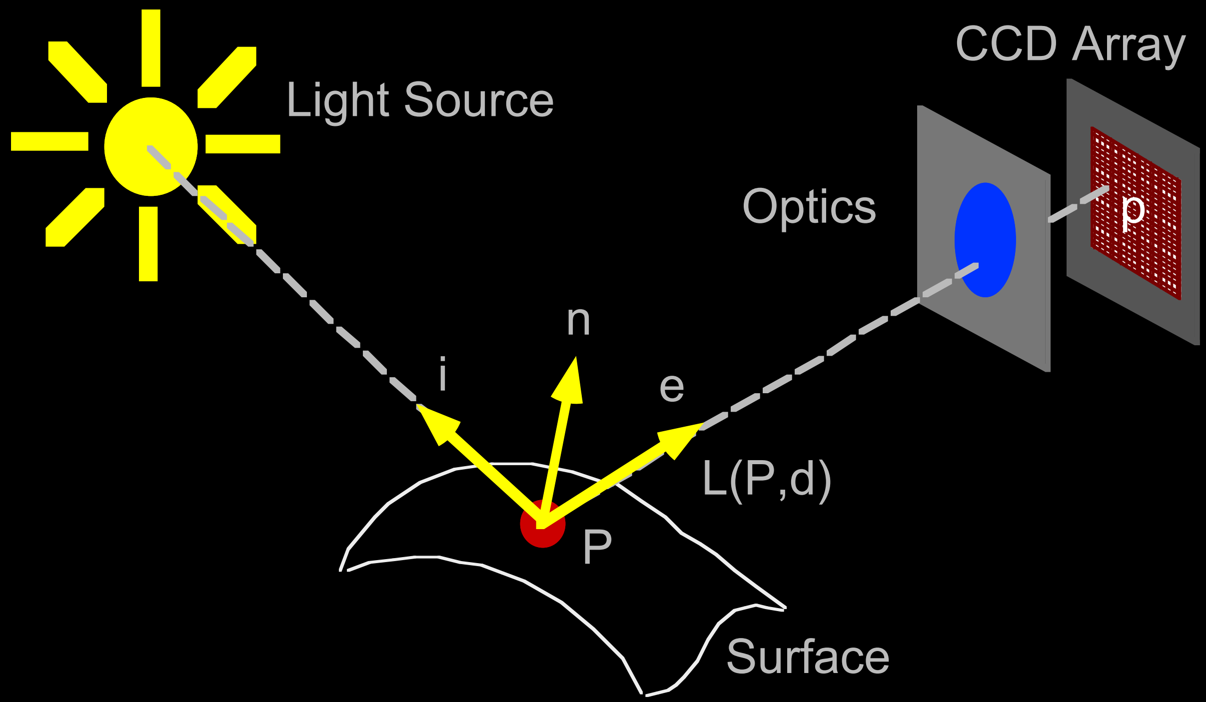
Overview of Image Formation

- Light and Optics
  - Pinhole camera model
  - Perspective projection
  - Thin lens model
  - Fundamental equation
  - Distortion: spherical & chromatic aberration, radial distortion
  - Reflection and Illumination: color, Lambertian and specular surfaces, Phong, BRDF
- Sensing Light
- Conversion to Digital Images
- Sampling Theorem
- Other Sensors: frequency, type, ….

Basic Radiometry

- Radiometry is the part of image formation concerned with the relation among the amounts of light energy emitted from light sources, reflected from surfaces, and registered by sensors.



Lecture Assumptions

- Typical imaging scenario:
  - visible light
  - ideal lenses
  - standard sensor (e.g. TV camera)
  - opaque objects
- Goal

To create 'digital' images which can be processed to recover some of the characteristics of the 3D world which was imaged.

Steps

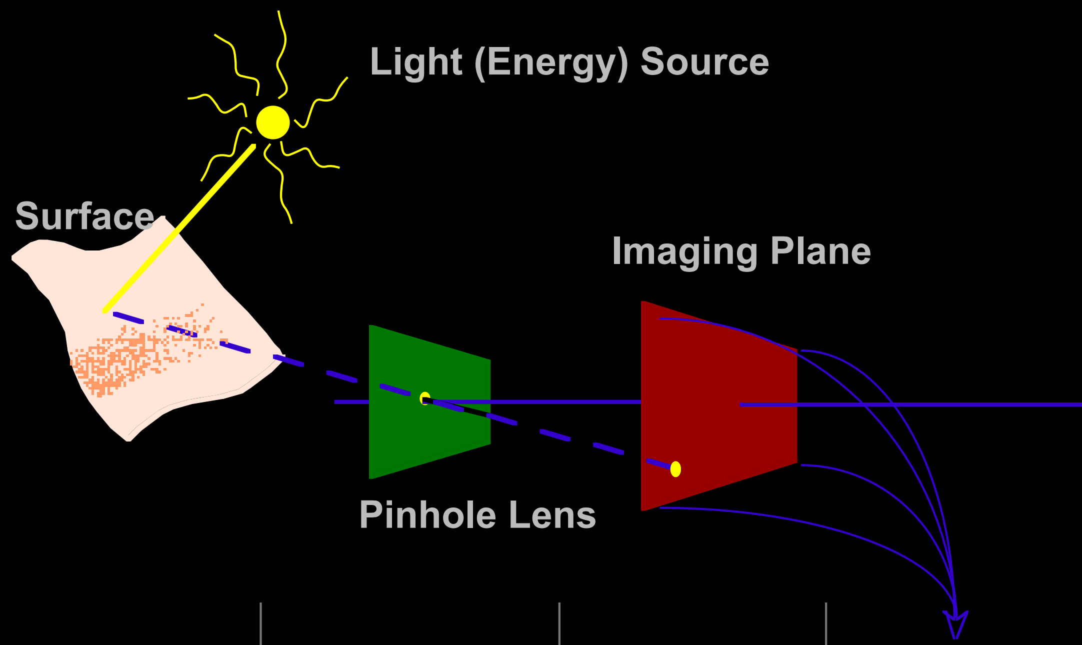
World → Optics → Sensor →
Signal → Digitizer →
Digital Representation

World	reality
Optics	focus light from world on sensor
Sensor	converts light to electrical energy
Signal	representation of incident light as continuous electrical energy
Digitizer	converts continuous signal to discrete signal
Digital Rep.	final representation of reality in computer memory

Factors in Image Formation

- Geometry
  - concerned with the relationship between points in the three-dimensional world and their images
- Radiometry
  - concerned with the relationship between the amount of light radiating from a surface and the amount incident at its image
- Photometry
  - concerned with ways of measuring the intensity of light
- Digitization
  - concerned with ways of converting continuous signals (in both space and time) to digital approximations

Image Formation



World	Optics	Sensor	Signal
		B&W Film	Silver Density
		Color Film	Silver density in three color layers
		TV Camera	Electrical

- Geometry describes the projection of:

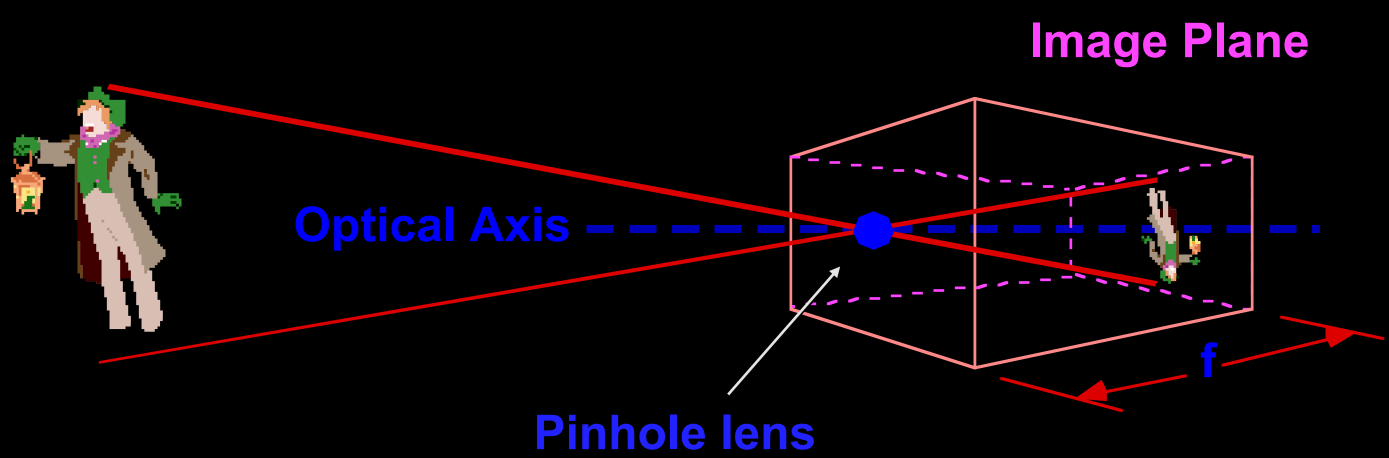
  three-dimensional (3D) world → two-dimensional (2D) image plane.

- Typical Assumptions
  - Light travels in a straight line
- Optical Axis: the perpendicular from the image plane through the pinhole (also called the central projection ray)
- Each point in the image corresponds to a particular direction defined by a ray from that point through the pinhole.
- Various kinds of projections:
  - - perspective - oblique
  - - orthographic - isometric
  - - spherical

Basic Optics

- Two camera models are commonly used:
  - Pin-hole camera
  - Optical system composed of lenses
- Pin-hole is the basis for most graphics and vision
  - Derived from physical construction of early cameras
  - Mathematics is very straightforward
- Thin lens model is first of the lens models
  - Mathematical model for a physical lens
  - Lens gathers light over area and focuses on image plane.

Pinhole Camera Model



- World projected to 2D Image
  - Image inverted
  - Size reduced (usually)
  - Image is dim
  - No direct depth information
- f called the focal length of the “lens”
- Known as perspective projection

Pinhole camera image

Amsterdam

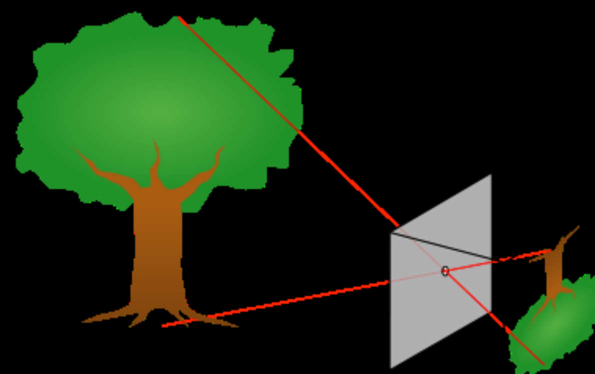
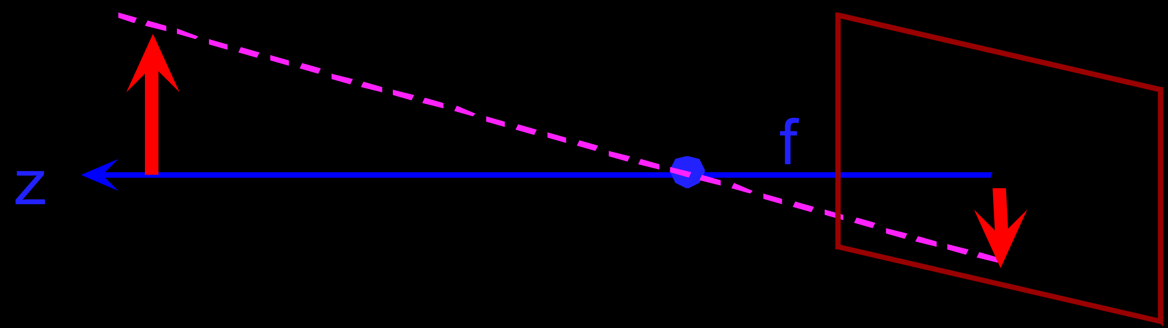


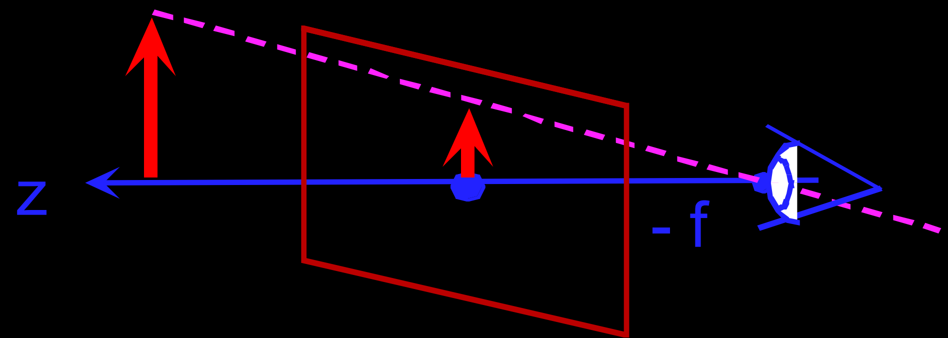
Photo by Robert Kosara, robert@kosara.net
http://www.kosara.net/gallery/pinholeamsterdam/pic01.html

Equivalent Geometry

- Consider case with object on the optical axis:



- More convenient with upright image:

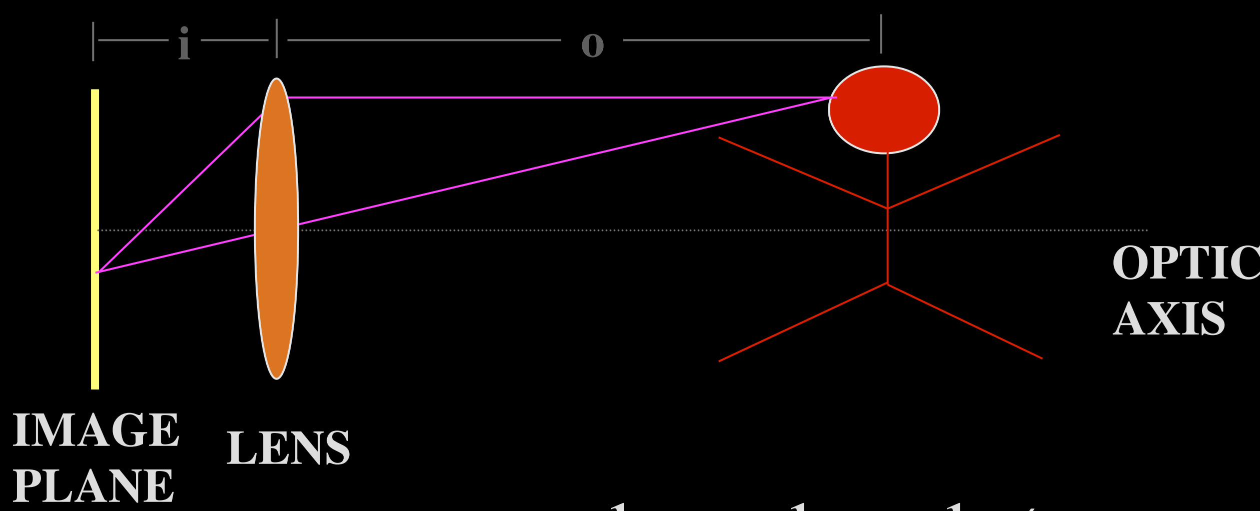


Projection plane z = 0

- Equivalent mathematically

Thin Lens Model

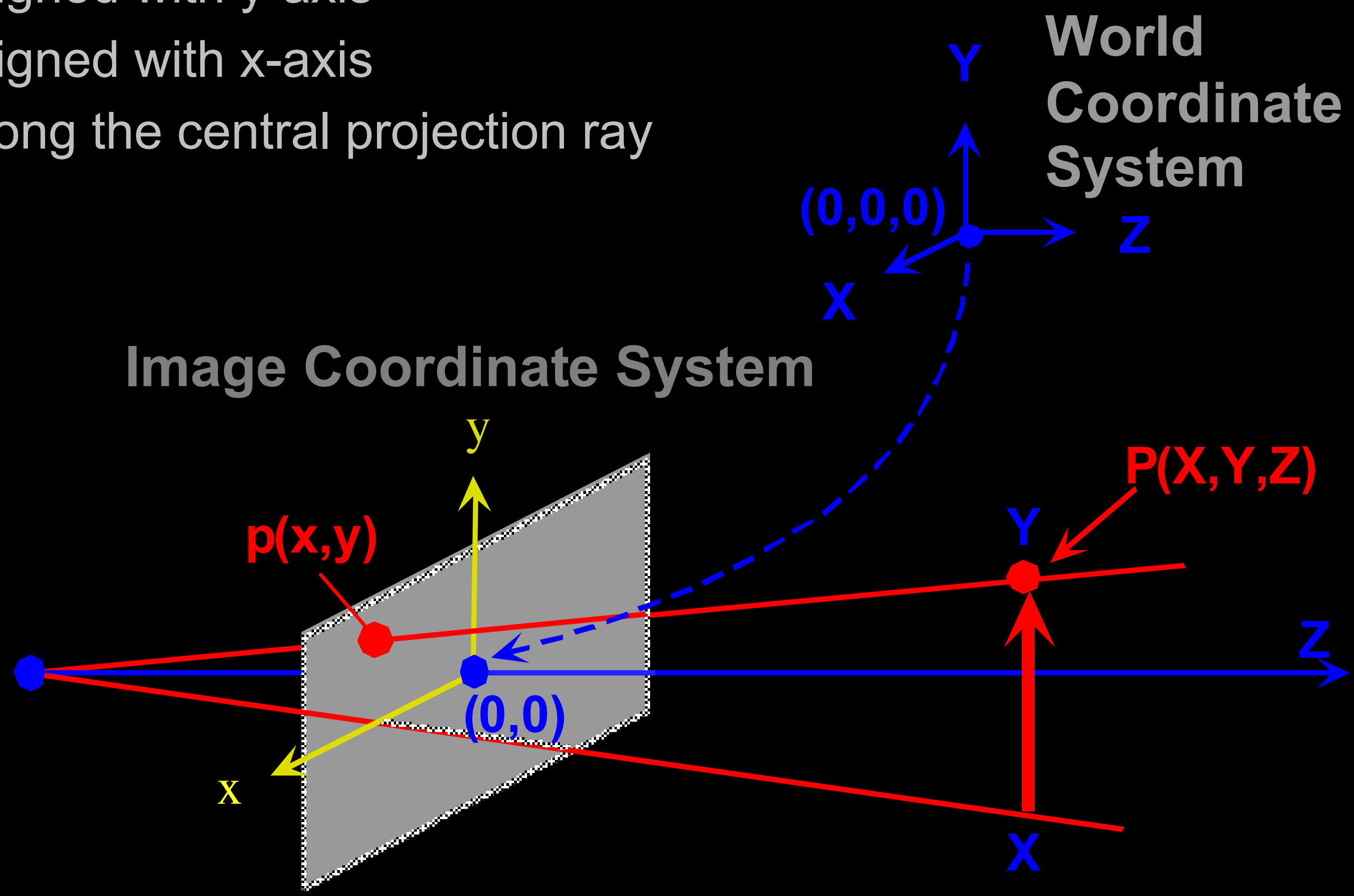
- Rays entering parallel on one side converge at focal point.
- Rays diverging from the focal point become parallel.



$$\frac{1}{f} = \frac{1}{i} + \frac{1}{o}$$ 'THIN LENS LAW'

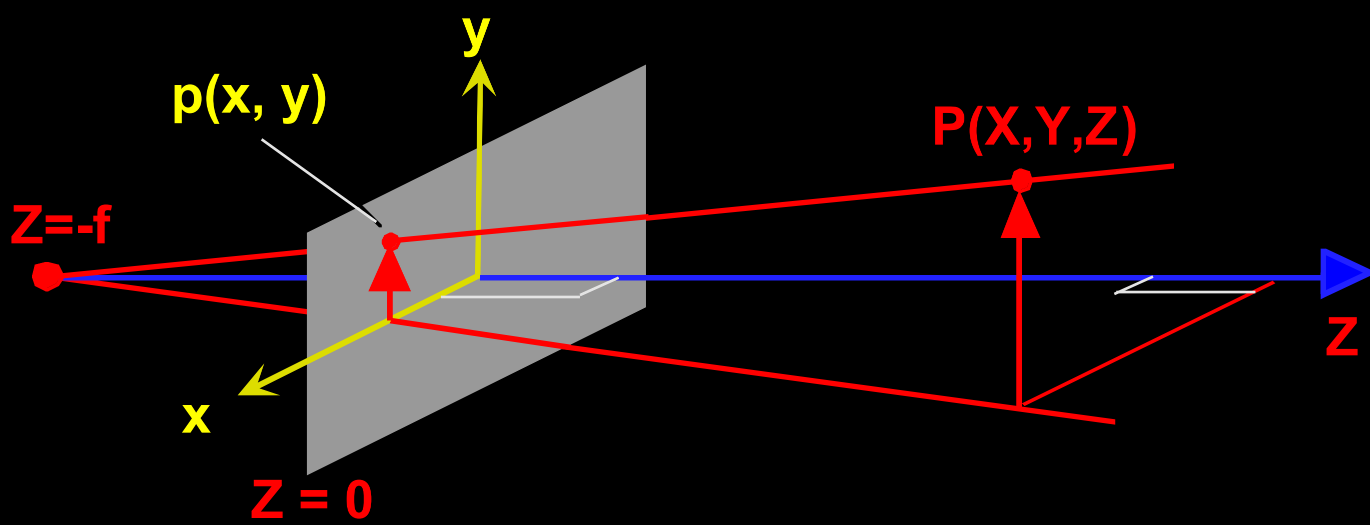
Coordinate System

- Simplified Case:
  - Origin of world and image coordinate systems coincide
  - Y-axis aligned with y-axis
  - X-axis aligned with x-axis
  - Z-axis along the central projection ray



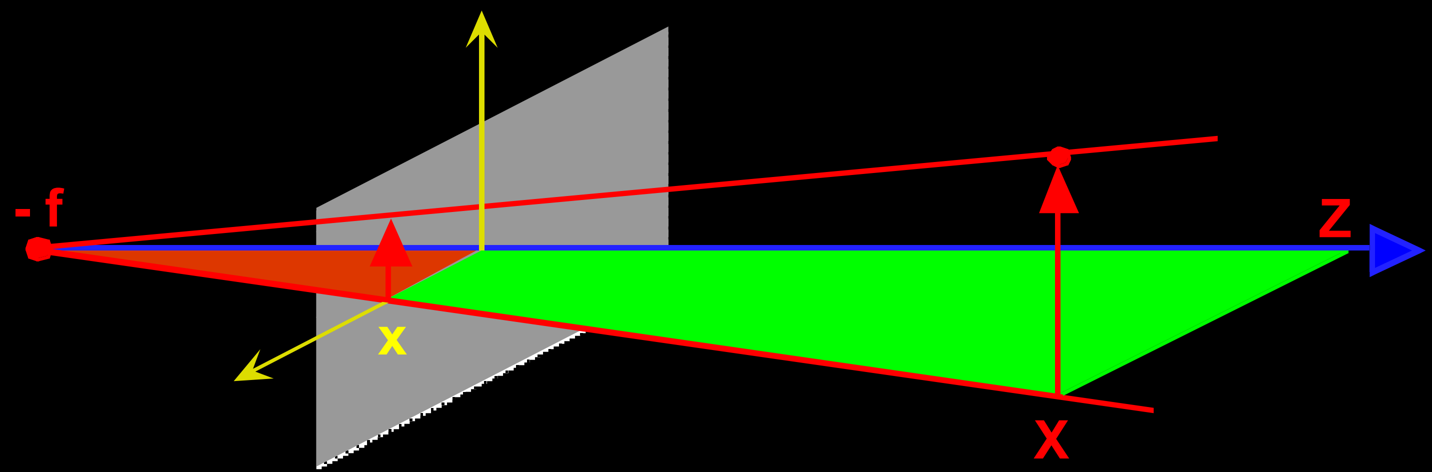
Perspective Projection

- Compute the image coordinates of p in terms of the world coordinates of P.



- Look at projections in x-z and y-z planes

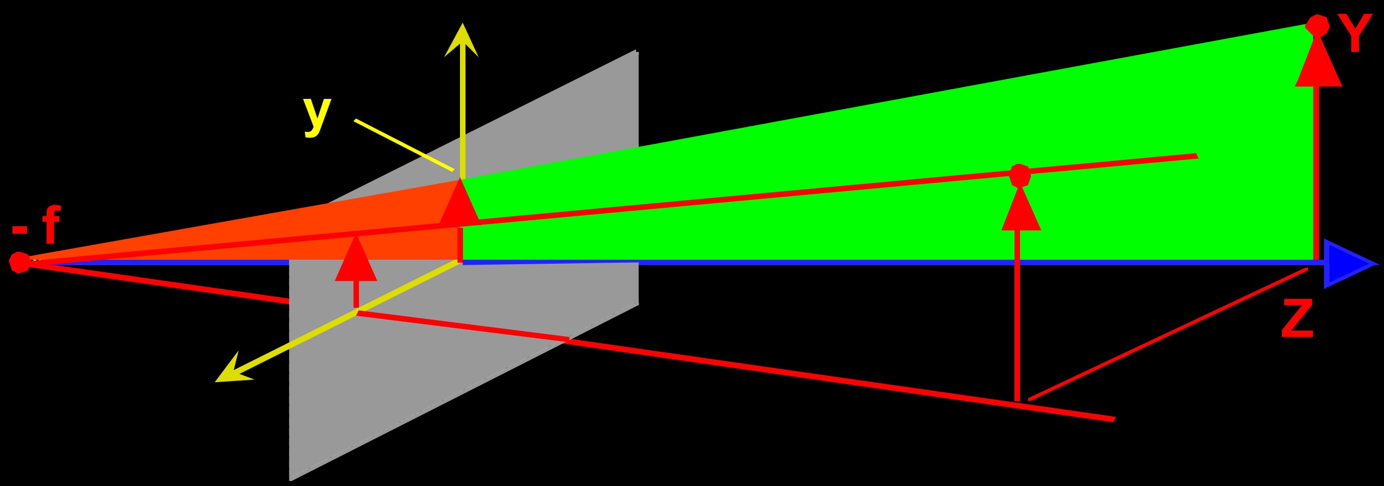
X-Z Projection



- By similar triangles: $\frac{x}{f} = \frac{X}{Z+f}$

$$x = \frac{fX}{Z+f}$$

Y-Z Projection



- By similar triangles: $\frac{y}{f} = \frac{Y}{Z+f}$

$$y = \frac{fY}{Z+f}$$

Perspective Equations

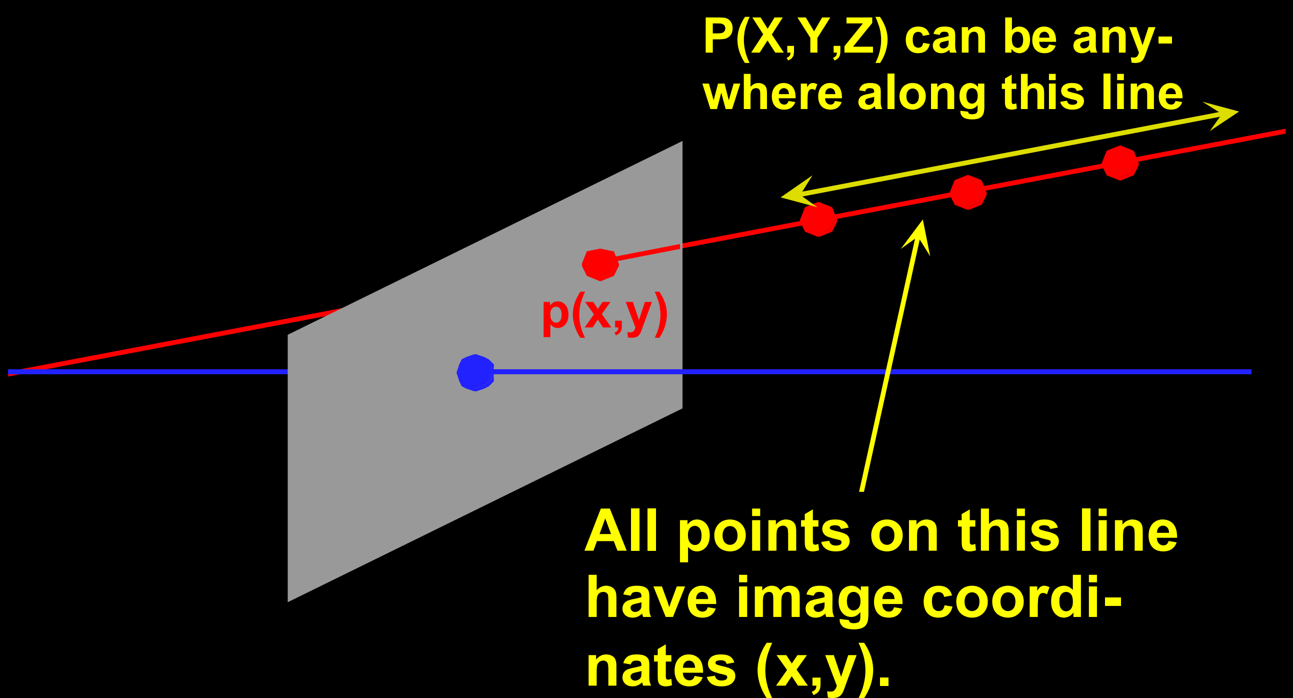
- Given point P(X,Y,Z) in the 3D world
- The two equations:

$$x = \frac{fX}{Z+f} \qquad y = \frac{fY}{Z+f}$$

- transform world coordinates (X,Y,Z)
  into image coordinates (x,y)

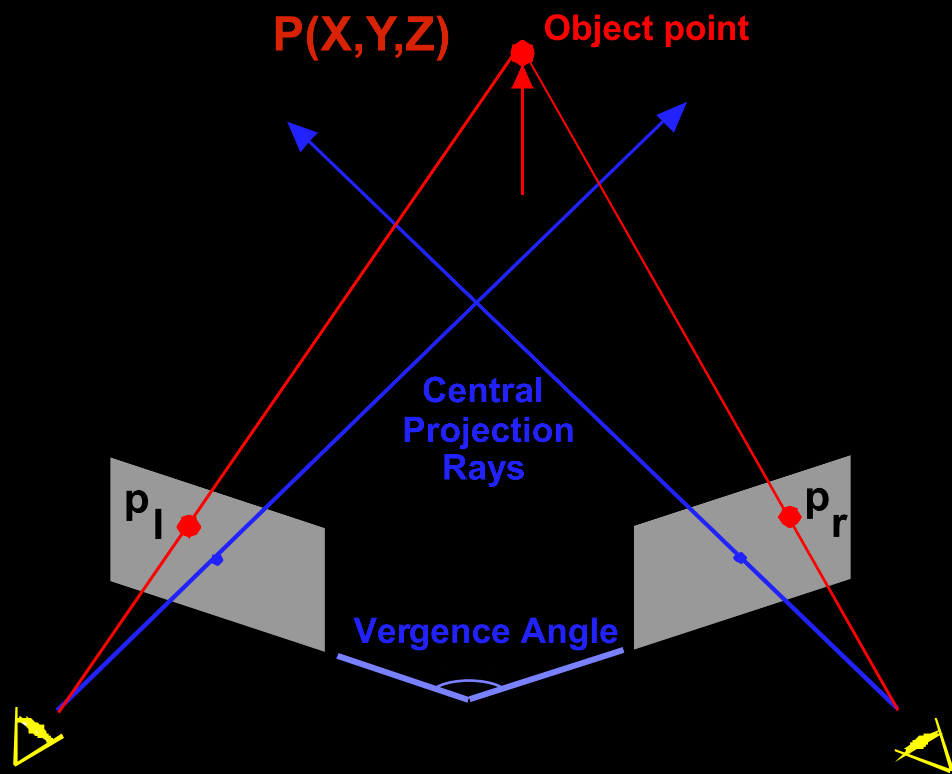
Reverse Projection

- Given a center of projection and image coordinates of a point, it is not possible to recover the 3D depth of the point from a single image.



In general, at least two images of the same point taken from two different locations are required to recover depth.

Stereo Geometry



- Depth obtained by triangulation
- Correspondence problem: p_l and p_r must correspond to the left and right projections of P, respectively.

Variability in appearance

- Consequences of image formation geometry for computer vision
  - What set of shapes can an object take on if it is:
    - rigid
    - non-rigid
    - planar
    - non-planar
  - SIFT features:
    - Deals with variability of rigid, planar shapes under perspective distortion, or piecewise rigid, planar shapes.
- Sensitivity to errors.

Radiometry and Light Sources

Radiometry

- Brightness: informal notion used to describe both scene and image brightness.
- Image brightness: related to energy flux incident on the image plane:

IRRADIANCE (illuminance)

"It's a bright day."

- Scene brightness: brightness related to energy flux emitted (radiated) from a surface.

RADIANCE (luminance)

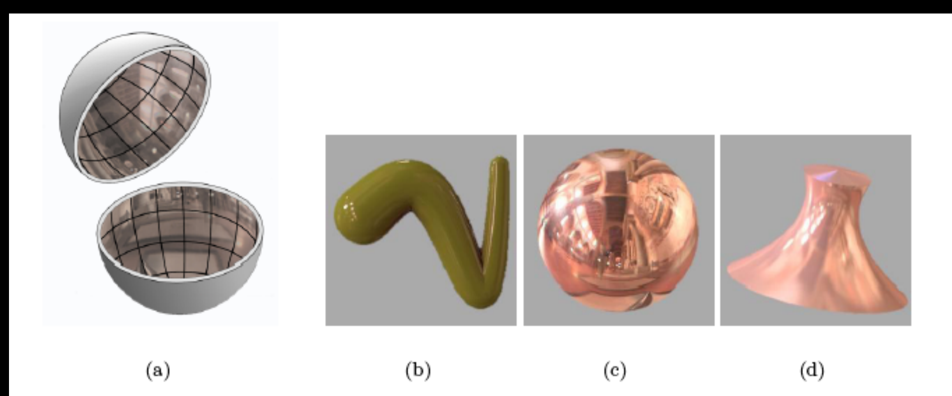
"Yikes, that shirt is way too bright!"

Light and Surfaces

- Reflection
  - mirrors
  - highlights
  - specularities
- Scattering
  - Lambertian
  - matte
  - diffuse

Light sources

- Point source
- Extended source
- Single wavelength
- Multi-wavelength
- Uniform
- Non-uniform

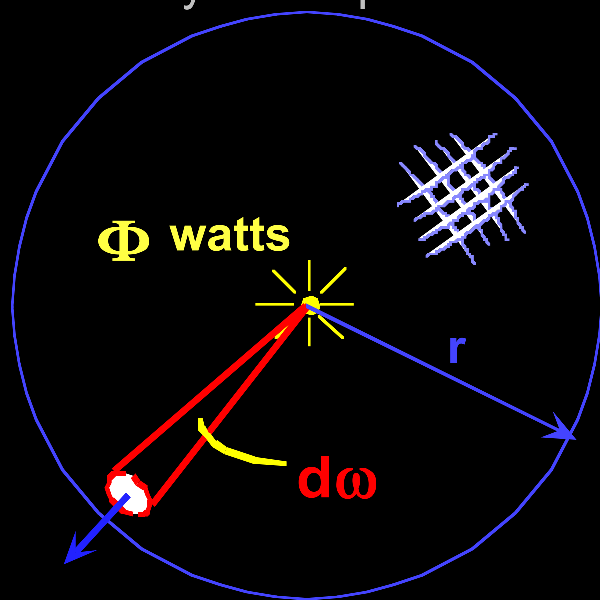


Linearity of Light

- Linearity
  - definition: For a linear function f(x), we have:
    - Additivity: $f(x+y) = f(x) + f(y)$.
    - Homogeneity : $f(ax) = a f(x)$.
- For extended sources
- For multiple wavelengths
- Across time

Power and P. S. Radiant Intensity

- Point source power: watts.
- F watts radiated into 4pi steradians
- Point source radiant intensity: watts per steradian.

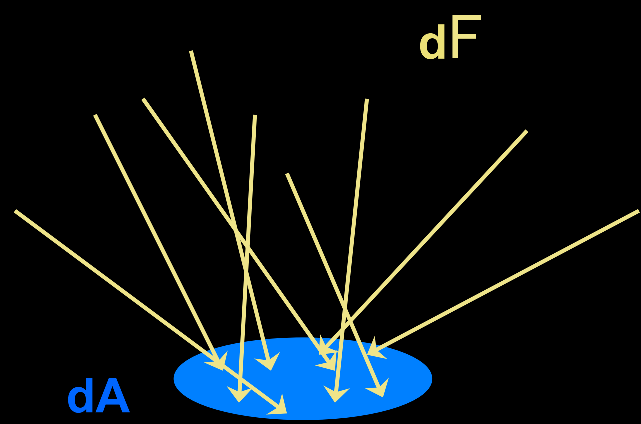


$$F = \int_{sphere} dF$$

R = Point Source Radiant Intensity = $\frac{dF}{dw}$ Watts/unit solid angle (steradian)
(of source)

Irradiance

- Light falling on a surface from all directions.
- How much?

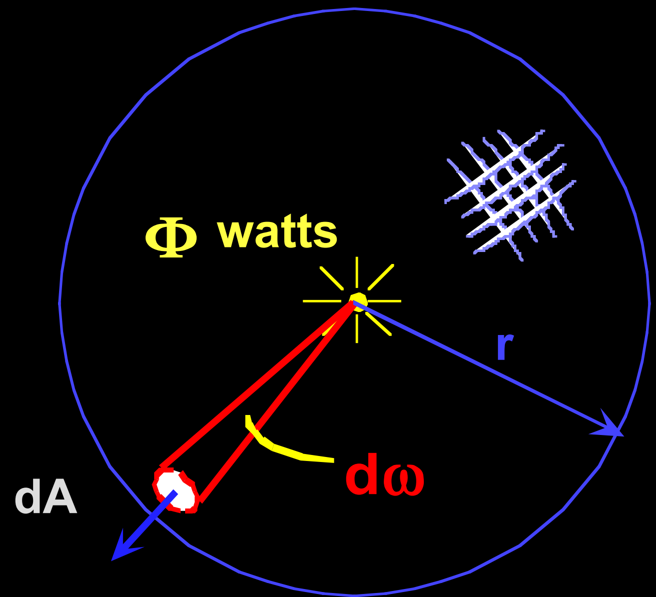


- Irradiance: power per unit area falling on a surface.

$$\text{Irradiance } E = \frac{dF}{dA} \quad \text{watts/m}^2$$

Inverse Square Law: Point Sources

- Relationship between point source radiance (radiant intensity) and irradiance



$$dw = \frac{dA}{r^2}$$

$$E = \frac{dF}{dA}$$

R: Radiant Intensity

E: Irradiance

F: Watts

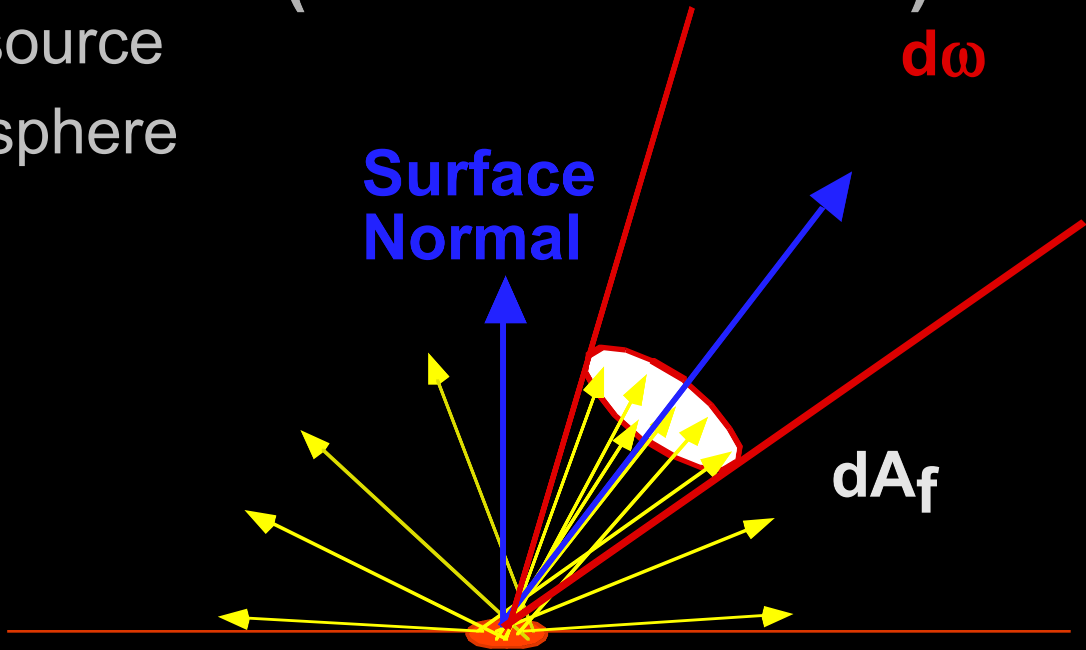
w : Steradians

$$R = \frac{dF}{dw} = \frac{r^2\ dF}{dA} = r^2 E$$

$$E = \frac{R}{r^2}$$

Surface Radiance (Extended source)

- Surface acts as light source
- Radiates over a hemisphere



R: Radiant Intensity

E: Irradiance

L: Surface radiance

- Surface Radiance: power per unit foreshortened area emitted into a solid angle

$$L = \frac{dF}{dA_f \, dw}$$

(watts/m^2 steradian)

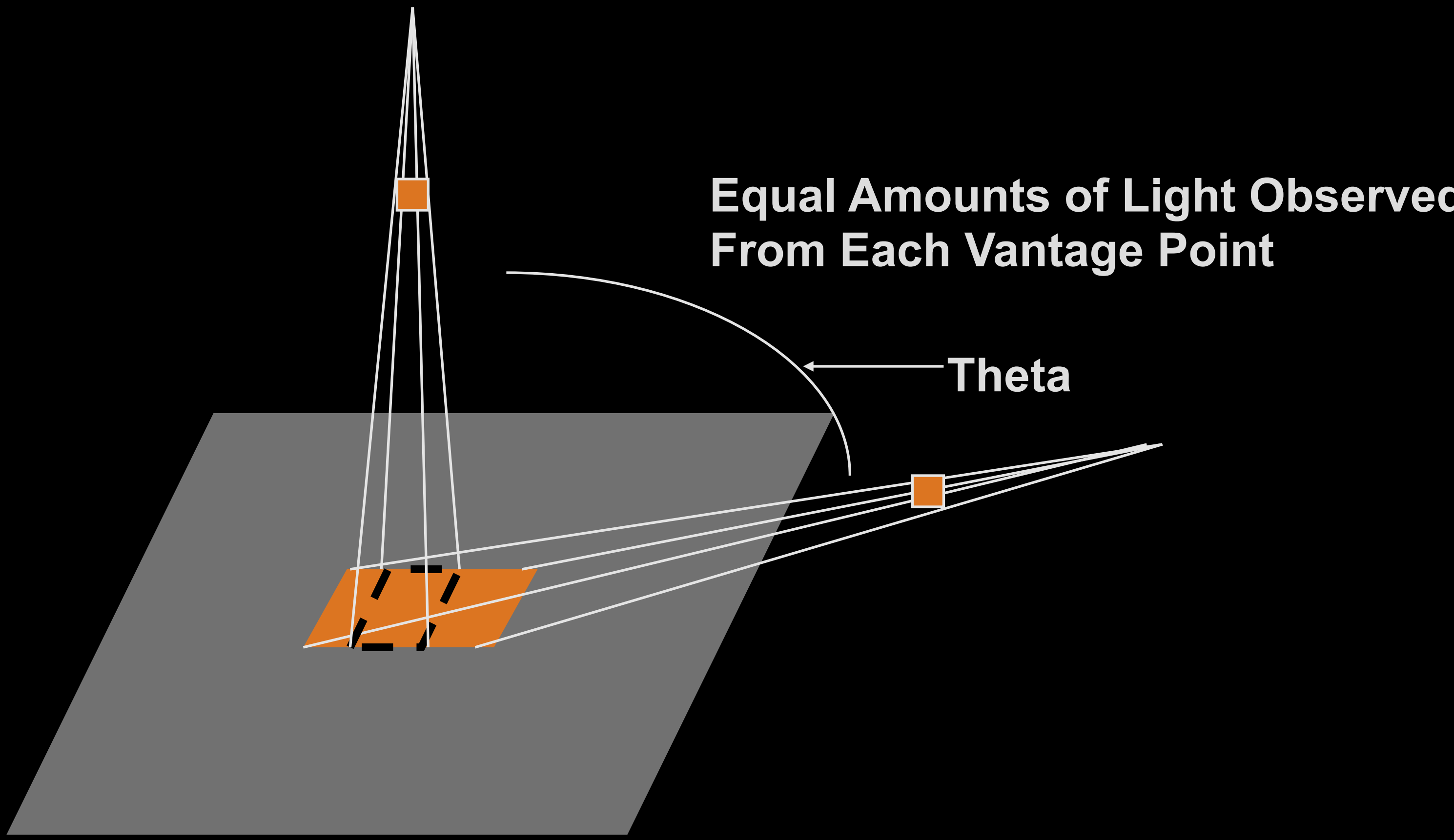
Pseudo-Radiance

- Consider two definitions:
  - Radiance:
    power per unit foreshortened area emitted into a solid angle
  - Pseudo-radiance
    power per unit area emitted into a solid angle
  - Why should we work with radiance rather than pseudo-radiance?
    - Only reason: Radiance is more closely related to our intuitive notion of “brightness”.

Lambertian Surfaces

- A particular point P on a Lambertian (perfectly matte) surface appears to have the same brightness no matter what angle it is viewed from.
  - Piece of paper
  - Matte paint
- Doesn’t depend upon incident light angle.
- What does this say about how they emit light?

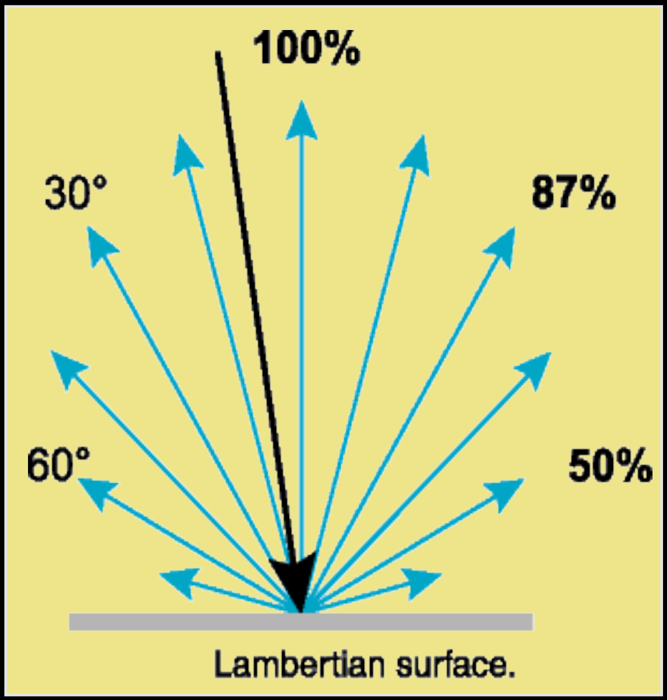
Lambertian Surfaces



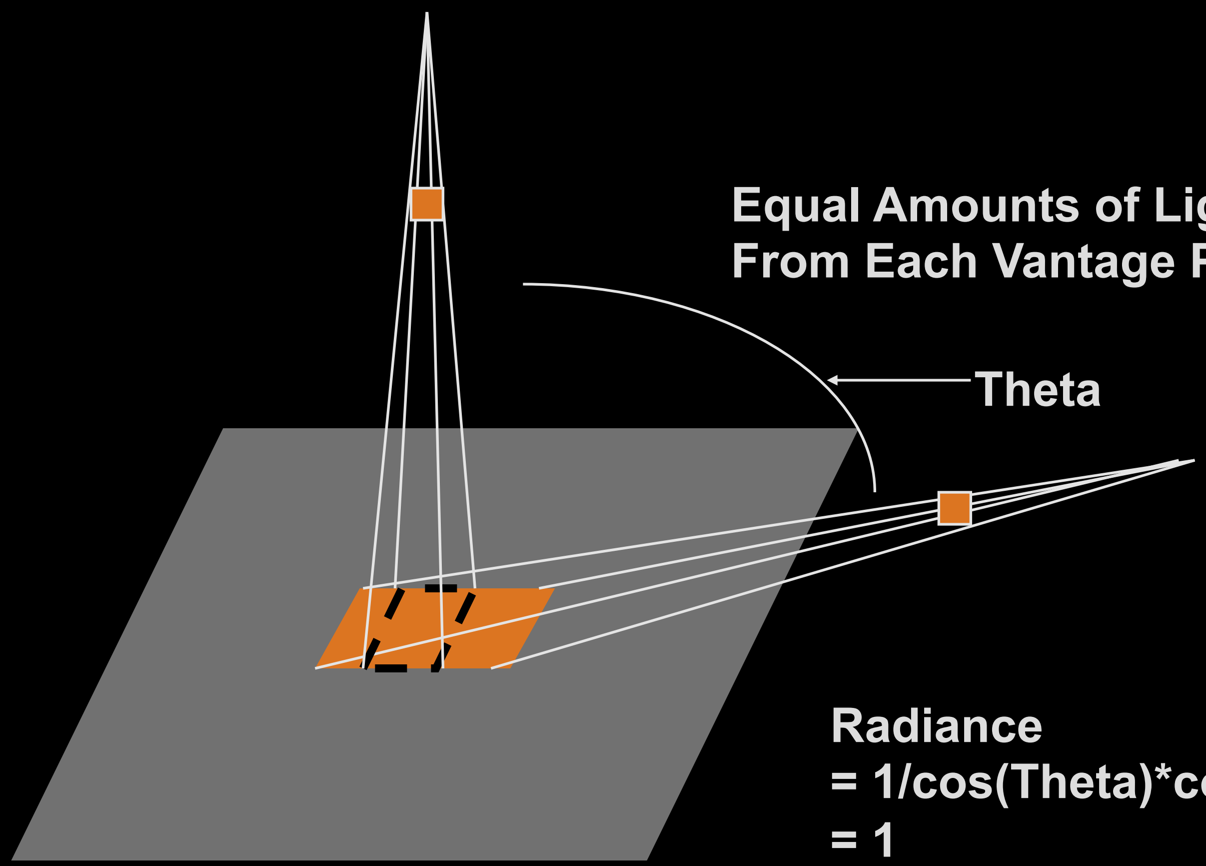
Area of black box = 1
Area of orange box = 1/cos(Theta)
Foreshortening rule.

Lambertian Surfaces

Relative magnitude of light scattered in each direction.
Proportional to cos (Theta).



Lambertian Surfaces

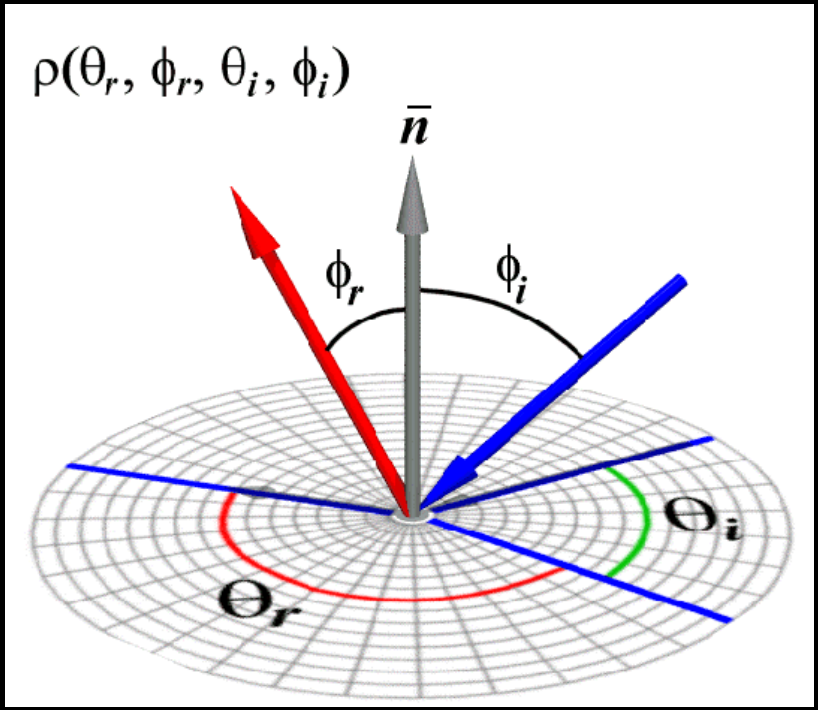


Radiance
= 1/cos(Theta)*cos(Theta)
= 1

Area of black box = 1
Area of orange box = 1/cos(Theta)
Foreshortening rule.

BRDF

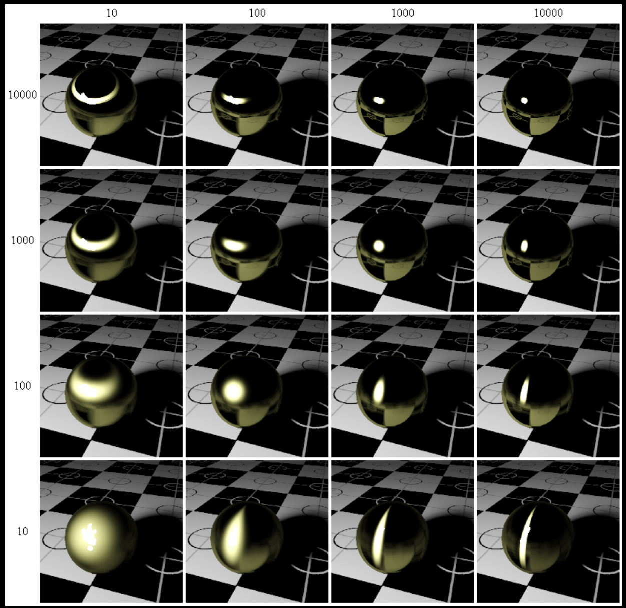
- Bidirectional Reflectance Distribution Function
  - a function of 2 directions (or 4 scalar values)



http://escience.anu.edu.au/lecture/cg/GlobalIllumination/BRDF.en.html

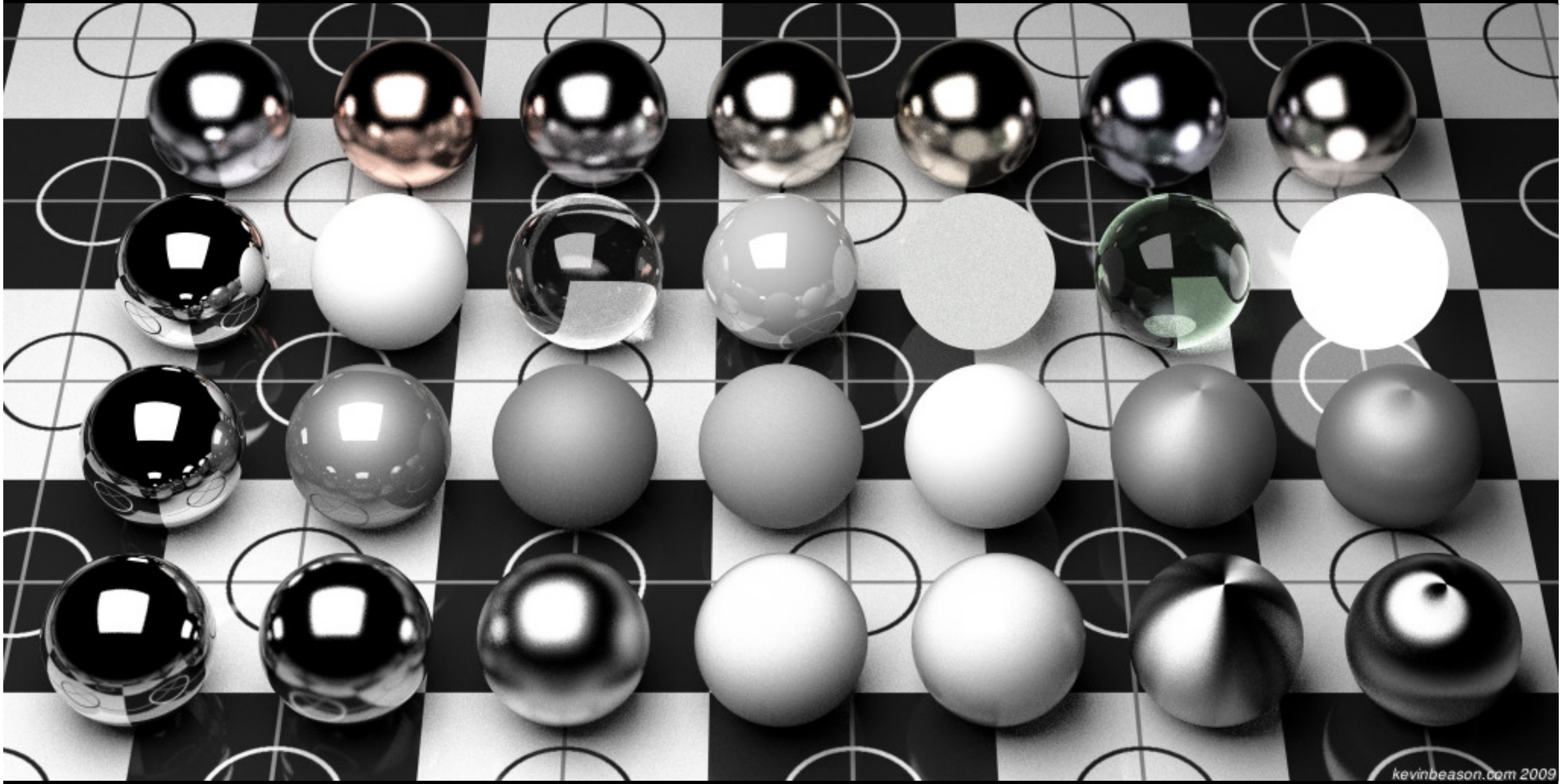
BRDF

- Bidirectional Reflectance Distribution Function
  - a function of 2 directions (or 4 scalar values)



http://escience.anu.edu.au/lecture/cg/GlobalIllumination/BRDF.en.html

BRDF



http://www.kevinbeason.com/worklog/

BRDF again

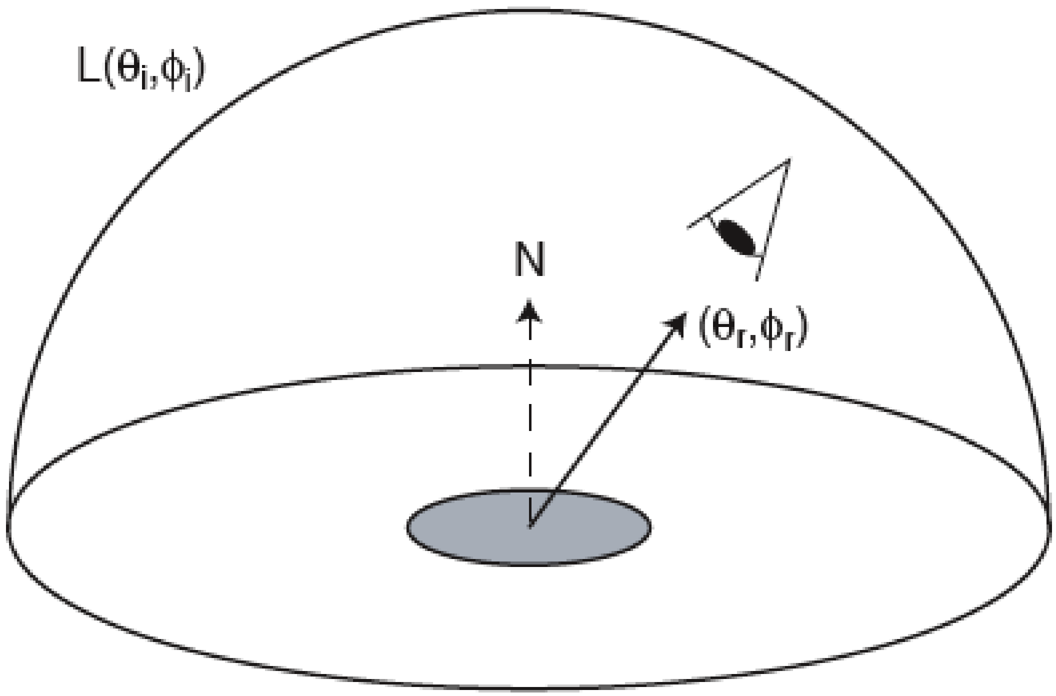
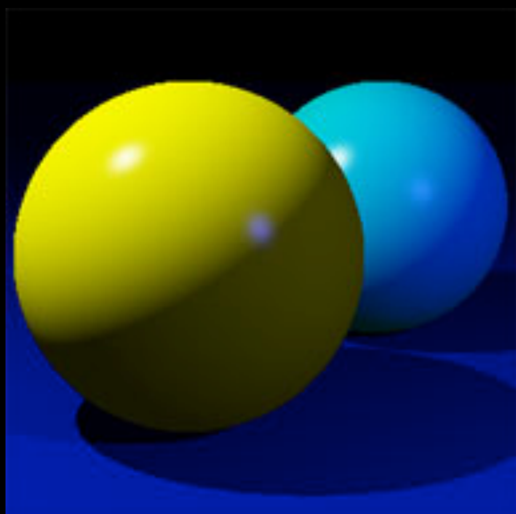


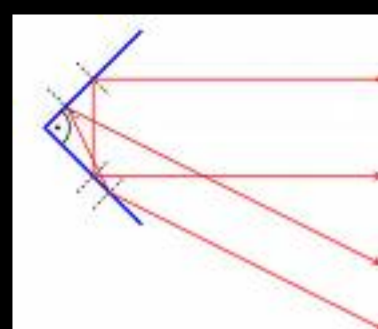
Figure 3.1. A viewer observes a surface patch with normal **N** from direction (θ_r, ϕ_r). $L(\theta_i, \phi_i)$ represents radiance of illumination from direction (θ_i, ϕ_i). The coordinate system is such that **N** points in direction $(0, 0)$.

$$B(\theta_r, \phi_r) = \int_{\phi_i=0}^{2\pi} \int_{\theta_i=0}^{\pi/2} L(\theta_i, \phi_i) f(\theta_i, \phi_i; \theta_r, \phi_r) \cos\theta_i \sin\theta_i \, d\theta_i \, d\phi_i,$$

The BRDF

The bidirectional reflectance distribution function.





BRDF

- Ron Dror’s thesis

Reflection parameters

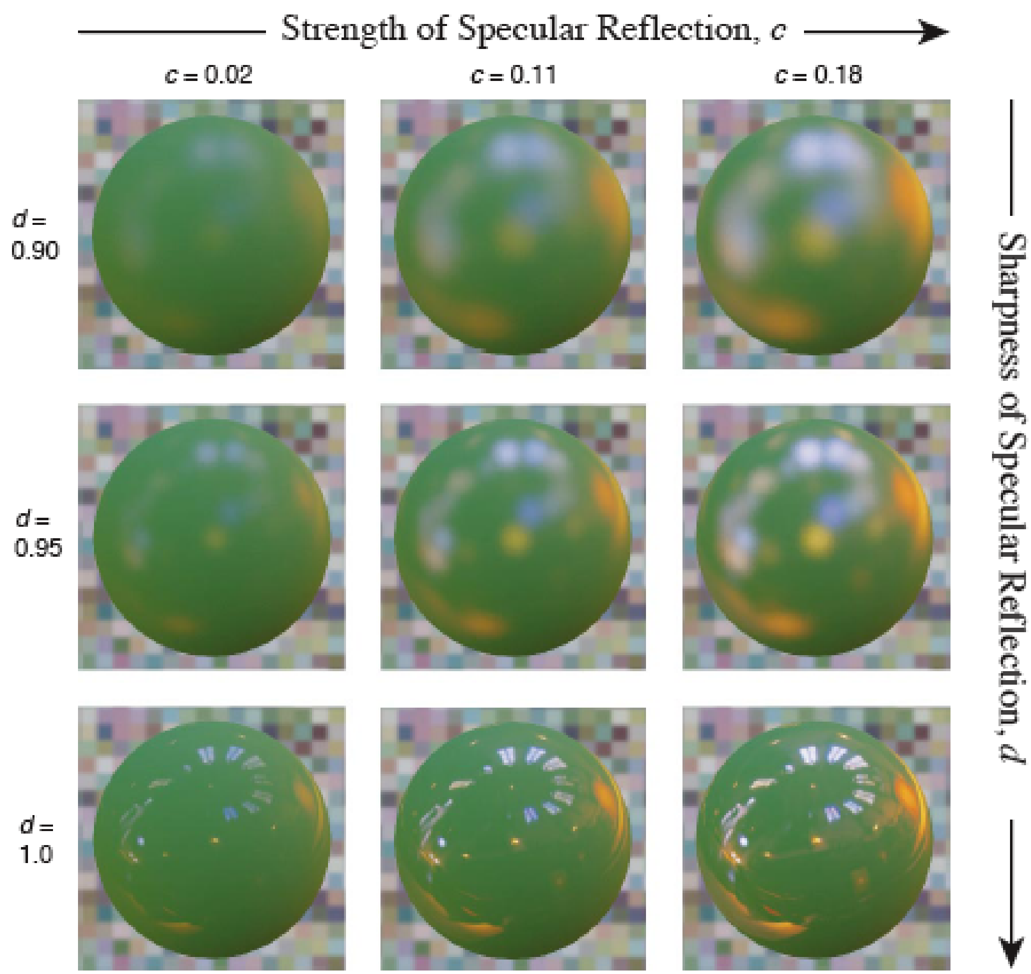
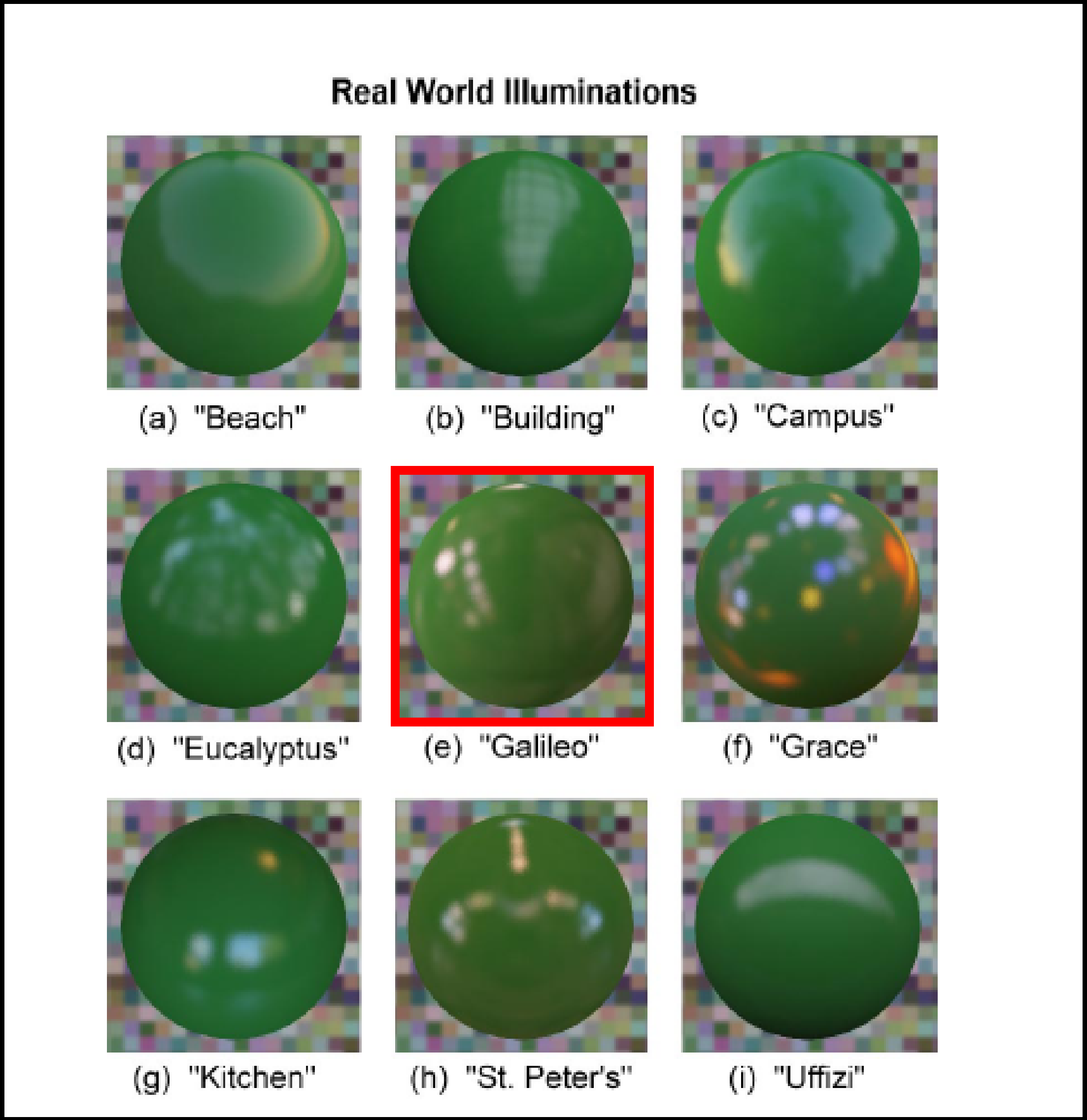
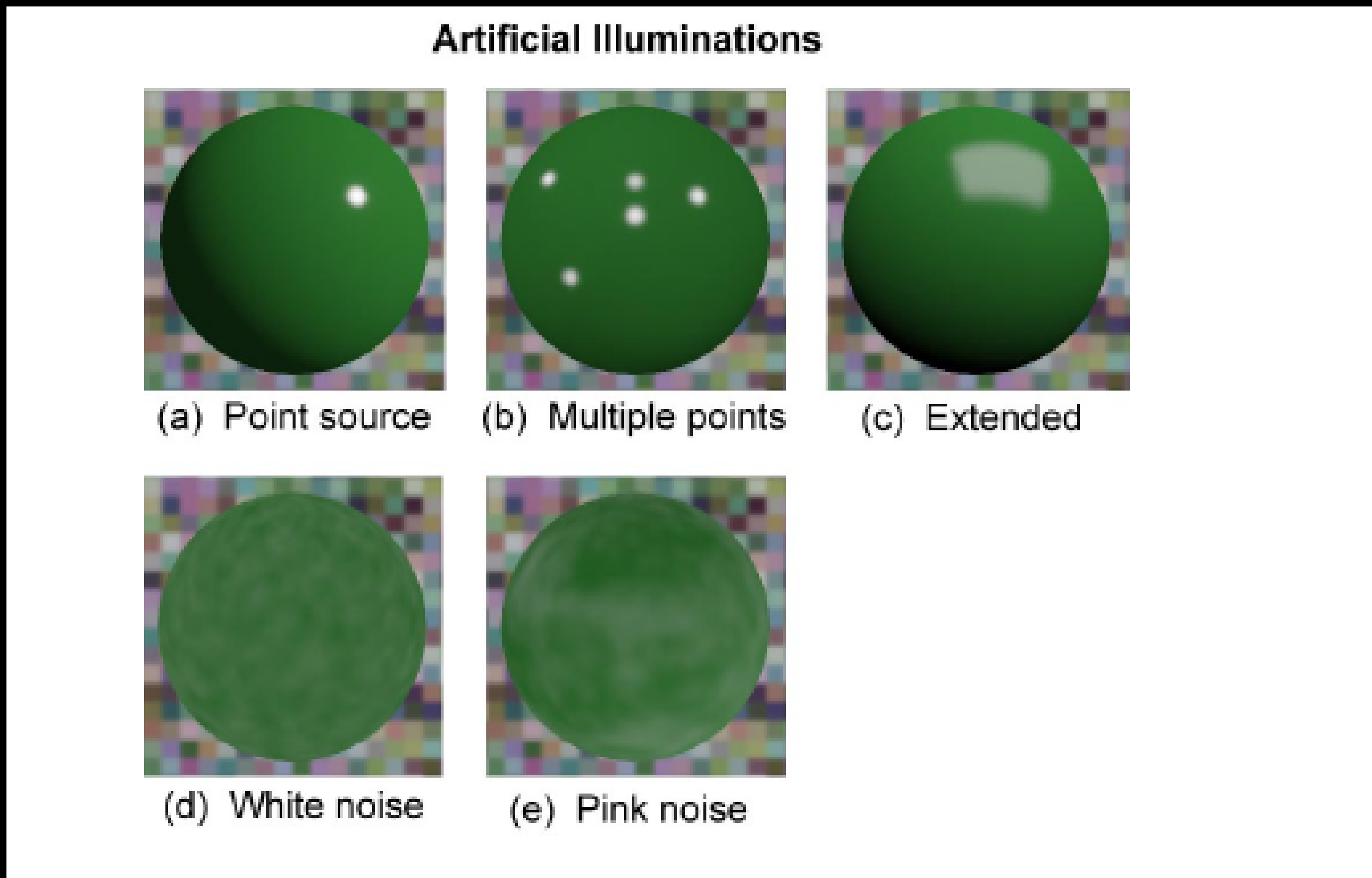


Figure 2.6. Grid showing range of reflectance properties used in the experiments for a particular real-world illumination map. All the spheres shown have an identical diffuse component. In Pellacini's reparameterization of the Ward model, the specular component depends on the c and d parameters. The strength of specular reflection, c, increases with ρ_s, while the sharpness of specular reflection, d, decreases with α. The images were rendered in *Radiance*, using the techniques described in Appendix B.

Real World Light Variation



Fake Light and Real Light



Simple and Complex Light

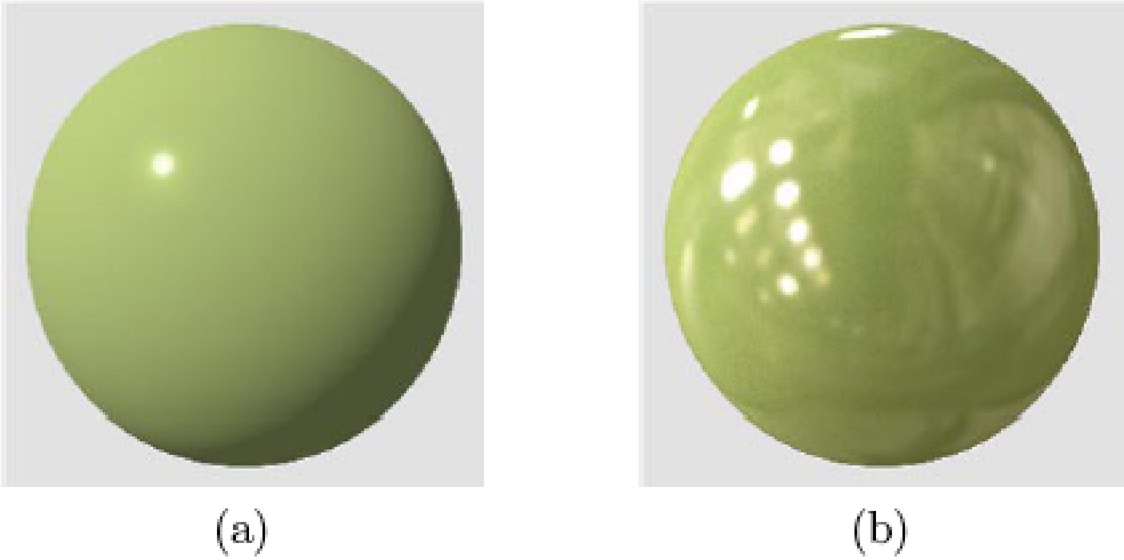


Figure 2.9. (a) A shiny sphere rendered under illumination by a point light source. (b) The same sphere rendered under photographically-acquired real-world illumination. Humans perceive reflectance properties more accurately in (b).

BRDF again

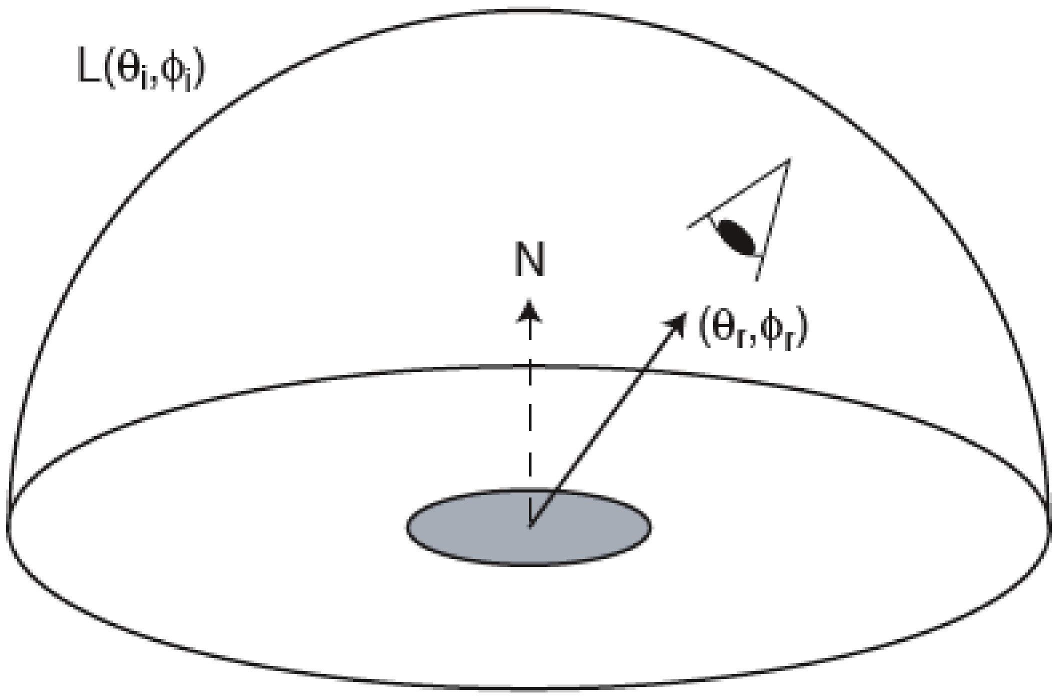


Figure 3.1. A viewer observes a surface patch with normal **N** from direction (θ_r, ϕ_r). $L(\theta_i, \phi_i)$ represents radiance of illumination from direction (θ_i, ϕ_i). The coordinate system is such that **N** points in direction $(0, 0)$.

$$B(\theta_r, \phi_r) = \int_{\phi_i=0}^{2\pi} \int_{\theta_i=0}^{\pi/2} L(\theta_i, \phi_i) f(\theta_i, \phi_i; \theta_r, \phi_r) \cos\theta_i \sin\theta_i \, d\theta_i \, d\phi_i,$$

Likelihood of material

Figure 3.6. A photograph of a matte sphere, shown against a uniform gray background. This image could also be produced by a chrome sphere under appropriate illumination, but that scenario is highly unlikely.

Classifying Under Uncertainty

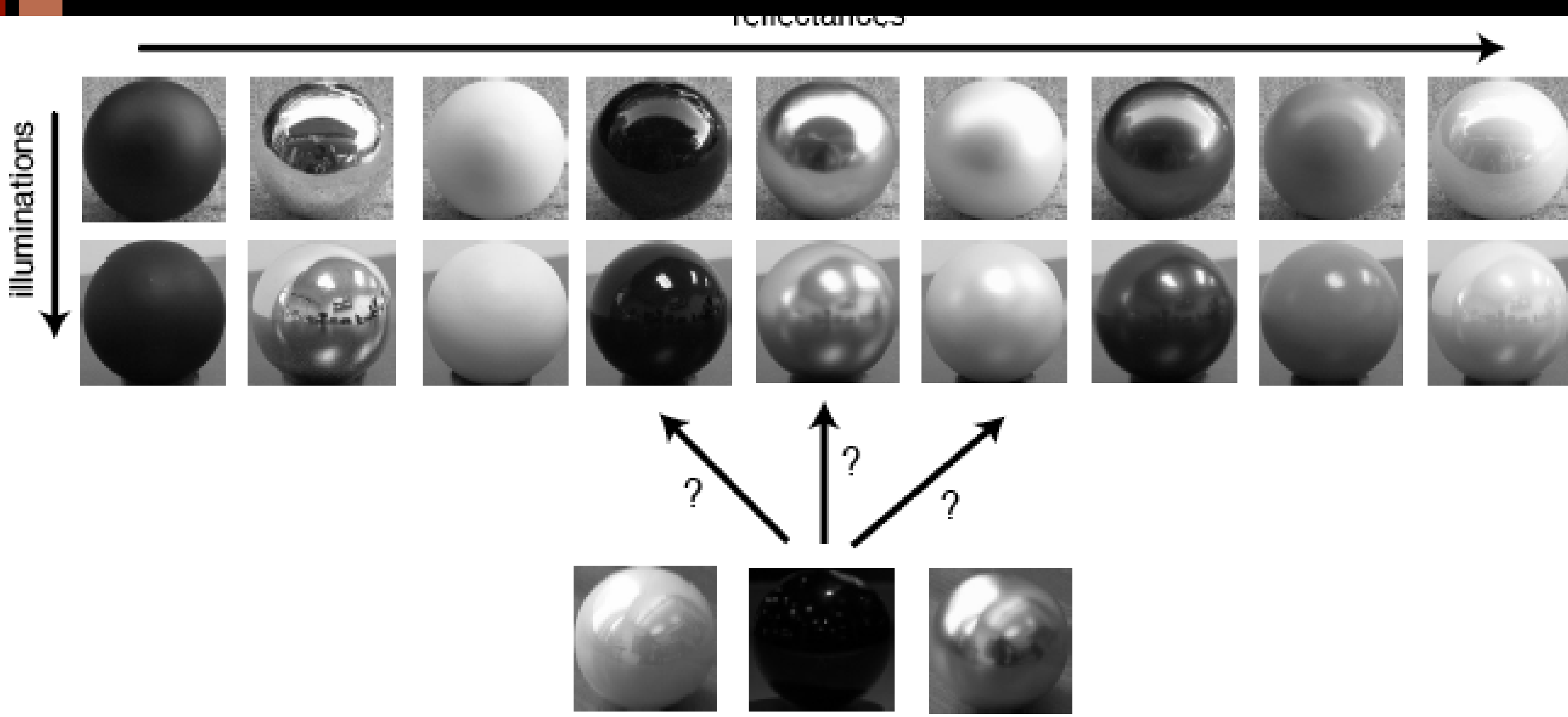
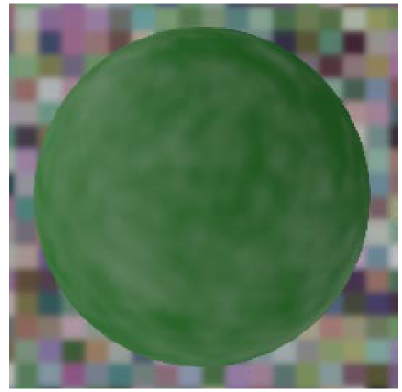


Figure 3.7. The problem addressed by a classifier of Chapter 6, illustrated using a database of photographs. Each of nine spheres was photographed under seven different illuminations. We trained a nine-way classifier using the images corresponding to several illuminations, and then used it to classify individual images under novel illuminations.

More Fake Light

(a) Original



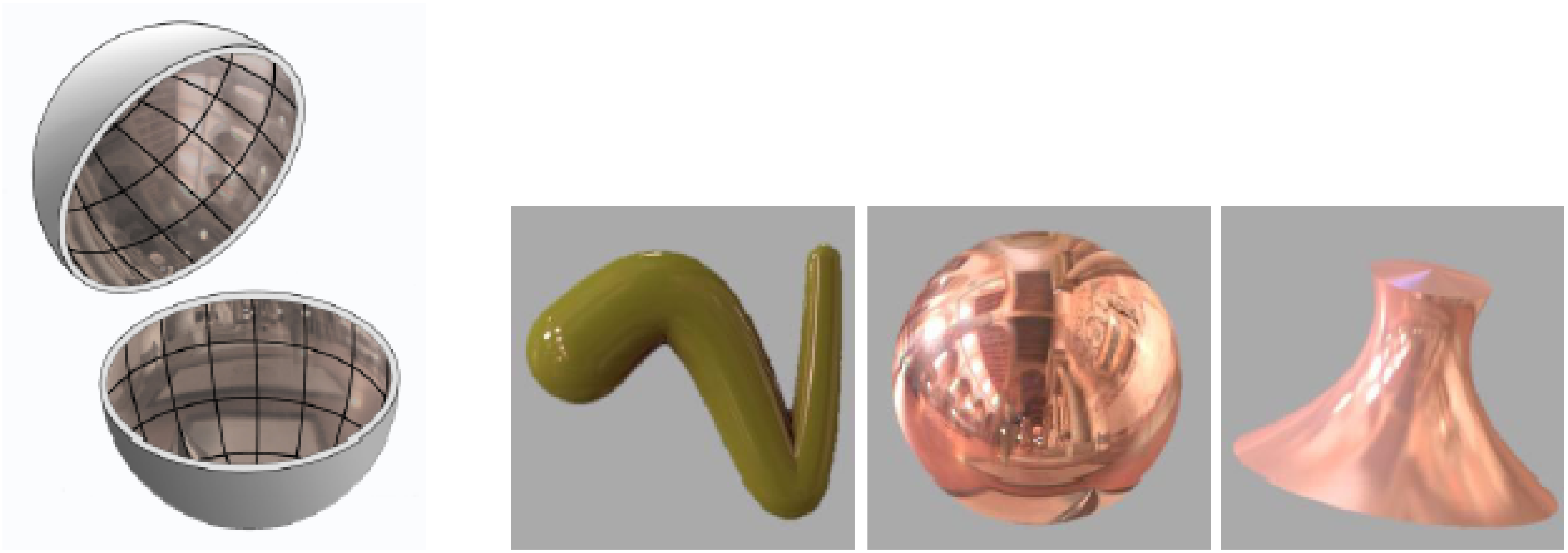
(b) $1/f^2$ power spectrum

(c) Heeger and Bergen texture

(d) Portilla and Simoncelli texture

Figure 4.14. Spheres of identical reflectance properties rendered under a photographically-acquired illumination map (a) and three synthetic illumination maps (b-d). The illumination in (b) is Gaussian noise with a $1/f^2$ power spectrum. The illumination in (c) was synthesized with the procedure of Heeger and Bergen [43] to match the pixel histogram and marginal wavelet histograms of the illumination in (a). The illumination in (d) was synthesized using the technique of Portilla and Simoncelli, which also enforces conditions on the joint wavelet histograms. The illumination map of (a) is due to Debevec [24].

Illumination Maps



(a) (b) (c) (d)

Figure 5.2. (a) A photographically-acquired illumination map, illustrated on the inside of a spherical shell. The illumination map is identical to that of Figure 4.1d. (b-d) Three surfaces of different geometry and reflectance rendered under this illumination map using the methods of Appendix B.

Classification

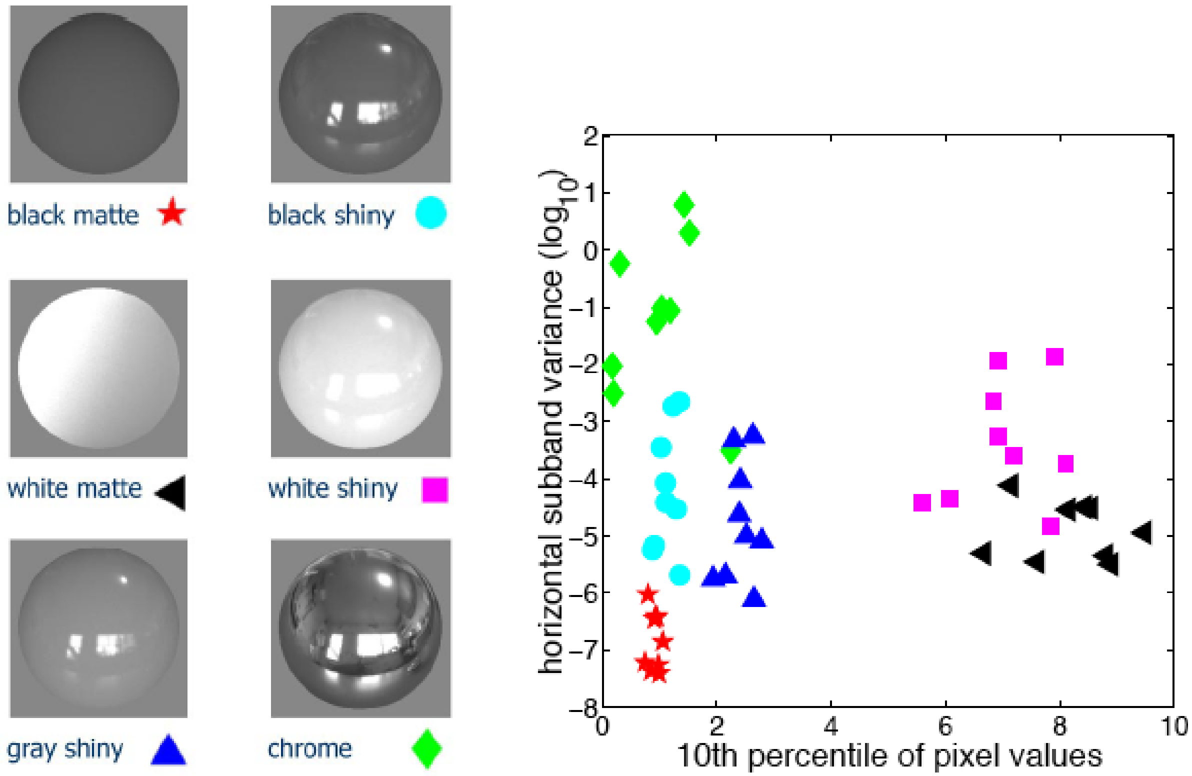


Figure 5.11. At left, synthetic spheres of 6 different reflectances, each rendered under one of Debevec's illumination maps. Ward model parameters are as follows: black matte, $\rho_d = .1$, $\rho_s = 0$; black shiny, $\rho_d = .1$, $\rho_s = .1$, $\alpha = .01$; white matte, $\rho_d = .9$, $\rho_s = 0$; white shiny, $\rho_d = .7$, $\rho_s = .25$, $\alpha = .01$; chrome, $\rho_d = 0$, $\rho_s = .75$, $\alpha = 0$; gray shiny, $\rho_d = .25$, $\rho_s = .05$, $\alpha = .01$. We rendered each sphere under the nine photographically-acquired illuminations depicted in Figure 2.7 and plotted a symbol corresponding to each in the two-dimensional feature space at right. The horizontal axis represents the 10th percentile of pixel intensity, while the vertical axis is the log variance of horizontally-oriented QMF wavelet coefficients at the second-finest scale, computed after geometrically distorting the original image as described in Section 6.1.2.

Classification Continued

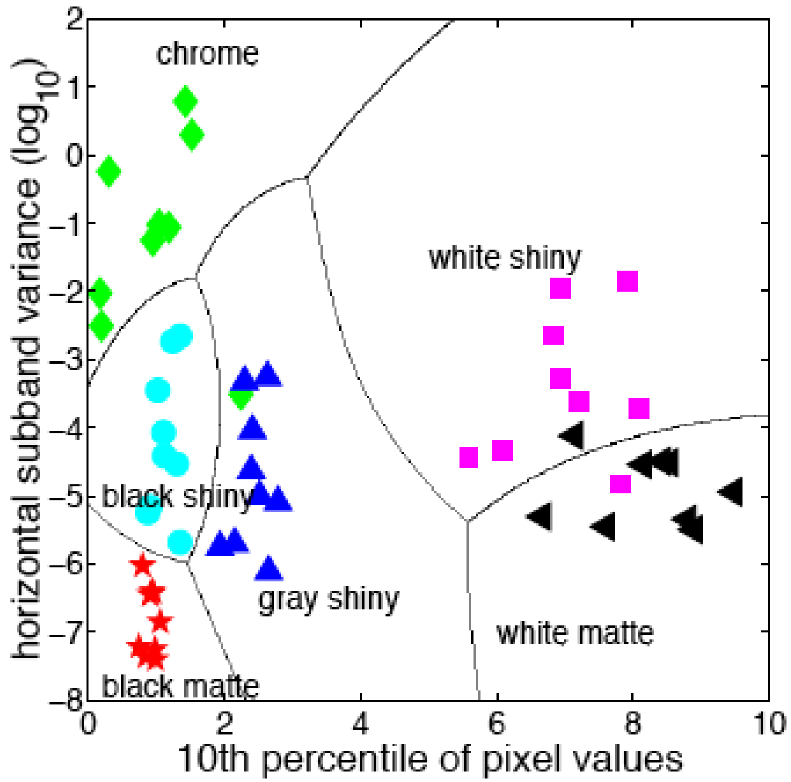


Figure 5.12. The curved lines separate regions assigned to different reflectances by a simple classifier based on two image features. The training examples are the images described in Figure 5.11. The classifier is a one-versus-all support vector machine, described in Section 6.1.1. Using additional image features improves classifier performance.

Lighting direction vs. Viewer direction

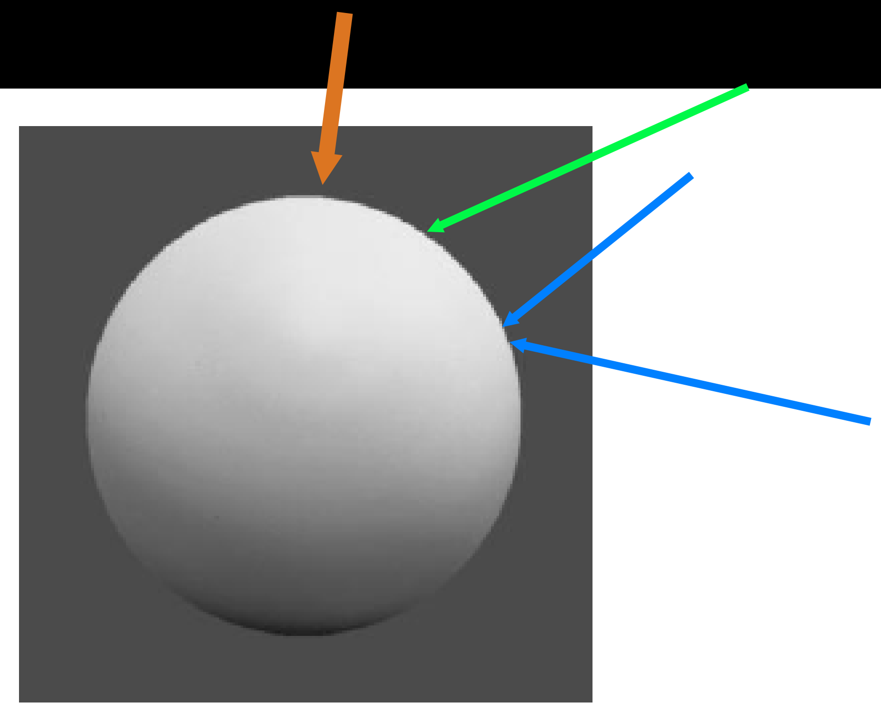


Figure 3.6. A photograph of a matte sphere, shown against a uniform gray background.

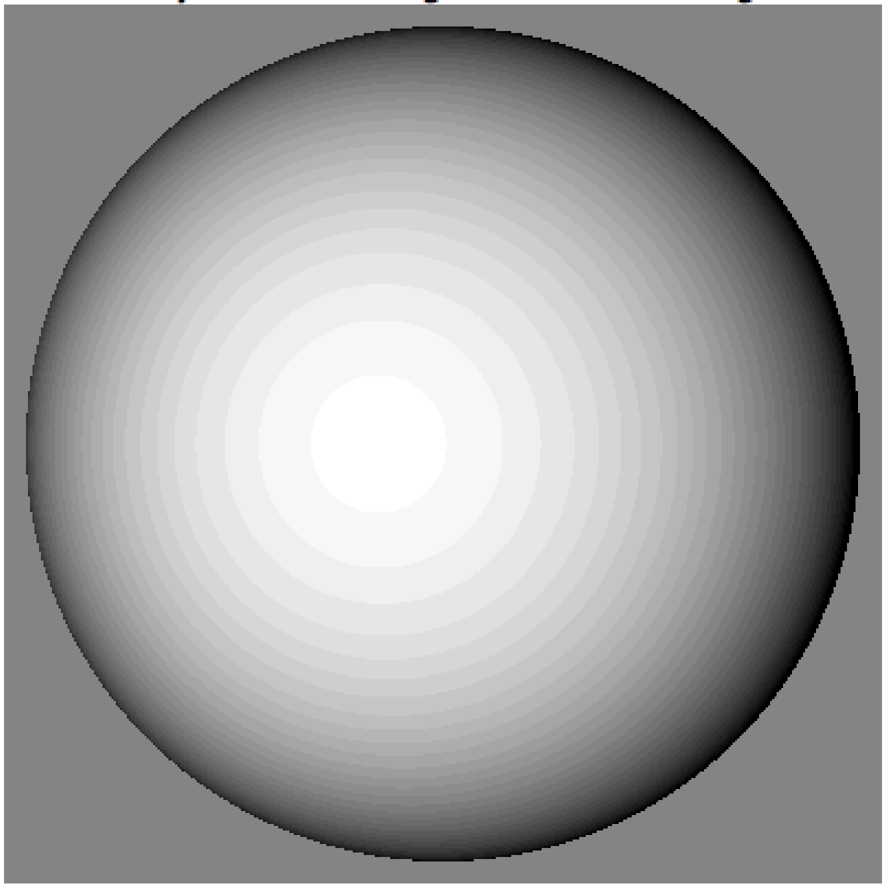
For Lambertian surface:

Viewer direction is irrelevant

Lighting direction is *very relevant*

Lambertian Sphere

Orthogonal Projection,
Infinitely Distant Point Light from -90 to +90 degrees



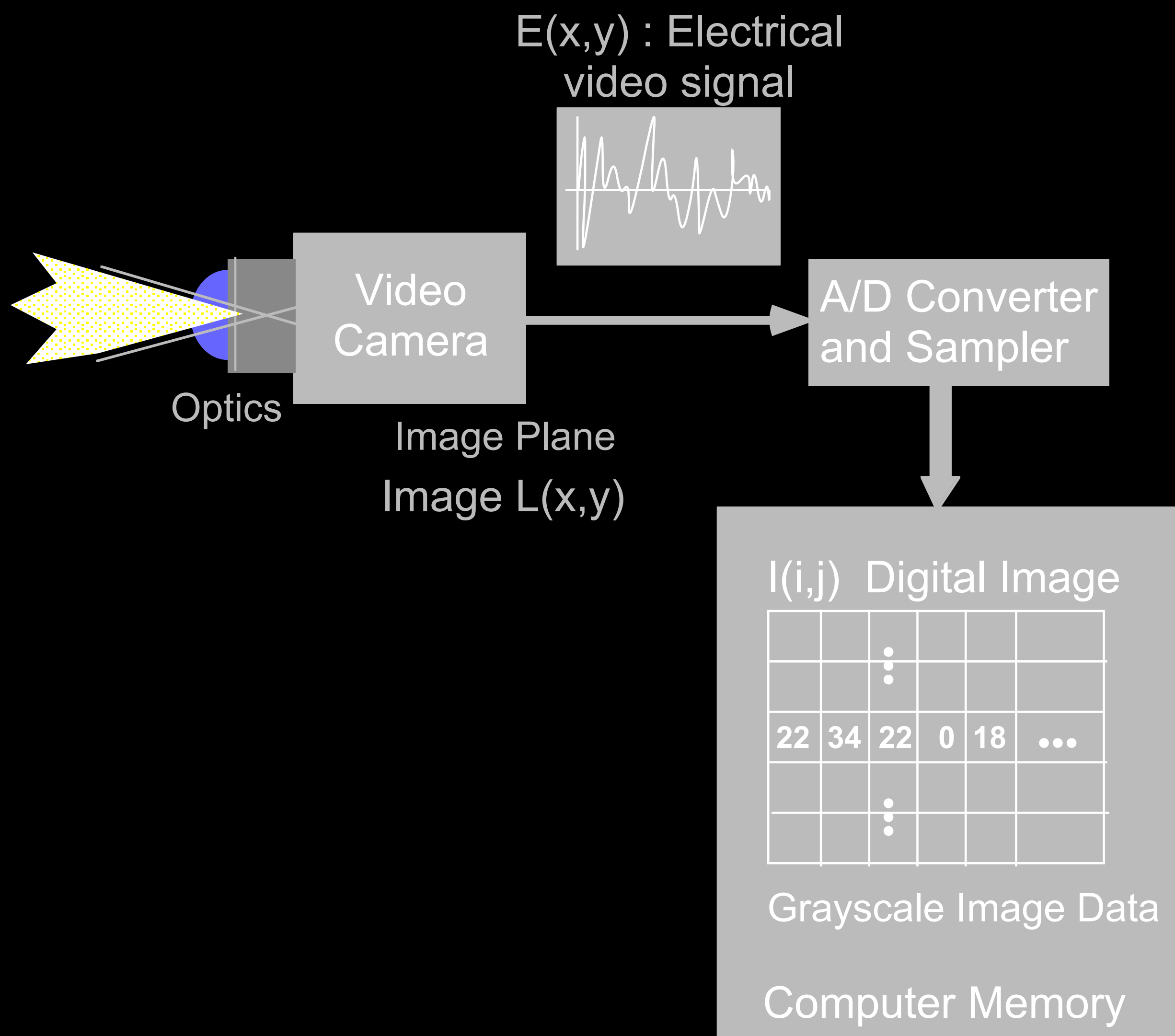
Photometry

- Photometry:

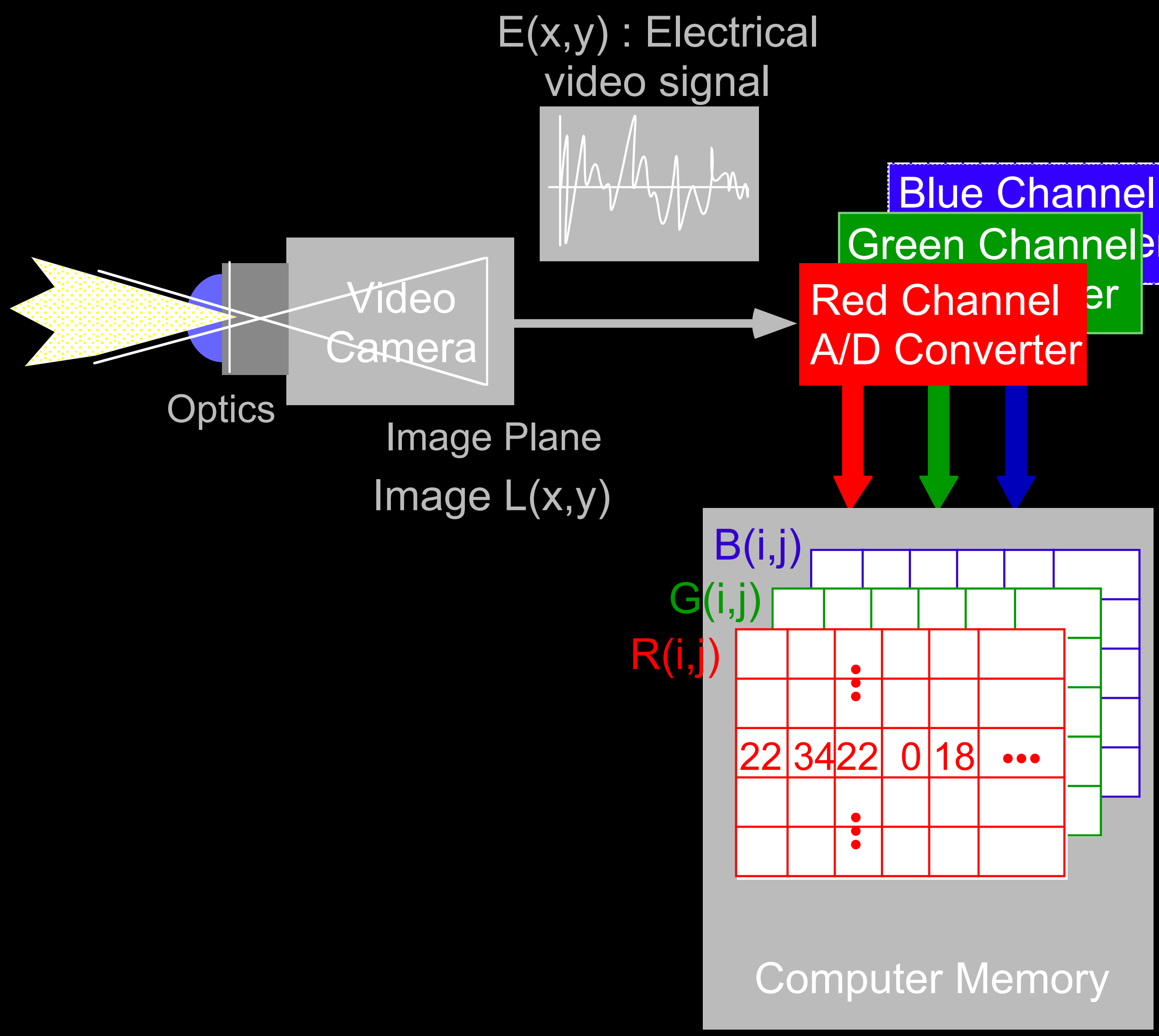
  Concerned with mechanisms for converting light energy into electrical energy.

World → Optics → Sensor → Signal → Digitizer → Digital Representation

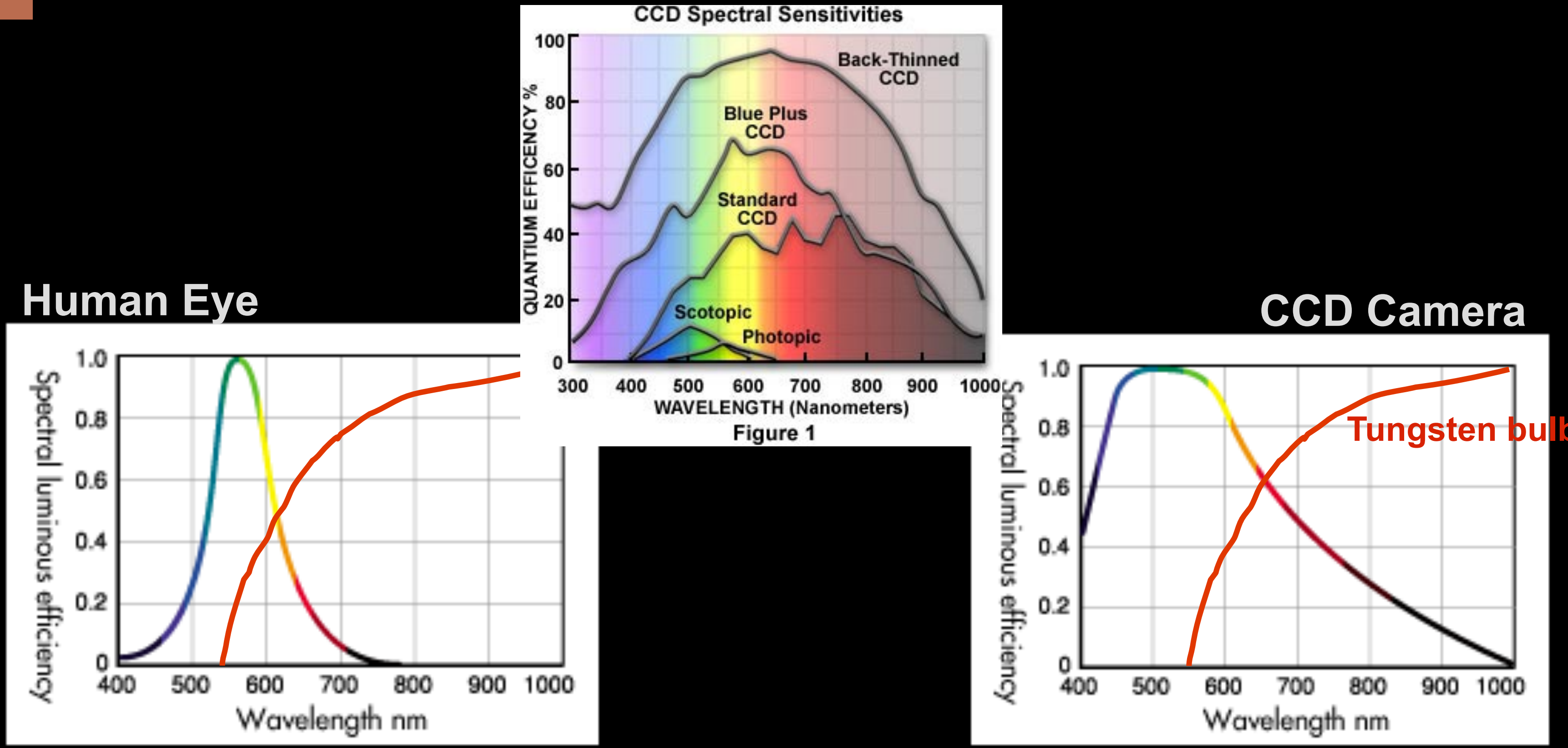
B&W Video System



Color Video System



Spectral Sensitivity



- Figure 1 shows relative efficiency of conversion for the eye (scotopic and photopic curves) and several types of CCD cameras. Note the CCD cameras are much more sensitive than the eye.
- Note the enhanced sensitivity of the CCD in the Infrared and Ultraviolet (bottom two figures)
- Both figures also show a handrawn sketch of the spectrum of a tungsten light bulb

Human Vision Regimes

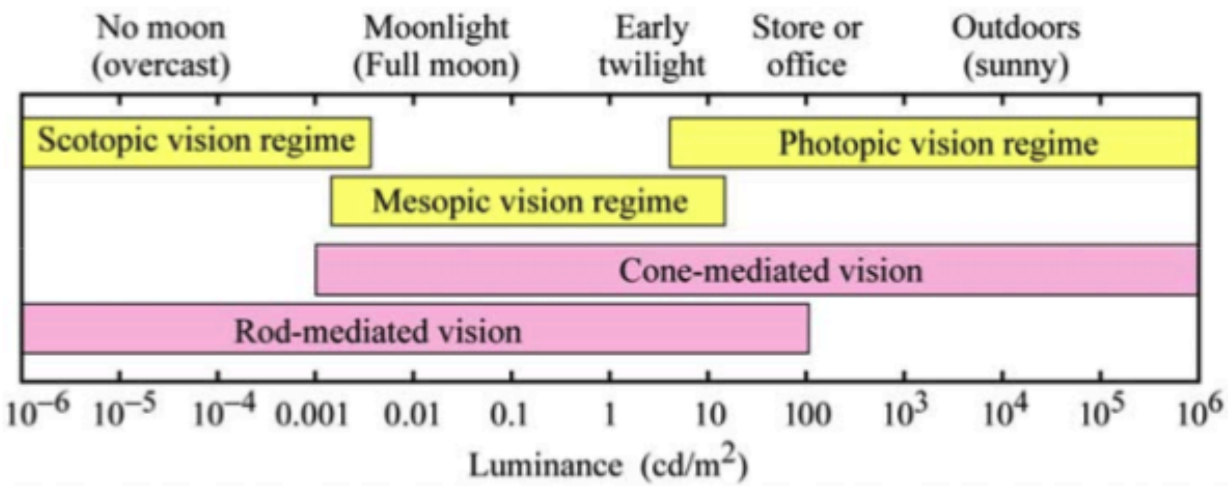


Fig. 16.2. Approximate ranges of vision regimes and receptor regimes (after Osram Sylvania, 2000).

Characteristics

- In general, $V(x,y) = k\,E(x,y)^{g}$ where
  - k is a constant
  - **g is a parameter of the type of sensor**
    - **g=1 (approximately) for a CCD camera**
    - **g=.65 for an old type vidicon camera**
- Factors influencing performance:
  - Optical distortion: pincushion, barrel, non-linearities
  - Sensor dynamic range (30:1 CCD, 200:1 vidicon)
  - Sensor Shading (nonuniform responses from different locations)
- **TV Camera pros: cheap, portable, small size**
- **TV Camera cons: poor signal to noise, limited dynamic range, fixed array size with small image (getting better)**

Sensor Performance

- Optical Distortion: pincushion, barrel, non-linearities
- Sensor Dynamic Range: (30:1 for a CCD, 200:1 Vidicon)
- Sensor Blooming: spot size proportional to input intensity
- Sensor Shading: (non-uniform response at outer edges of image)
- Dead CCD cells

There is no "universal sensor". Sensors must be selected/tuned for a particular domain and application.

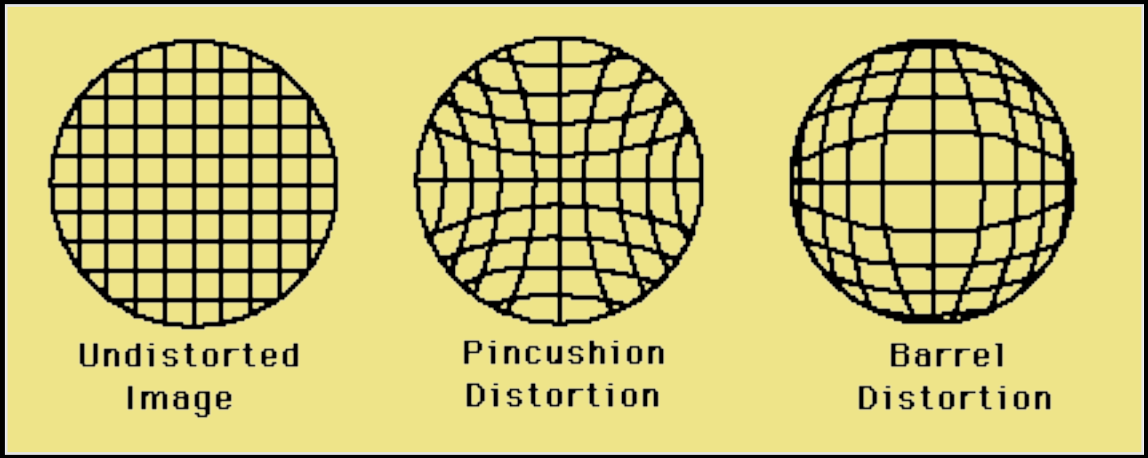
Lens Aberrations

- In an ideal optical system, all rays of light from a point in the object plane would converge to the same point in the image plane, forming a clear image.
- The lens defects which cause different rays to converge to different points are called aberrations.
  - Distortion: barrel, pincushion
  - Curvature of field
  - Chromatic aberration
  - Spherical aberration
  - Coma
  - Astigmatism

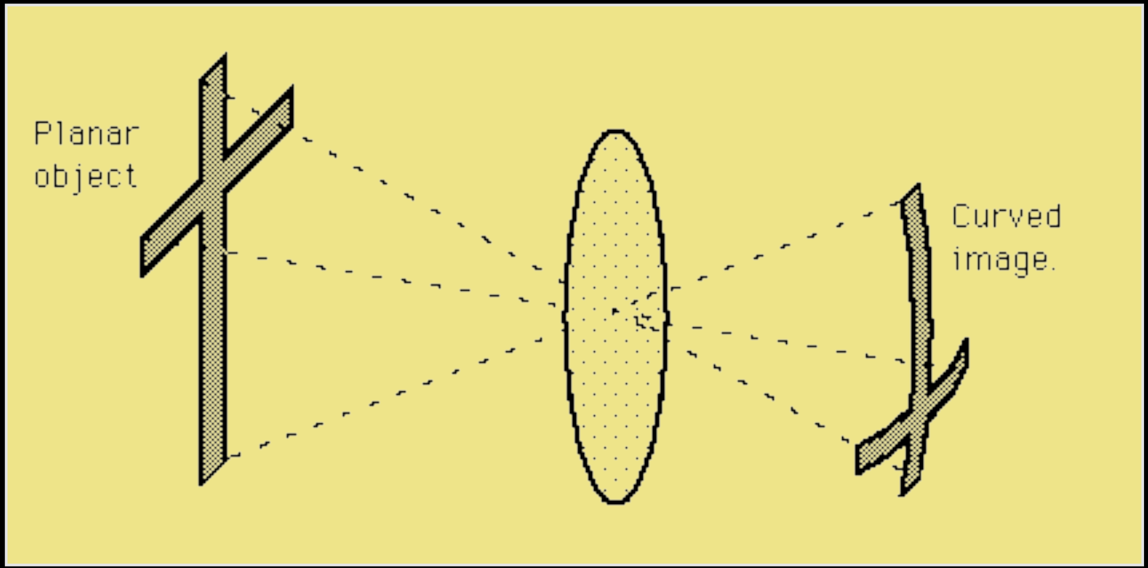
Aberration slides after http://hyperphysics.phy-astr.gsu.edu/hbase/geoopt/aberrcon.html#c1

Lens Aberrations

- Distortion

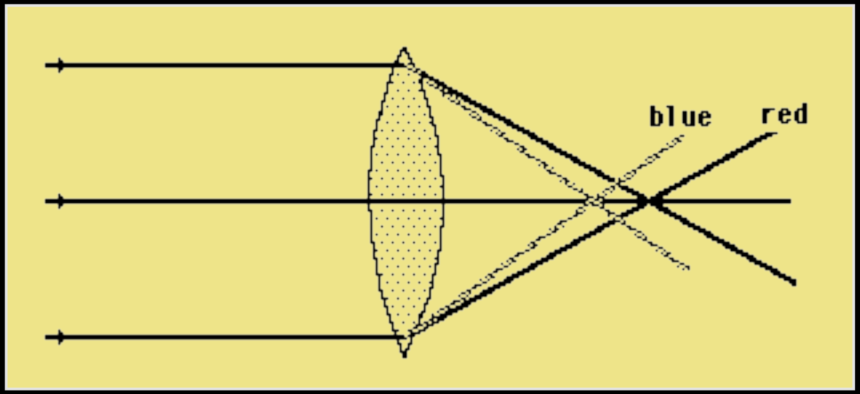


- Curved Field



Lens Aberrations

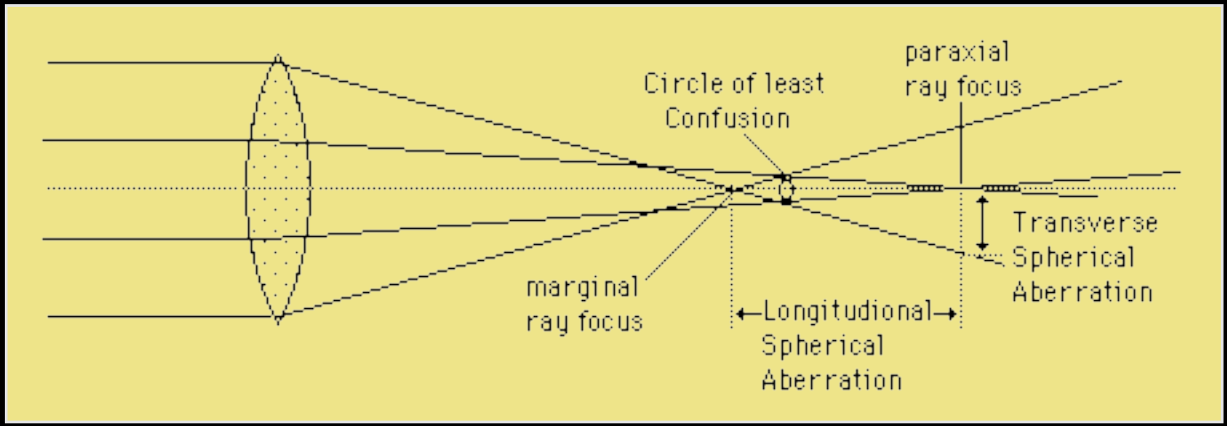
- Chromatic Aberration



- Focal Length of lens depends on refraction and
- The index of refraction for blue light (short wavelengths) is larger than that of red light (long wavelengths).
- Therefore, a lens will not focus different colors in exactly the same place
- The amount of chromatic aberration depends on the dispersion (change of index of refraction with wavelength) of the glass.

Lens Aberration

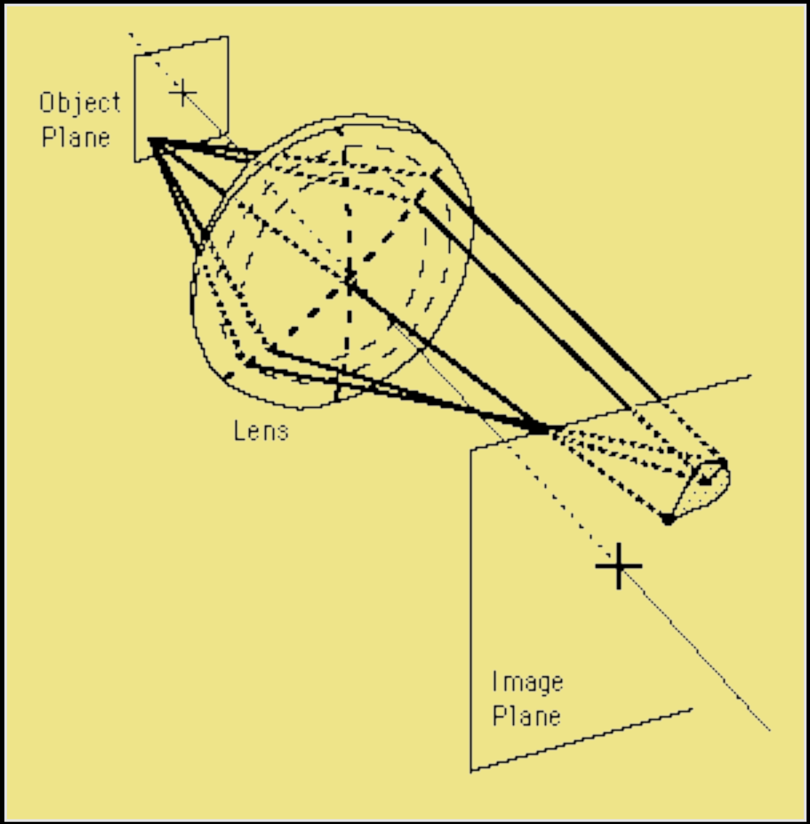
- Spherical Aberration



- Rays which are parallel to the optic axis but at different distances from the optic axis fail to converge to the same point.

Lens Aberrations

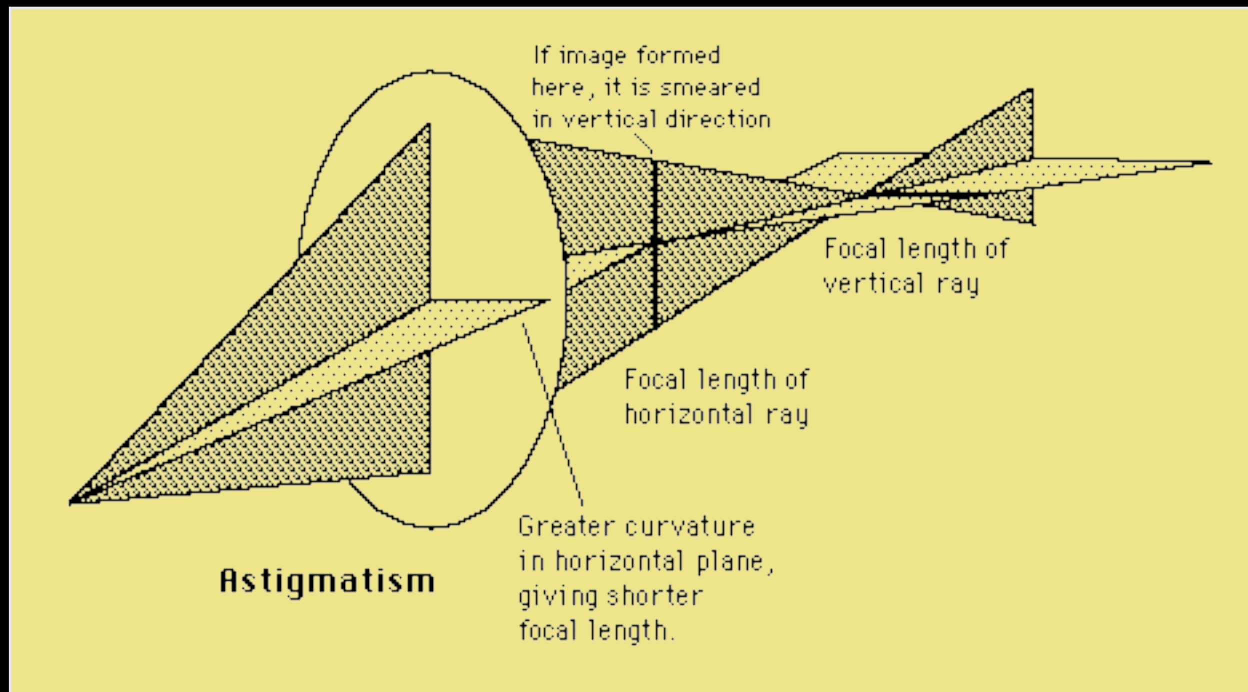
- Coma



- Rays from an off-axis point of light in the object plane create a trailing "comet-like" blur directed away from the optic axis
- Becomes worse the further away from the central axis the point is

Lens Aberrations

- Astigmatism

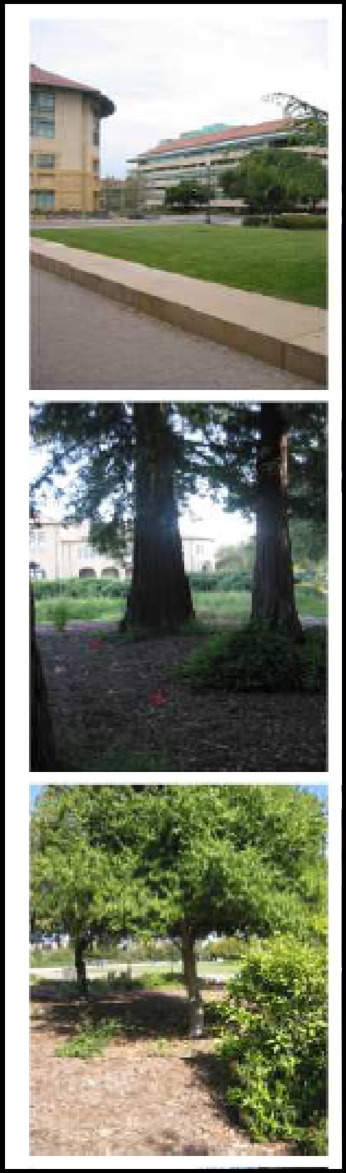


- Results from different lens curvatures in different planes.

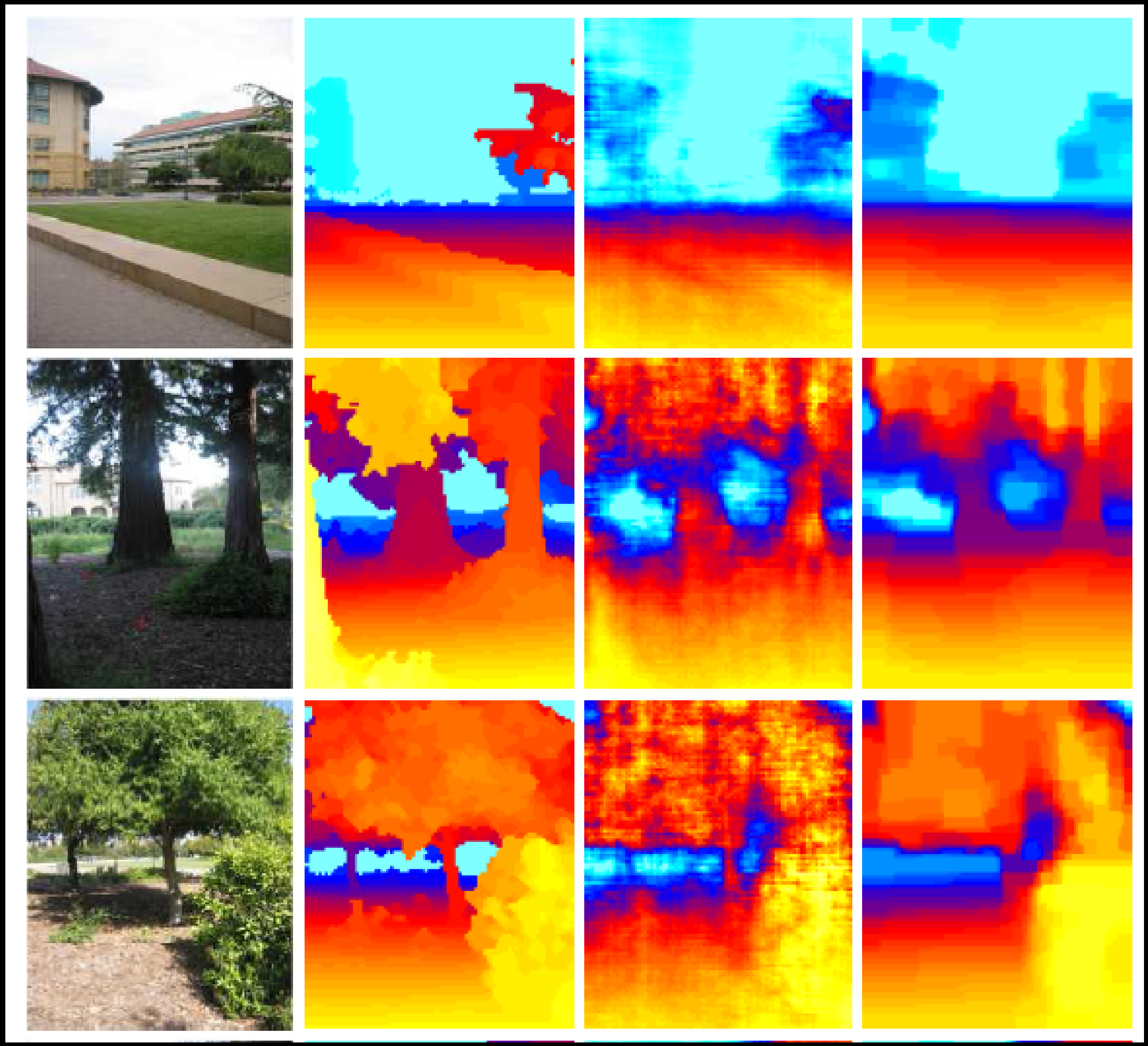
Sensor Summary

- Visible Light/Heat
  - Camera/Film combination
  - Digital Camera
  - Video Cameras
  - FLIR (Forward Looking Infrared)
- Range Sensors
  - Radar (active sensing)
    - sonar
    - laser
  - Triangulation
    - stereo
    - structured light
      - – striped, patterned
    - Moire
    - Holographic Interferometry
    - Lens Focus
    - Fresnel Diffraction
- Others
- Almost anything which produces a 2d signal that is related to the scene can be used as a sensor

Depth Images (Stanford)



Depth Images

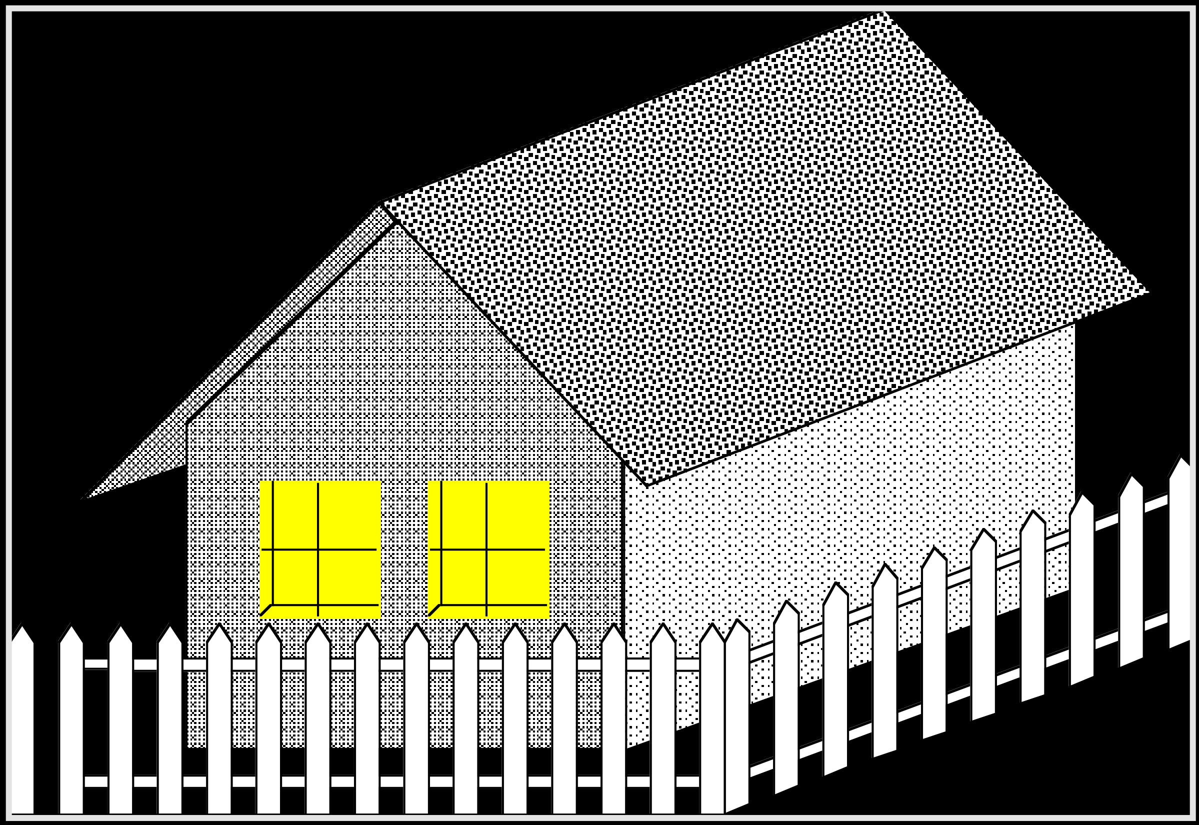


Digitization

World → Optics → Sensor →
Signal → Digitizer →
Digital Representation

- Digitization: conversion of the continuous (in space and value) electrical signal into a digital signal (digital image)
- Three decisions must be made:
  - Spatial resolution (how many samples to take)
  - Signal resolution (dynamic range of values)
  - Tessellation pattern (how to 'cover' the image with sample points)

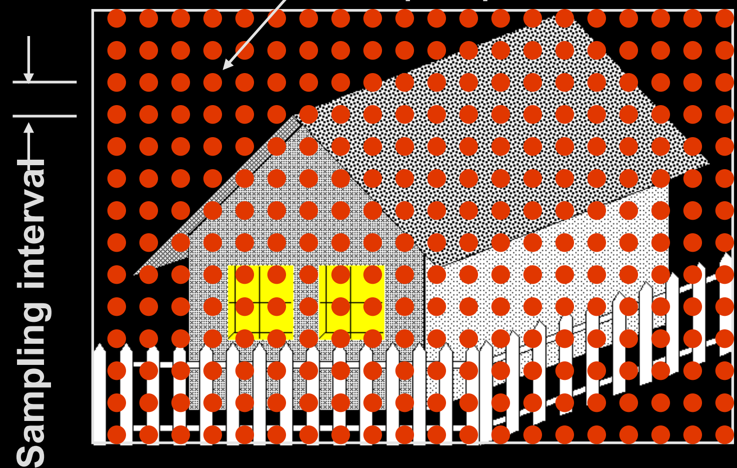
Digitization: Spatial Resolution



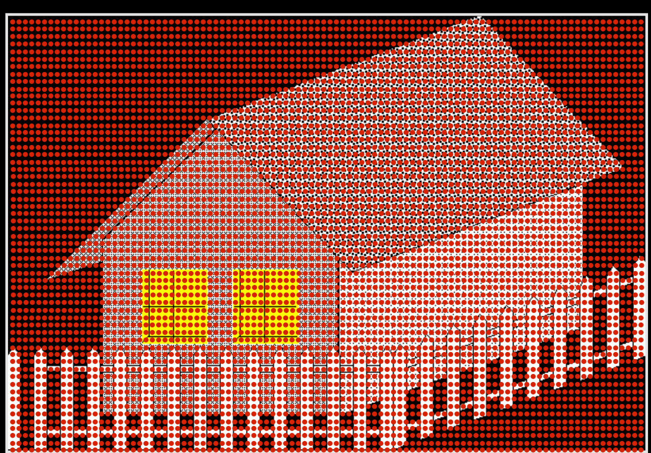
- Let's digitize this image
  - Assume a square sampling pattern
  - Vary density of sampling grid

Spatial Resolution

Sample picture at each red point



Coarse Sampling: 20 points per row by 14 rows

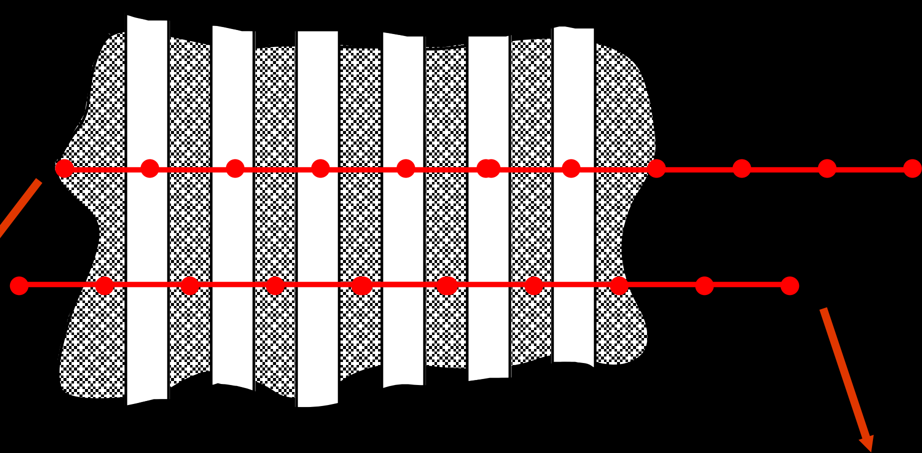


Finer Sampling: 100 points per row by 68 rows

Effect of Sampling Interval - 1

- Look in vicinity of the picket fence:

Sampling Interval: 



100	100	100	100	100	100
100	100	100	100	100	100
100	100	100	100	100	100
100	100	100	100	100	100

White Image!

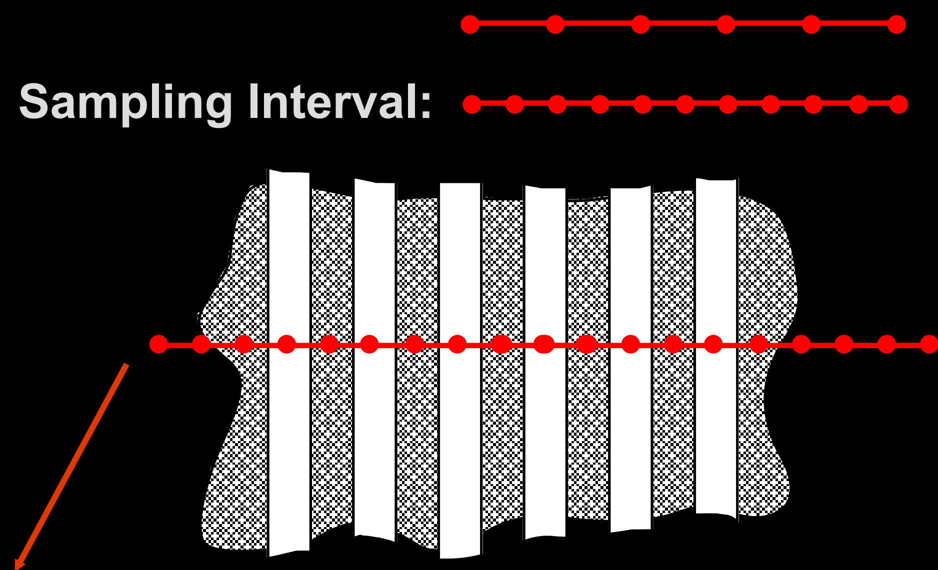
NO EVIDENCE OF THE FENCE!

40	40	40	40	40	40
40	40	40	40	40	40
40	40	40	40	40	40
40	40	40	40	40	40

Dark Gray Image!

Effect of Sampling Interval - 2

- Look in vicinity of picket fence:



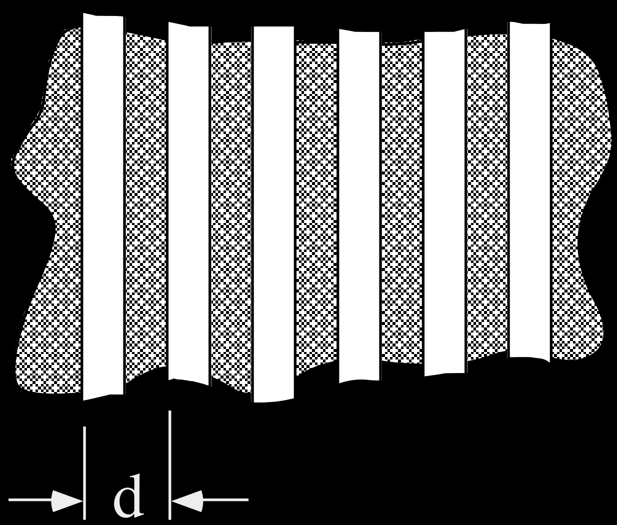
40	100	40	100	40
40	100	40	100	40
40	100	40	100	40
40	100	40	100	40

Now we've got a fence!

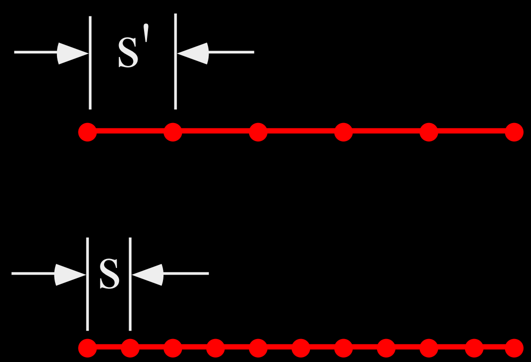
What's the difference between this attempt and the last one?

The Missing Fence Found

- Consider the repetitive structure of the fence:



Sampling Intervals



Case 1: s' = d	The sampling interval is equal to the size of the repetitive structure	**NO FENCE**
Case 2: s = d/2	The sampling interval is one-half the size of the repetitive structure	**FENCE**

The Sampling Theorem

- IF: the size of the smallest structure to be preserved is d
- THEN: the sampling interval must be smaller than d/2

- Can be shown to be true mathematically
- Repetitive structure has a certain frequency ('pickets/foot')
  - To preserve structure must sample at **greater** than twice the frequency
  - Holds for images, audio CDs, digital television….
- Leads naturally to Fourier Analysis

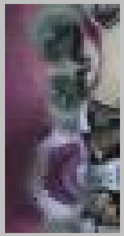
Human Eye Sampling

- Fine near the center of the retina (fovea)
- Coarse at the edges

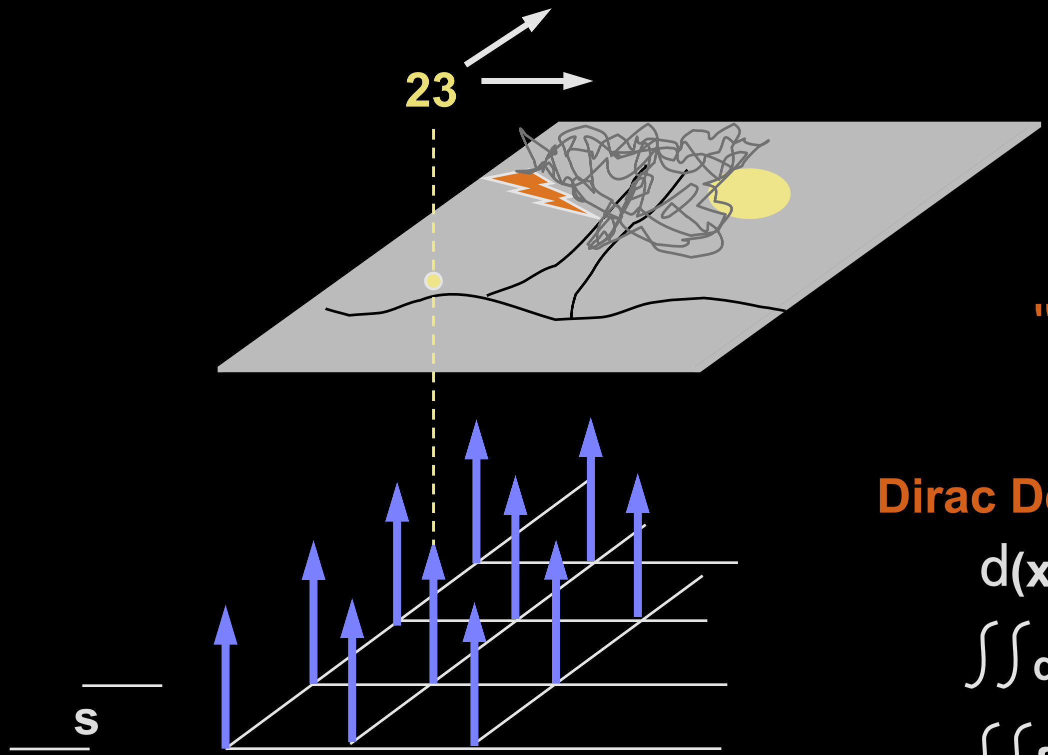
- Strategy:
  - Detect points of interest with low resolution sampling
  - “Foveate” to point of interest and use high resolution sampling.

Human Eye Sampling

Cartesian image ----- Log-Polar representation ----- Retinal representation



- Rough Idea: Ideal Case



"Digitized Image"

"Continuous Image"

Dirac Delta Function 2D "Comb"

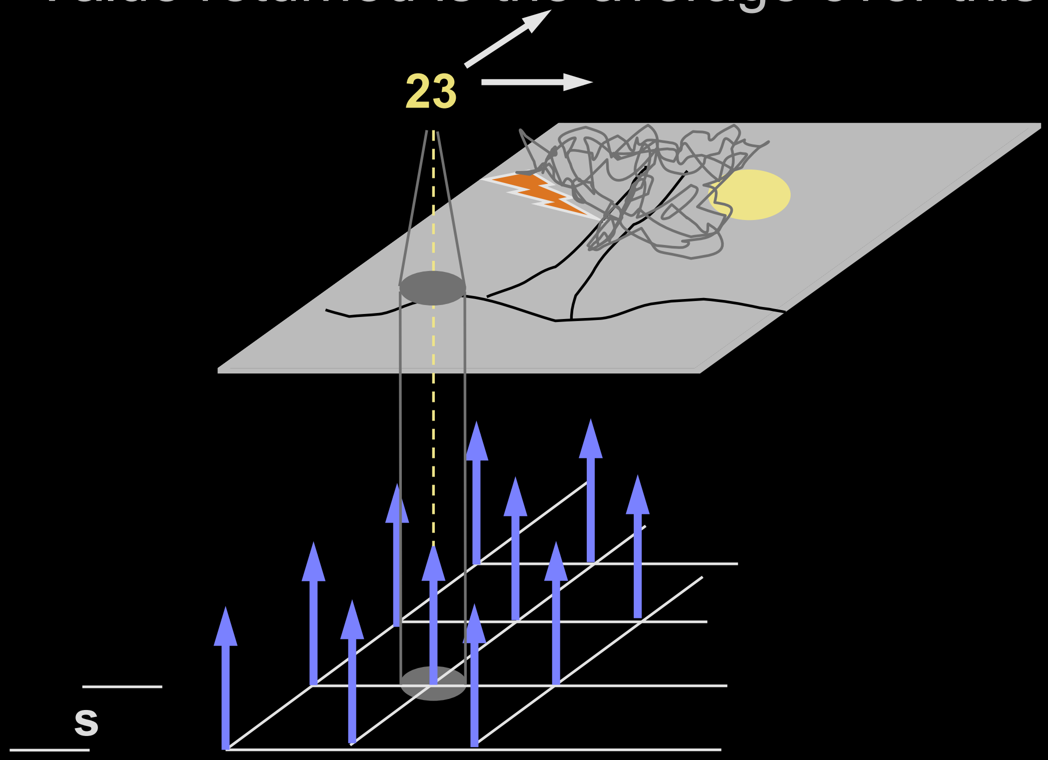
d(x,y) = 0 for x ≠ 0, y≠ 0

∫∫d(x,y) dx dy = 1

∫∫f(x,y)d(x-a,y-b) dx dy = f(a,b)

d(x-ns,y-ns) for n = 1….32 (e.g.)

- Rough Idea: Actual Case
  - Can't realize an ideal point function in real equipment
  - "Delta function" equivalent has an area
  - Value returned is the average over this area



Projection through a pixel

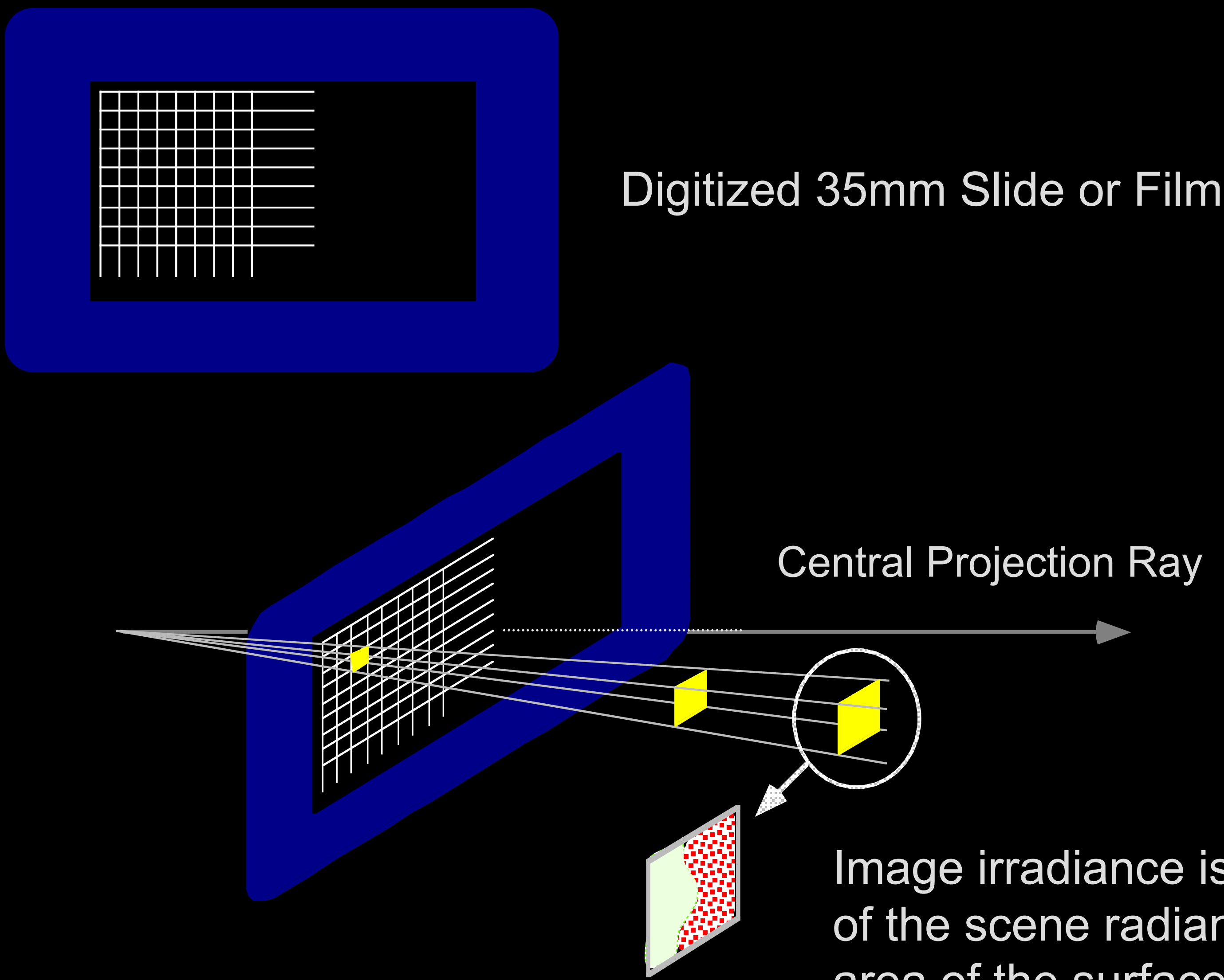


Image irradiance is the average of the scene radiance over the area of the surface intersecting the solid angle!

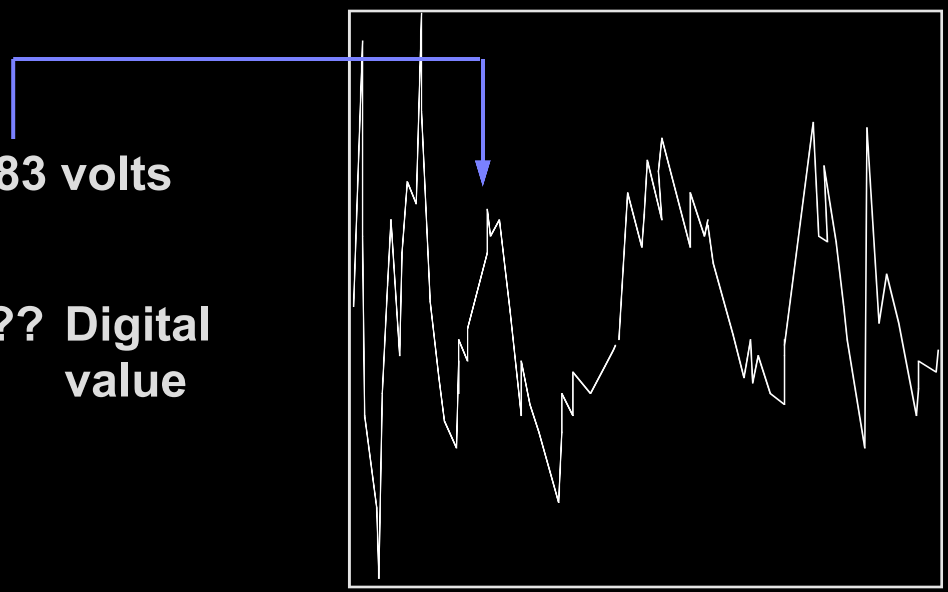
Mixed Pixel Problem

Signal Quantization

- Goal: determine a mapping from a continuous signal (e.g. analog video signal) to one of K discrete (digital) levels.

I(x,y) = .1583 volts

= ???? Digital value



Quantization

- I(x,y) = continuous signal: $0 \leq I \leq M$
- Want to quantize to K values 0,1,....K-1
- K usually chosen to be a power of 2:

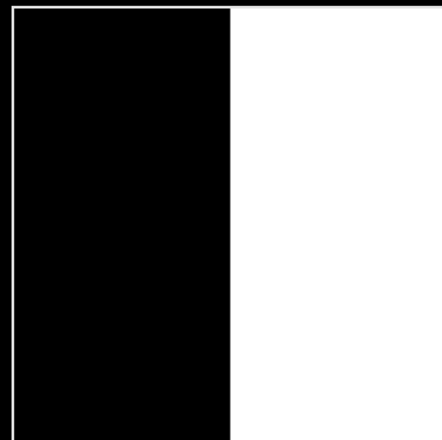
K: #Levels	#Bits
2	1
4	2
8	3
16	4
32	5
64	6
128	7
256	8

- Mapping from input signal to output signal is to be determined.
- Several types of mappings: uniform, logarithmic, etc.

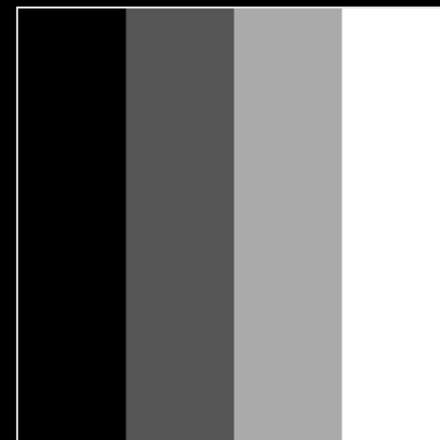
Choice of K

Original

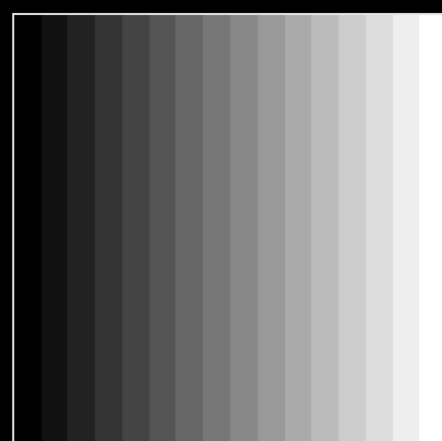
Linear Ramp



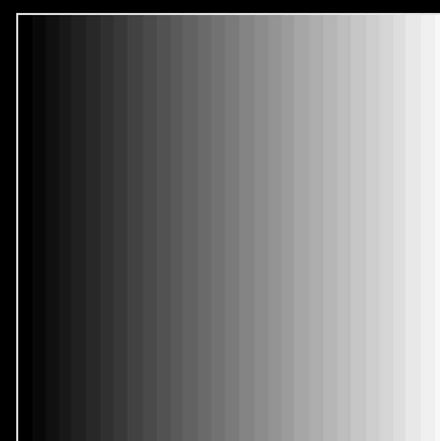
K=2



K=4

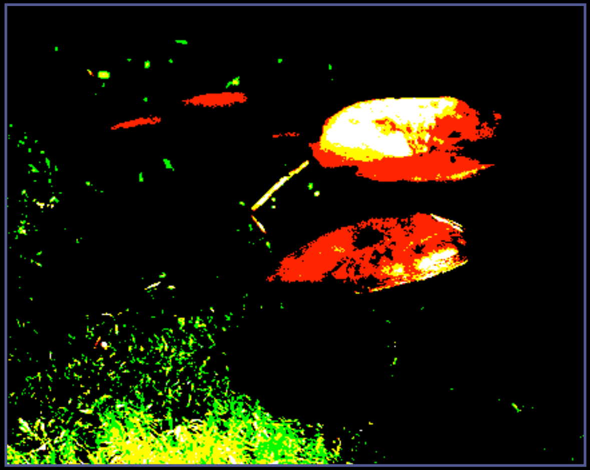


K=16



K=32

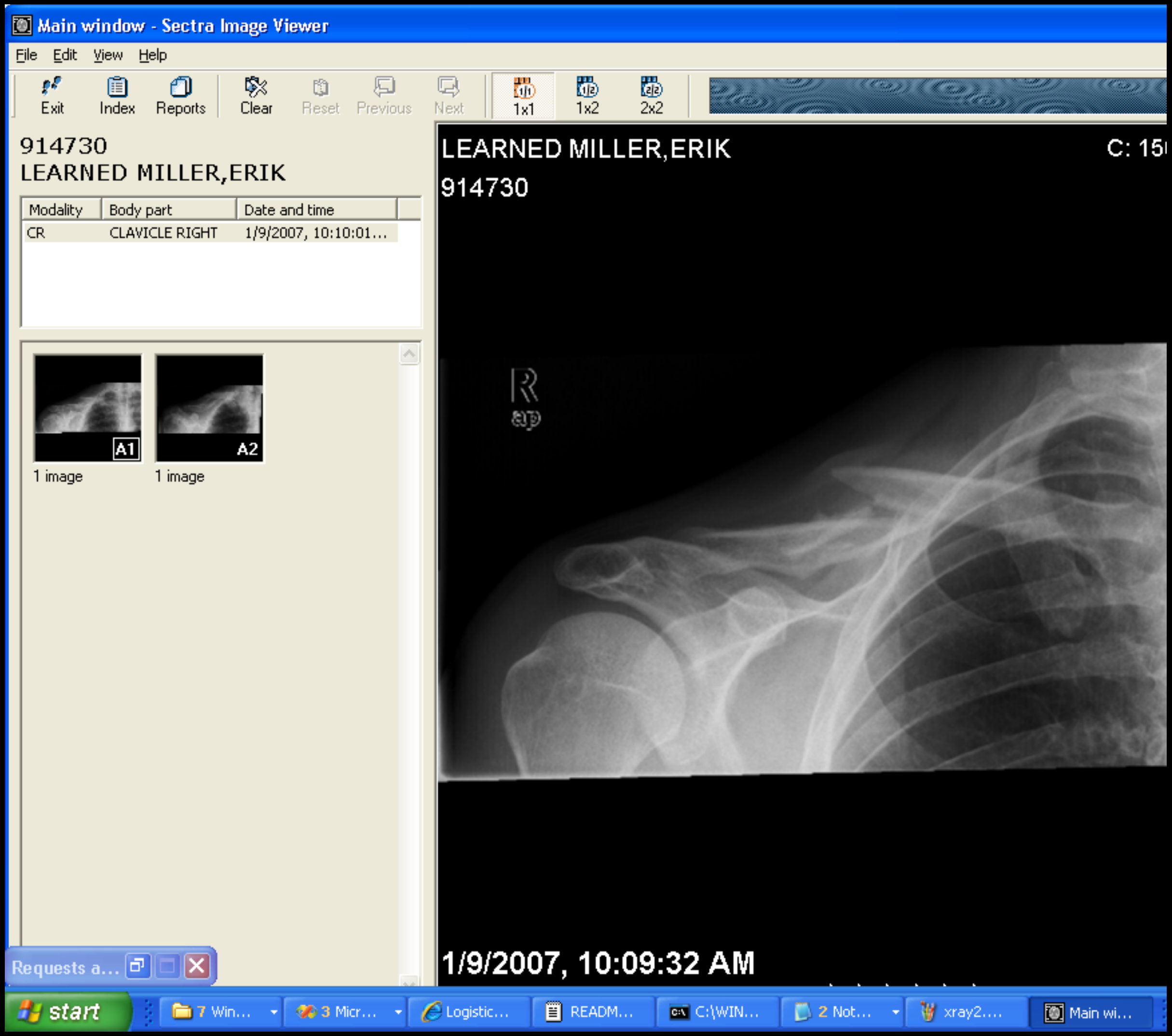
Choice of K



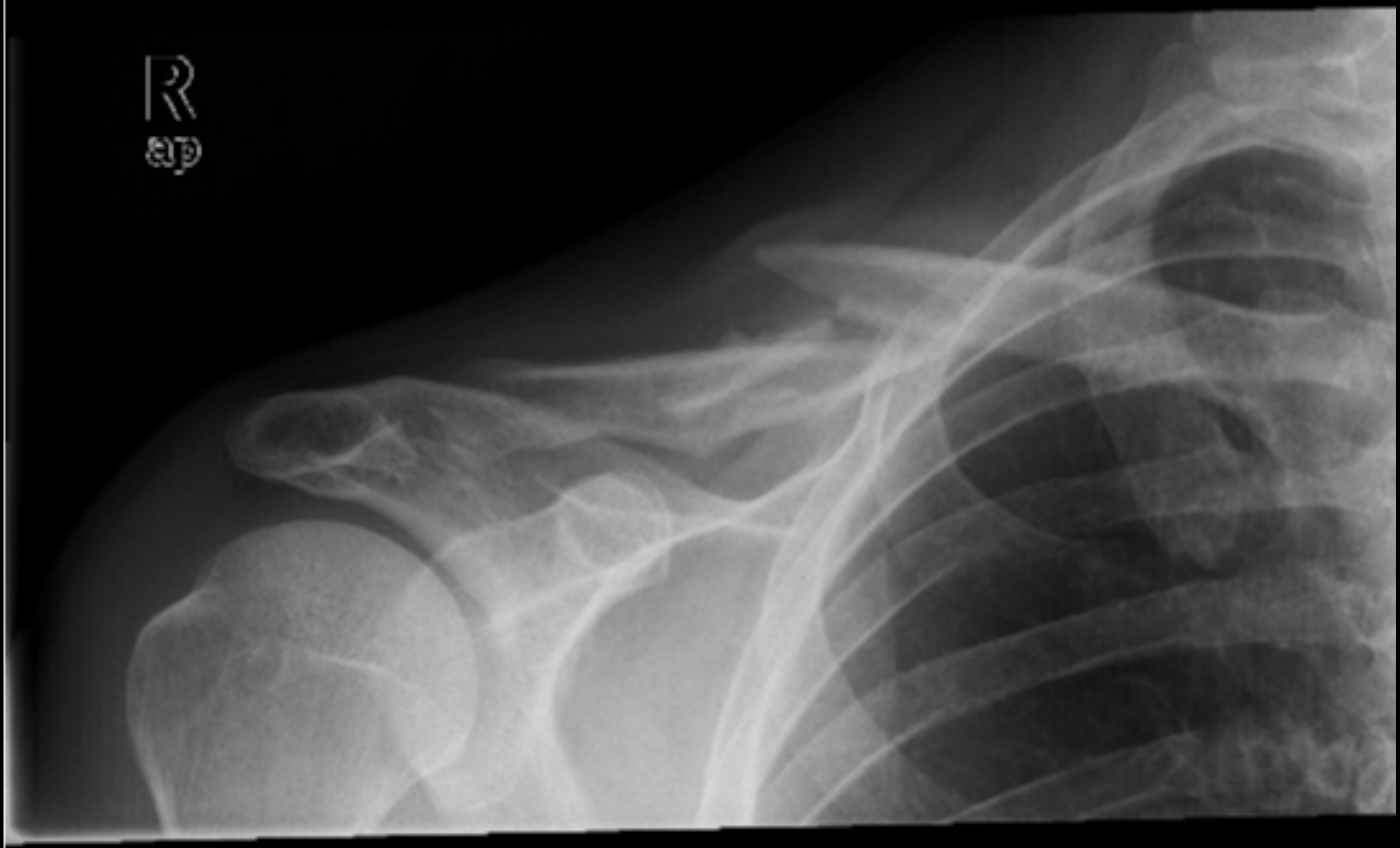
K=2 (each color)

K=4 (each color)

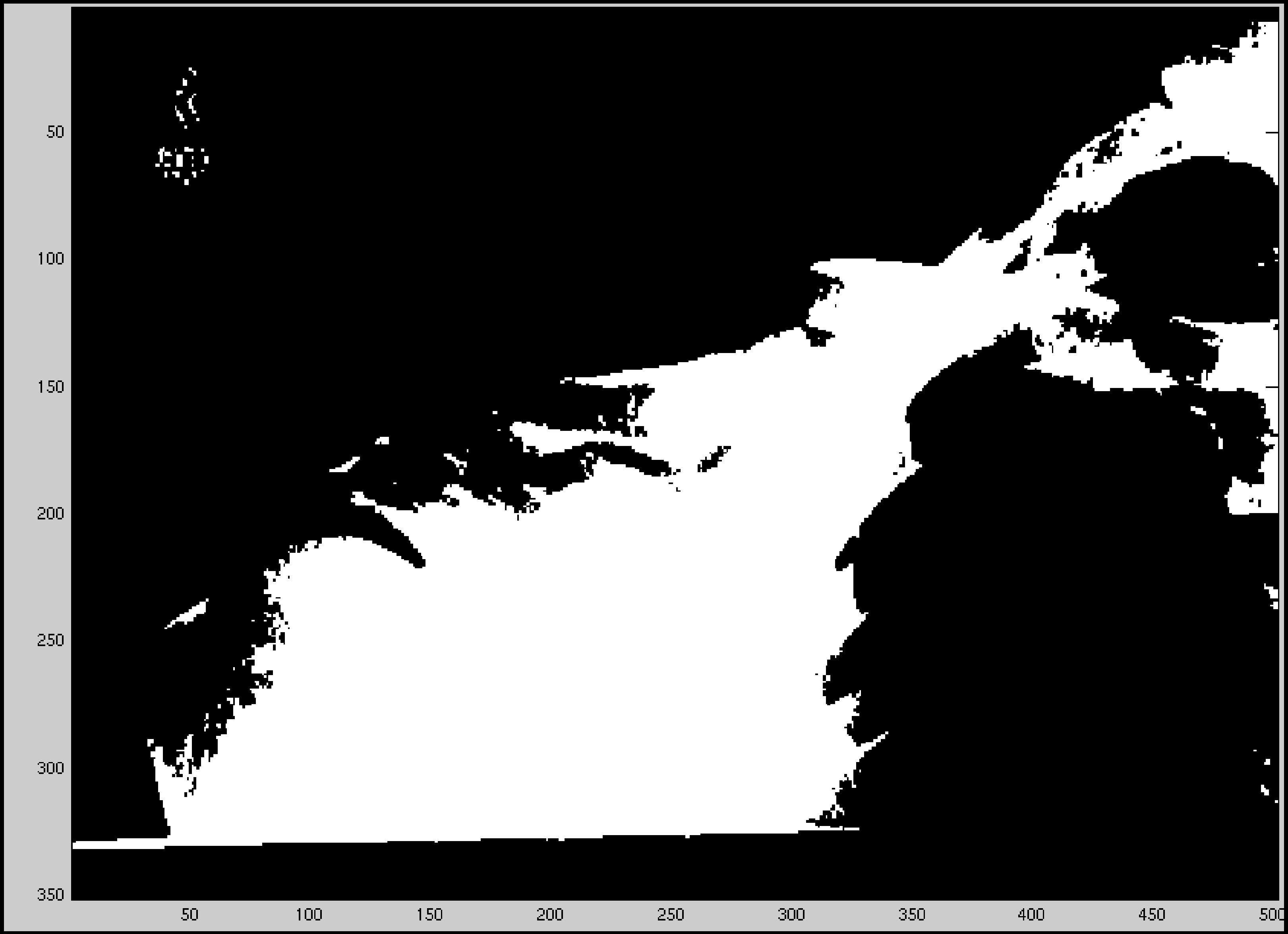
Digital X-rays



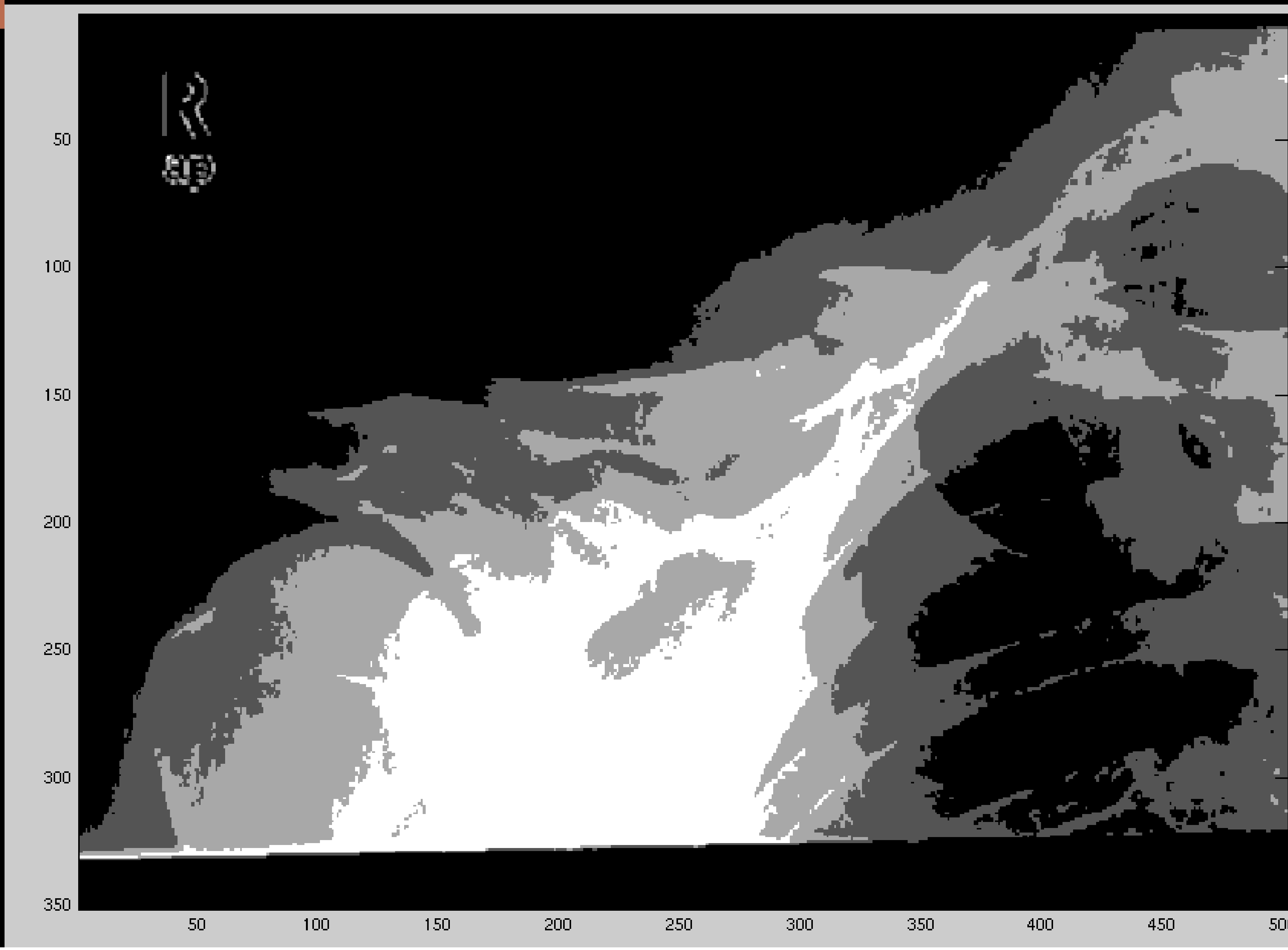
Digital X-rays: 8 is enough?



Digital X-rays: 1 bit

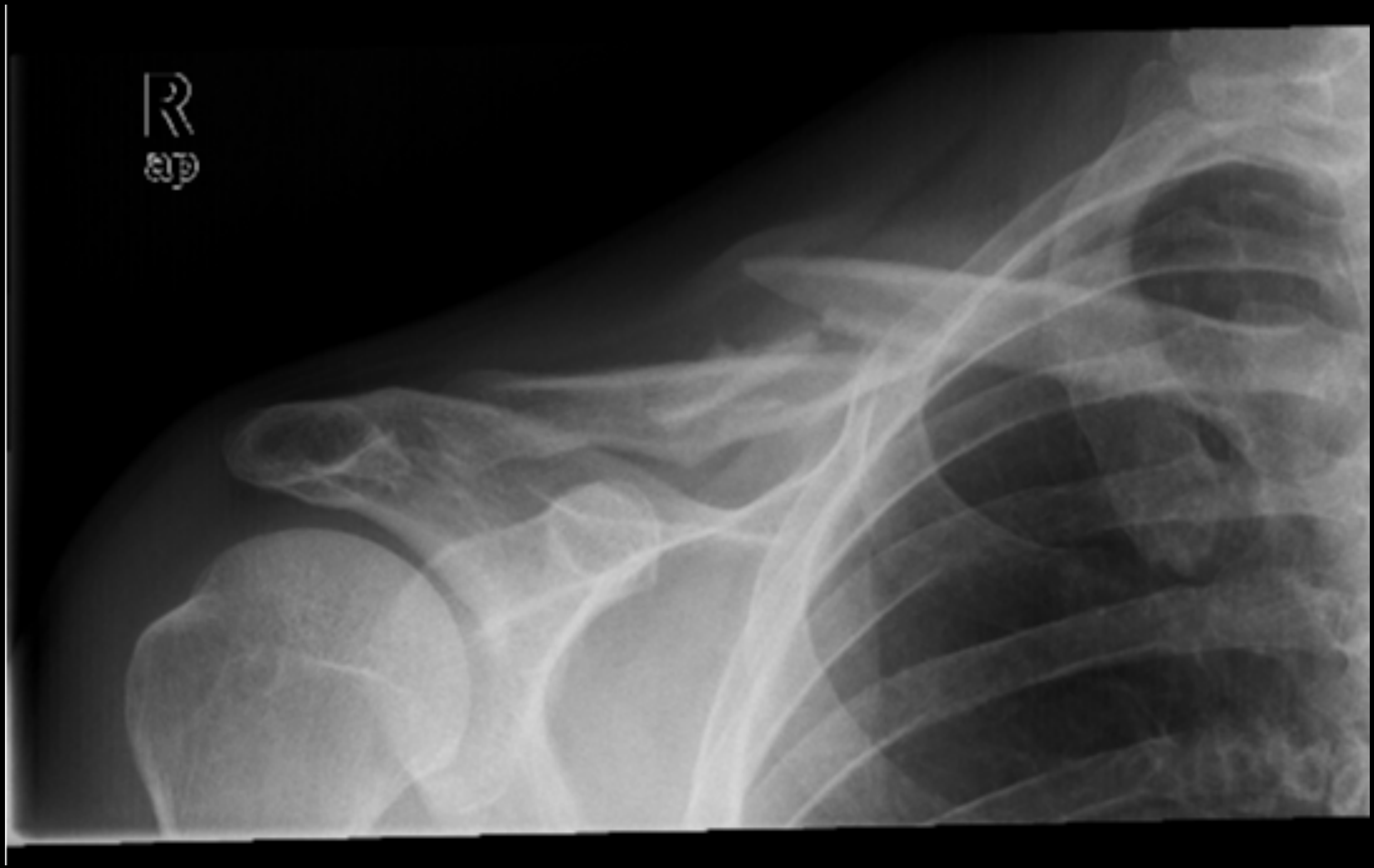


Digital X-rays: 2 bits



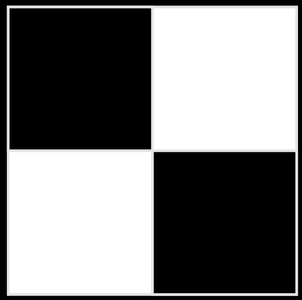
Digital X-rays: 3 bit

Digital X-rays: 8 is enough?



Gray Levels-Resolution Trade Off

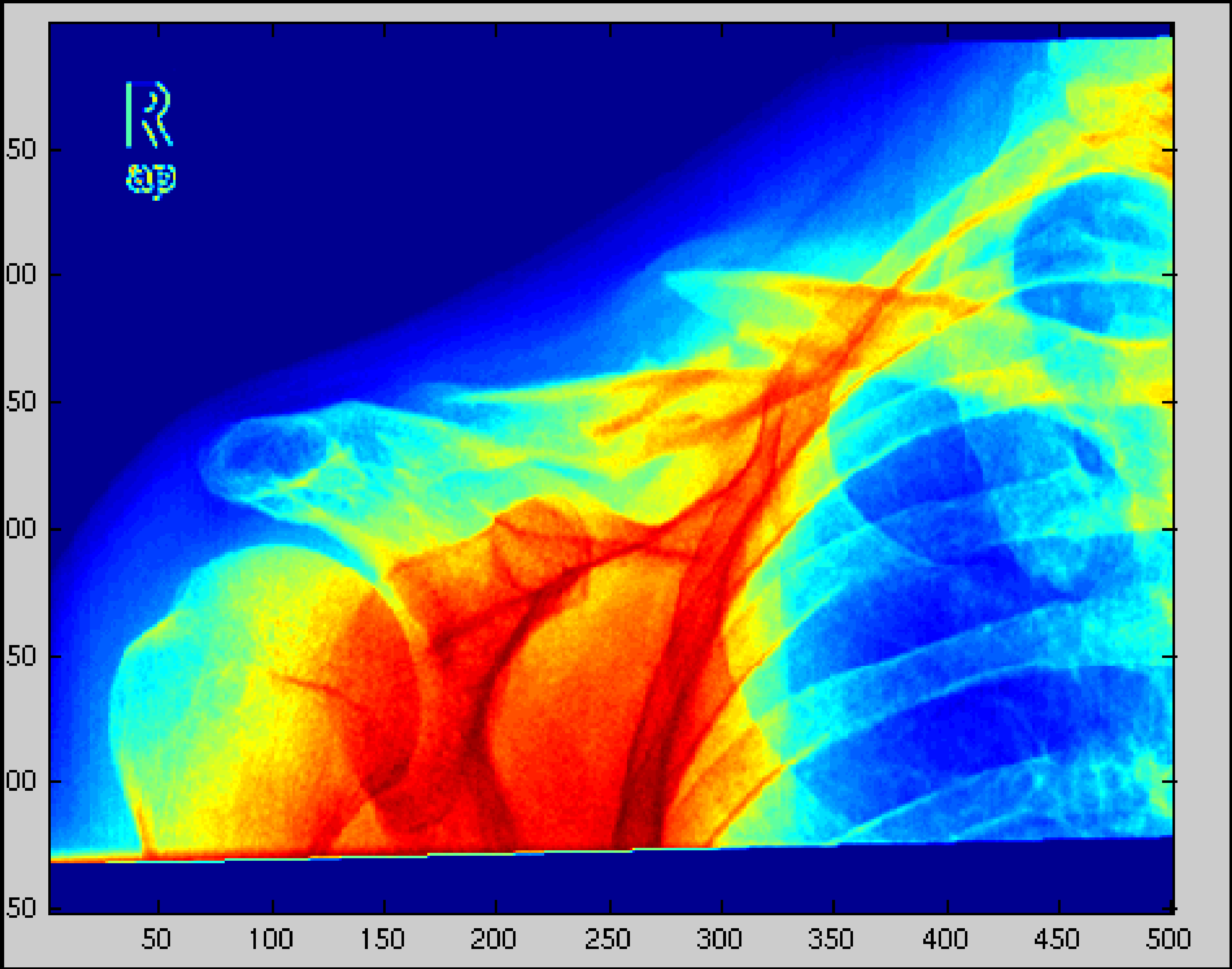
- More gray levels can be simulated with more resolution.
- A “gray” pixel:



- Doubling the resolution in each direction adds at least four new gray levels. But maybe more?

Introduction to
Computer Vision

Pseudocolor



Digital X-rays: 8 is enough?