D transformations	35
2.1.3	3D transformations	39
2.1.4	3D rotations	41
2.1.5	3D to 2D projections	46
2.1.6	Lens distortions	58
2.2	Photometric image formation	60
2.2.1	Lighting	60
2.2.2	Reflectance and shading	62
2.2.3	Optics	68
2.3	The digital camera	73
2.3.1	Sampling and aliasing	77
2.3.2	Color	80
2.3.3	Compression	90
2.4	Additional reading	93
2.5	Exercises	93

Image Formation



Digital Camera

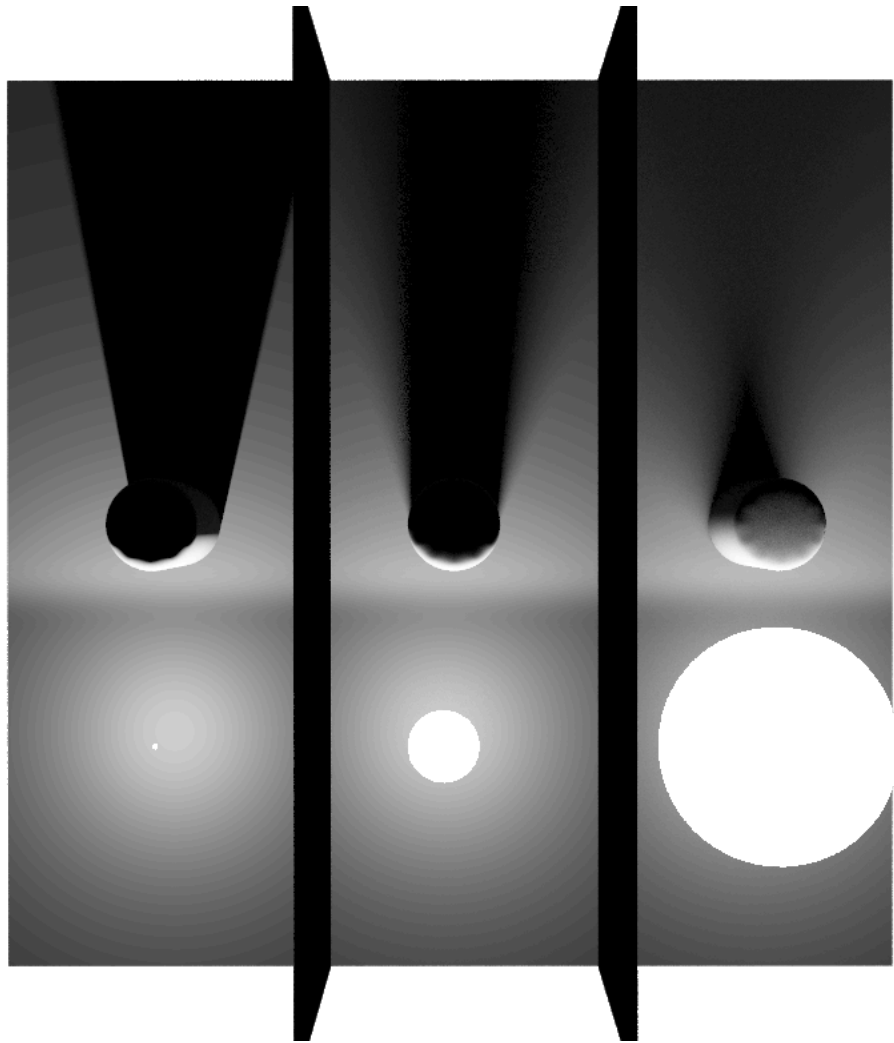


The Eye

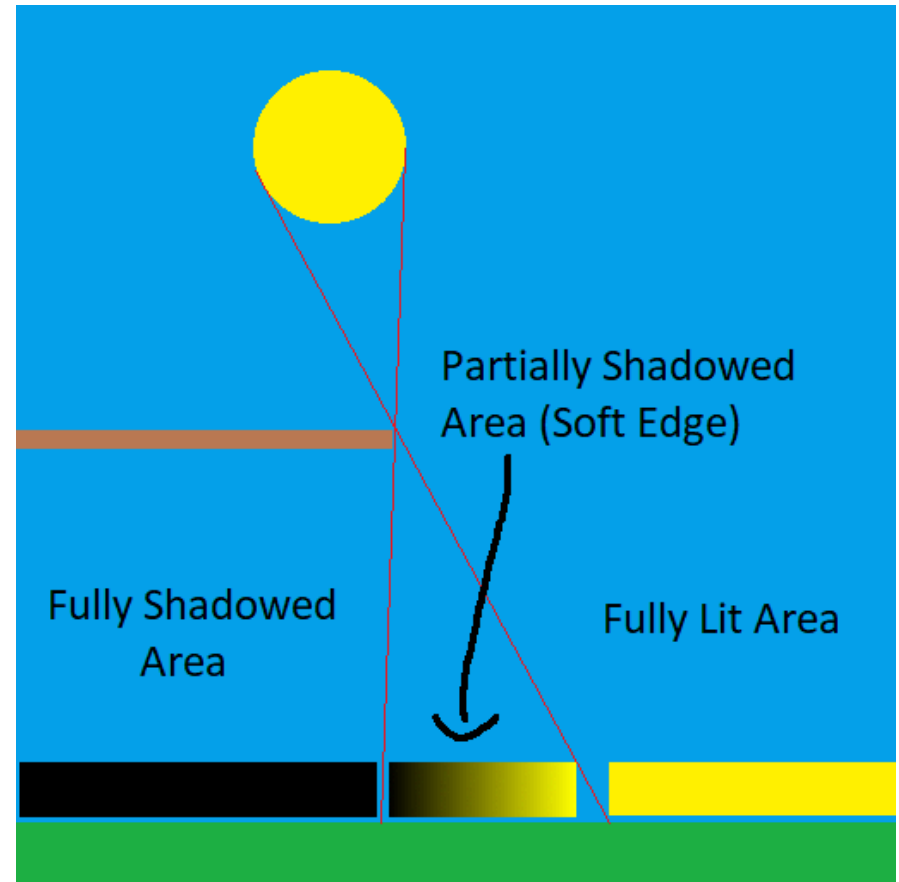
2.2.1 Lighting

- 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

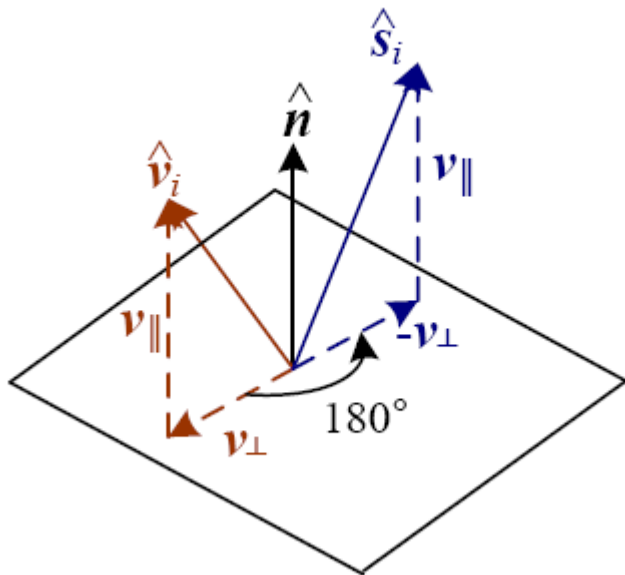
$$L(\hat{\mathbf{v}}; \lambda),$$



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 :

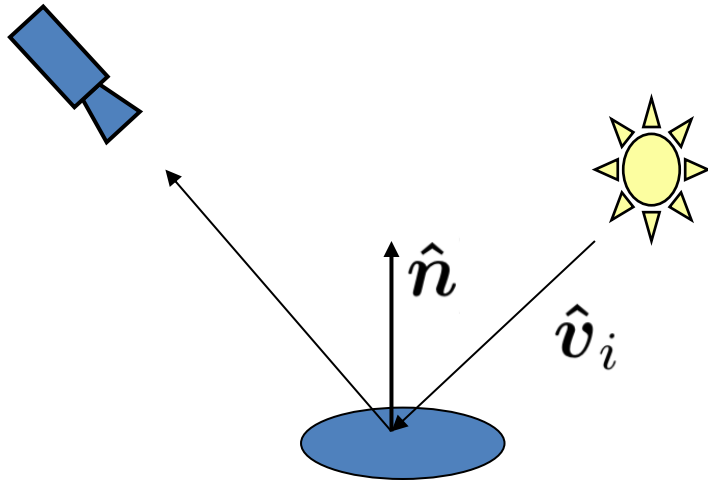
$$\hat{s}_i = v_{\parallel} - v_{\perp} = (2\hat{n}\hat{n}^T - \mathbf{I})v_i.$$

Diffuse Reflection

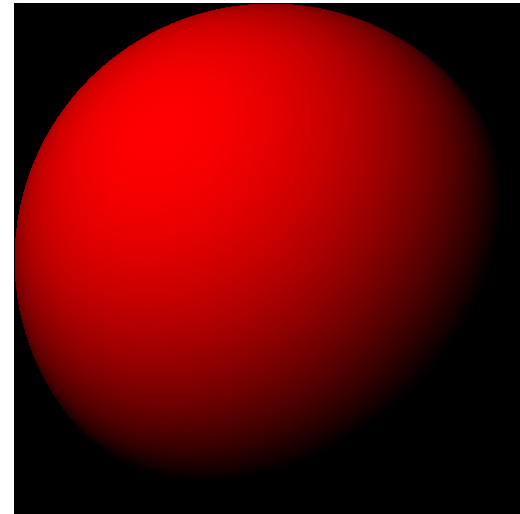


- Lambertian
- Oren-Nayar
- Fully general: BRDF

Lambertian Reflectance Model



Surface normal \hat{n}
Direction of illumination \hat{v}_i

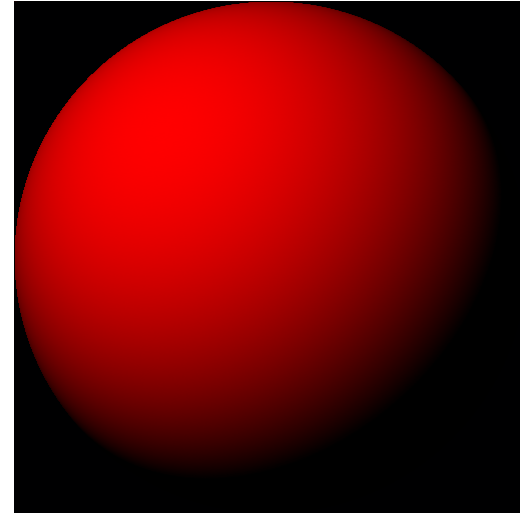
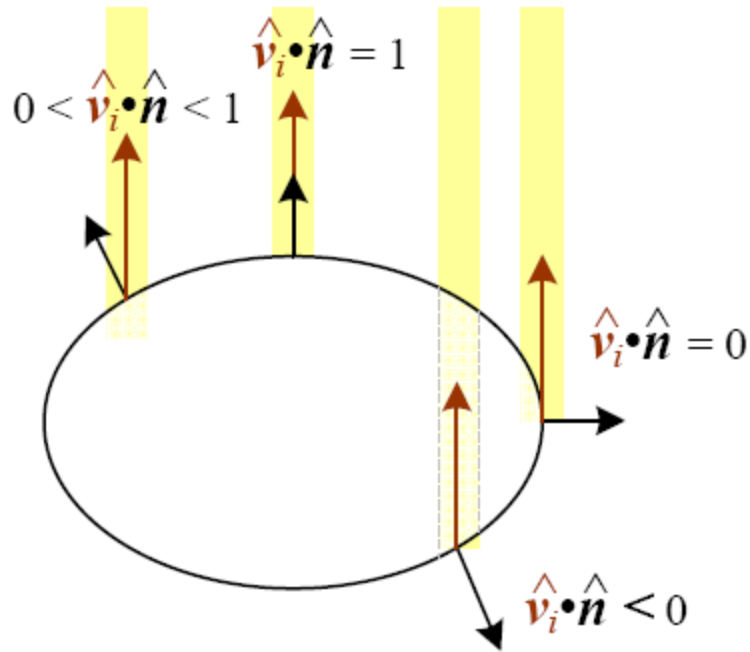


A Lambertian sphere

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

Commonly used in
computer vision and
computer graphics

Foreshortening



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

Confusion around Lambert's cosine law

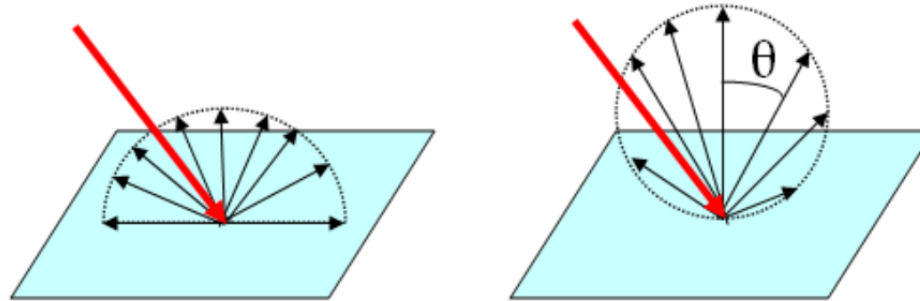


Figure: 1. The mental images corresponding to the two descriptions of Lambertian surfaces.

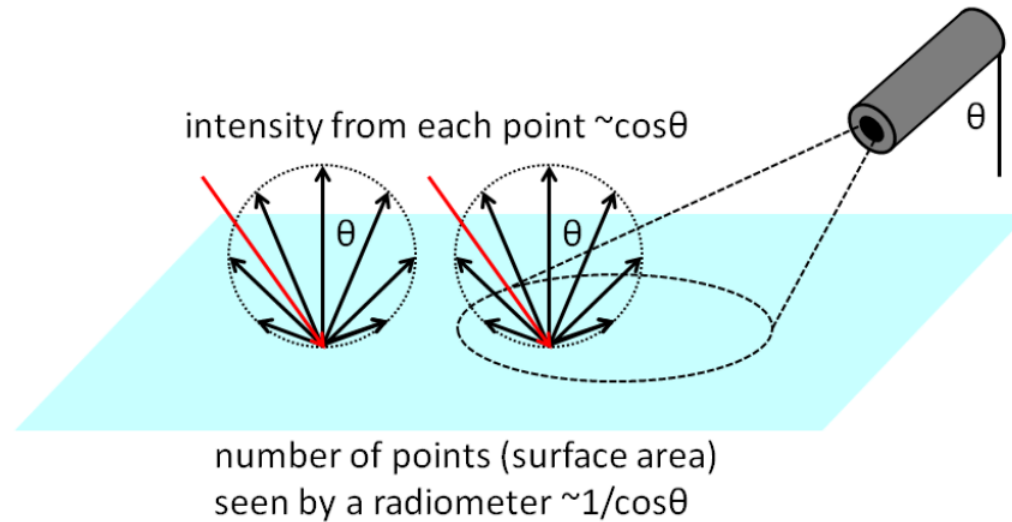


Figure: 2. Resolution of the paradoxical statements about how Lambertian surfaces reflect light. png

Oren-Nayar Reflectance Model

Roughness \longrightarrow

Lambertian



Rendered Spheres



Image of Full Moon

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

Phong Reflectance Model

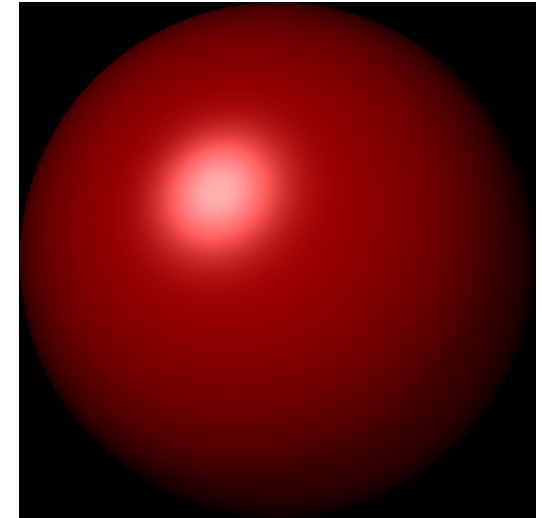
Combines diffuse (Lambertian) and specular lobe

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

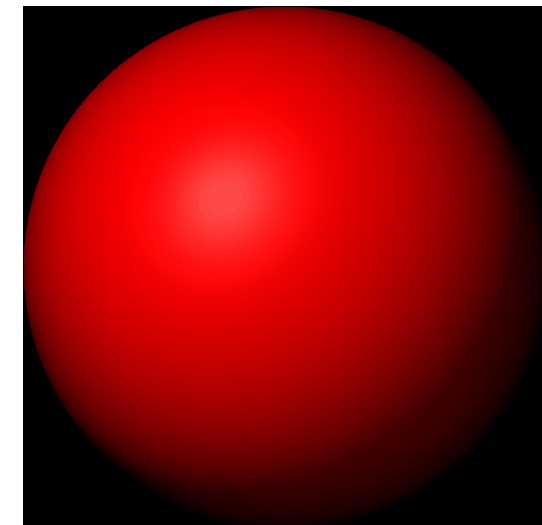
$$k_d(\lambda) \sum_i L_i(\lambda) [\hat{\mathbf{v}}_i \cdot \hat{\mathbf{n}}]^+ + k_s(\lambda) \sum_i L_i(\lambda) (\hat{\mathbf{v}}_r \cdot \hat{\mathbf{s}}_i)^{k_e}$$

↑
diffuse

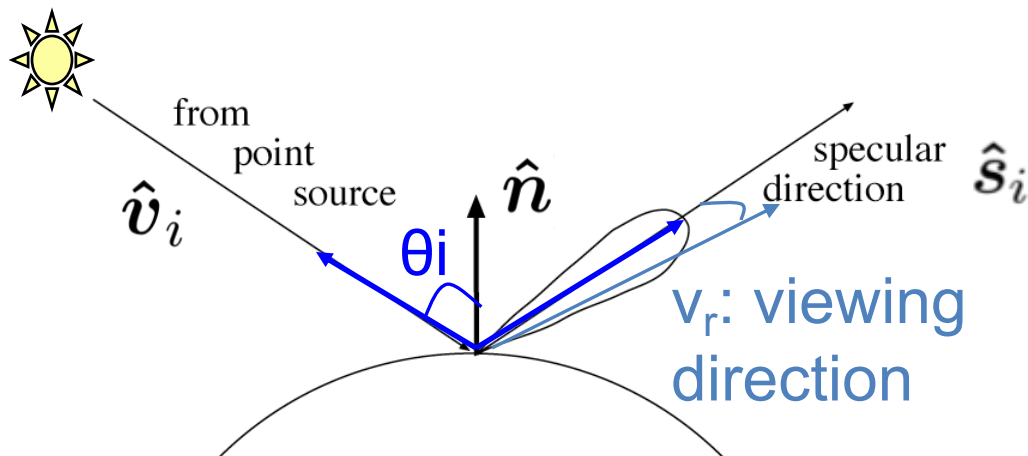
↑
specular



$k_d=0.3, k_s=0.7, k_e=2$



$k_d=0.7, k_s=0.3, k_e=0.5$



BRDF

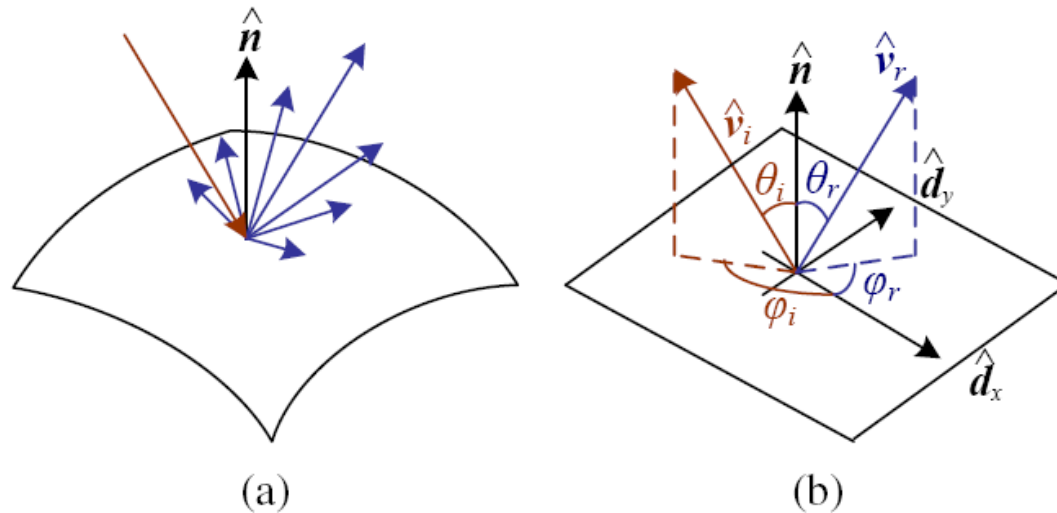


Figure 2.15: (a) Light scattering when hitting a surface. (b) The bidirectional reflectance distribution function (BRDF) $f(\theta_i, \phi_i, \theta_r, \phi_r)$ is parameterized by the angles the incident \hat{v}_i and reflected \hat{v}_r light ray directions make with the local surface coordinate frame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

For an isotropic material, we can simplify the BRDF to

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

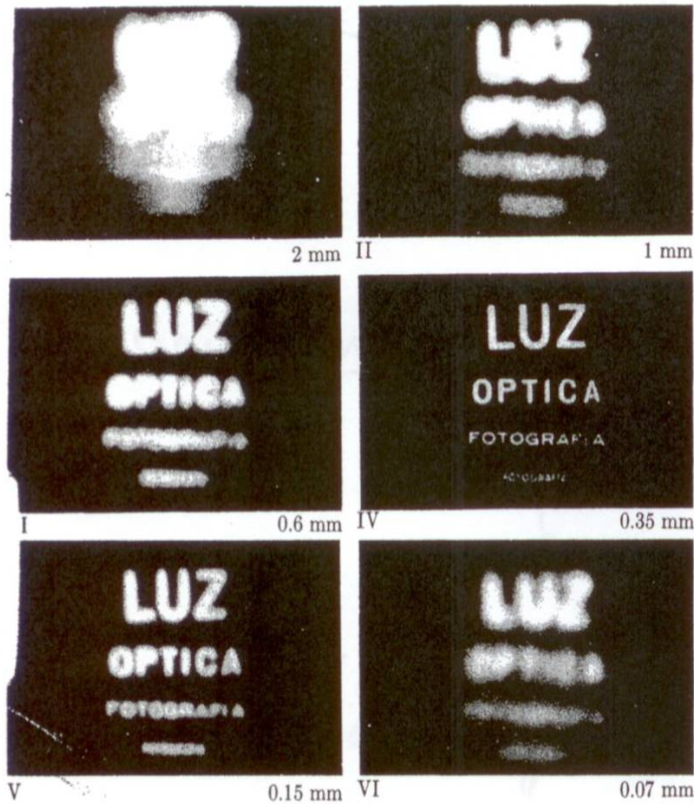
While light is scattered uniformly in all directions, i.e., the BRDF is constant,

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

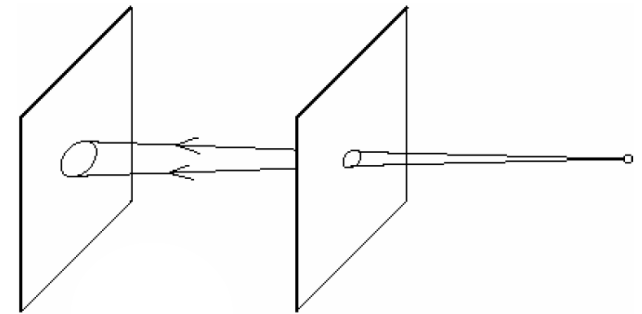
2.2.3 Optics

Pinhole size / aperture

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



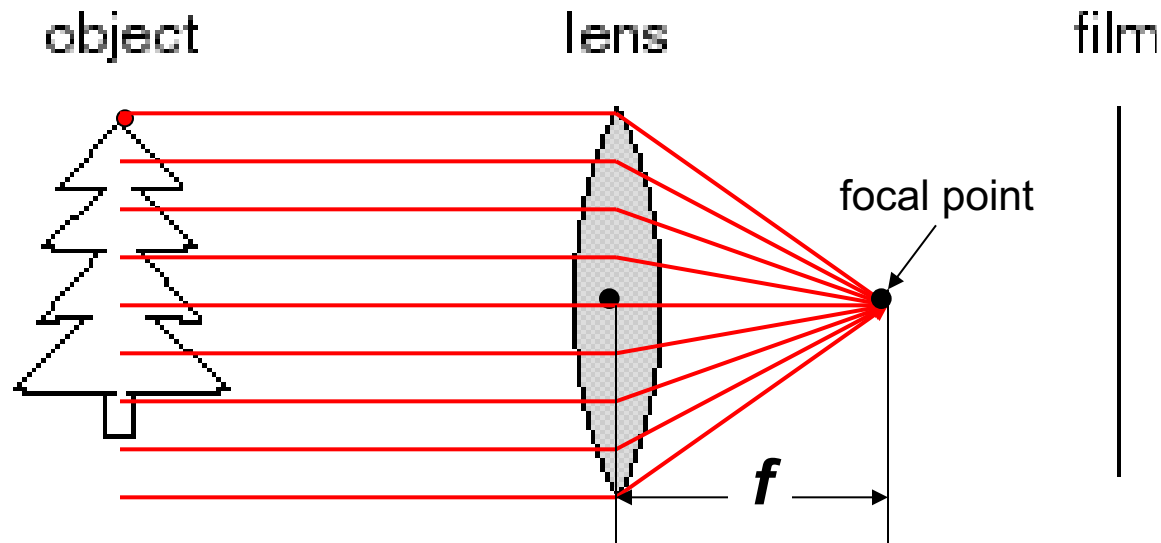
Larger



Smaller

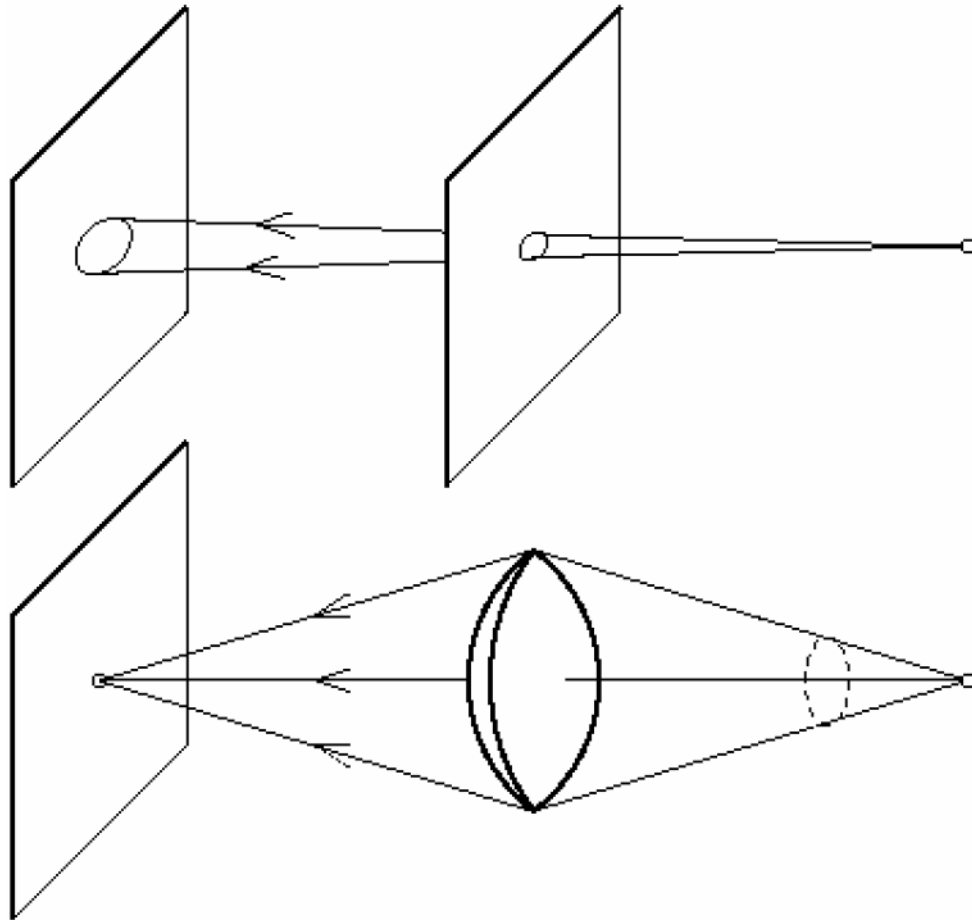
Fig. 5.96 The pinhole camera. Note the variation in image clarity as the hole diameter decreases. [Photos courtesy Dr. N. Joel, UNESCO.]

Adding a lens

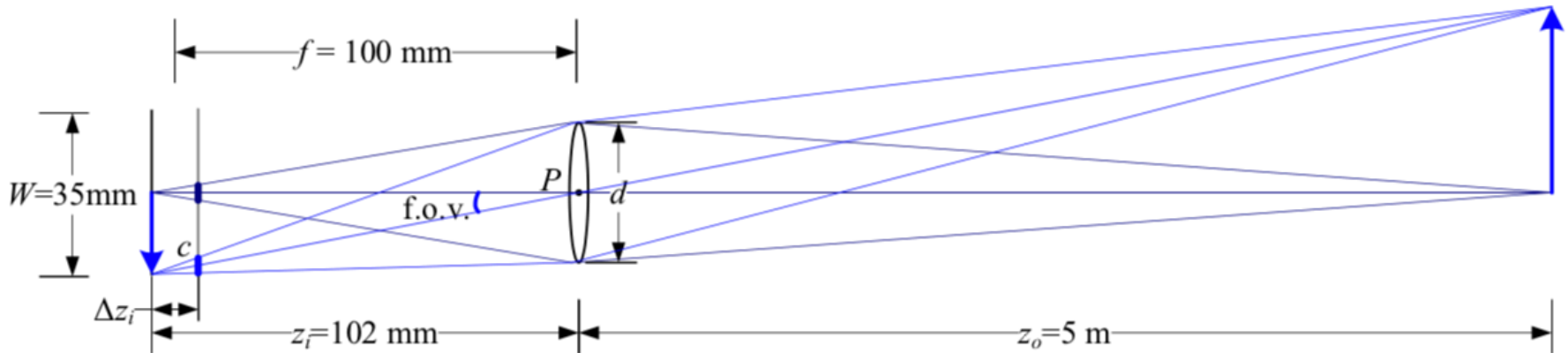


- A lens focuses light onto the film
 - Rays passing through the center are not deviated
 - All parallel rays converge to one point on a plane located at the *focal length* f

Pinhole vs. lens



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 $z_i > f$, we bring the focal plane back from infinity.
 - E.g.: $z_i = 102 \text{ mm}$, $f = 100 \text{ mm} \Rightarrow z_o = 5 \text{ m}$

Focus and depth of field



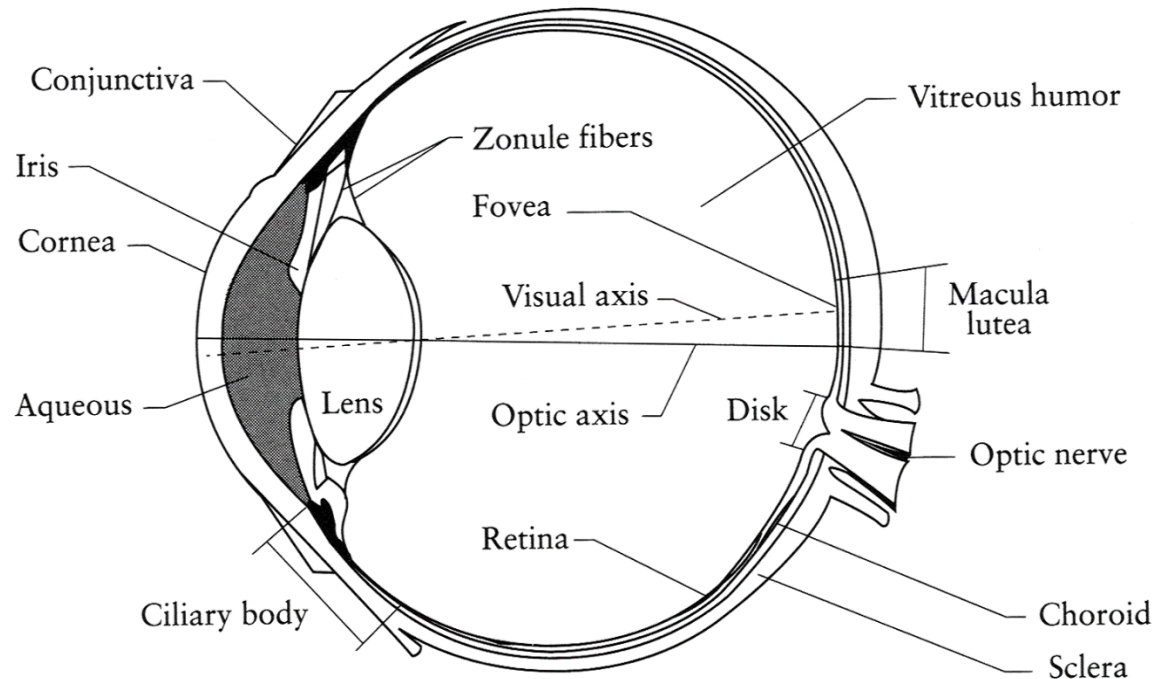
The smaller the aperture (area that lets light through), the more a lens behaves as a pinhole, the more everything is in focus.

Phones: fake “depth of field” with deep learning and stereo 😊

Image Formation

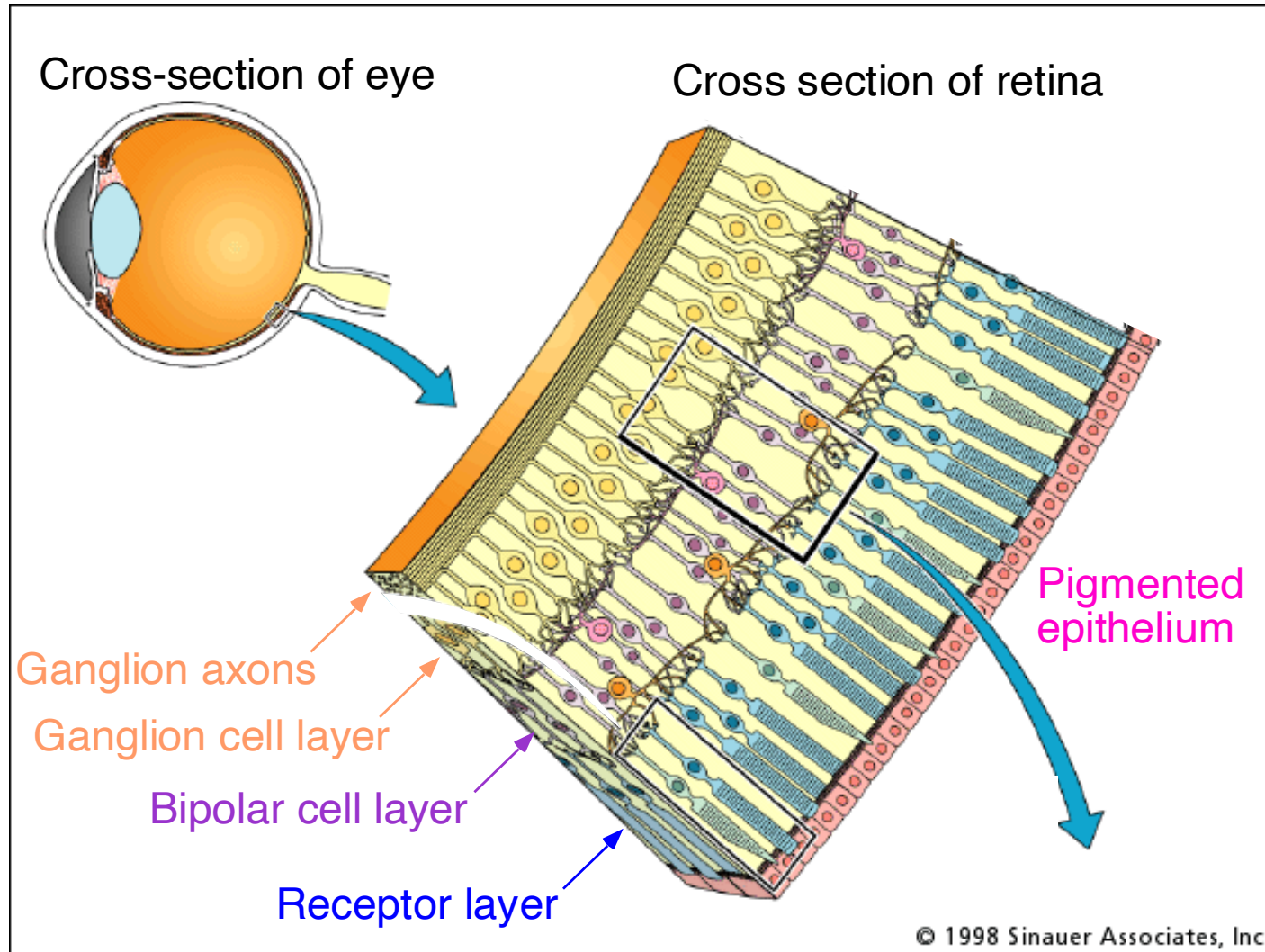
2.1	Geometric primitives and transformations	31
2.1.1	Geometric primitives	32
2.1.2	2D transformations	35
2.1.3	3D transformations	39
2.1.4	3D rotations	41
2.1.5	3D to 2D projections	46
2.1.6	Lens distortions	58
2.2	Photometric image formation	60
2.2.1	Lighting	60
2.2.2	Reflectance and shading	62
2.2.3	Optics	68
2.3	The digital camera	73
2.3.1	Sampling and aliasing	77
2.3.2	Color	80
2.3.3	Compression	90
2.4	Additional reading	93
2.5	Exercises	93

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**

The Retina



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

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

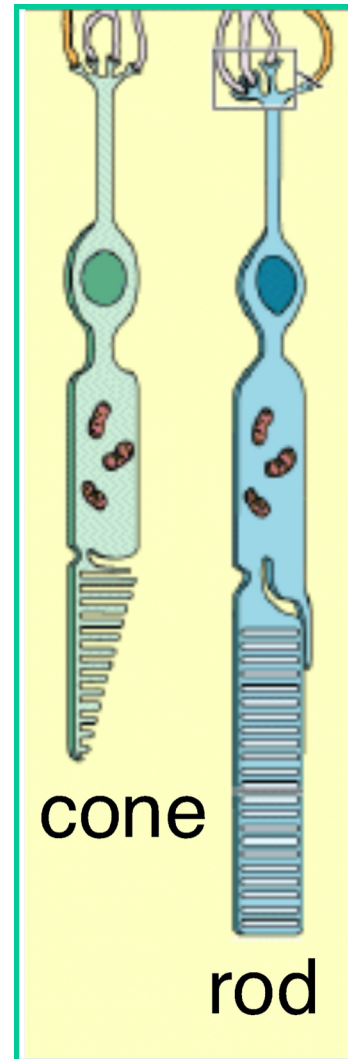
Two types of light-sensitive receptors

Cones

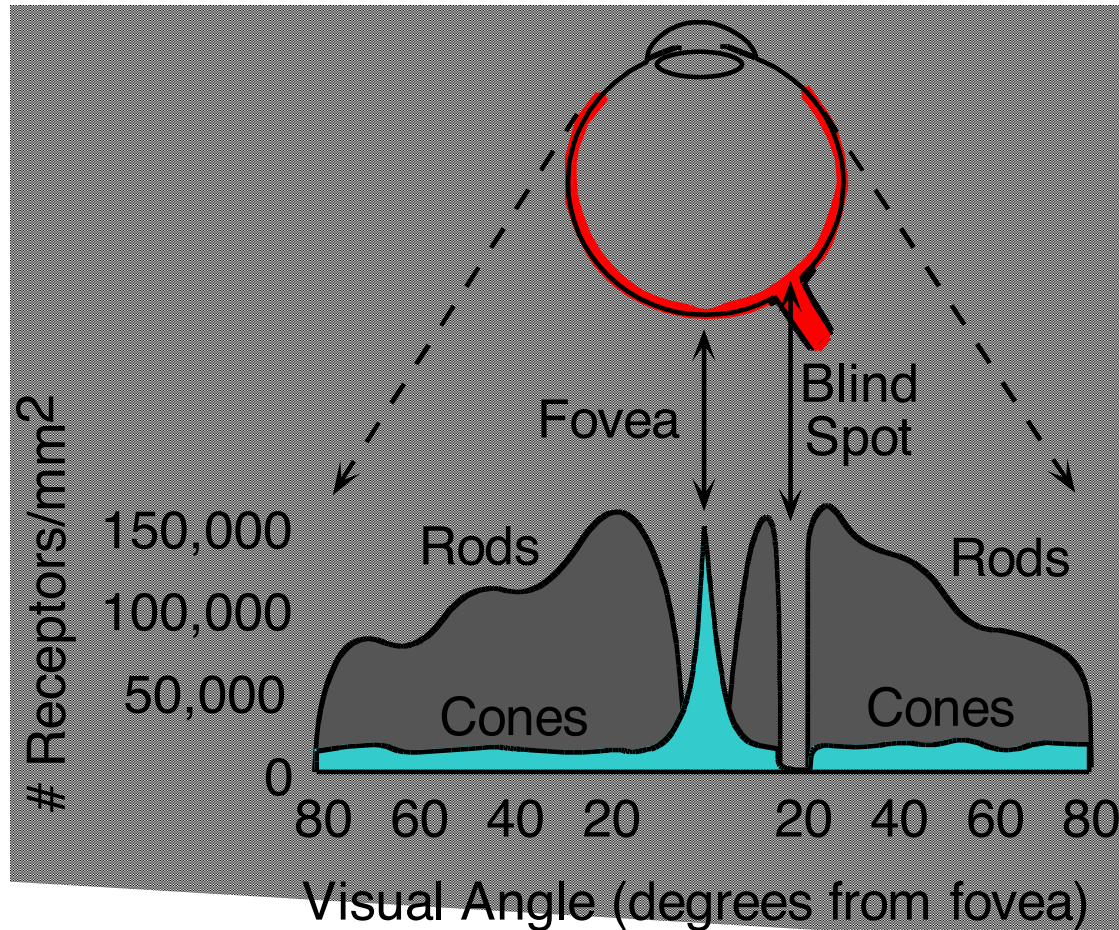
cone-shaped
less sensitive
operate in high light
color vision

Rods

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



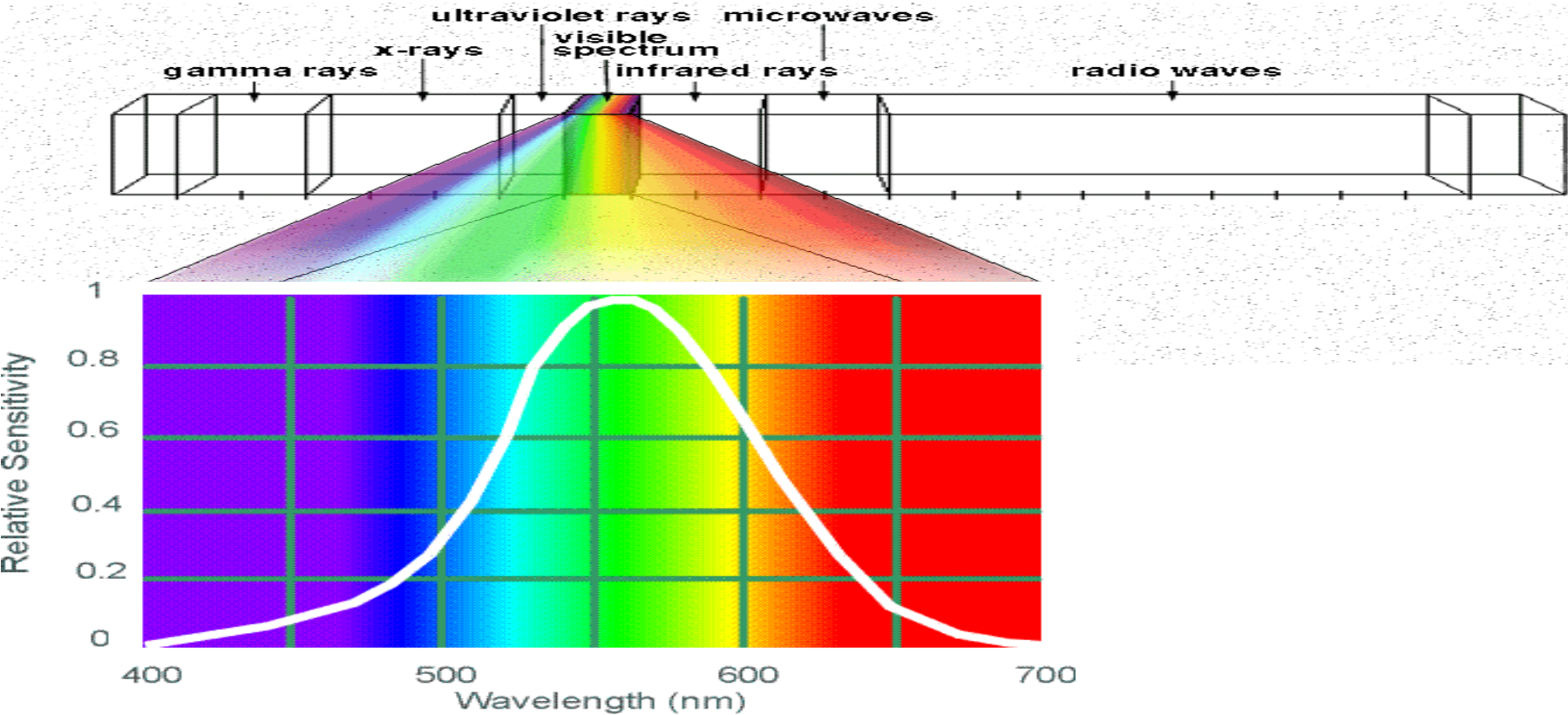
Distribution of Rods and Cones



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

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

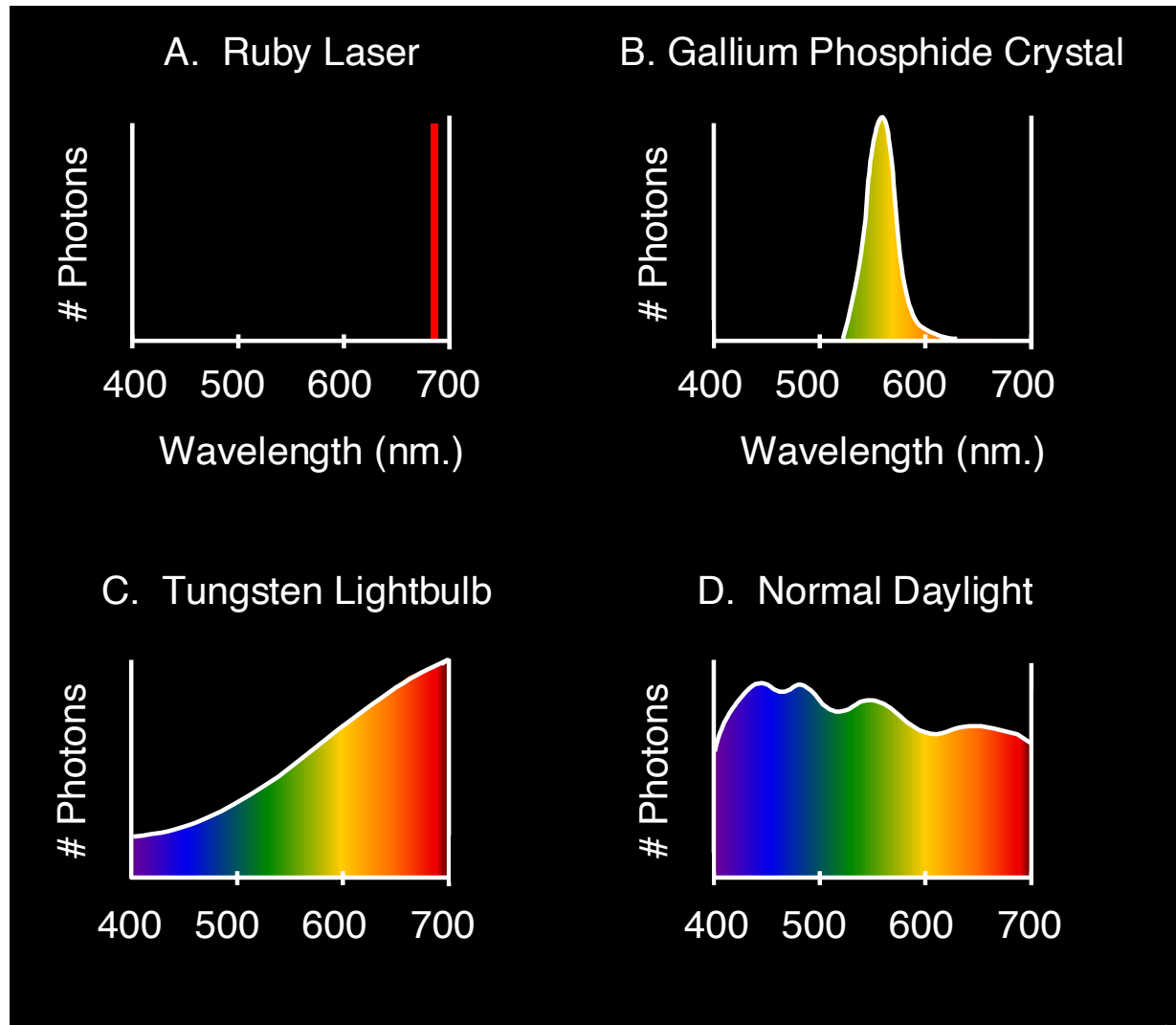
Electromagnetic Spectrum



Human Luminance Sensitivity Function

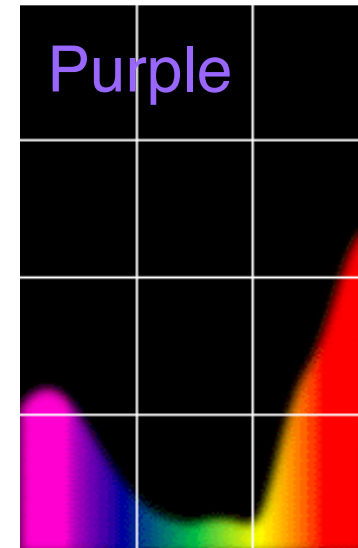
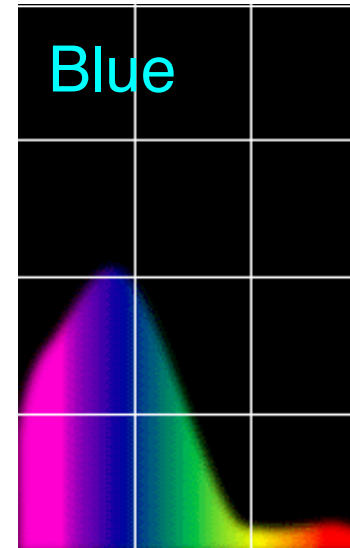
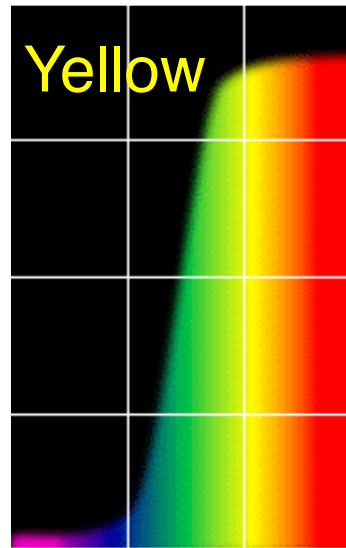
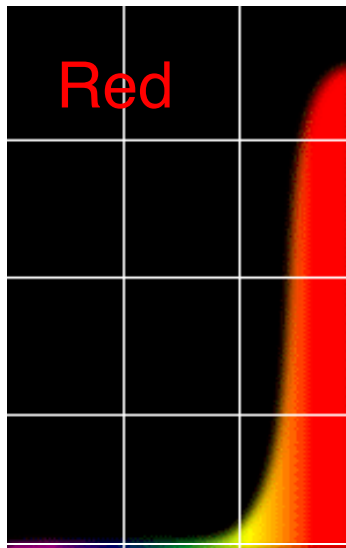
The Physics of Light

Some examples of the spectra of light sources



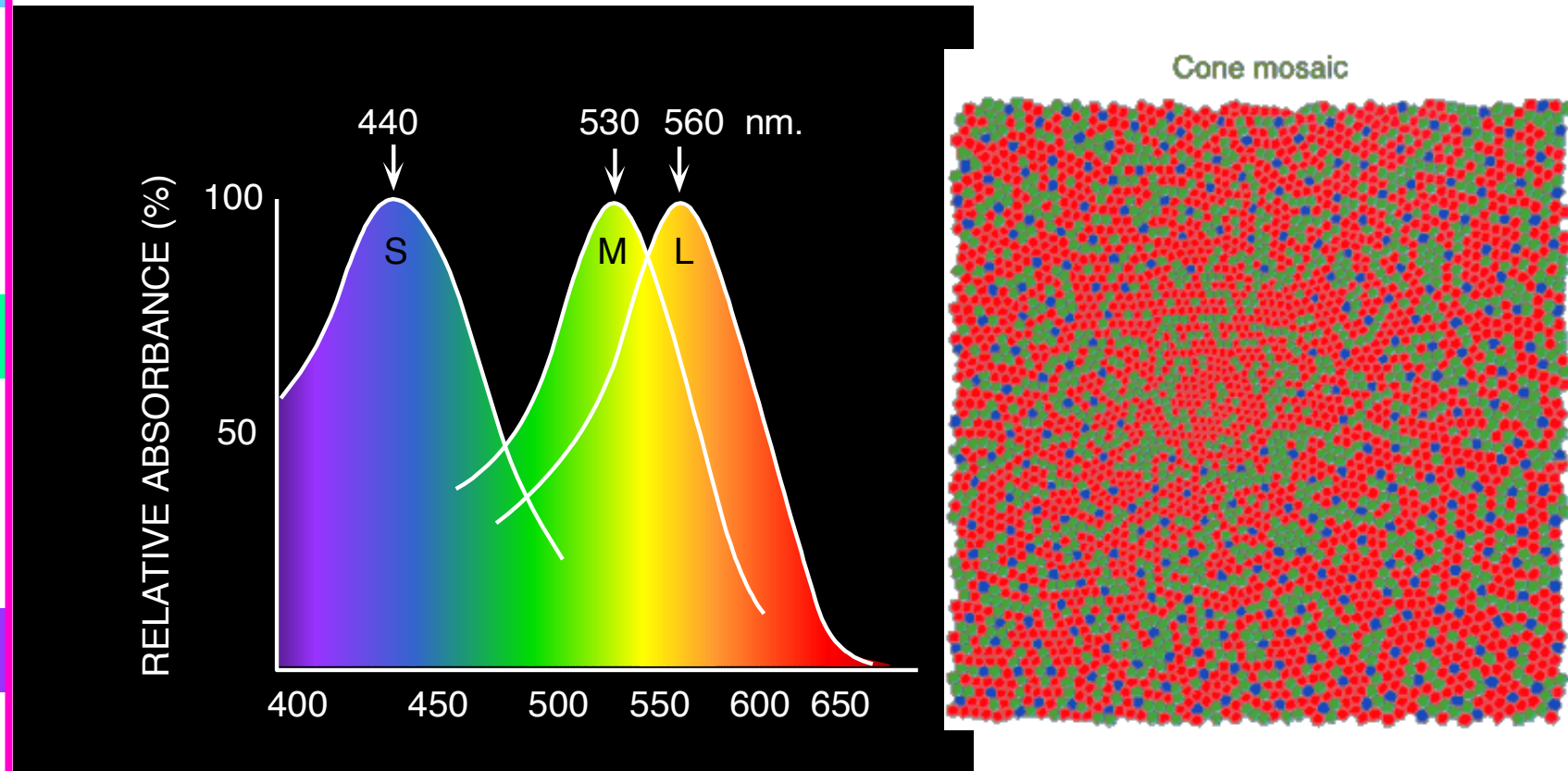
The Physics of Light

Some examples of the reflectance spectra of surfaces



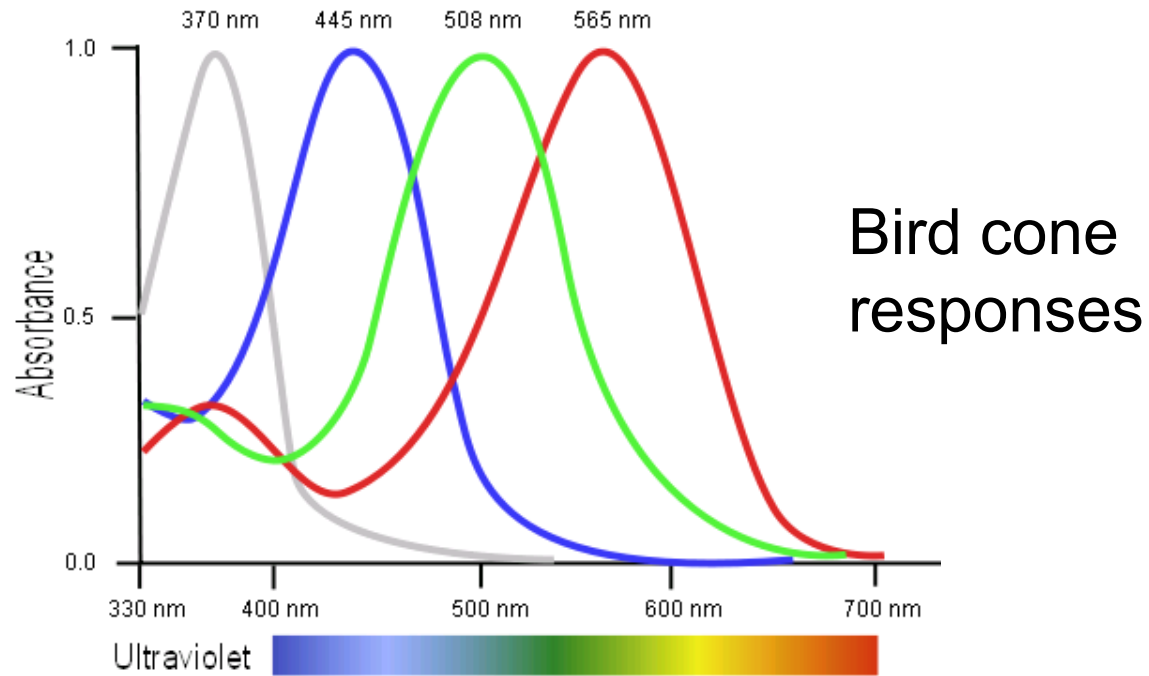
Physiology of Color Vision

Three kinds of cones:



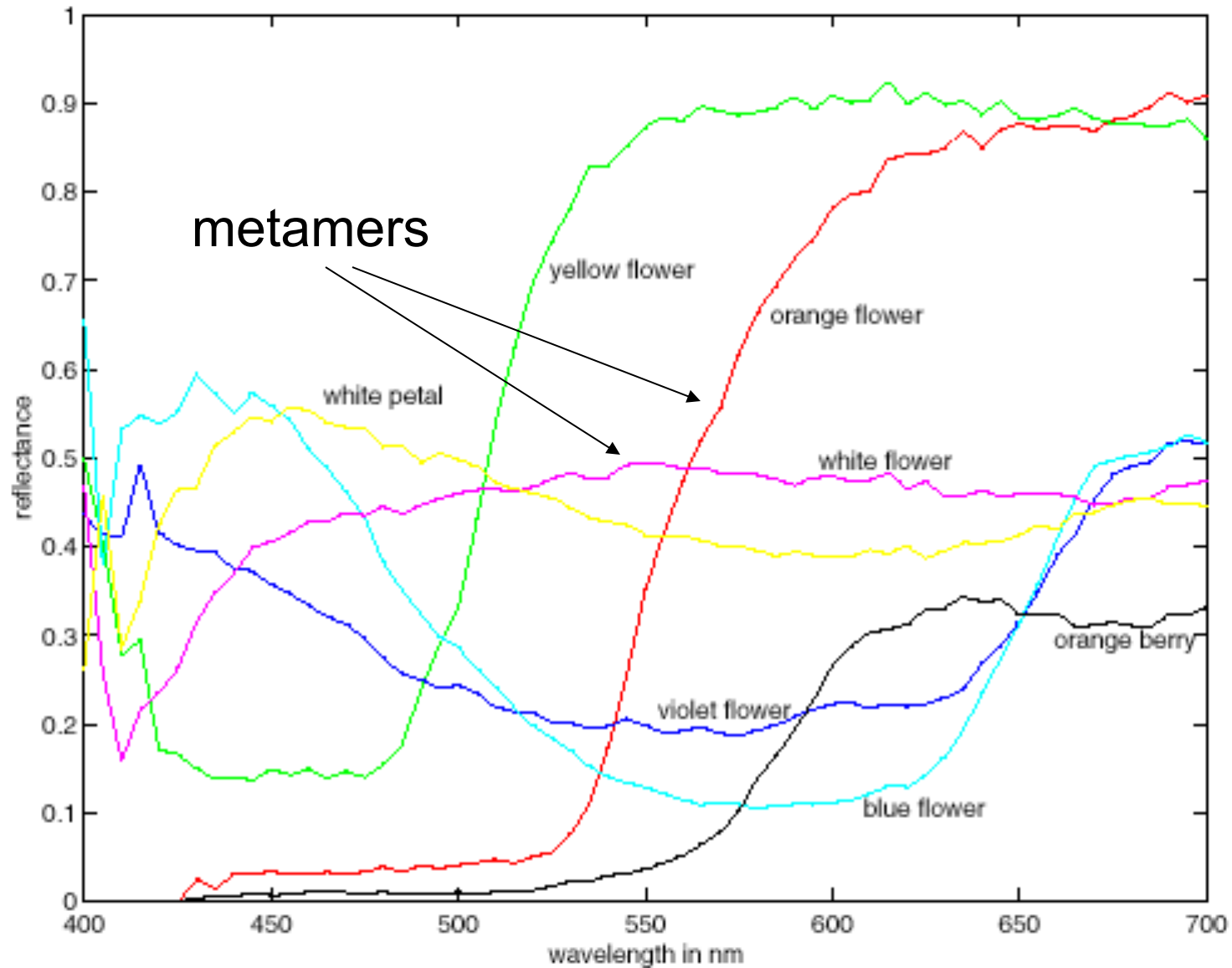
- Why are M and L cones so close?
- Why are there 3?

Tetrachromatism



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

More Spectra

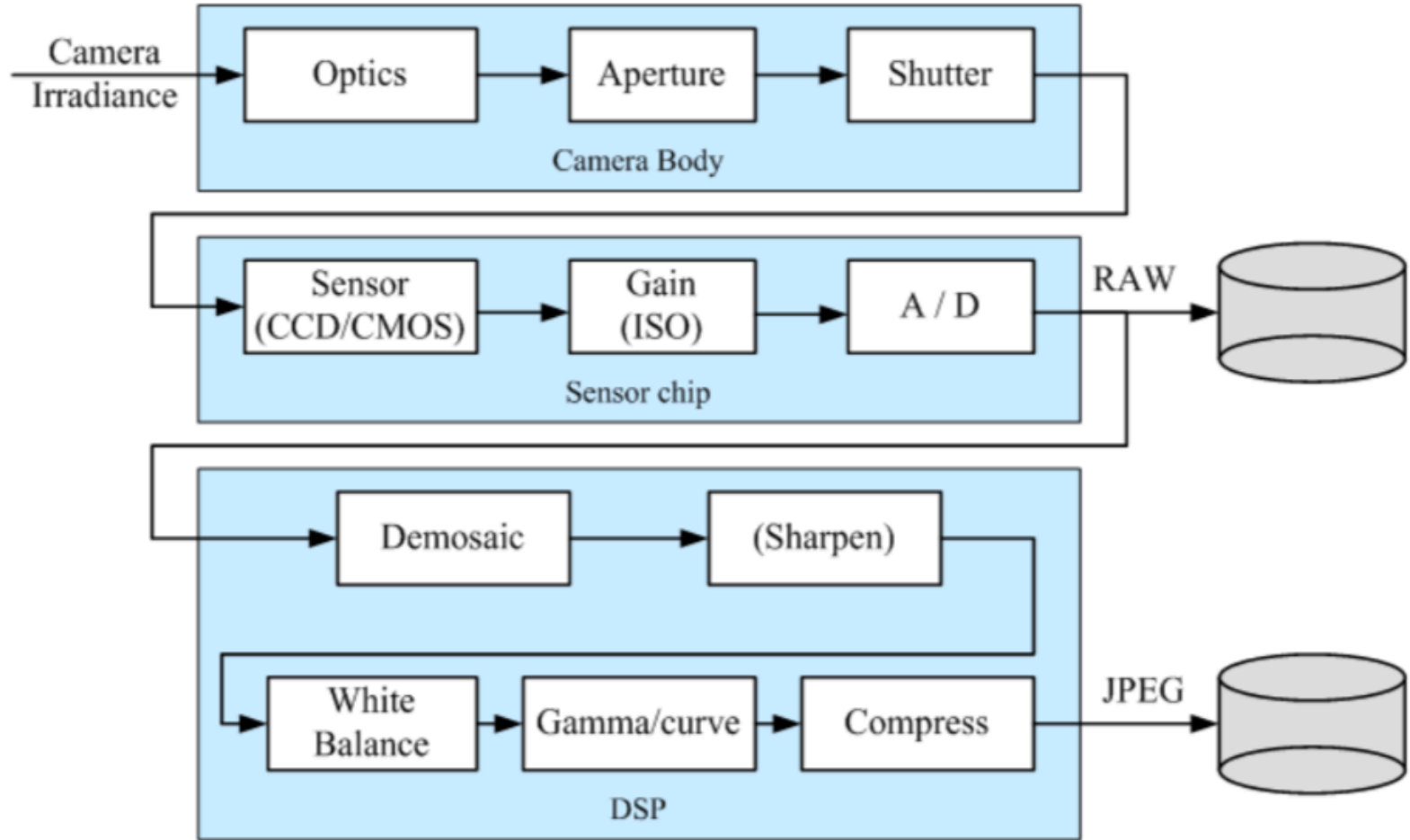


2.3 The Digital camera

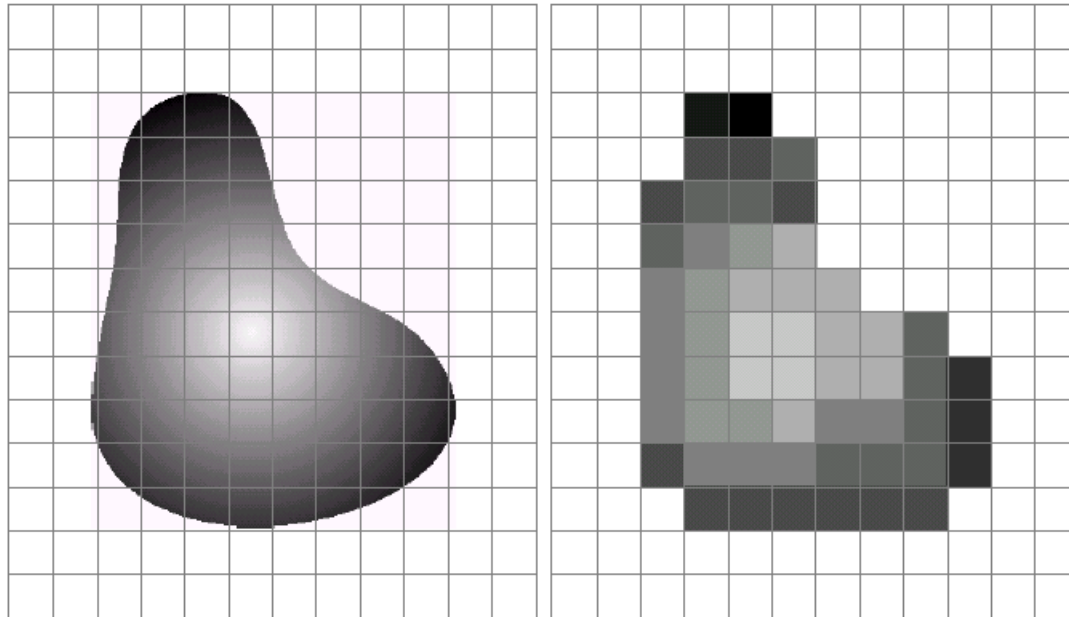


- A digital camera replaces film with a sensor array
 - Each cell in the array is light-sensitive diode that converts photons to electrons
 - Two common types:
 - Charge Coupled Device (CCD)
 - CMOS
 - <http://electronics.howstuffworks.com/digital-camera.htm>

The sensing pipeline



Sensor Array



a b

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



CMOS sensor

Sampling and Quantization

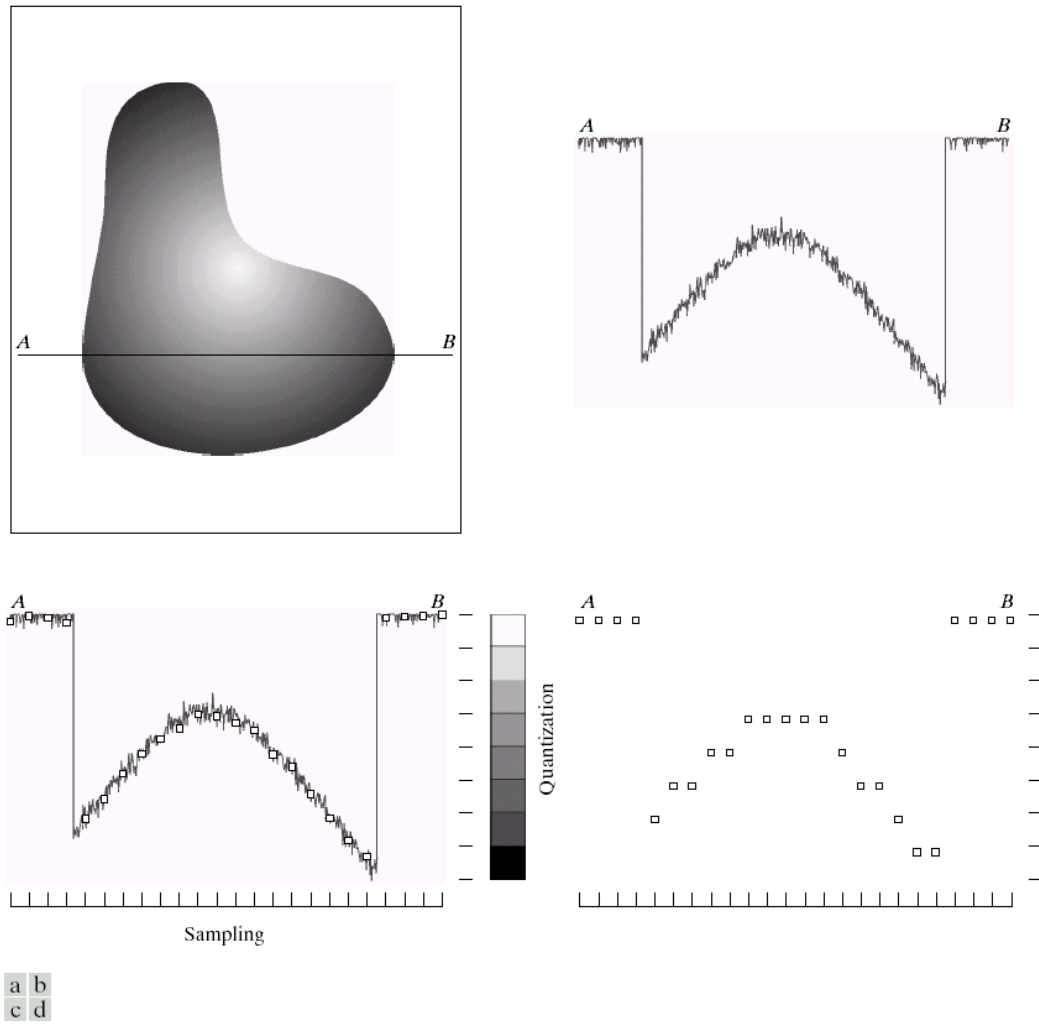
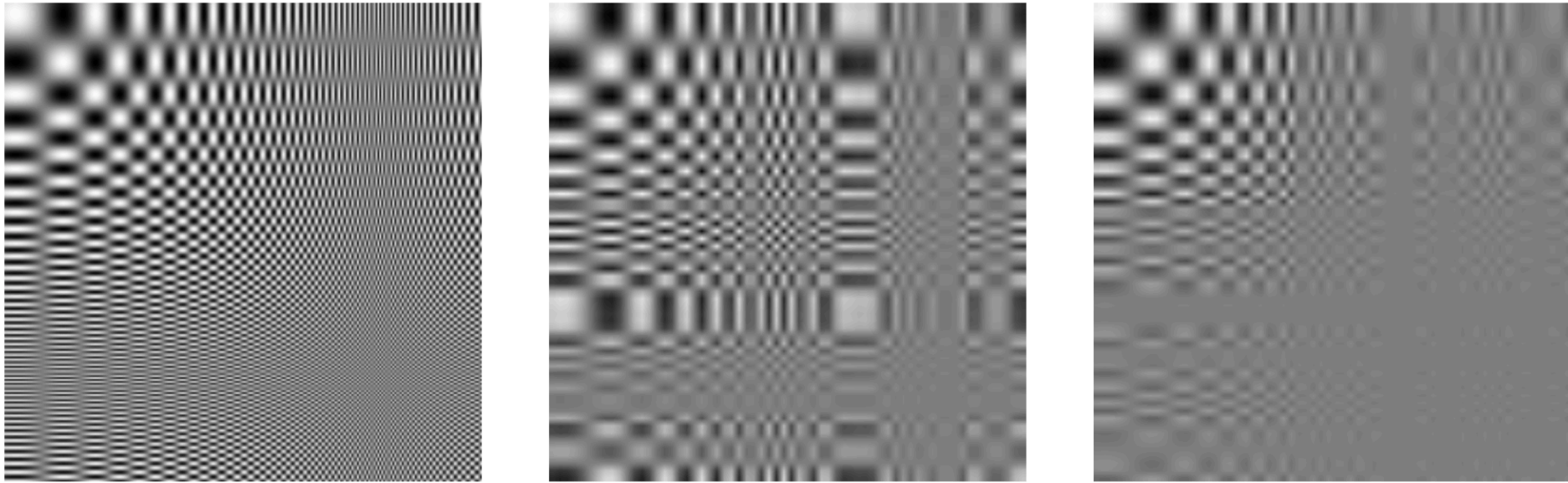
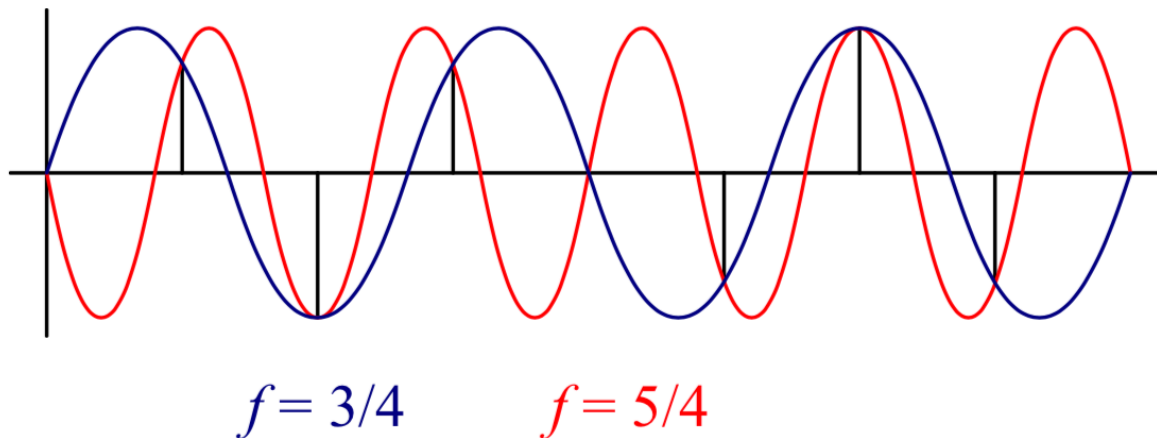


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

2.3.1 Sampling and Aliasing



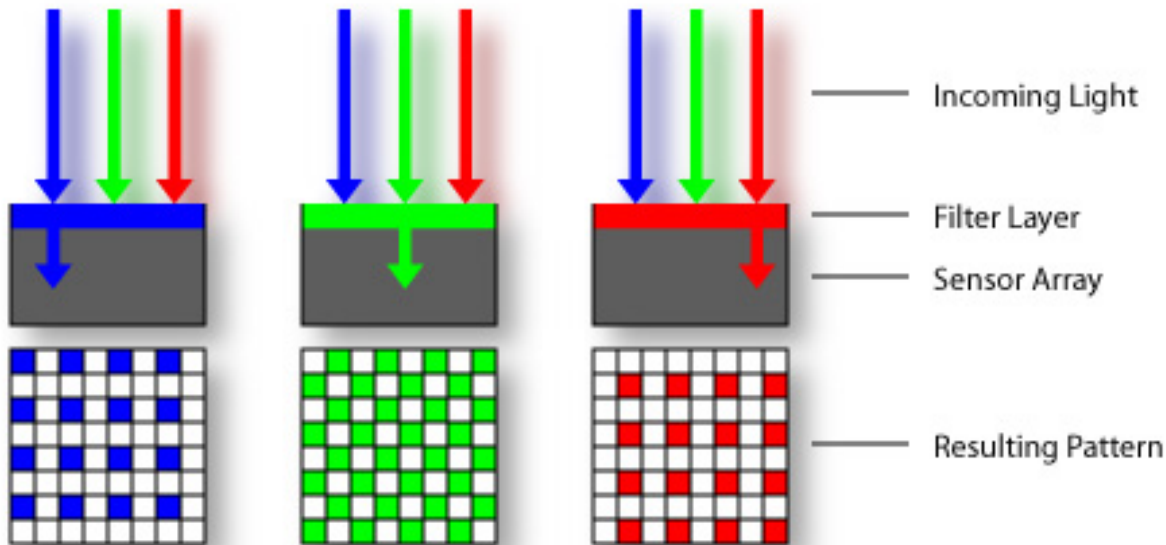
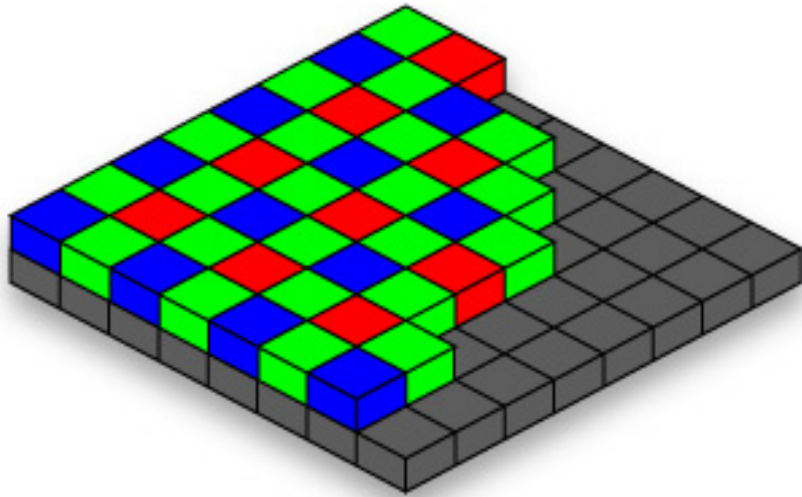
Violation of Shannon's sampling theorem: $f_s \geq 2 f_{\max}$



Rolling Shutter

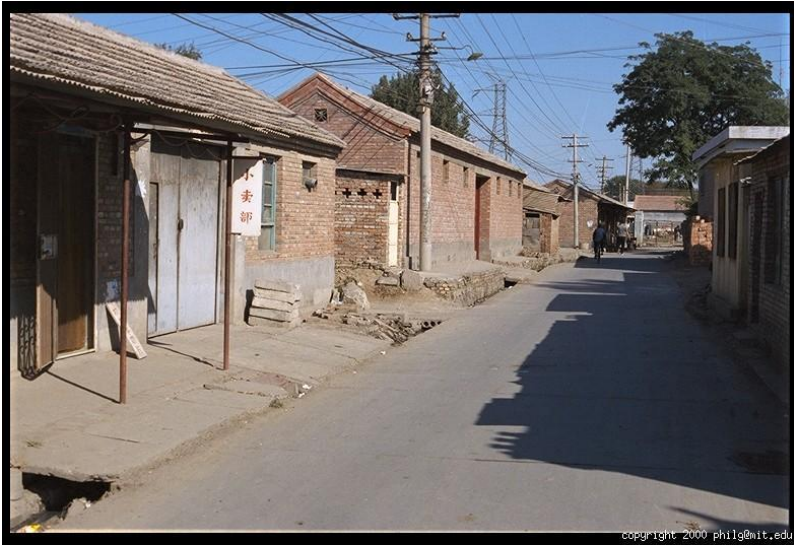


2.3.2 Color: the Bayer grid



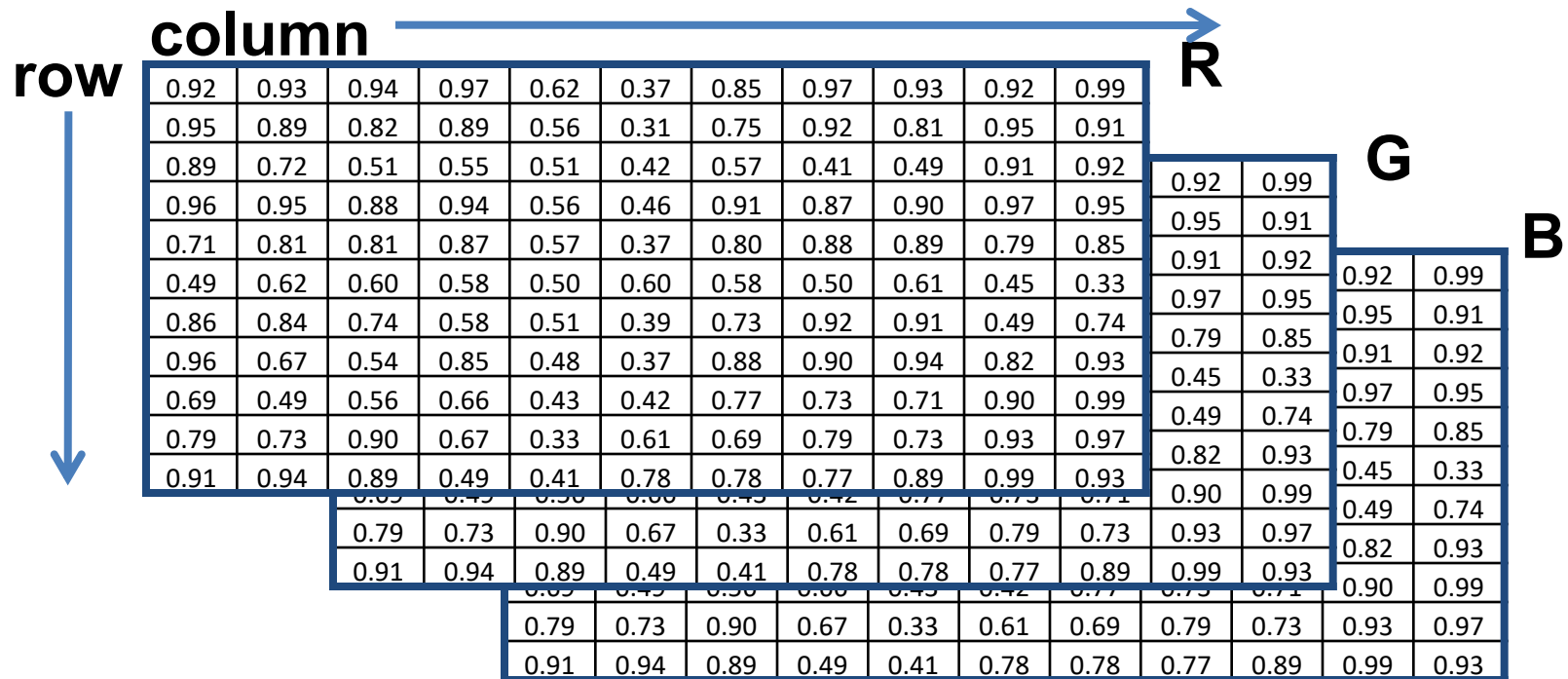
- Estimate RGB at 'G' cells from neighboring values

Color Image



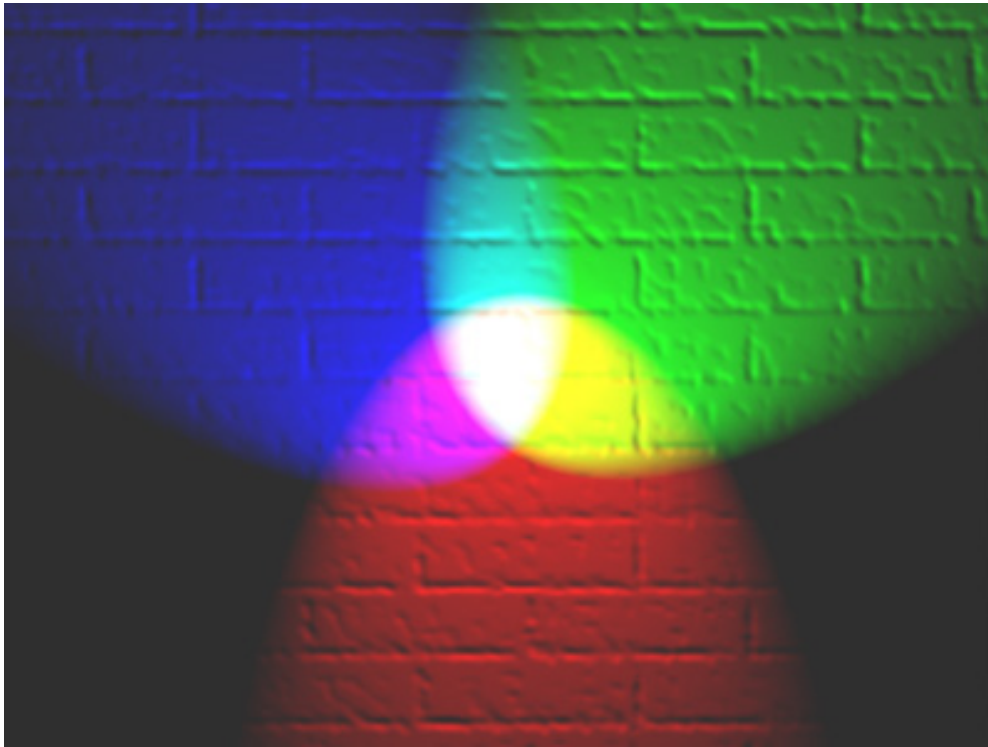
Images in ~~Matlab~~ Python

- 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 bth channel
 - $im(N-1, M-1, 2)$ = bottom-right pixel in B-channel



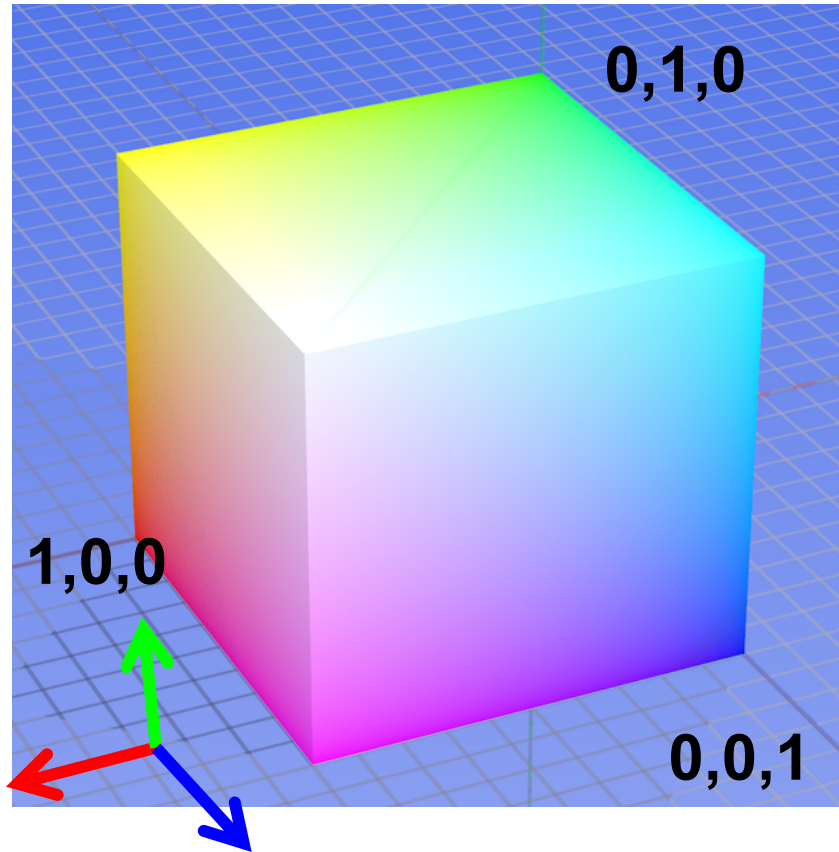
Color spaces

- How can we represent color?



Color spaces: RGB

Default color space



R
(G=0,B=0)



G
(R=0,B=0)



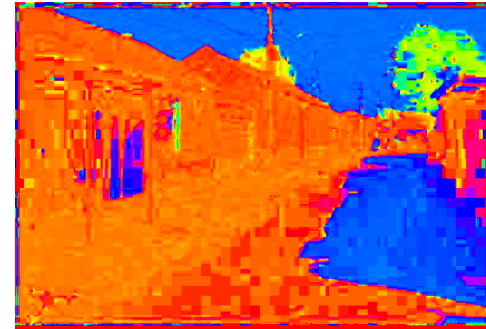
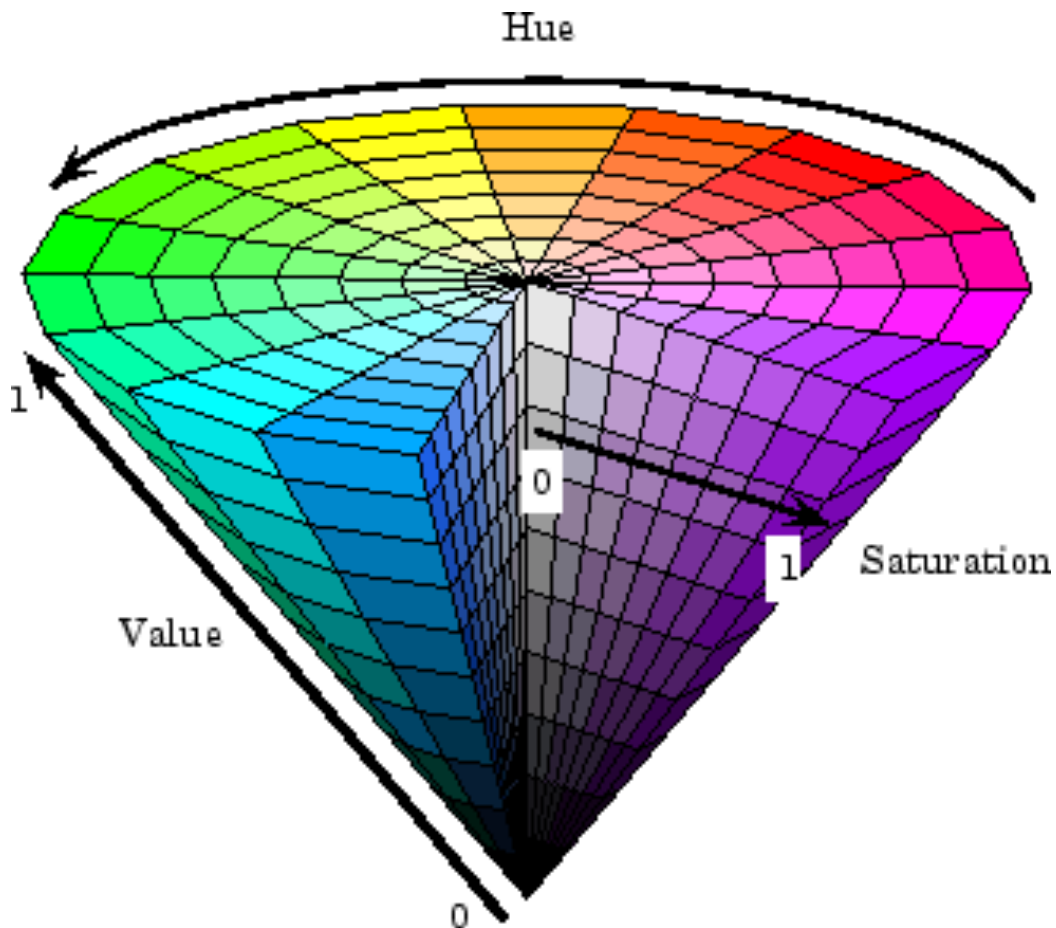
B
(R=0,G=0)

Some drawbacks

- Strongly correlated channels
- Non-perceptual

Color spaces: HSV

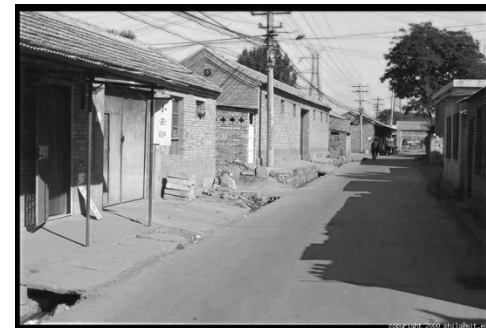
Intuitive color space



H
(S=1,V=1)



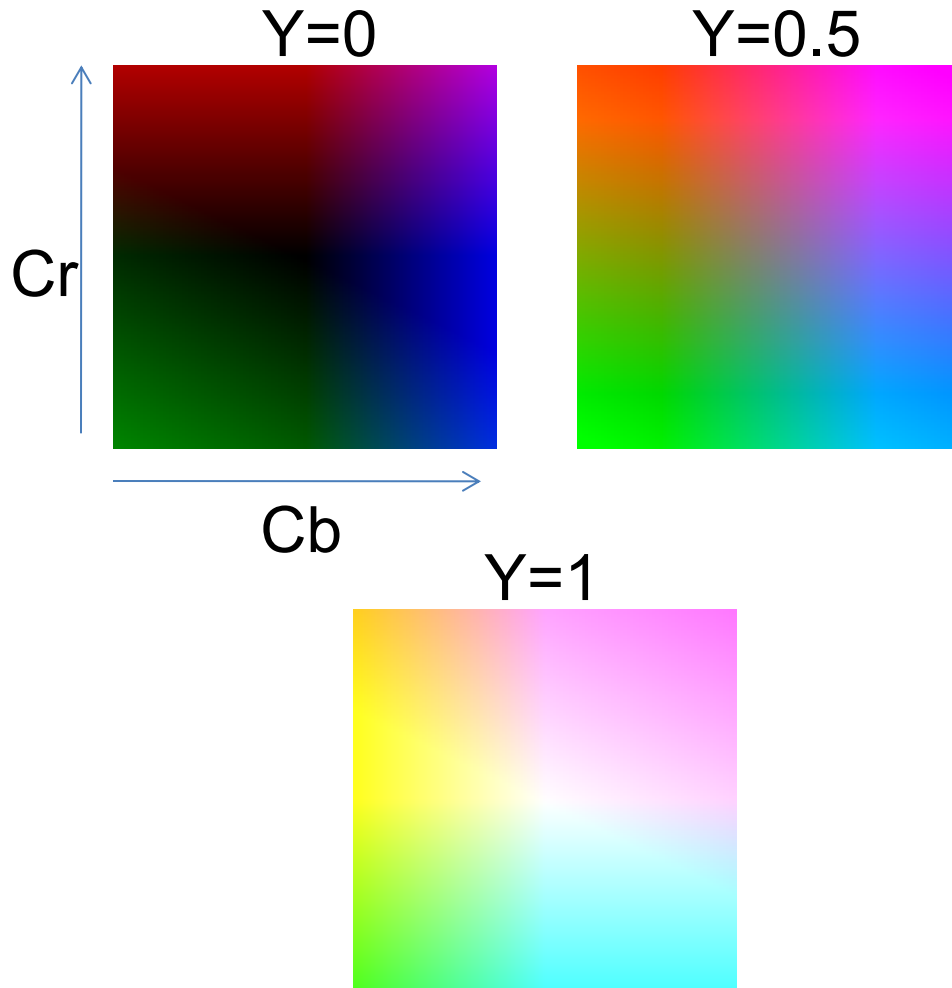
S
(H=1,V=1)



V
(H=1,S=0)

Color spaces: YCbCr

Fast to compute, good for compression, used by TV



Y
(Cb=0.5,Cr=0.5)



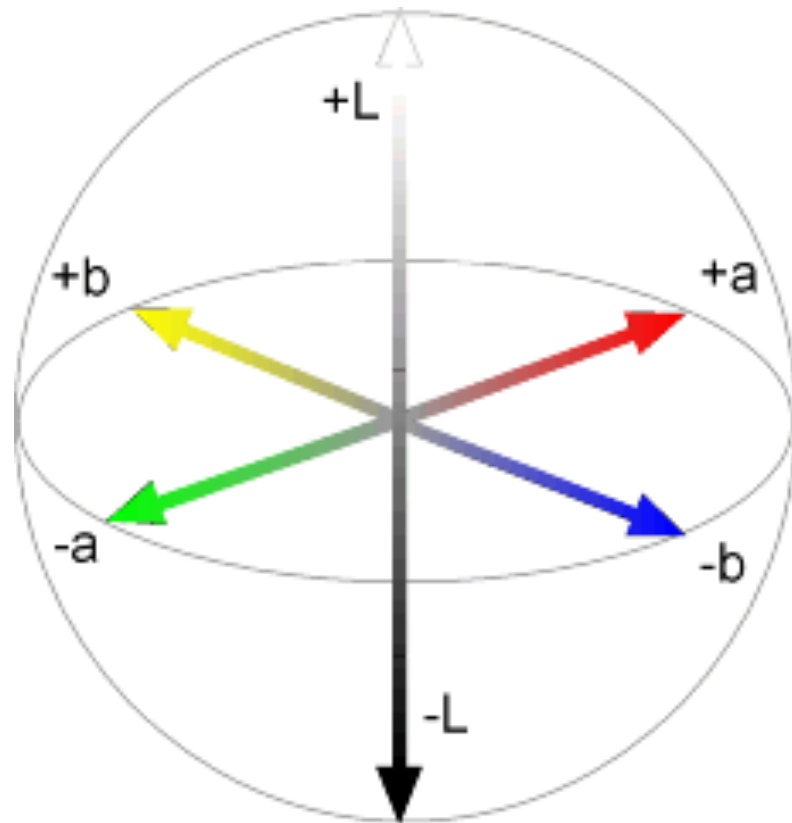
Cb
(Y=0.5,Cr=0.5)



Cr
(Y=0.5,Cb=0.5)

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



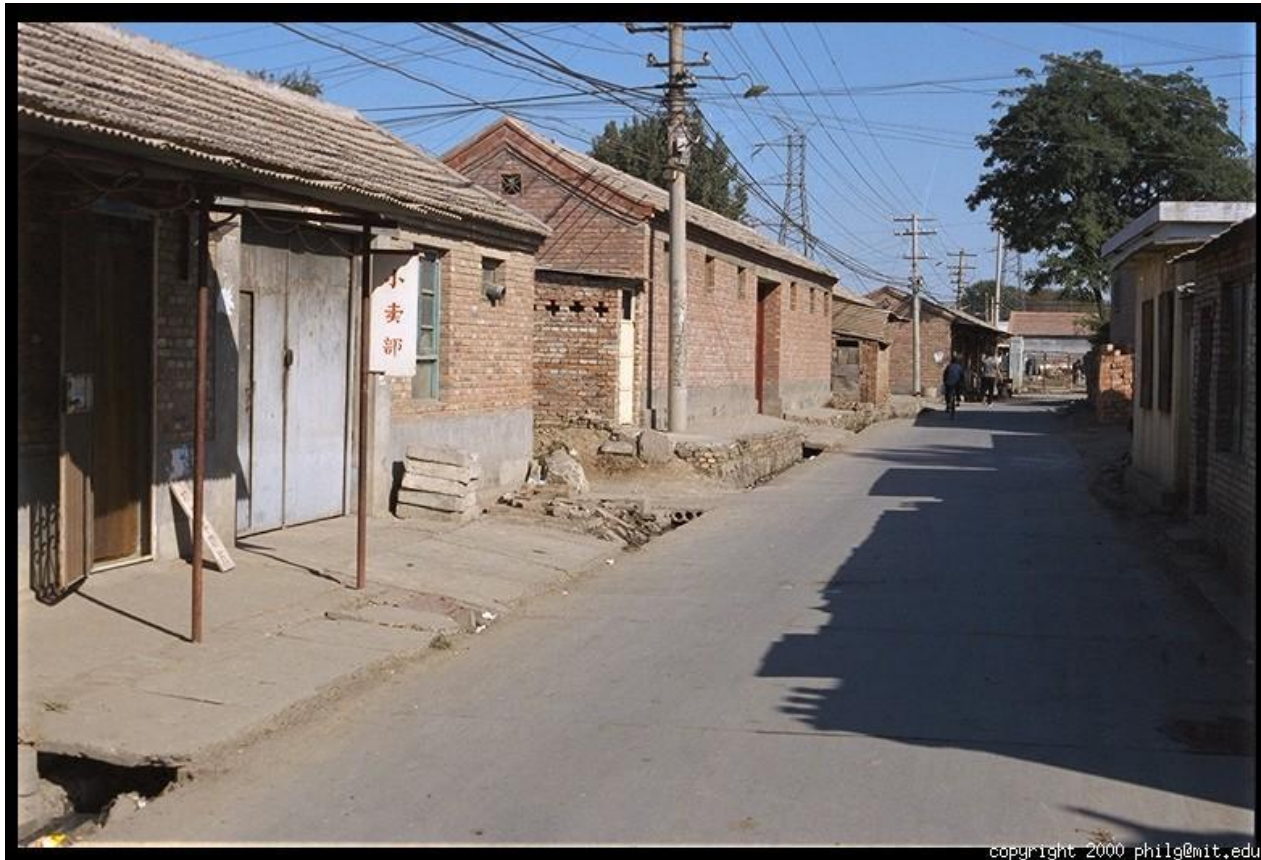
Only color shown – constant intensity

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.