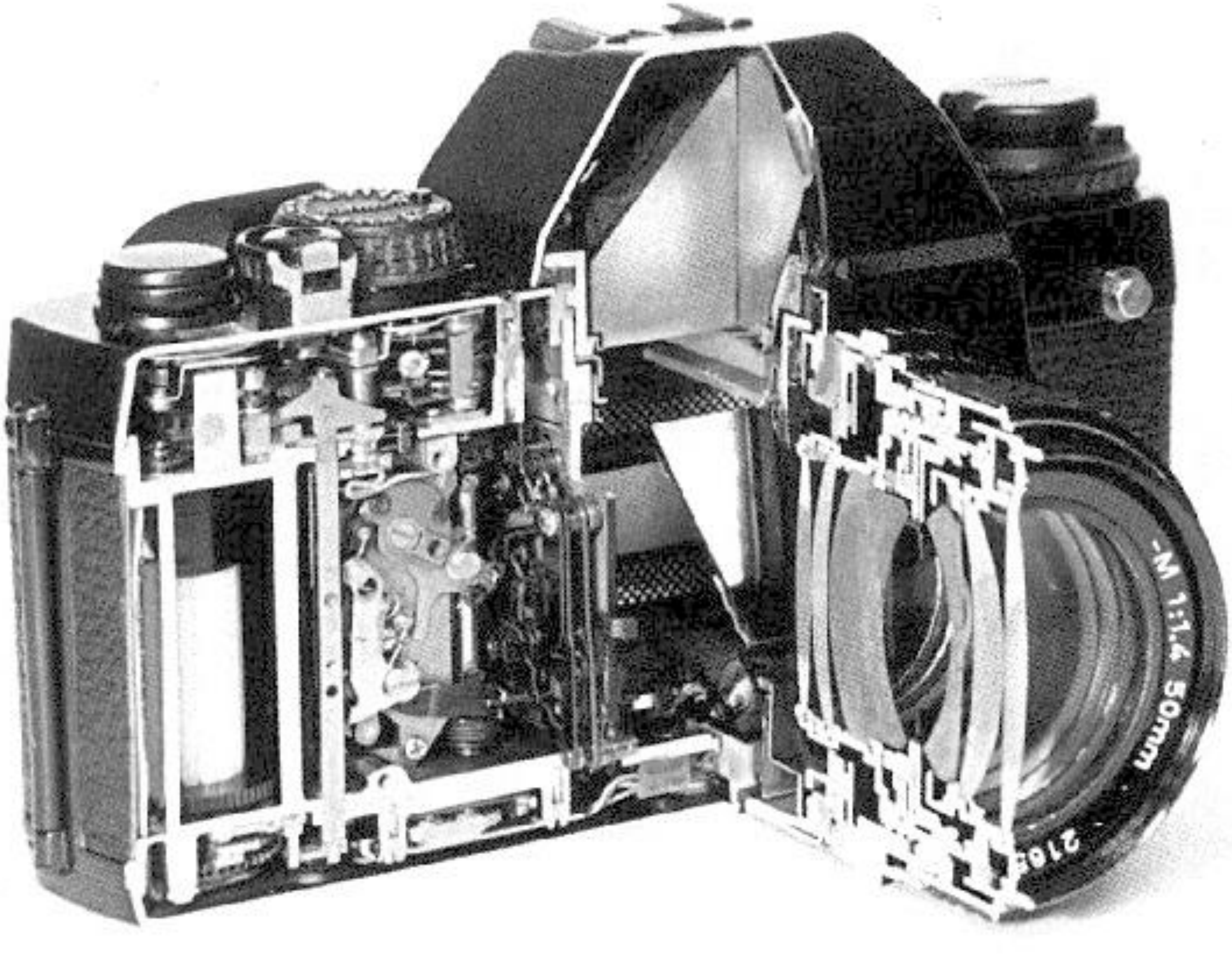
Image Processing - Intro

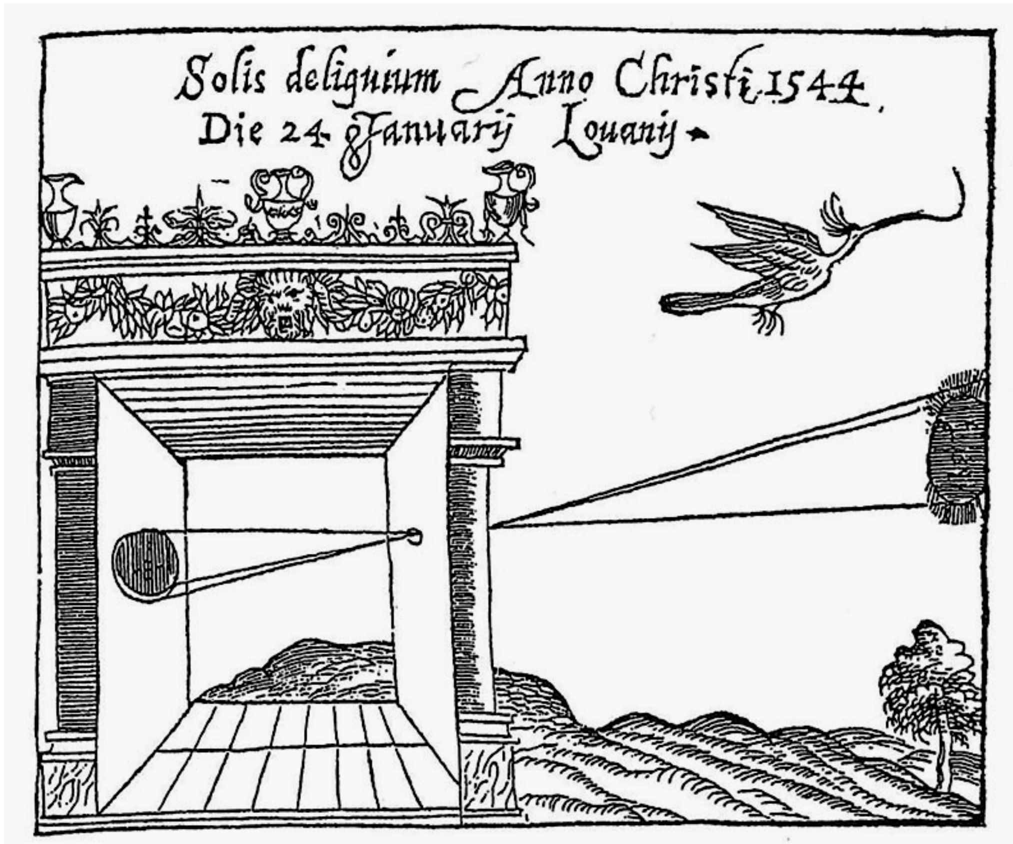
Tamás Szirányi

The path of light through optics



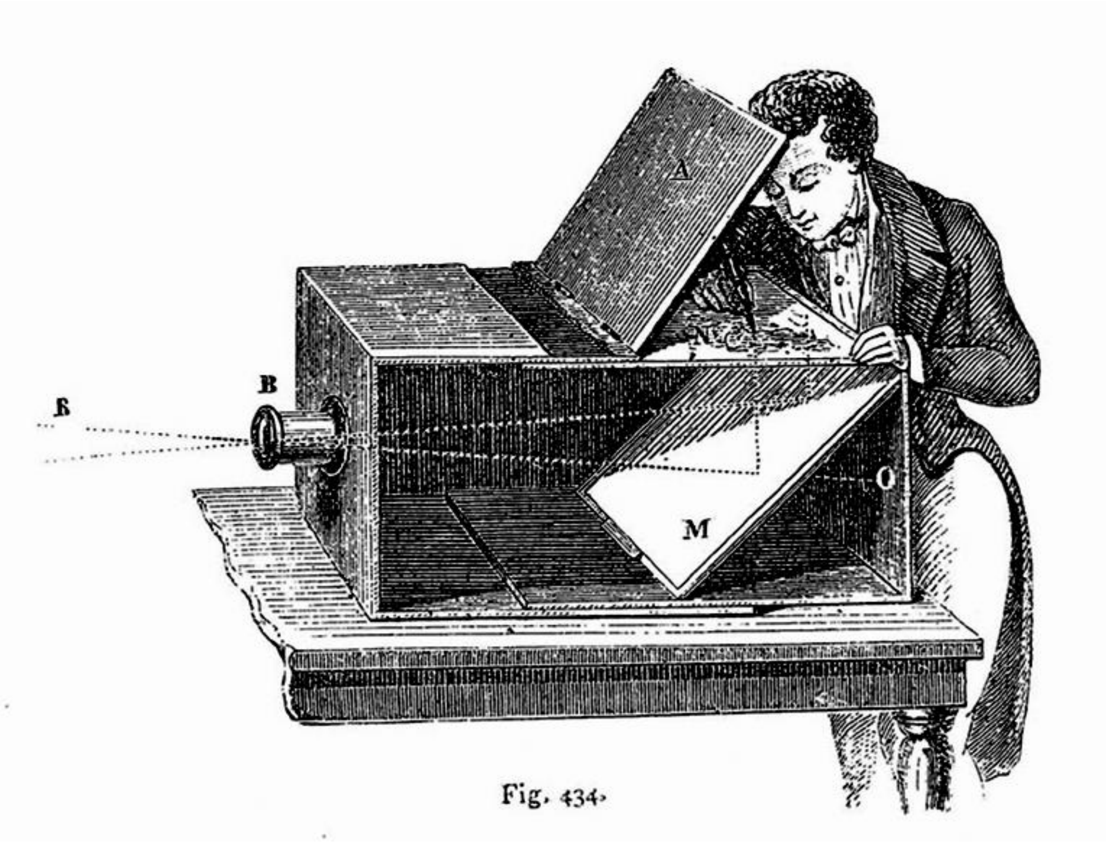
A Brief History of Images

1558

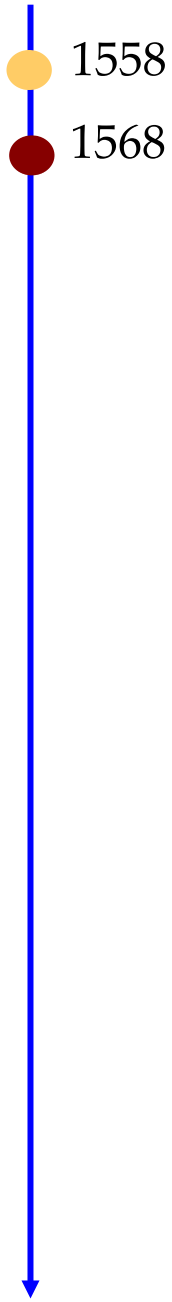


Camera Obscura, Gemma Frisius, 1558

A Brief History of Images

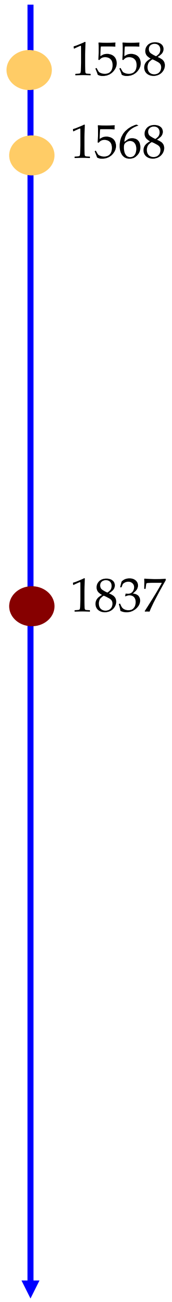


Lens Based Camera Obscura, 1568

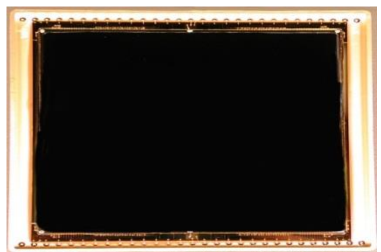


A Brief History of Images

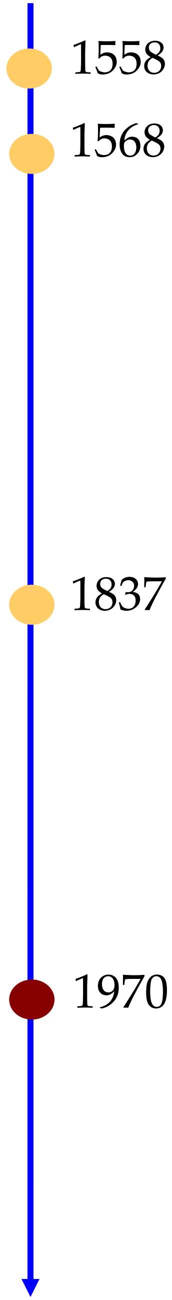
Still Life, Louis Jaques Mande Daguerre, 1837



A Brief History of Images



Silicon Image Detector, 1970



A Brief History of Images

Digital Cameras

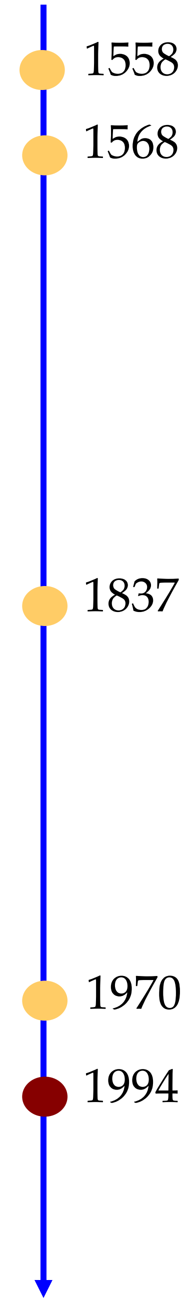
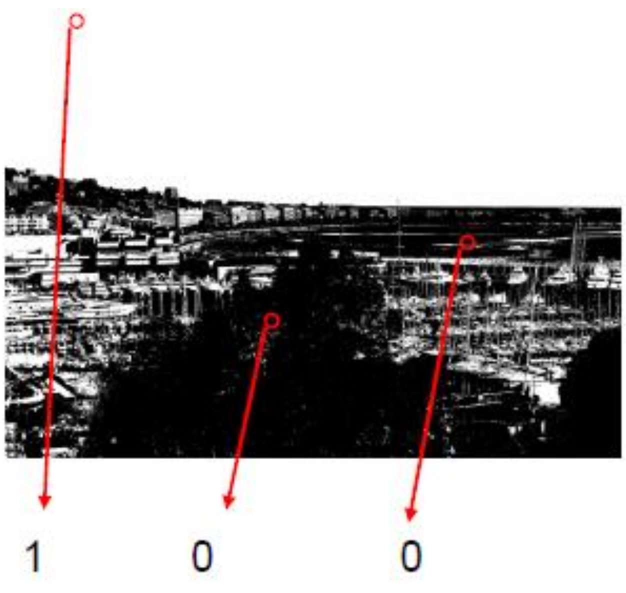


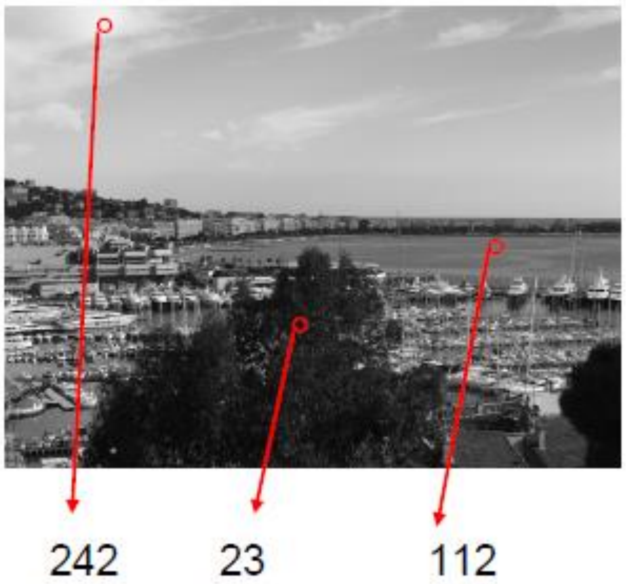
Image matrix at the input

- Digital image = 2D pixel array; (x,y): pixel coordinates

Binary: b(x,y)



Grayscale: f(x,y)
f(x,y) : 0 … 255

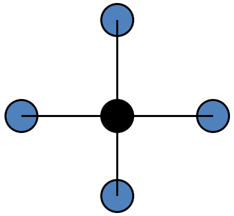


Color image array

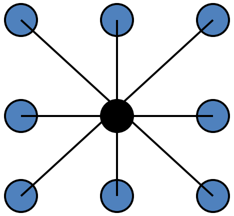
- Color image: 2D pixelarray (RGB)

Commonly–used Terminology

Neighbors of a pixel $p=(i,j)$



$N_4(p)=\{(i-1,j),(i+1,j),(i,j-1),(i,j+1)\}$



$N_8(p)=\{(i-1,j),(i+1,j),(i,j-1),(i,j+1),$
$(i-1,j-1),(i-1,j+1),(i+1,j-1),(i+1,j+1)\}$

Adjacency

4-adjacency: p,q are 4-adjacent if p is in the set $N_4(q)$

8-adjacency: p,q are 8-adjacent if p is in the set $N_8(q)$

Note that if p is in $N_{4/8}(q)$, then q must be also in $N_{4/8}(p)$

Common Distance Definitions

Euclidean distance
(2-norm)

$2\sqrt{2}$	$\sqrt{5}$	2	$\sqrt{5}$	$2\sqrt{2}$
$\sqrt{5}$	$\sqrt{2}$	1	$\sqrt{2}$	$\sqrt{5}$
2	1	0	1	2
$\sqrt{5}$	$\sqrt{2}$	1	$\sqrt{2}$	$\sqrt{5}$
$2\sqrt{2}$	$\sqrt{5}$	2	$\sqrt{5}$	$2\sqrt{2}$

D_4 distance
(city-block distance)

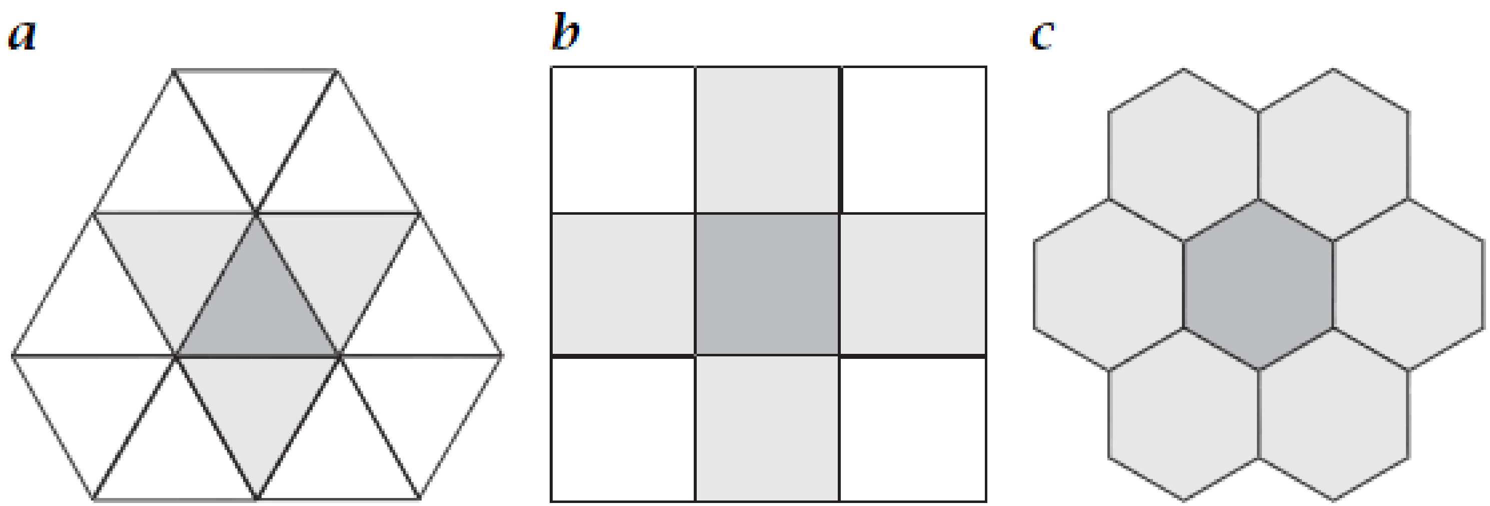
4	3	2	3	4
3	2	1	2	3
2	1	0	1	2
3	2	1	2	3
4	3	2	3	4

D_8 distance
(checkboard distance)

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

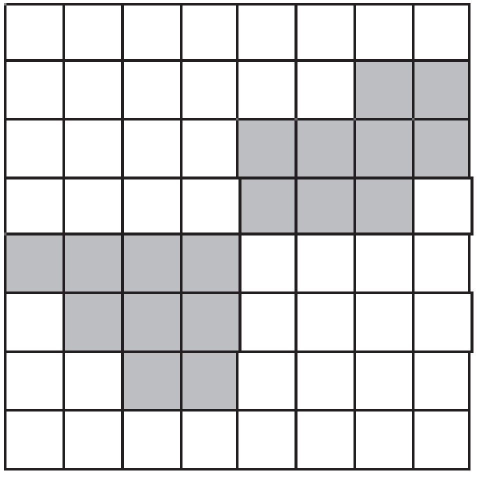
Picture elements

- Picture in 2D and 3D:
  - 2D: pixel (picture element)
  - 3D: voxel (volume element)
- Pixel geometry:



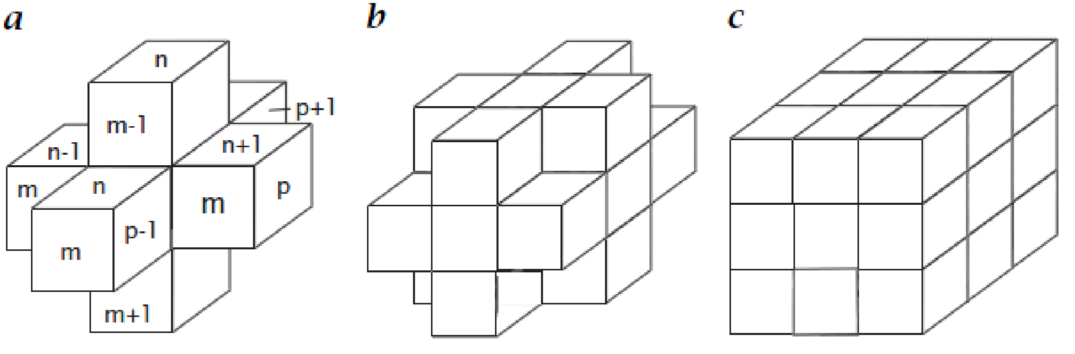
Neighborhood

- Neighborhood in 2D :
  - 4 or 8 connections

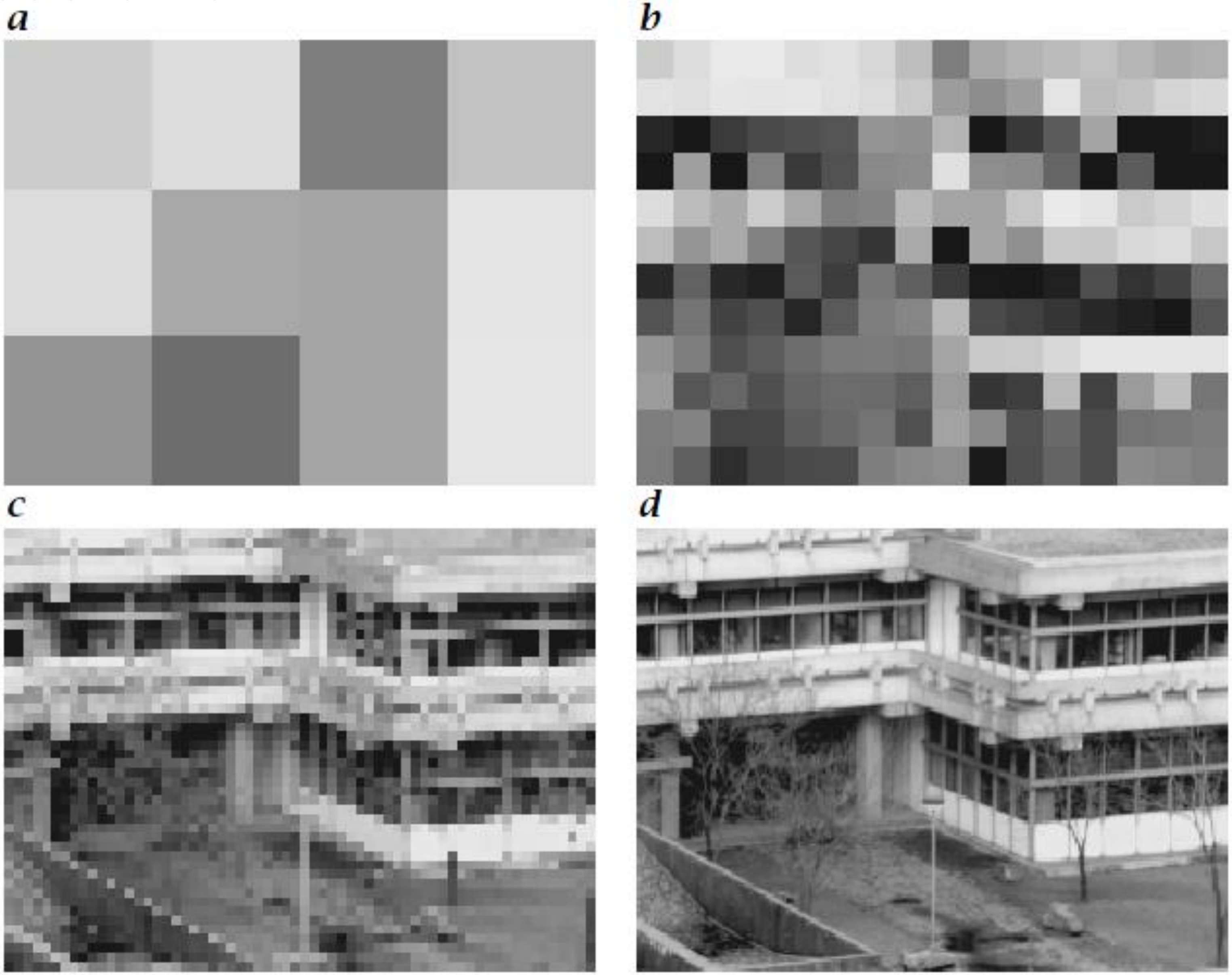


For 3D we have more cases:

– Side (6), edge(18), corner(26)



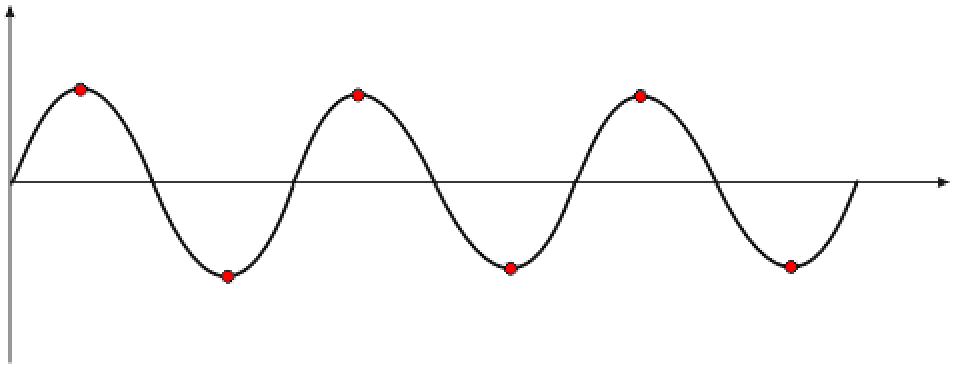
Resolution



Resolution

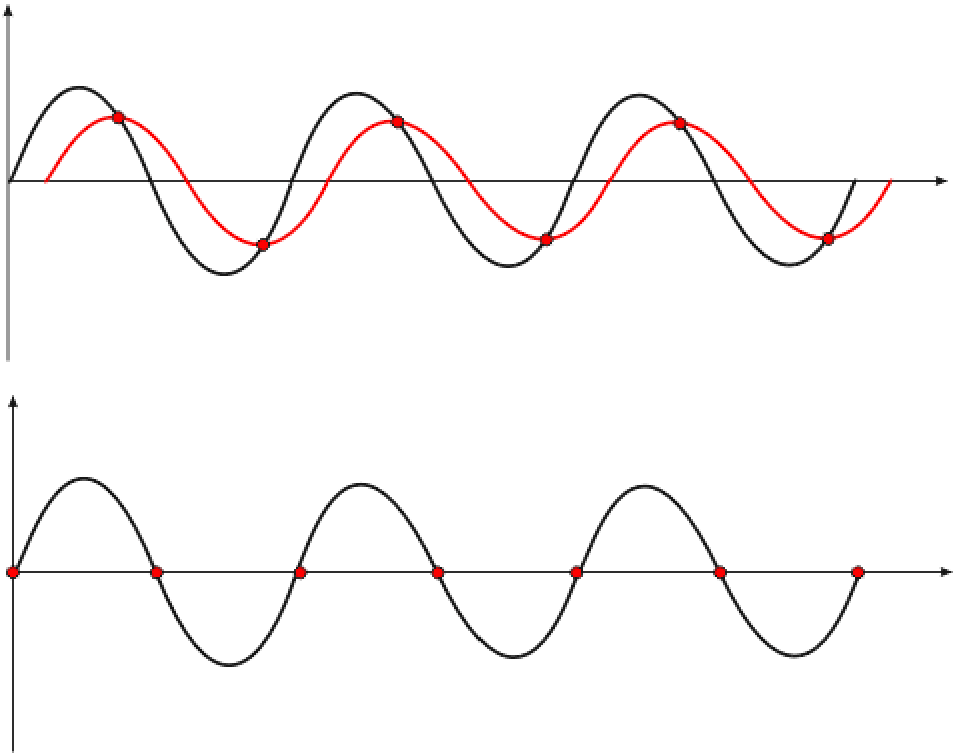
Sampling theorem

– Double highest frequency fits the sampling rate



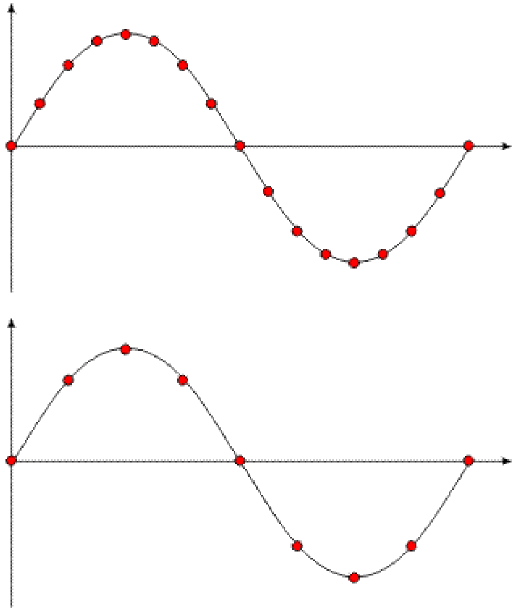
Sampling theory

- Sampling is just the double:



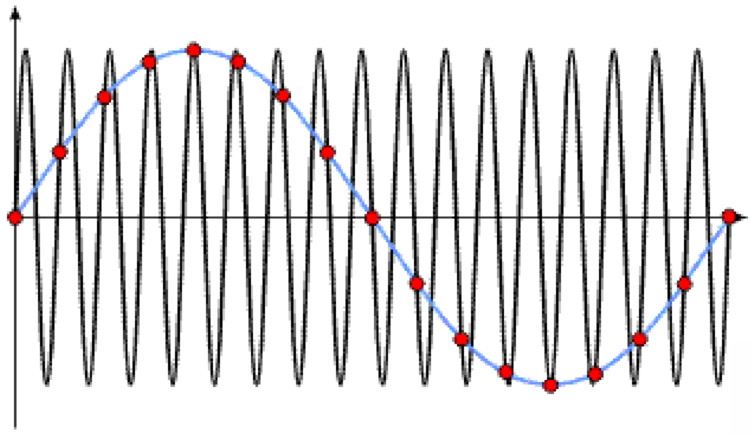
Nyquist frequency

- OK



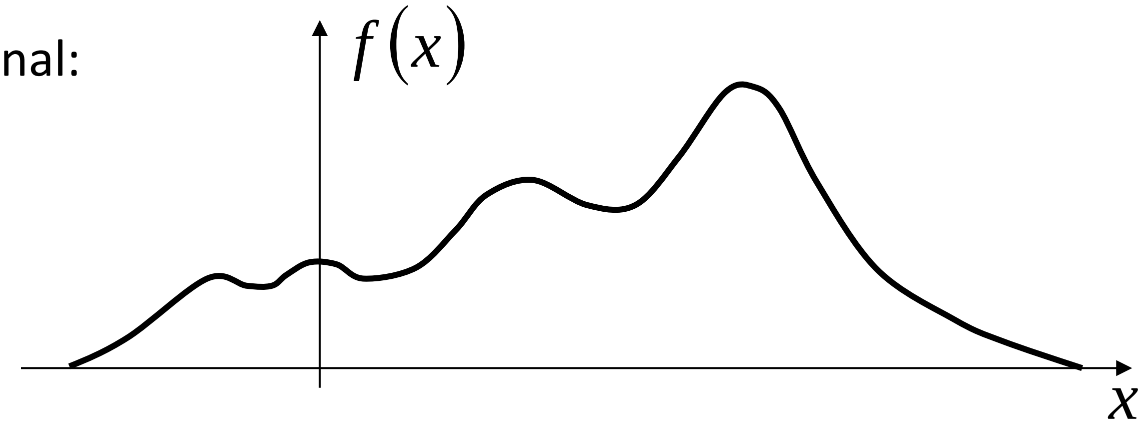
Nyquist sampling rate

- Bad rating:

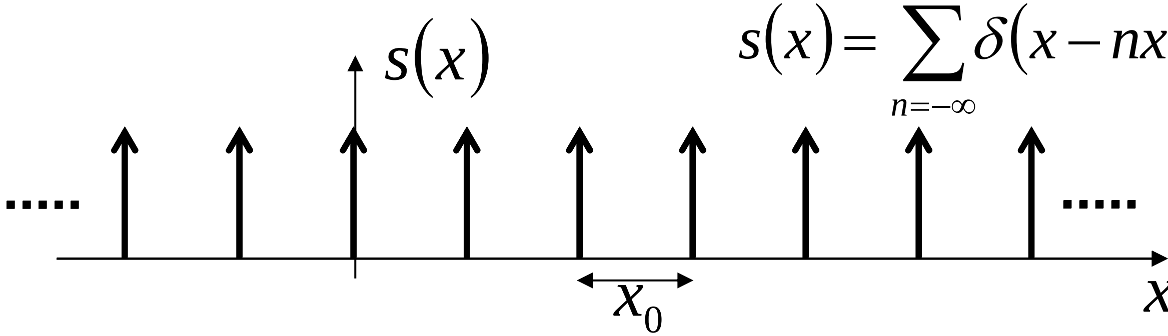


Sampling Theorem

Continuous signal:



Shah function (Impulse train):



$$s(x) = \sum_{n=-\infty}^{\infty} \delta(x - nx_0)$$

Sampled function:

$$f_s(x) = f(x)s(x) = f(x)\sum_{n=-\infty}^{\infty} \delta(x - nx_0)$$

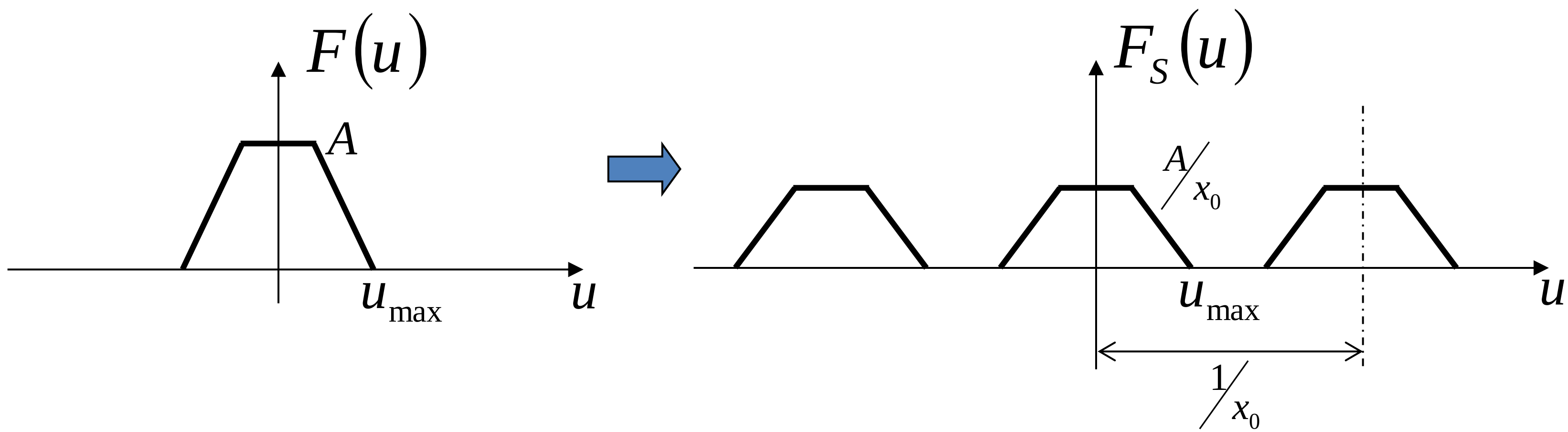
Sampling Theorem

Sampled function:

$$f_s(x) = f(x)s(x) = f(x)\sum_{n=-\infty}^{\infty}\delta(x - nx_0)$$

Sampling frequency $\frac{1}{x_0}$

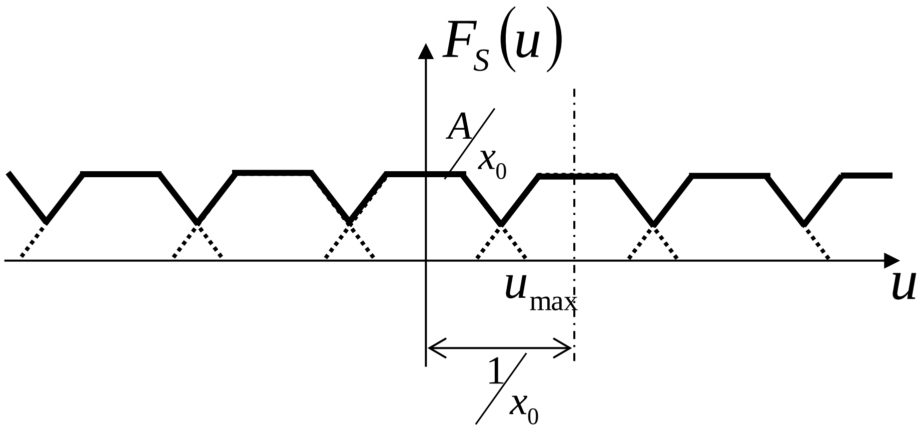
$$F_S(u) = F(u) * S(u) = F(u) * \frac{1}{x_0}\sum_{n=-\infty}^{\infty}\delta\left(u - \frac{n}{x_0}\right)$$



Only if $u_{max} \leq \frac{1}{2x_0}$

Nyquist Theorem

If $u_{max} > \frac{1}{2x_0}$



Aliasing

When can we recover $F(u)$ from $F_S(u)$?

Only if $u_{max} \leq \frac{1}{2x_0}$ (Nyquist Frequency)

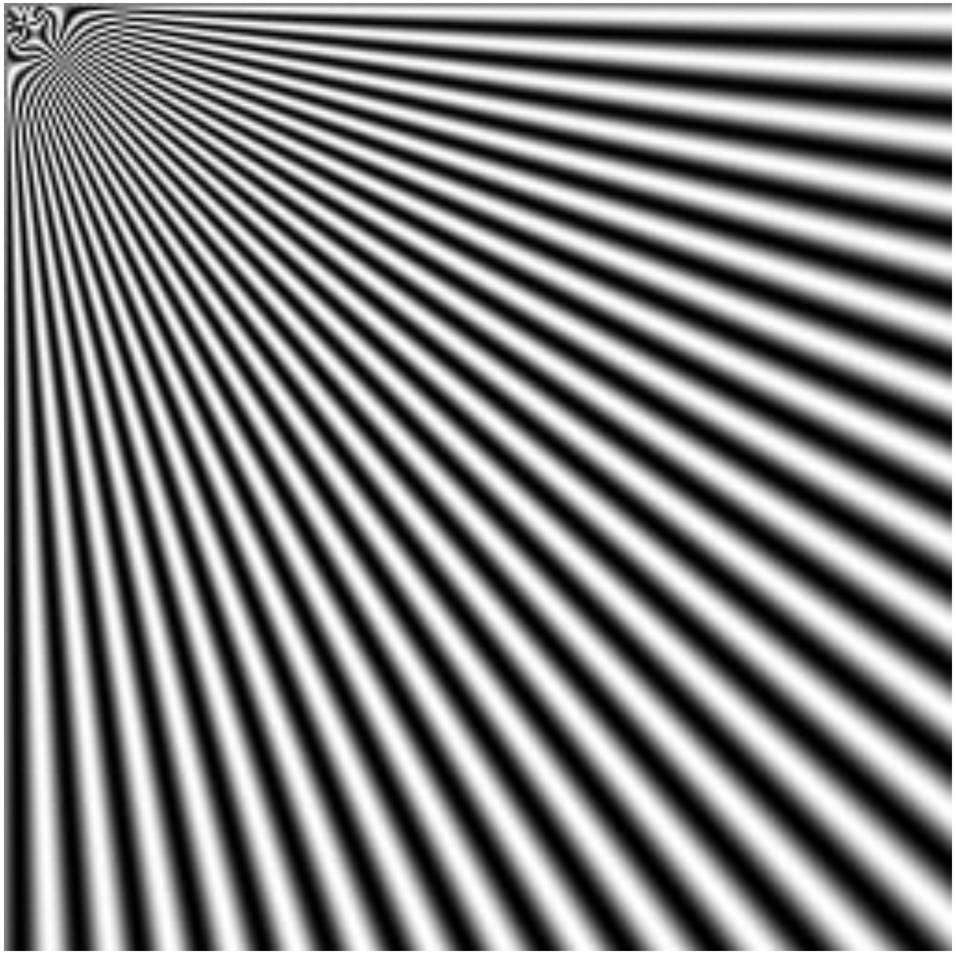
We can use

$$C(u) = \begin{cases} x_0 & |u| < 1/2x_0 \\ 0 & \text{otherwise} \end{cases}$$

Then $F(u) = F_S(u)C(u)$ and $f(x) = \text{IFT}[F(u)]$

Sampling frequency must be greater than $2u_{max}$

Aliasing



Aliasing in Digital Images

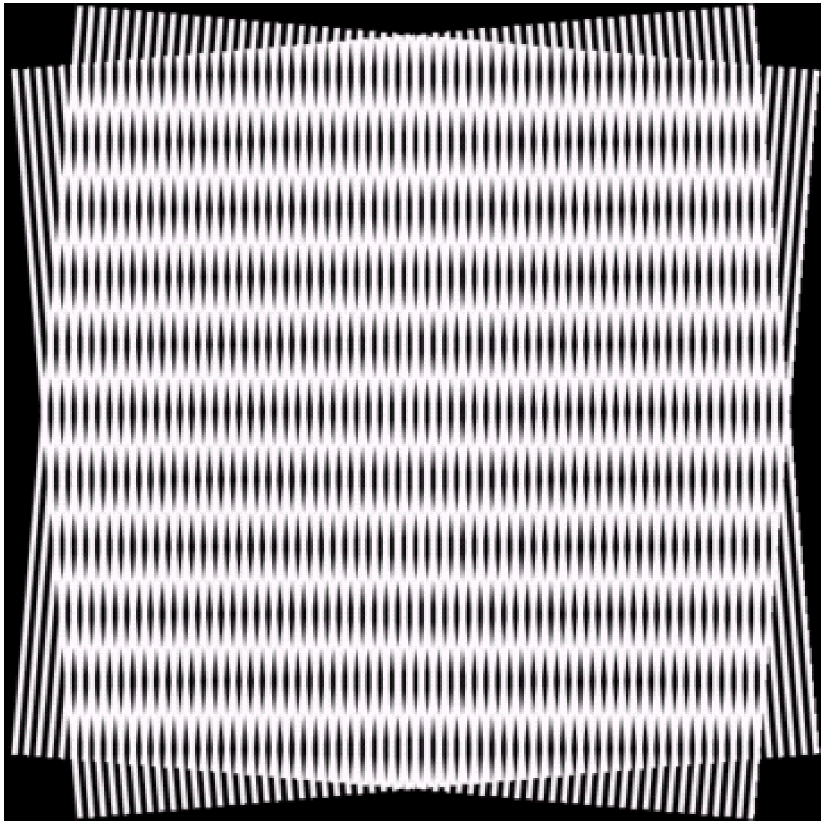
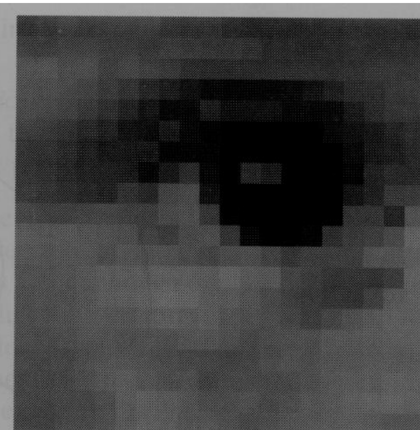


FIGURE 2.24 Illustration of the Moiré pattern effect.

Image Formation Fundamentals

How are images represented in the computer?



117	125	133	127	130	130	133	121	116	115	100	91	93	94	99	103	112	105	109	106
134	133	138	138	132	134	130	133	128	123	121	113	106	102	99	106	113	109	109	113
146	147	138	140	125	134	124	115	102	96	93	94	99	96	99	100	103	110	109	110
144	141	136	130	120	108	88	74	53	37	31	37	35	39	53	79	93	100	109	116
139	136	129	119	102	85	58	31	41	77	51	53	53	33	37	41	69	94	105	108
132	127	117	102	87	57	49	77	42	28	17	15	13	13	17	41	53	69	88	100
124	120	108	94	72	74	72	31	35	31	15	13	15	11	15	13	46	75	83	96
125	115	102	93	88	82	42	79	113	41	19	100	82	11	11	17	31	91	99	100
124	116	109	99	91	113	99	140	144	57	20	20	15	11	15	17	63	87	119	124
136	133	133	135	138	133	132	144	150	120	24	17	15	15	17	20	115	113	88	150
158	157	157	154	149	145	133	127	146	150	116	35	20	19	28	105	124	128	141	171
155	154	156	155	146	155	154	154	147	139	148	150	138	120	128	129	130	151	156	165
150	151	154	162	166	167	169	174	172	167	177	166	164	140	134	120	121	120	127	172
145	149	151	157	165	169	173	179	176	166	166	157	145	136	129	124	120	136	163	168
144	148	153	160	159	158	165	172	165	169	157	151	149	141	130	140	151	162	169	167
144	141	147	155	154	149	156	151	157	157	151	144	147	147	149	159	158	159	166	165
139	140	140	150	153	151	150	146	140	139	138	140	145	151	149	156	156	162	162	161
136	134	138	146	156	164	153	146	145	136	139	139	140	141	149	157	159	161	169	166
136	133	136	135	144	159	168	159	151	142	141	145	139	146	153	156	164	167	172	168
133	129	140	142	146	159	167	165	154	151	146	141	147	154	156	160	161	157	153	154

Color images

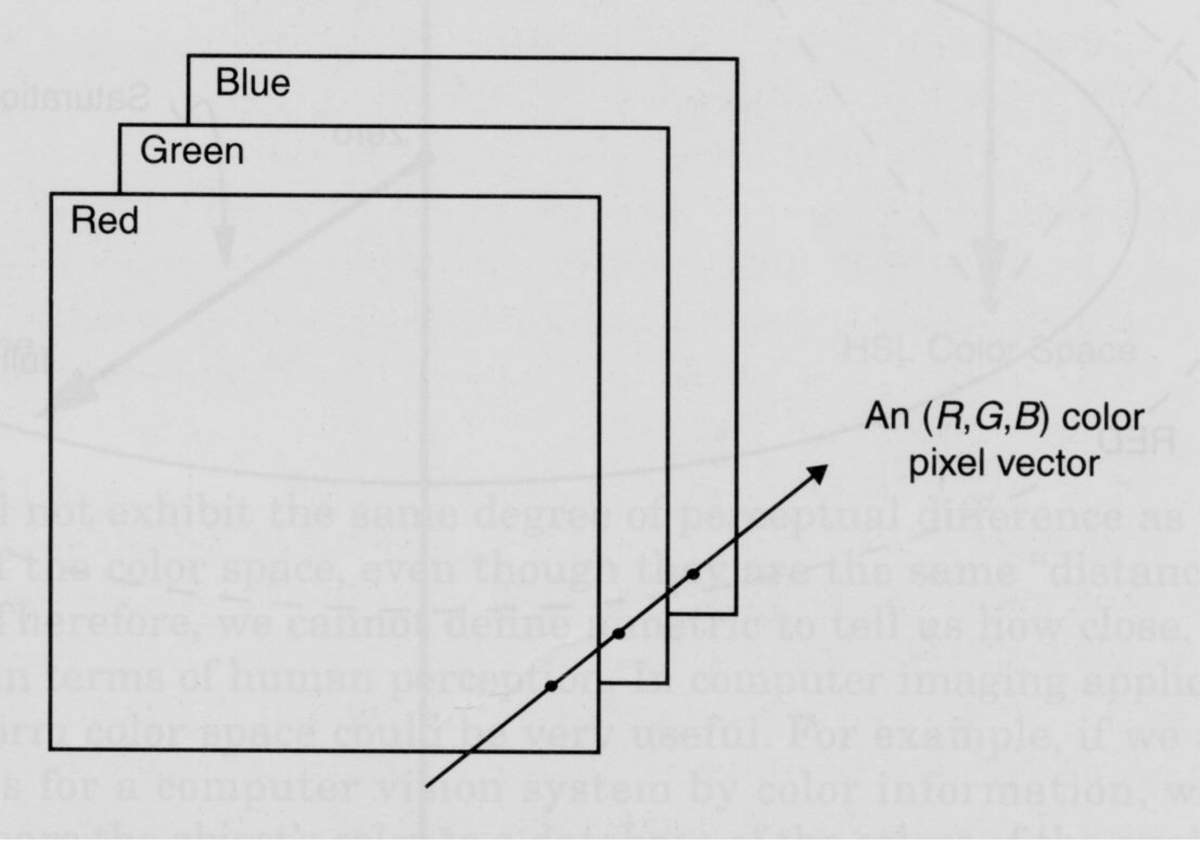
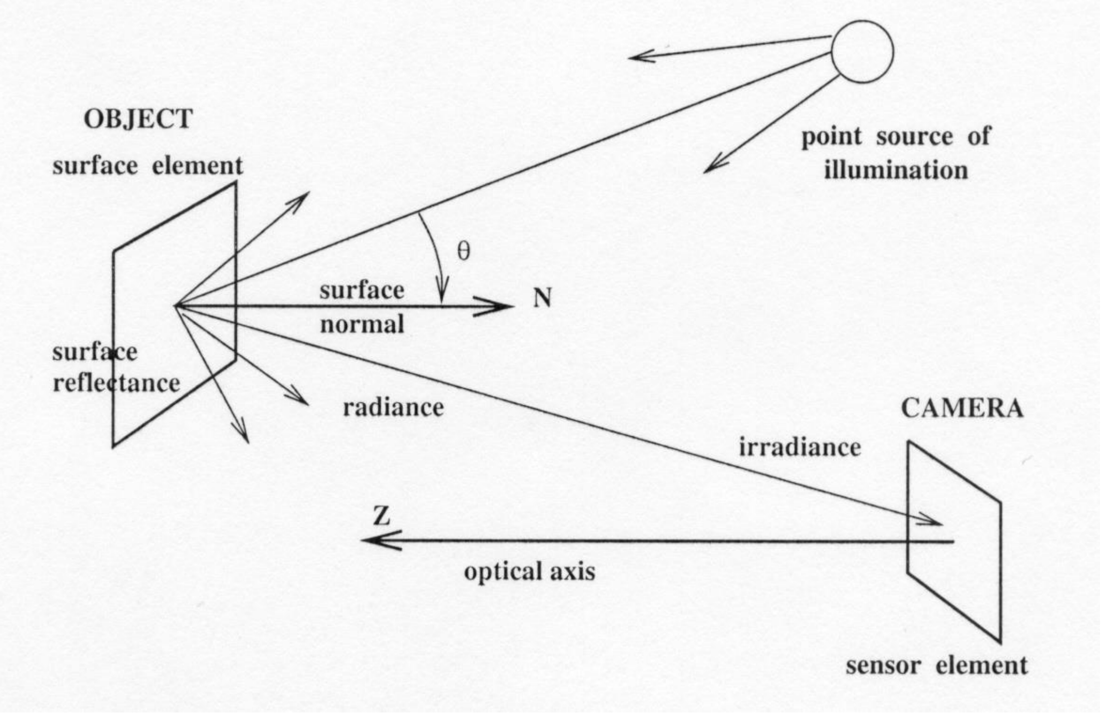


Image formation

- There are two parts to the image formation process:
  - The **geometry of image formation**, which determines where in the image plane the projection of a point in the scene will be located.
  - The **physics of light**, which determines the brightness of a point in the image plane as a function of illumination and surface properties.

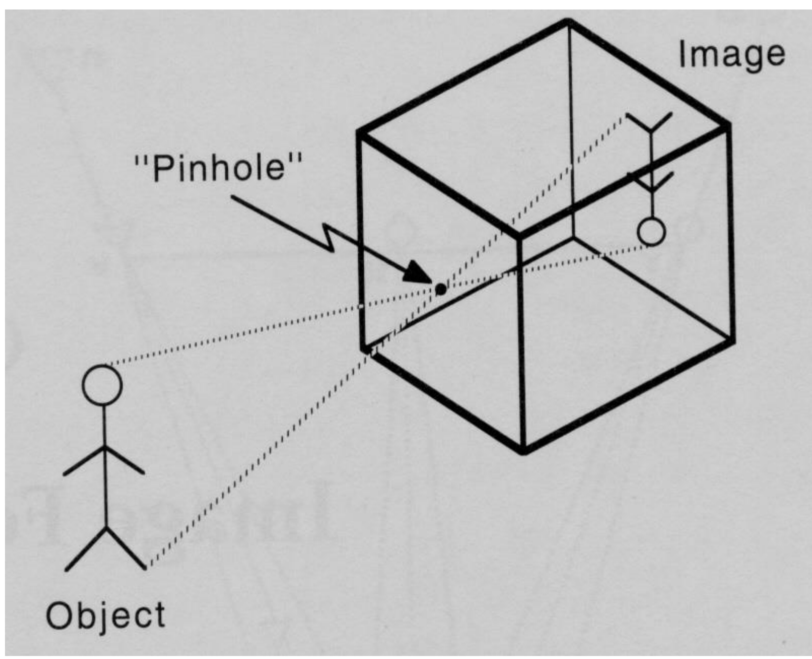
A Simple model of image formation

- The scene is illuminated by a single source.
- The scene reflects radiation towards the camera.
- The camera senses it via chemicals on film.



Pinhole camera

- This is the simplest device to form an image of a 3D scene on a 2D surface.
- Straight rays of light pass through a “pinhole” and form an inverted image of the object on the image plane.



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

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

Camera optics

- In practice, the aperture must be larger to admit more light.
- Lenses are placed to in the aperture to **<u>focus</u>** the bundle of rays from each scene point onto the corresponding point in the image plane

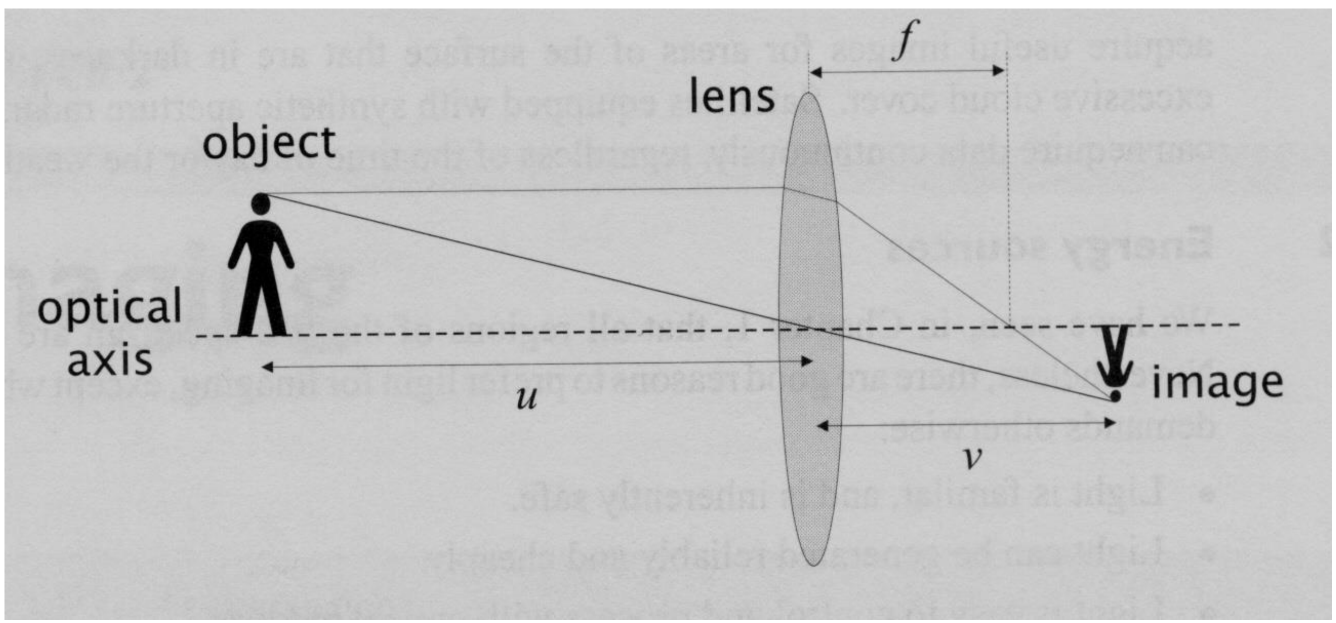


Image formation (cont’d)

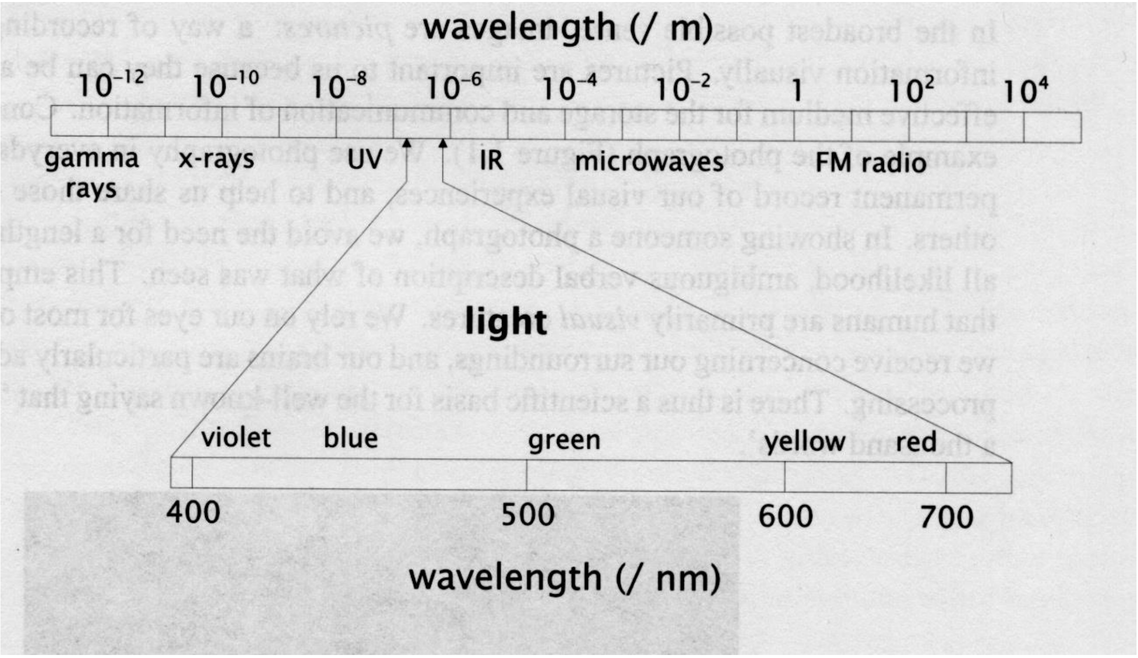
- Optical parameters of the lens
  - lens type
  - focal length
  - field of view
- Photometric parameters
  - type, intensity, and direction of illumination
  - reflectance properties of the viewed surfaces
- Geometric parameters
  - type of projections
  - position and orientation of camera in space
  - perspective distortions introduced by the imaging process

Image distortion

- Distortion (barrel, cushion)

What is light?

- The visible portion of the **electromagnetic** (EM) spectrum.
- It occurs between wavelengths of approximately 400 and 700 nanometers.



Short wavelengths

- Different wavelengths of radiation have different properties.
- The **x-ray** region of the spectrum, it carries sufficient energy to penetrate a significant volume or material.

Long wavelengths

- Copious quantities of **infrared** (IR) radiation are emitted from warm objects (e.g., locate people in total darkness).

Long wavelengths (cont'd)

- "**Synthetic aperture radar**" (SAR) imaging techniques use an artificially generated source of microwaves to probe a scene.
- SAR is unaffected by weather conditions and clouds (e.g., has provided us images of the surface of Venus).

Range images

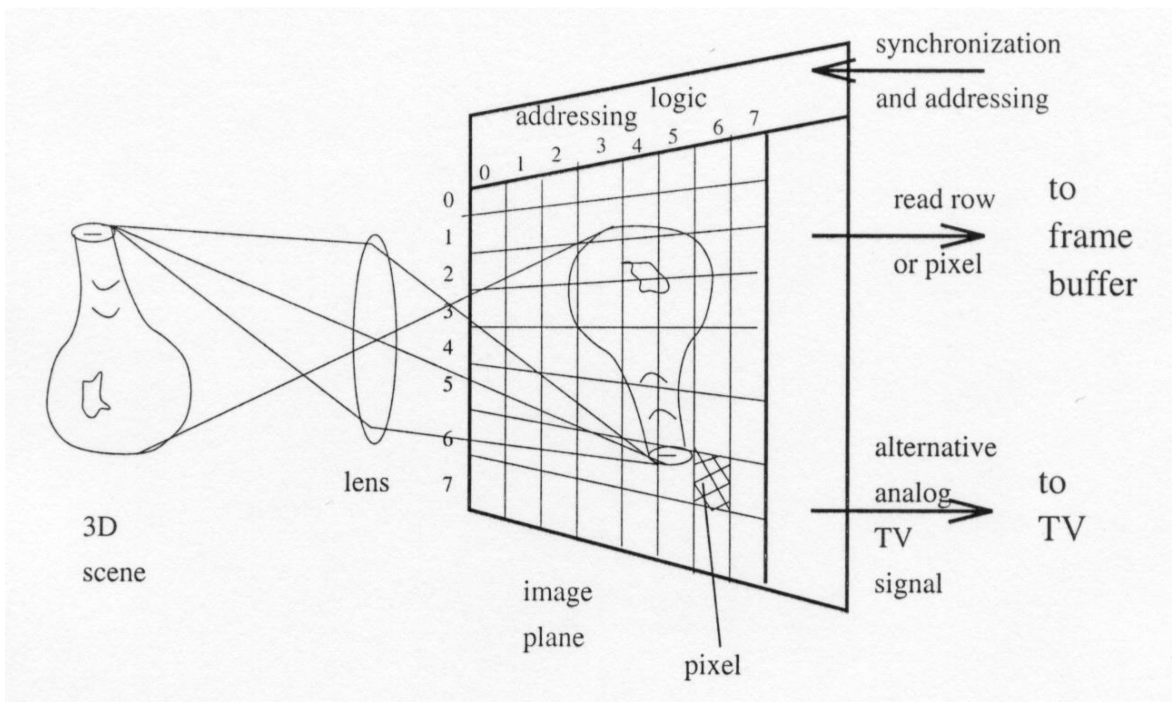
- An array of distances to the objects in the scene.
- They can be produced by sonar or by using laser rangefinders.

Sonic images

- Produced by the reflection of sound waves off an object.
- High sound frequencies are used to improve resolution.

CCD (Charged-Coupled Device) cameras

- Tiny solid state cells convert light energy into electrical charge.
- The image plane acts as a digital memory that can be read row by row by a computer.



Frame grabber

- Usually, a CCD camera plugs into a computer board (**frame grabber)**.
- The frame grabber digitizes the signal and stores it in its memory (**frame buffer**).

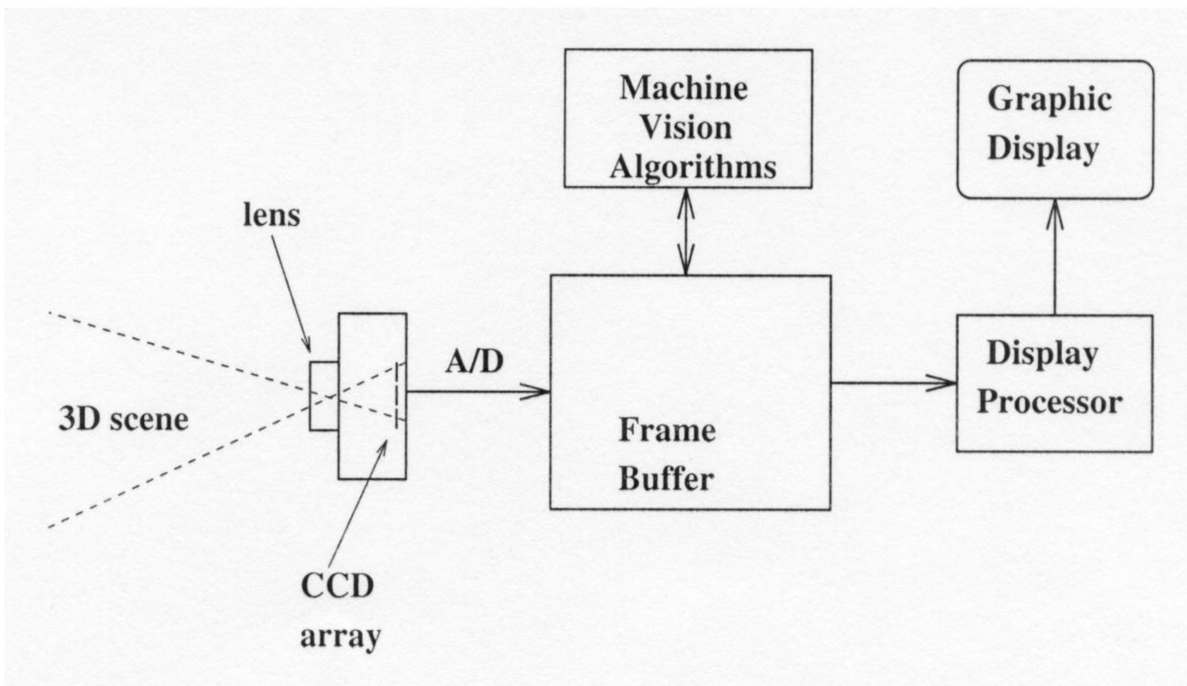
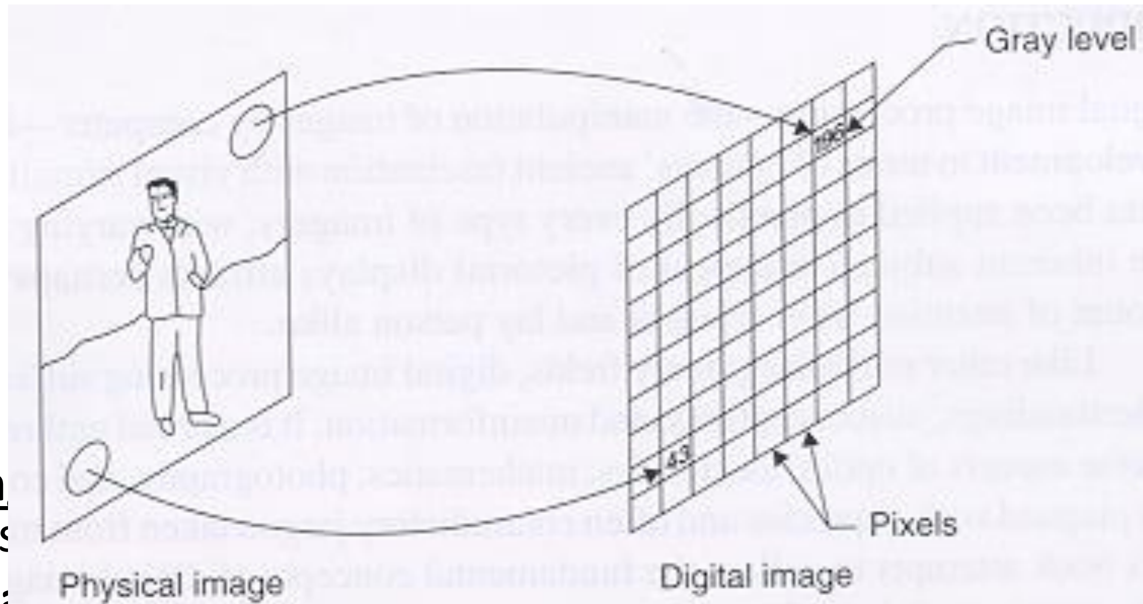


Image digitization



- **Samplin**... number of points...
- **Quantiza**... is the representation of the measured value at the sampled point by an integer.

Image digitization (cont'd)

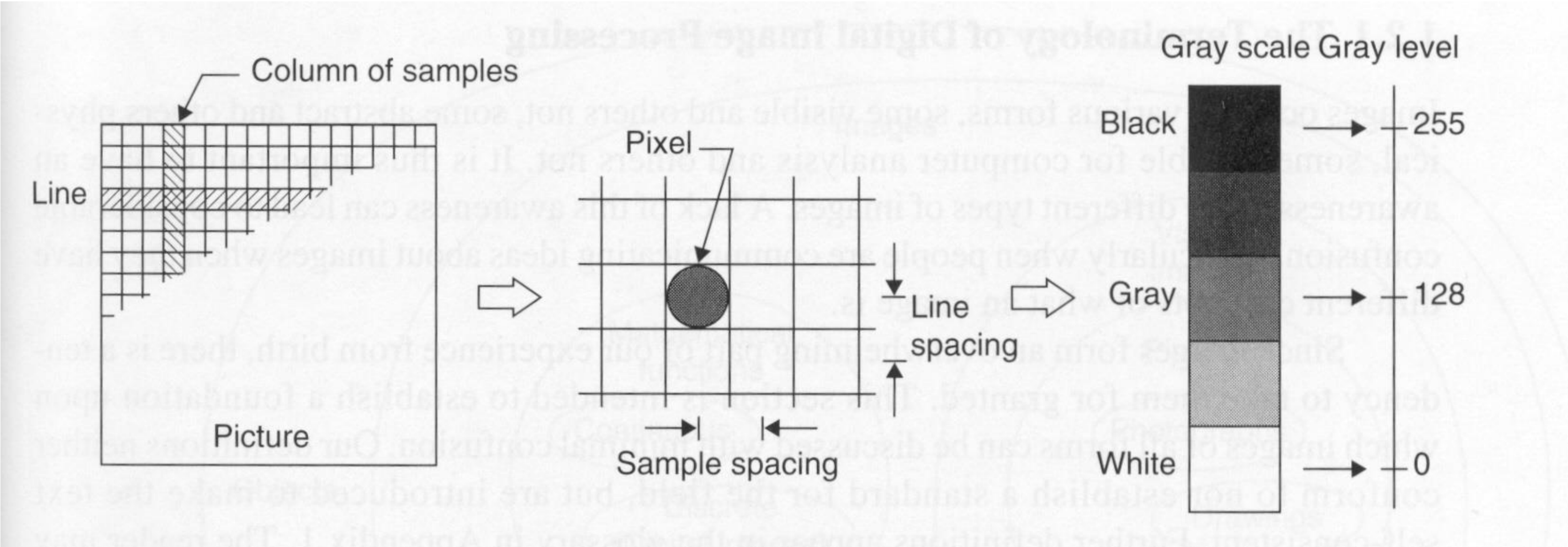


Image quantization(example)

- 256 gray levels (8bits/pixel)   32 gray levels (5 bits/pixel)   16 gray levels (4 bits/pixel)

- 8 gray levels (3 bits/pixel)   4 gray levels (2 bits/pixel)   2 gray levels (1 bit/pixel)

Electromagnetic spectrum

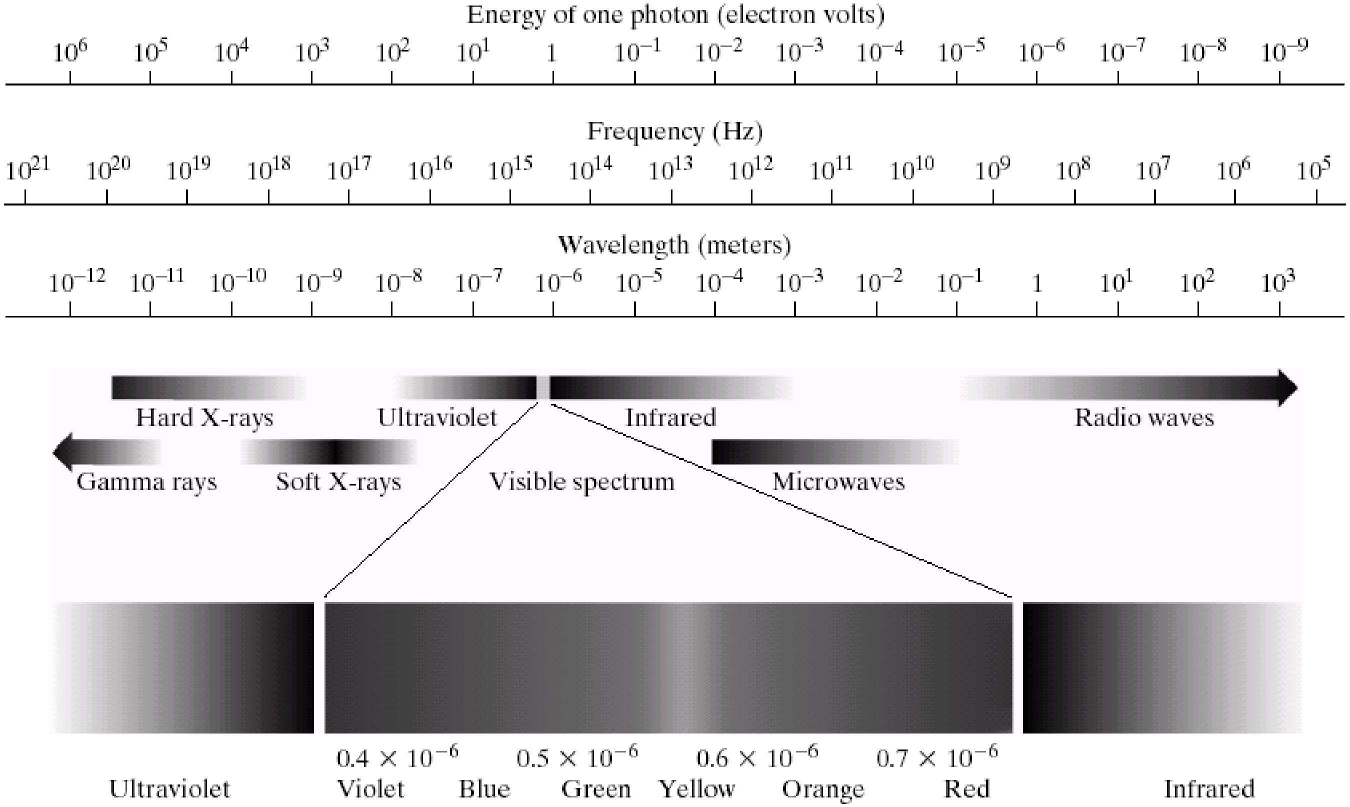


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

Light: the Visible Spectrum

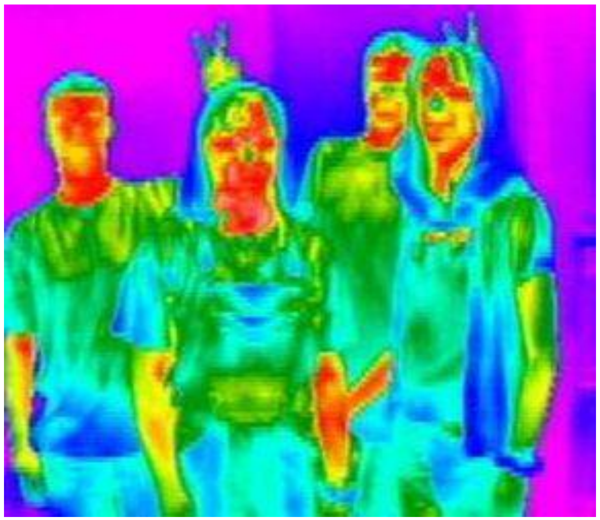
- Visible range: 0.43μm(violet)-0.78μm(red)
- Six bands: violet, blue, green, yellow, orange, red
- The color of an object is determined by the nature of the light *reflected* by the object
- Monochromatic light (gray level)
- Three elements measuring chromatic light
  - Radiance, luminance and brightness

Beyond Visible

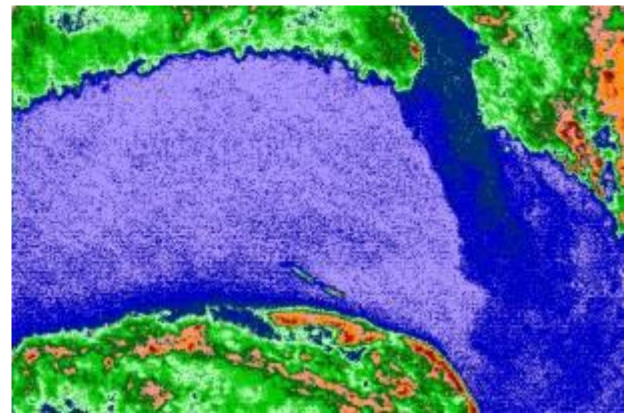
- Gamma-ray and X-ray: medical and astronomical applications
- Infrared (thermal imaging): near-infrared and far-infrared
- Microwave imaging:
- Radio-frequency: MRI and astronomic applications

Thermal Imaging

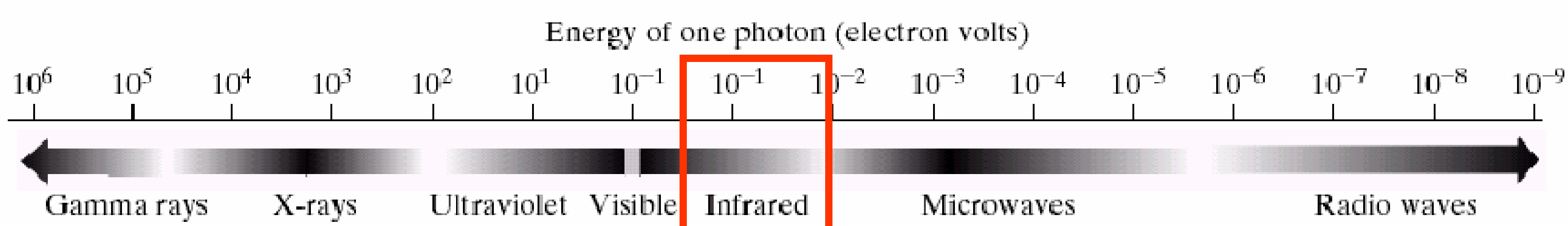
Operate in infrared frequency



Human body disperses heat (red pixels)

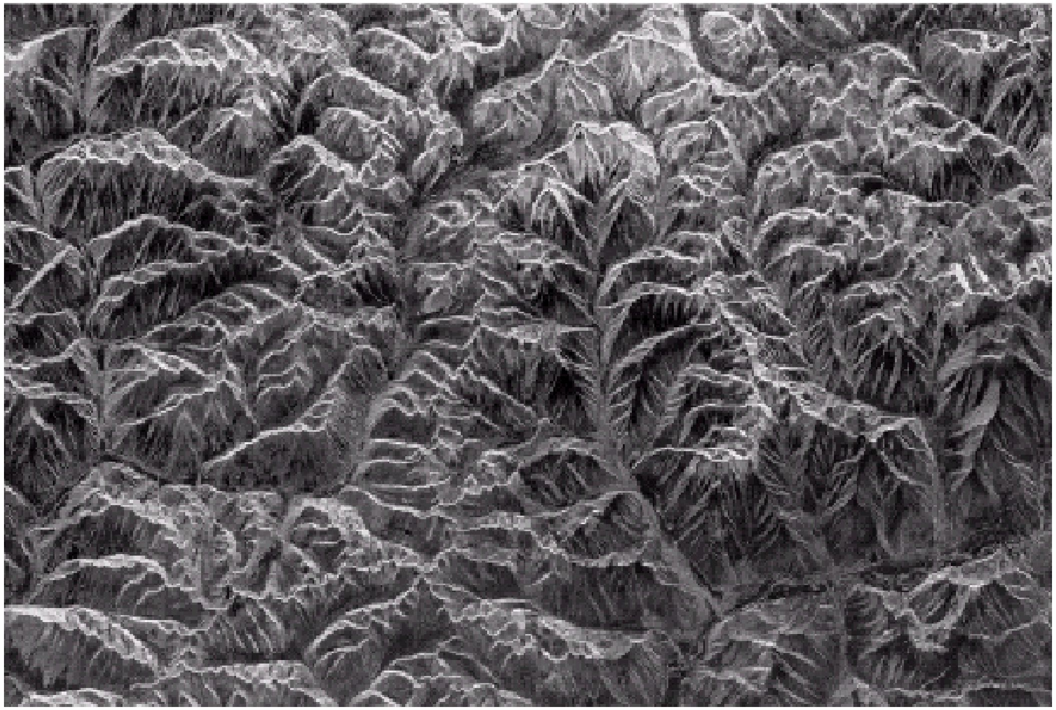


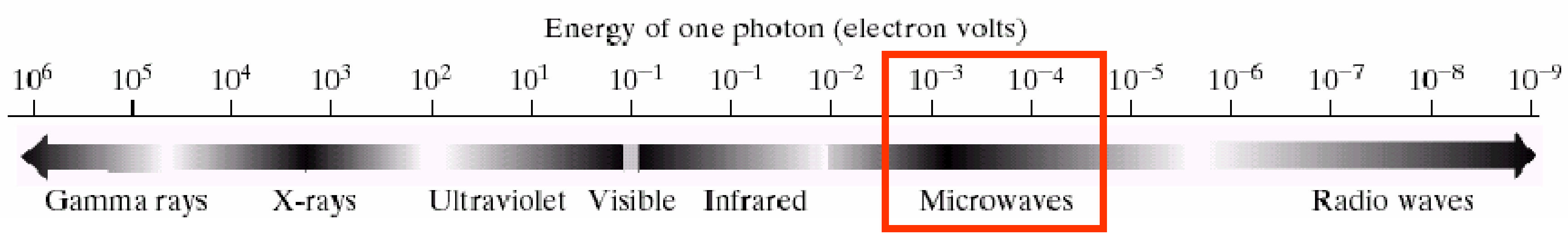
Different colors indicate varying temperatures



Radar Imaging

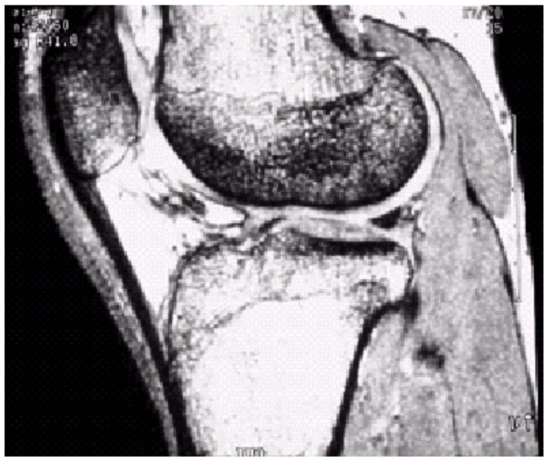
Operate in microwave frequency





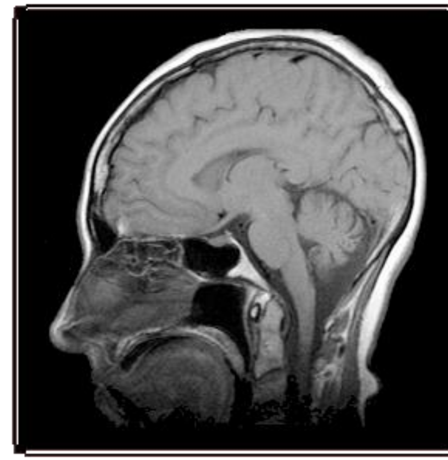
Magnetic Resonance Imaging (MRI)

Operate in radio frequency

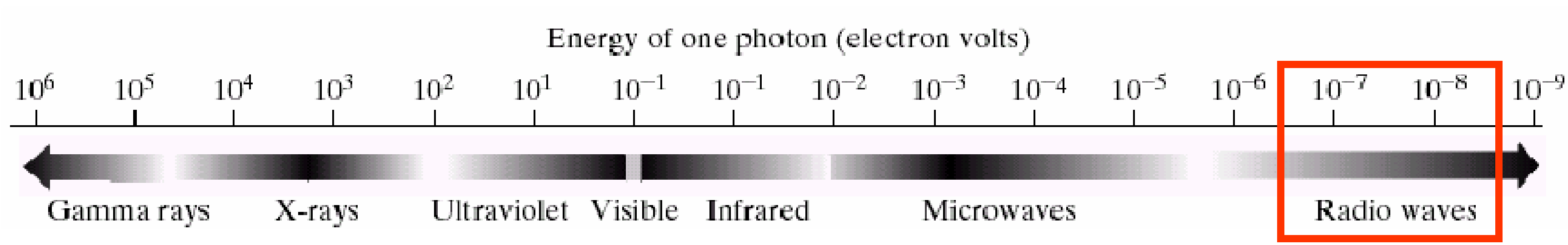


knee

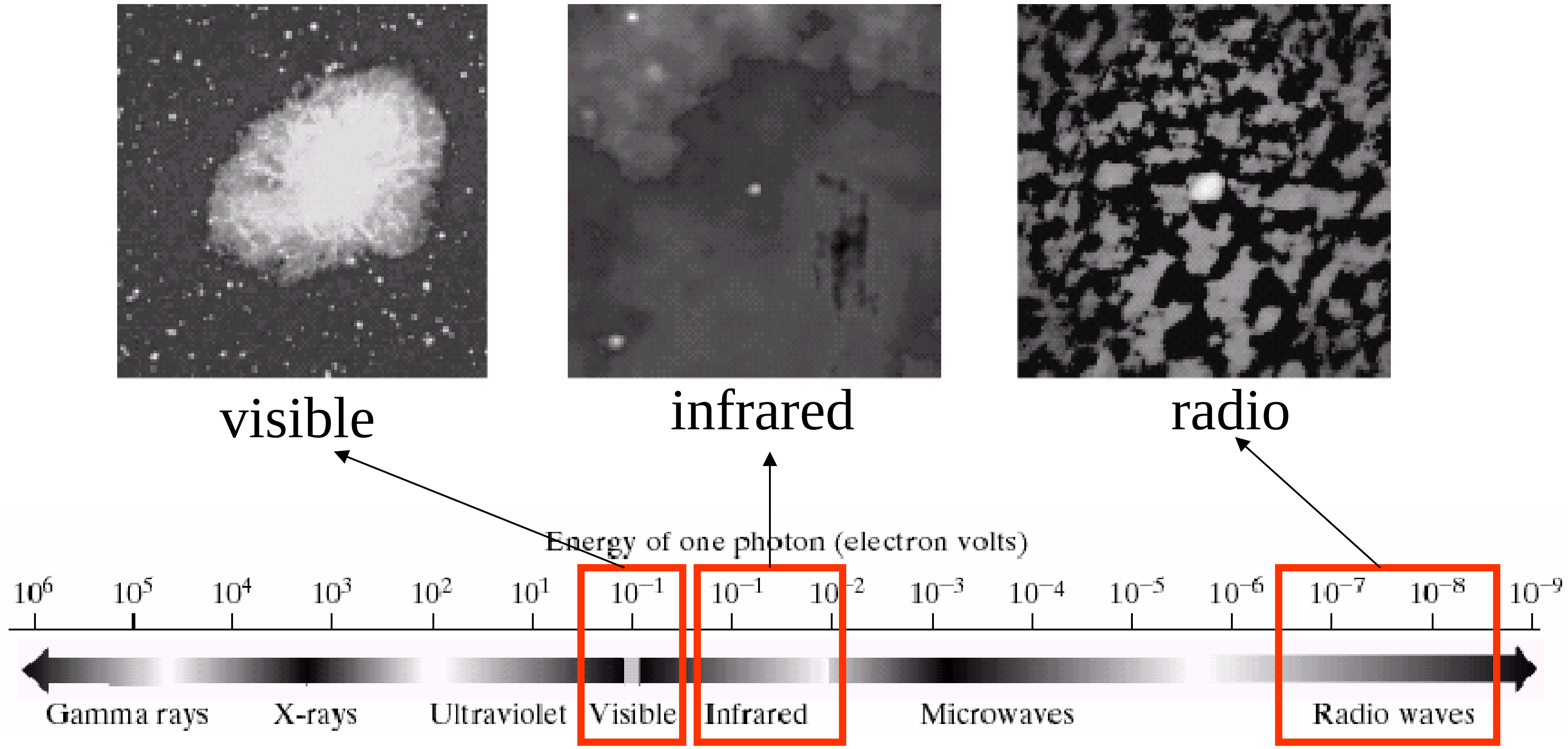
spine



head

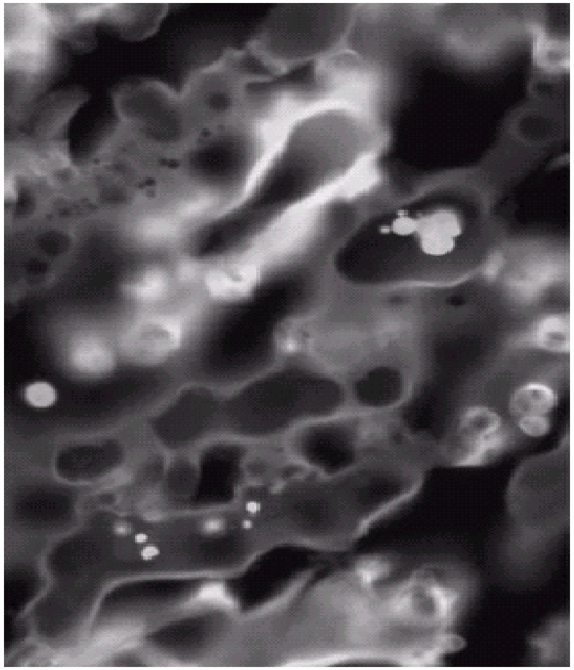


Comparison of Different Imaging Modalities

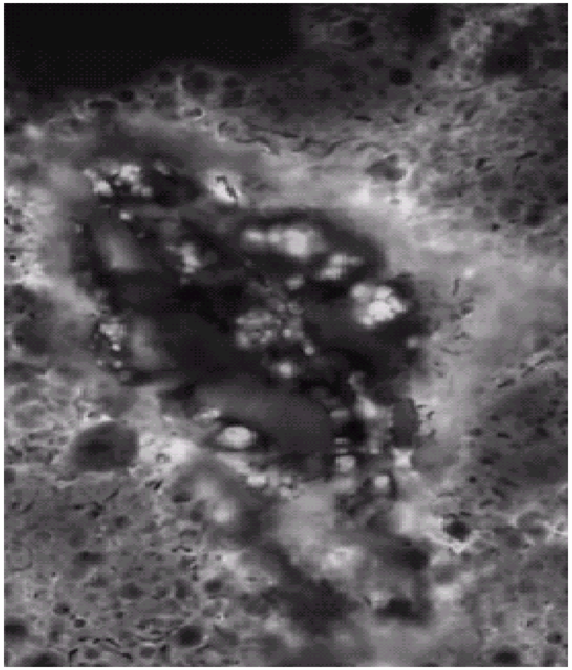


Fluorescence Microscopy Imaging

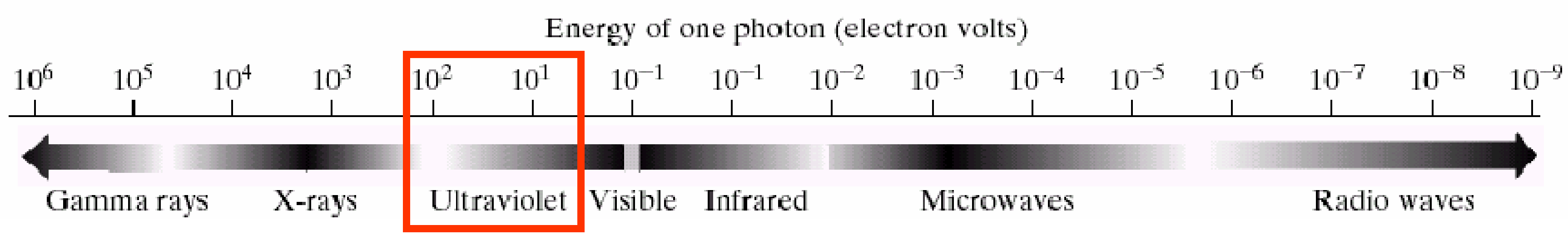
Operate in ultraviolet frequency



normal corn

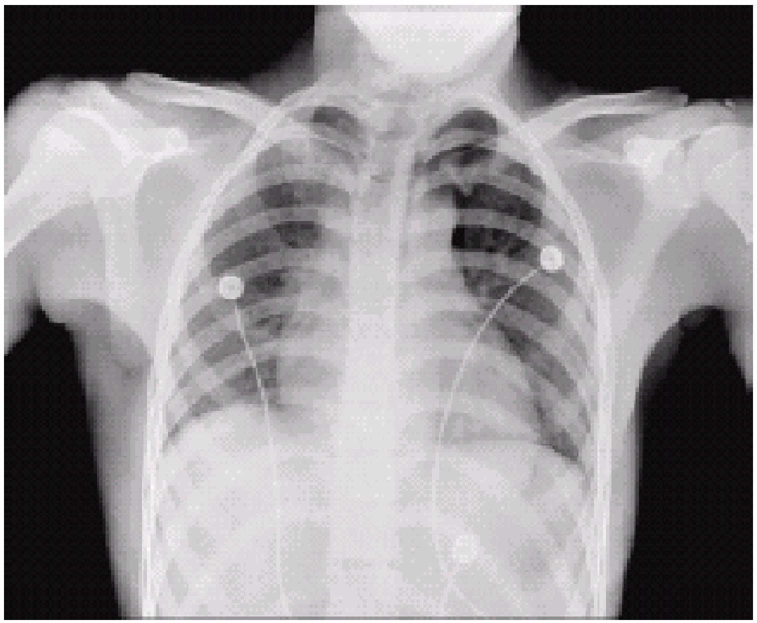


smut corn

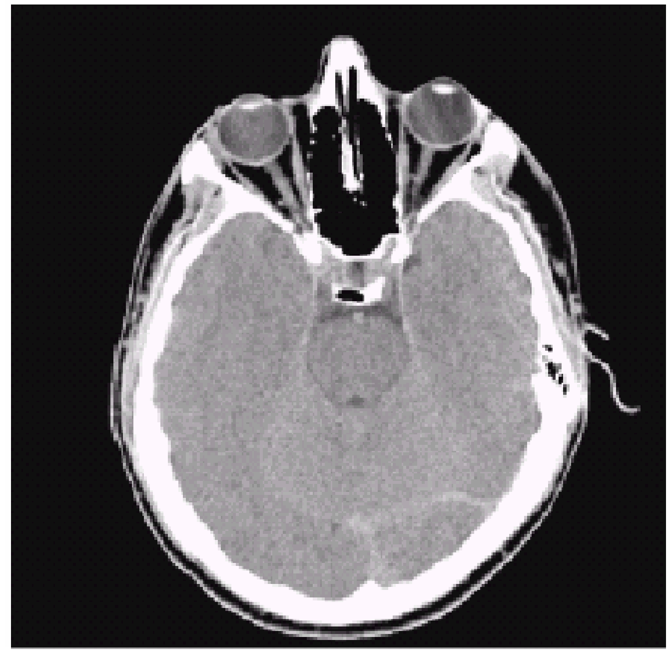


X-ray Imaging

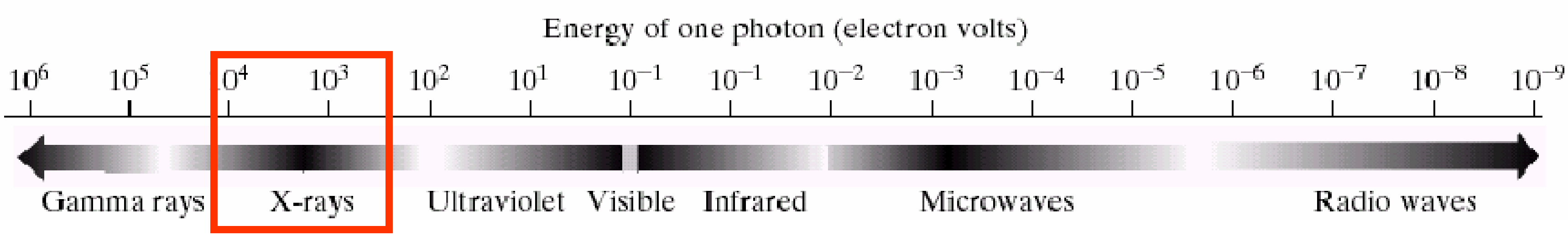
Operate in X-ray frequency



chest

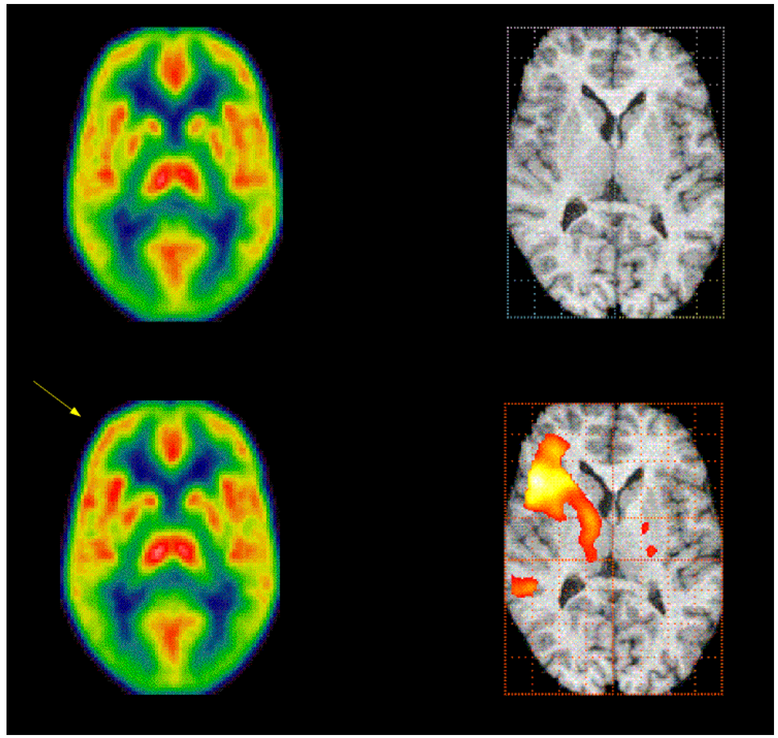


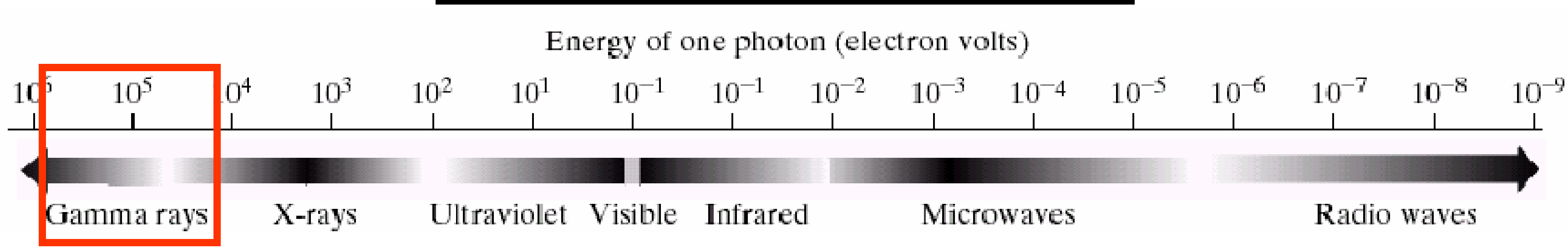
head



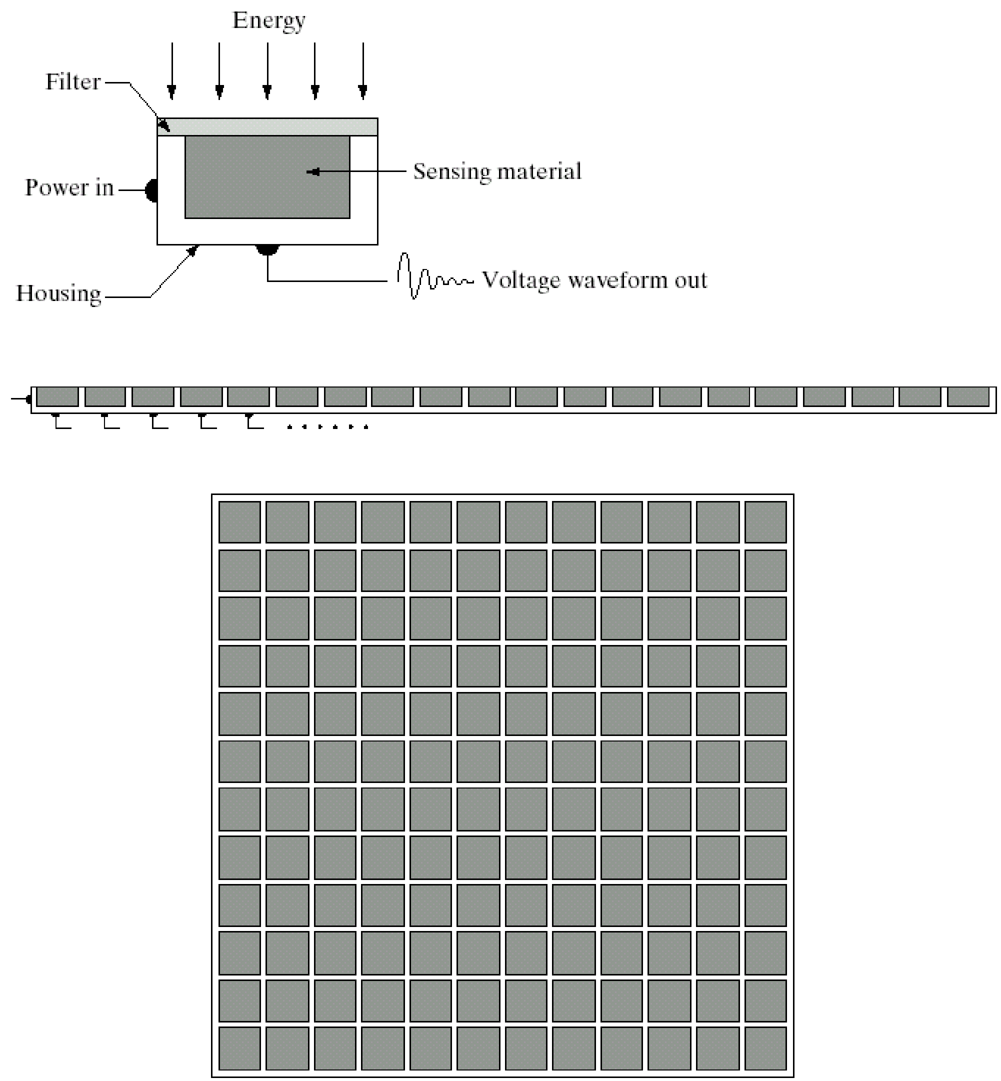
Positron Emission Tomography

Operate in gamma-ray frequency





Single-sensor Imaging



a
b
c

FIGURE 2.12
(a) Single imaging sensor.
(b) Line sensor.
(c) Array sensor.

“Motion” Aids Imaging

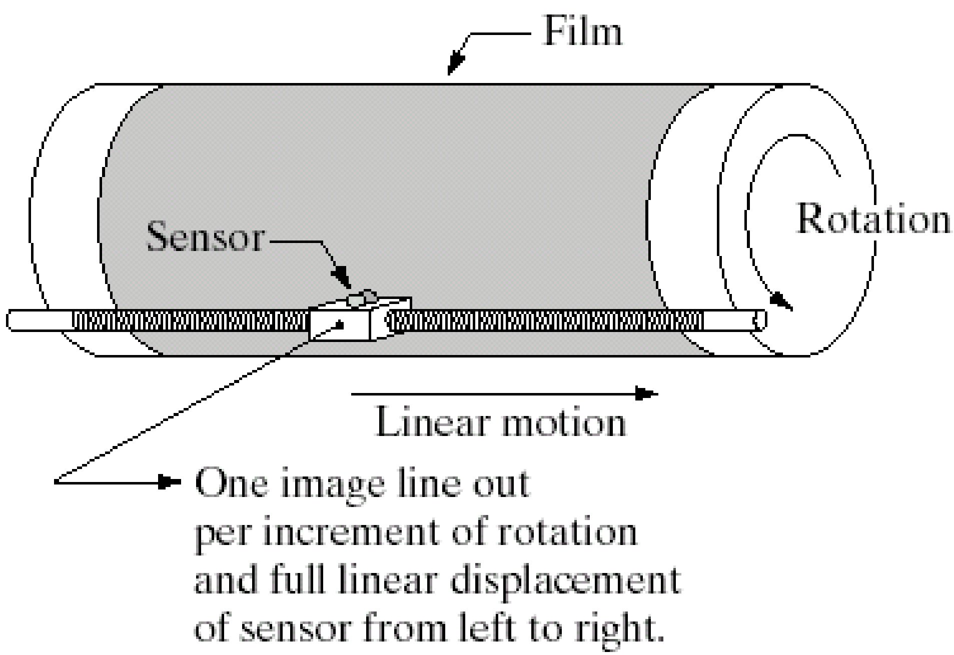


FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

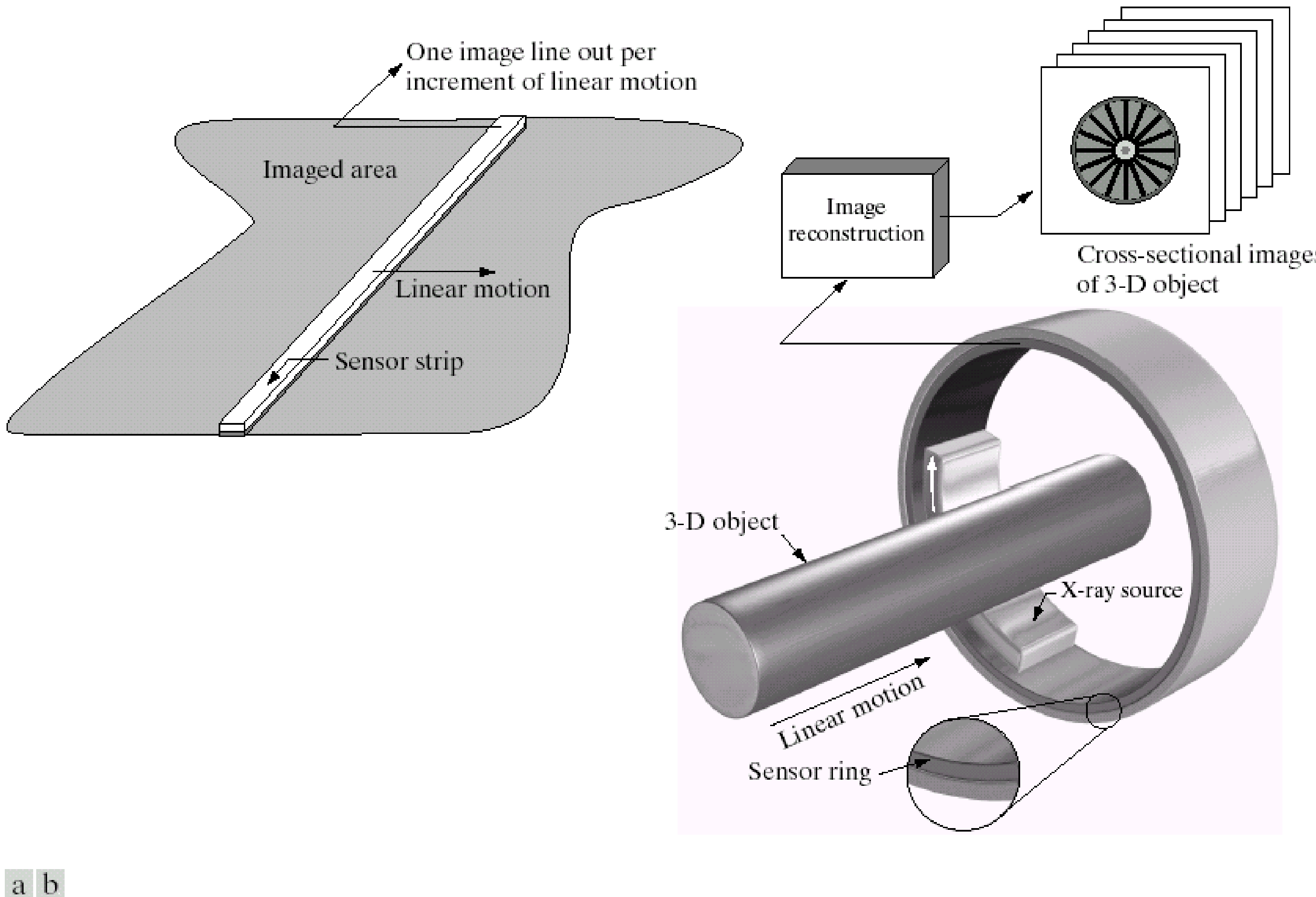


FIGURE 2.14 (a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.

Sensor Array: CCD Imaging

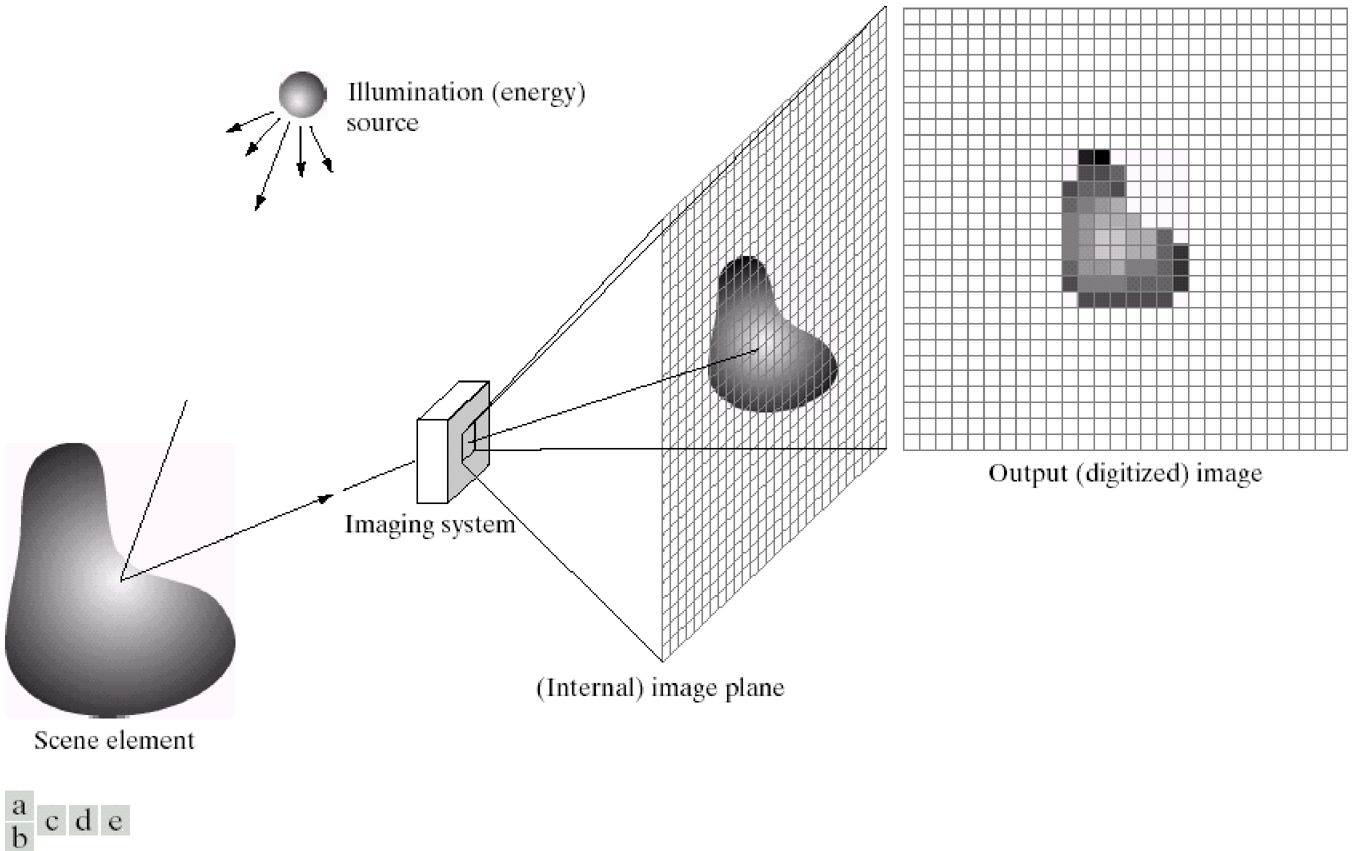


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

Image Formation Model

$$f(x,y)=i(x,y)r(x,y)$$

$$0<f(x,y)<\infty$$

Intensity – proportional to energy radiated by a physical source

$$0<i(x,y)<\infty$$

illumination

$$0<r(x,y)<1$$

reflectance

Sampling and Quantization: 1D Case

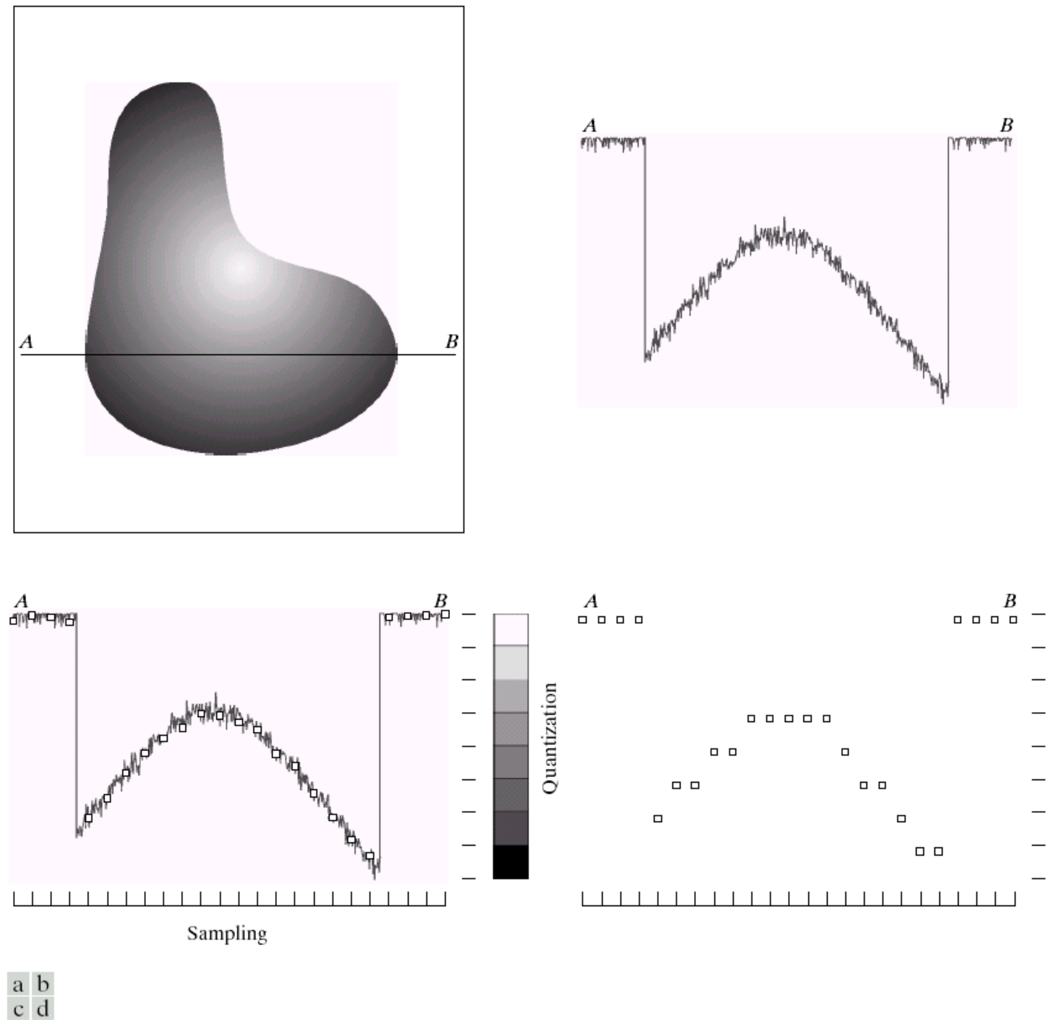


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

2D Sampling and Quantization

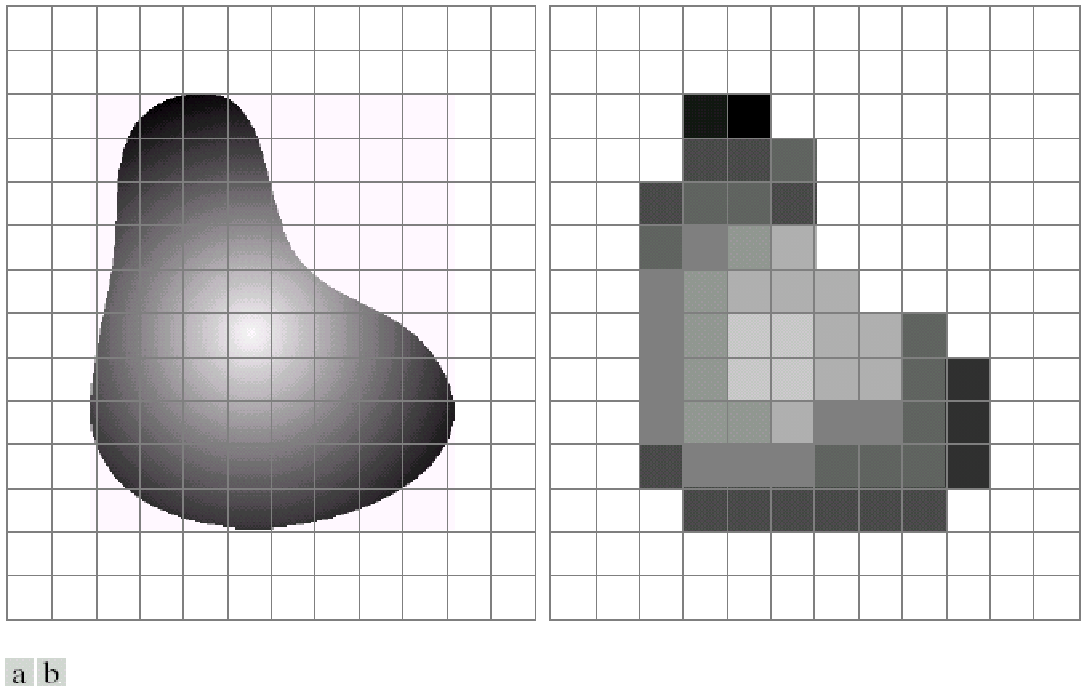


FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

Introduction to Grayscale Images

- Image acquisition
  - Light and Electromagnetic spectrum
  - Sampling and Quantization
- Image perception
  - Structure of human eyes
  - Image formation in human eyes
  - Human vision system
- Image representation
  - Spatial and bit-depth resolution
  - Local neighborhood

Human Eye Structure

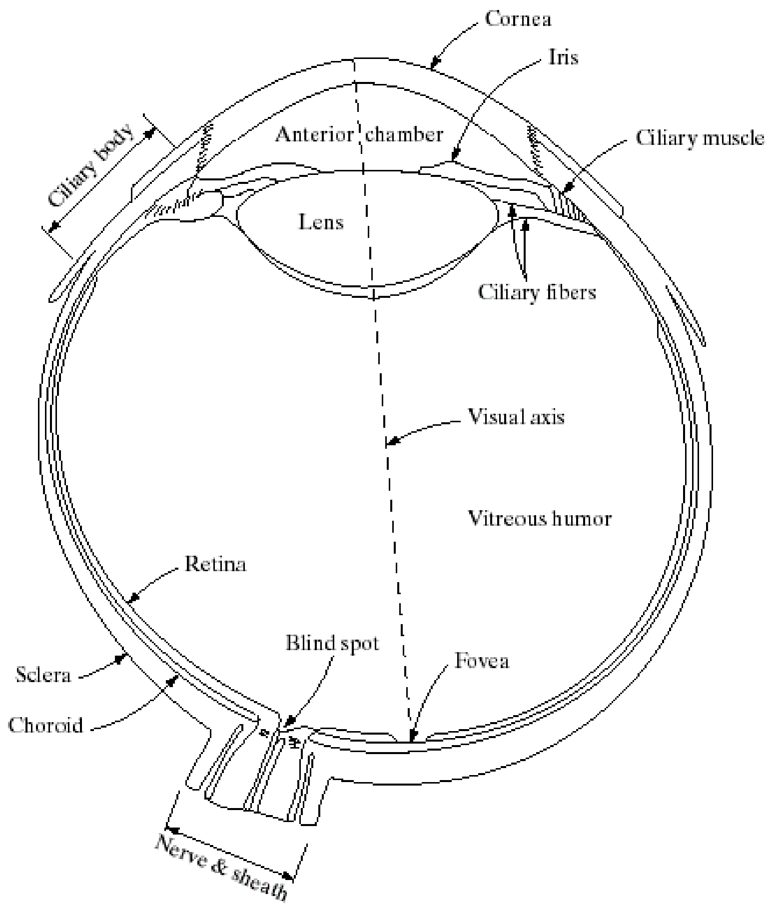


FIGURE 2.1 Simplified diagram of a cross section of the human eye.

Three membranes enclose the eye:
Cornea and sclera, Choroid, Retina

ciliary body

iris diaphragm

Pupil size: 2-8mm

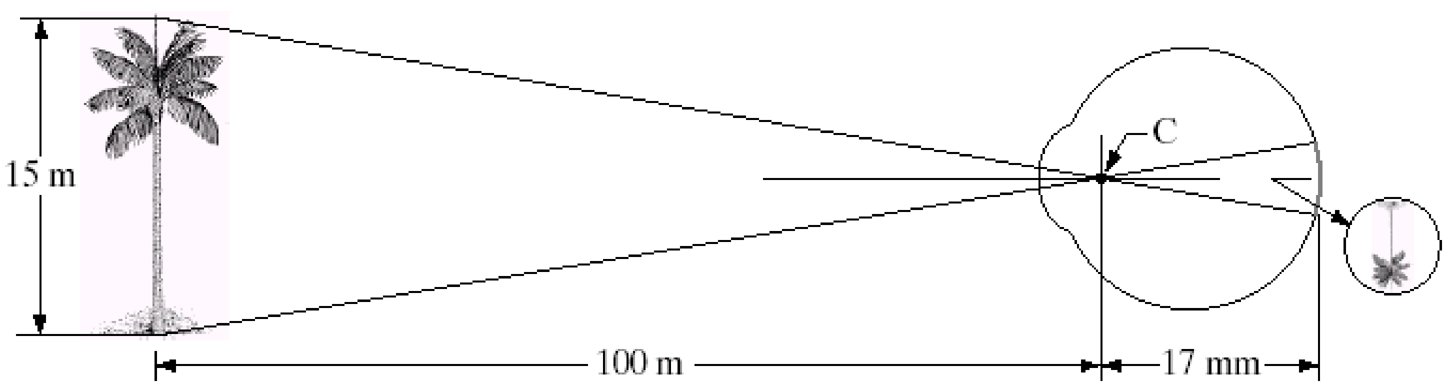
Eye color: melanin (pigment) in iris

Retina

- When the eye is properly focused, light from an outside object is imaged on the retina
- Two classes of receptors are located over the surface of retina: cones and rods
  - Cone: 6-7 million in each eye, central part of retina (fovea) and highly sensitive to color
  - Rod: 75-150 million, all over the retina surface and sensitive to low levels of illumination

Image Formation in the Eye

FIGURE 2.3 Graphical representation of the eye looking at a palm tree. Point *C* is the optical center of the lens.



Focal length: 14-17mm

Length of tree image≅2.55mm

For distant objects (>3m), lens exhibits the least refractive power (flattened)

For nearby objects (<1m), lens is most strongly refractive (curved)

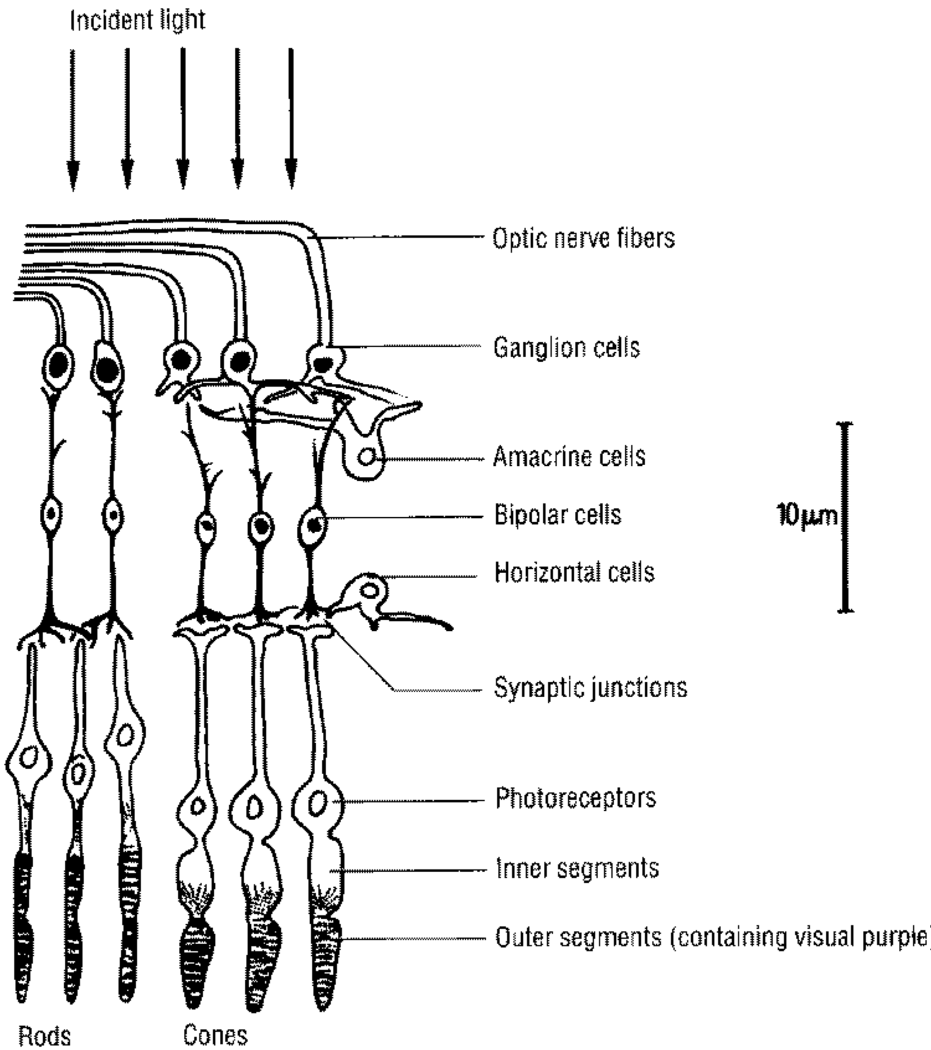


Fig. 1.9
Diagrammatic representation of the structure and relative position of the active elements of the retina

Eye Physiology

- Rods are more sensitive to light than the cones.

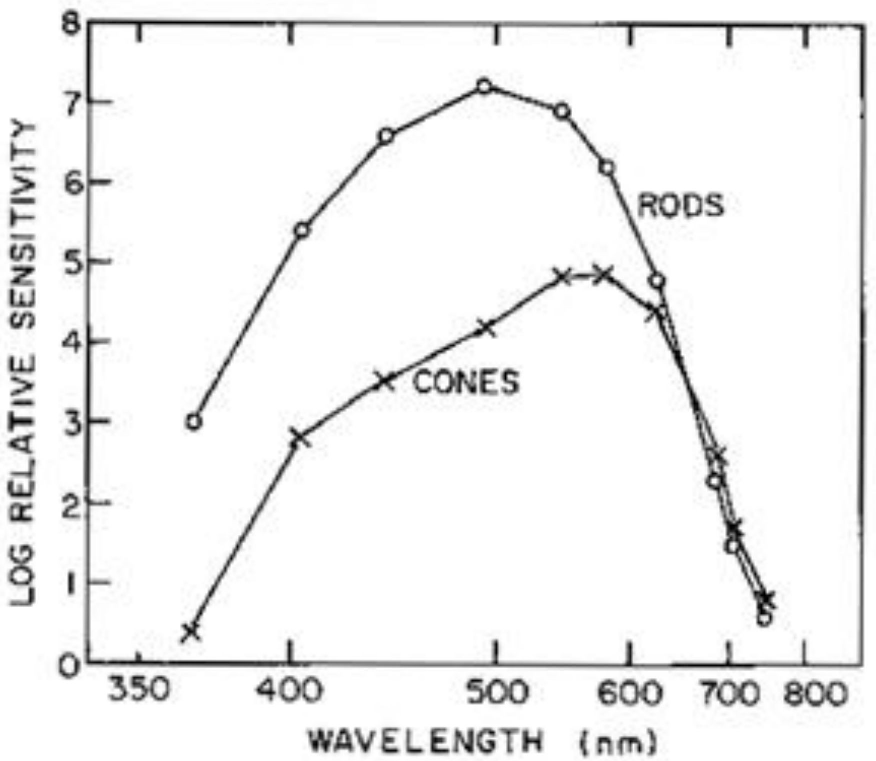


FIGURE 2.2-2. Sensitivity of rods and cones (7) [based upon measurements by Wald (8)].

Eye Physiology

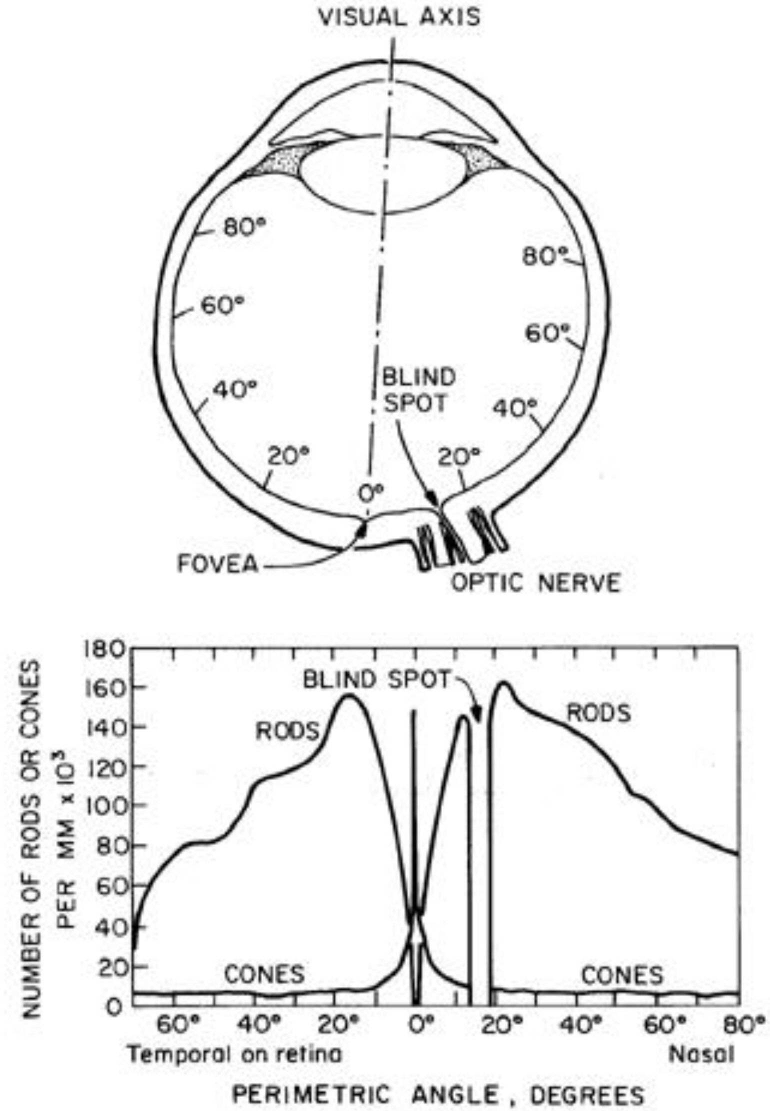


FIGURE 2.2-3. Distribution of rods and cones on the retina (4).

- The eye contains about 6.5 million cones and 100 million rods distributed over the retina.
- The density of the cones is greatest at the fovea, this is the region of sharpest photopic vision.

Rods and Cones in Retina

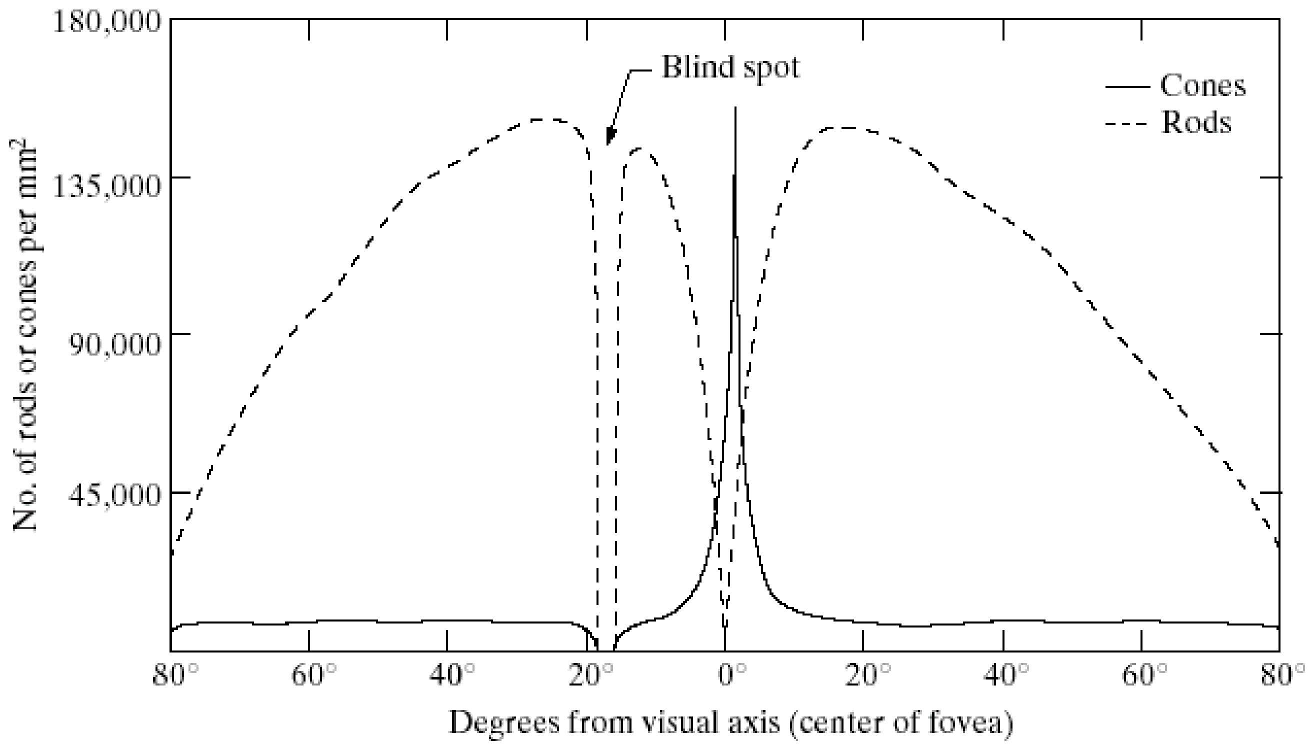


FIGURE 2.2 Distribution of rods and cones in the retina.

Eye Physiology

- There are three basic types of cones in the retina
- These cones have different absorption characteristics as a function of wavelength with peak absorptions in the red, green, and blue regions of the optical spectrum.

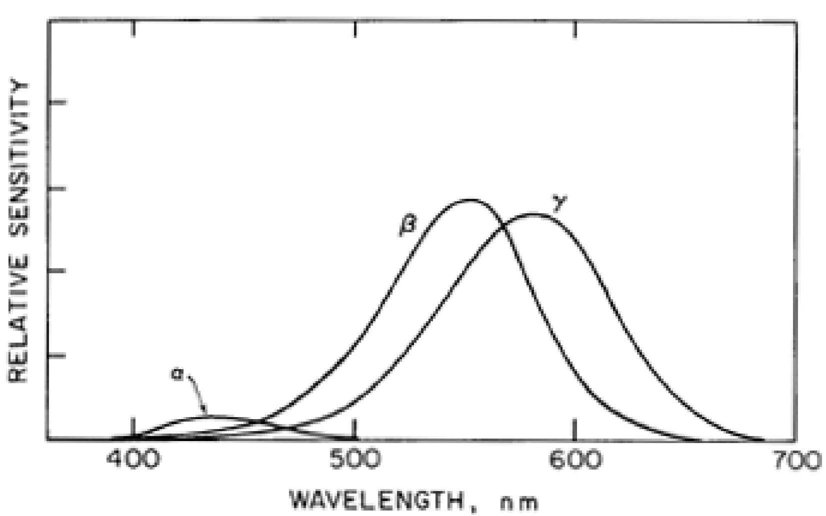


FIGURE 2.2-4. Typical spectral absorption curves of pigments of the retina (10).

α is blue, β is green, and γ is red

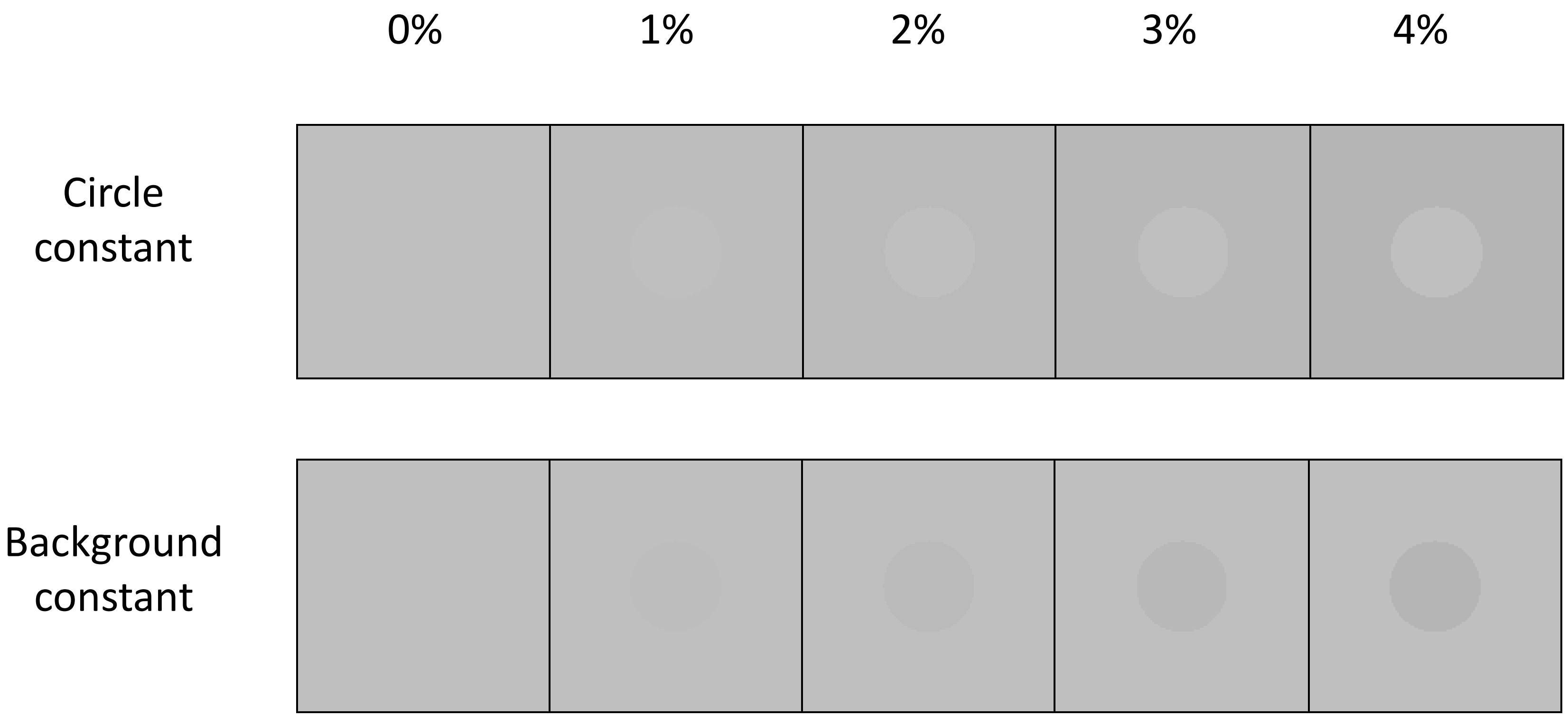
There is a relatively low sensitivity to blue light

There is a lot of overlap

Eye Physiology

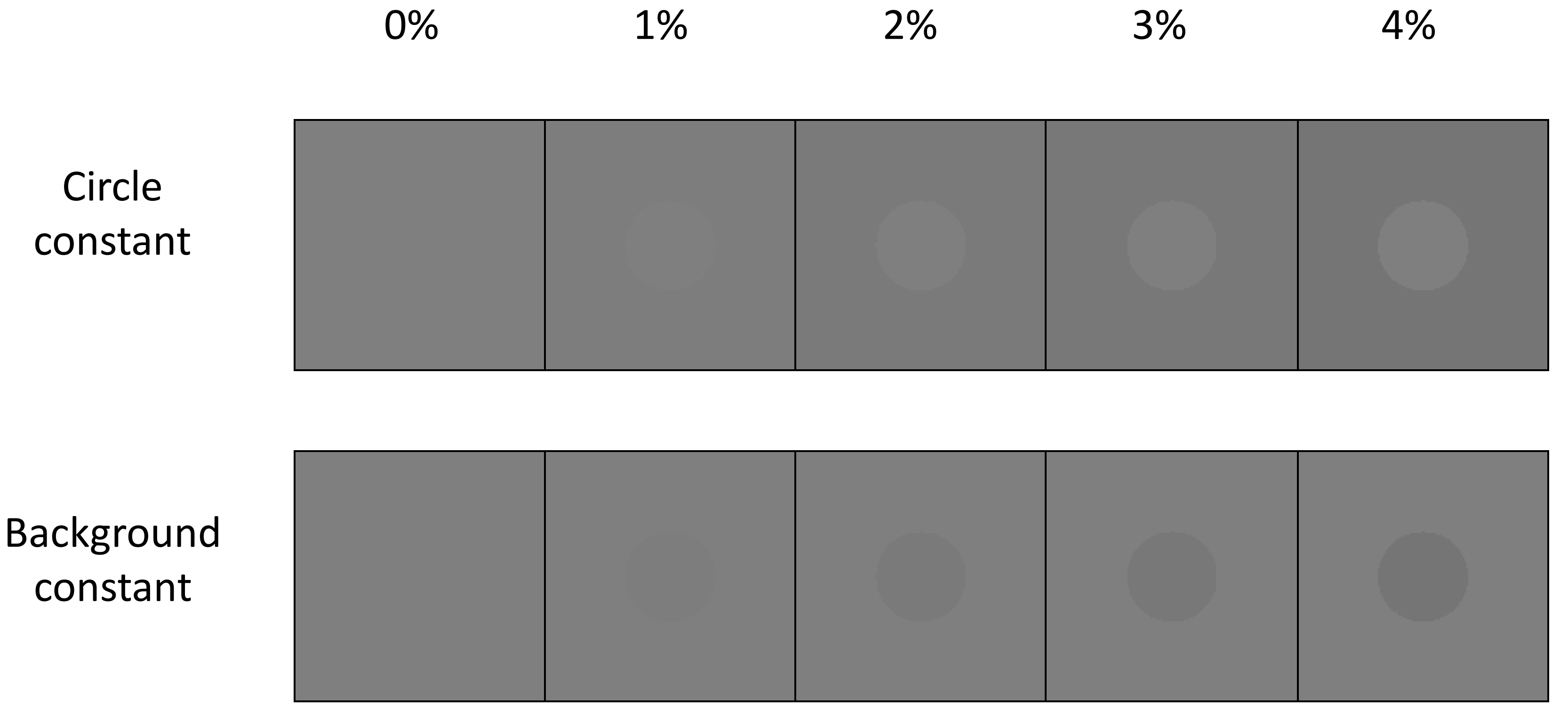
- The optic nerve bundle contains on the order of 800,000 nerve fibers.
- There are over 100,000,000 receptors in the retina.
- Therefore, the rods and cones must be interconnected to nerve fibers on a many-to-one basis.

Contrast Sensitivity



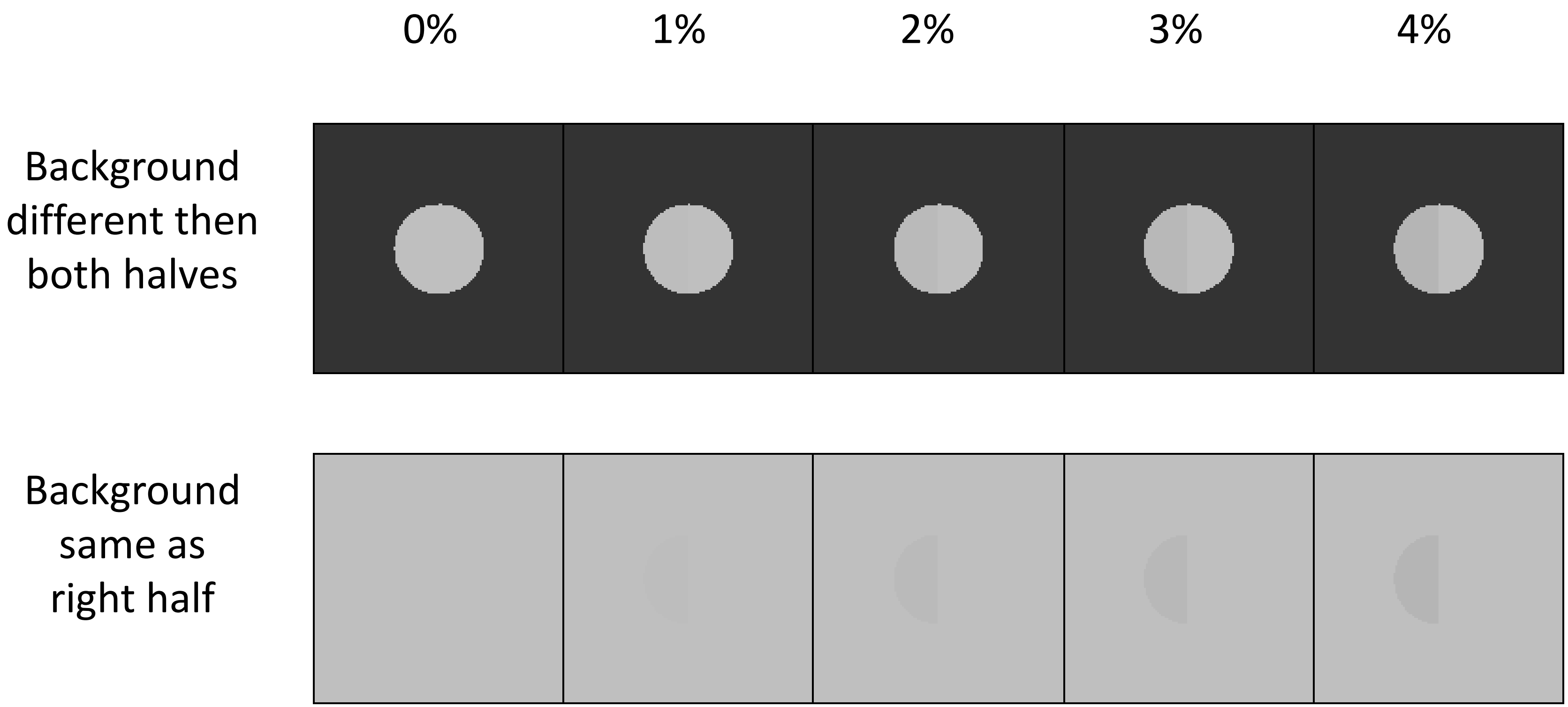
Just noticeable difference (JND) at 2%

Contrast Sensitivity



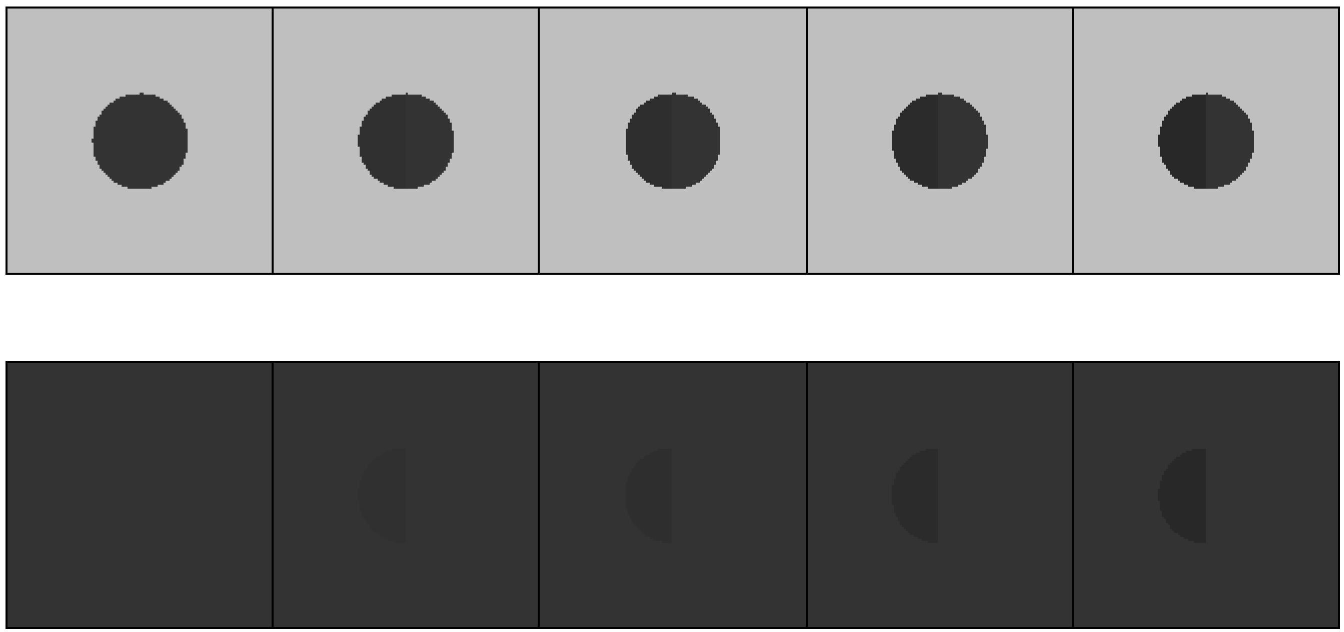
Just noticeable difference (JND) at 2%

Contrast Sensitivity



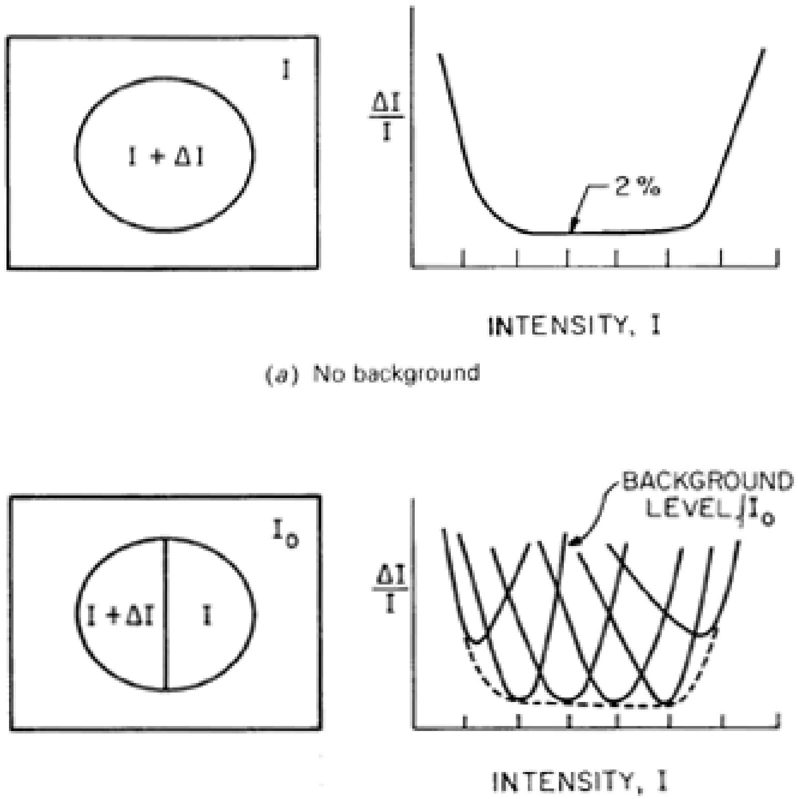
Just noticeable difference (JND): 4% (top) and 2% (bottom)

Contrast Sensitivity



Just noticeable difference (JND): 4% (top) and 2% (bottom)

Contrast Sensitivity



(a) No background

(b) With background

FIGURE 2.3-1. Contrast sensitivity measurements.

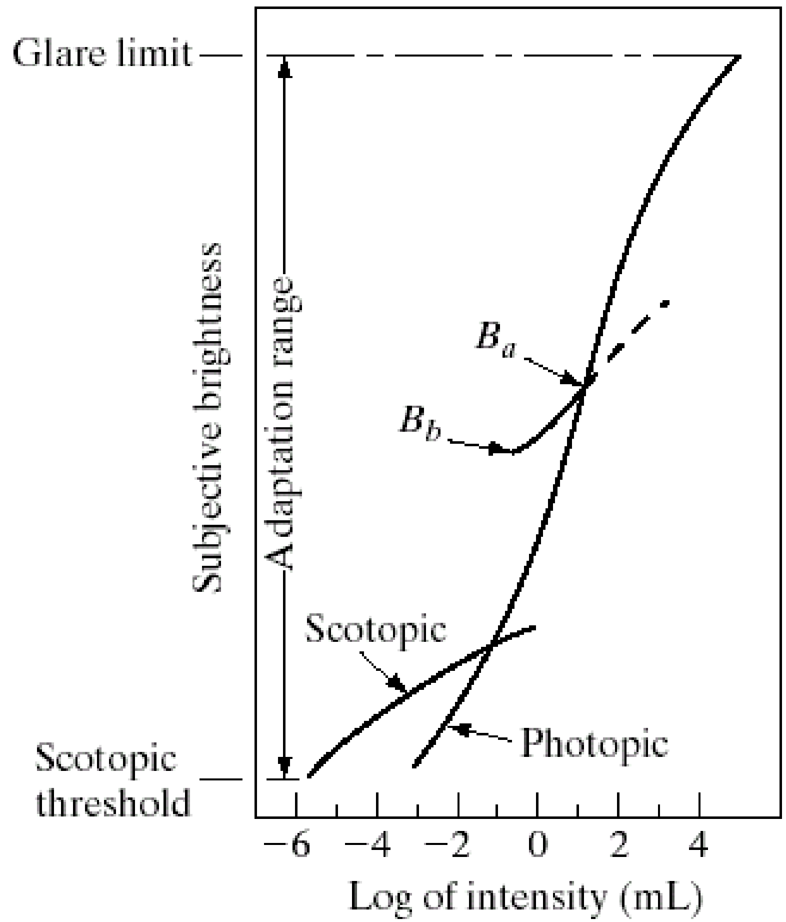
- The response of the eye to changes in the intensity of illumination is nonlinear
- Consider a patch of light of intensity i+dI surrounded by a background intensity I as shown in the previous figure

Contrast Sensitivity

- Over a wide range of intensities, it is found that the ratio dI/I, called the Weber fraction, is nearly constant at a value of about 0.02.
- This does not hold at very low or very high intensities
- Furthermore, contrast sensitivity is dependent on the intensity of the surround. Consider the second panel of the previous figure.

Brightness Adaptation

FIGURE 2.4
Range of subjective brightness sensations showing a particular adaptation level.



Human visual system cannot operate over such a high dynamic range simultaneously, But accomplish such large variation by changes in its overall sensitivity, a phenomenon called “brightness adaptation”

Brightness Discrimination

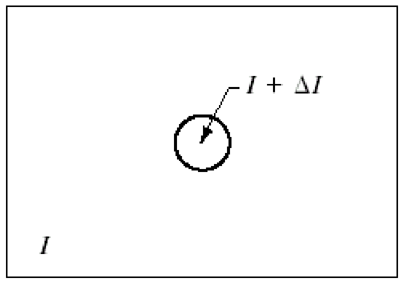
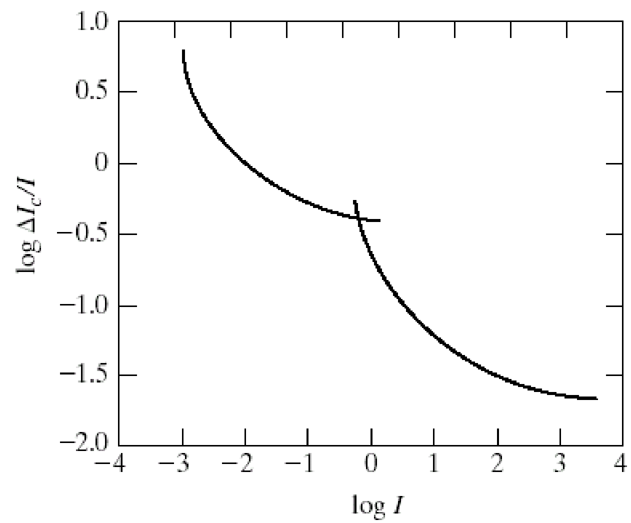


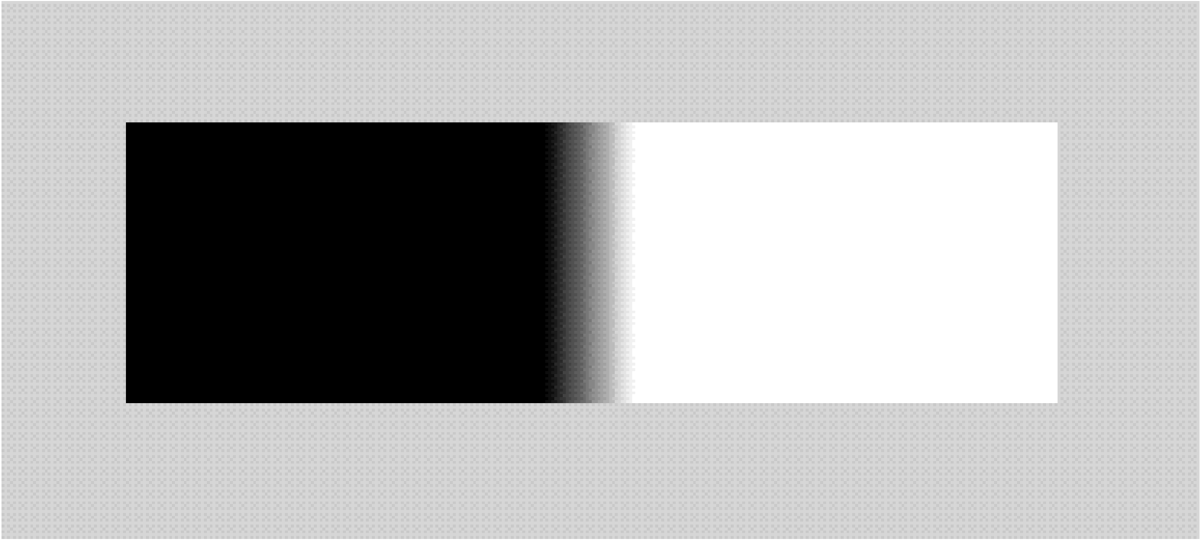
FIGURE 2.5 Basic experimental setup used to characterize brightness discrimination.

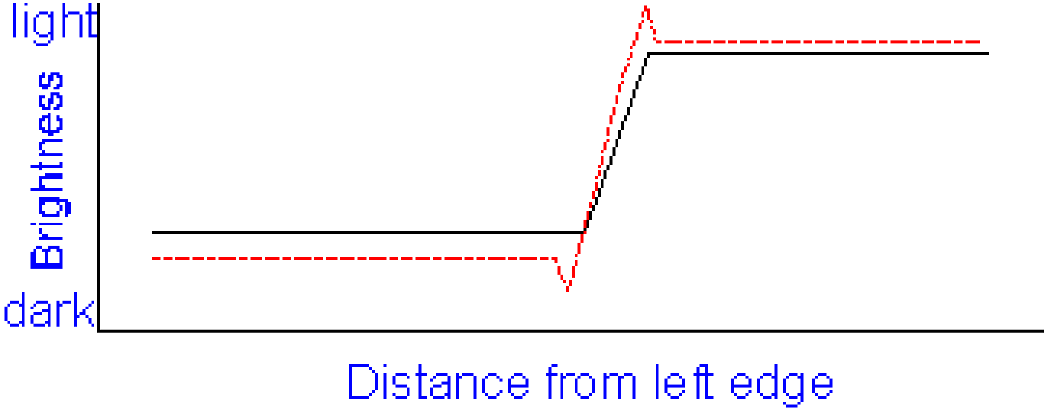
Weber ratio=$\Delta I/I$

FIGURE 2.6 Typical Weber ratio as a function of intensity.

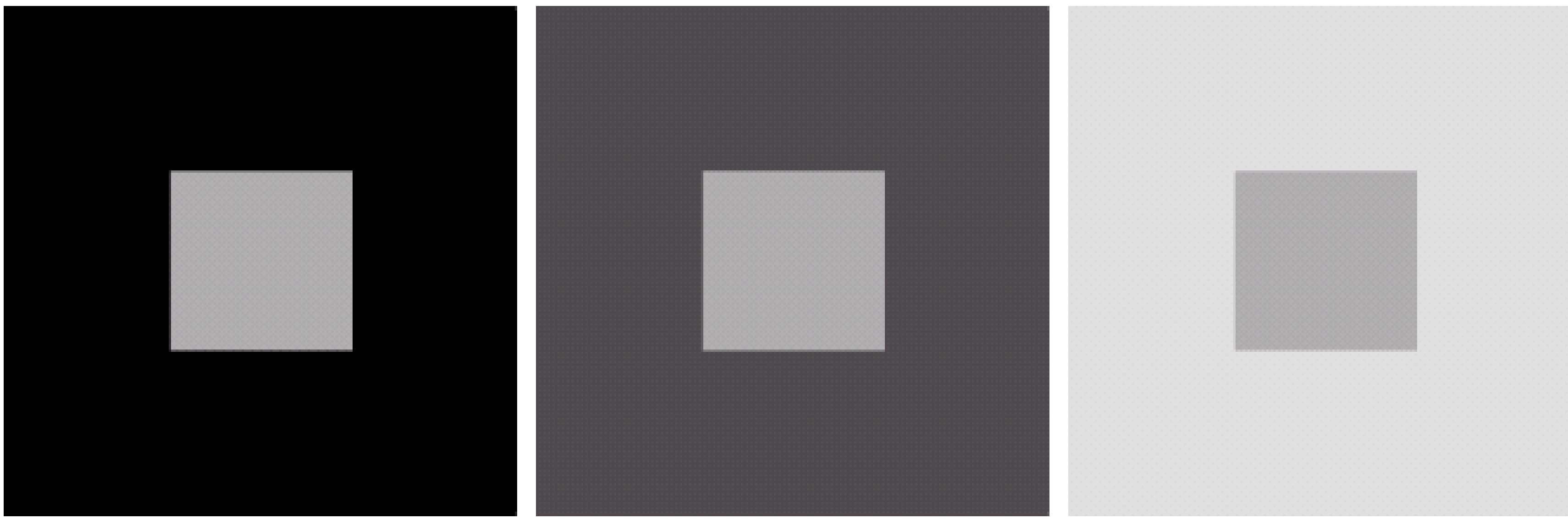


Mach Bands





Simultaneous Contrast



a b c

FIGURE 2.8 Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

Optical Illusions

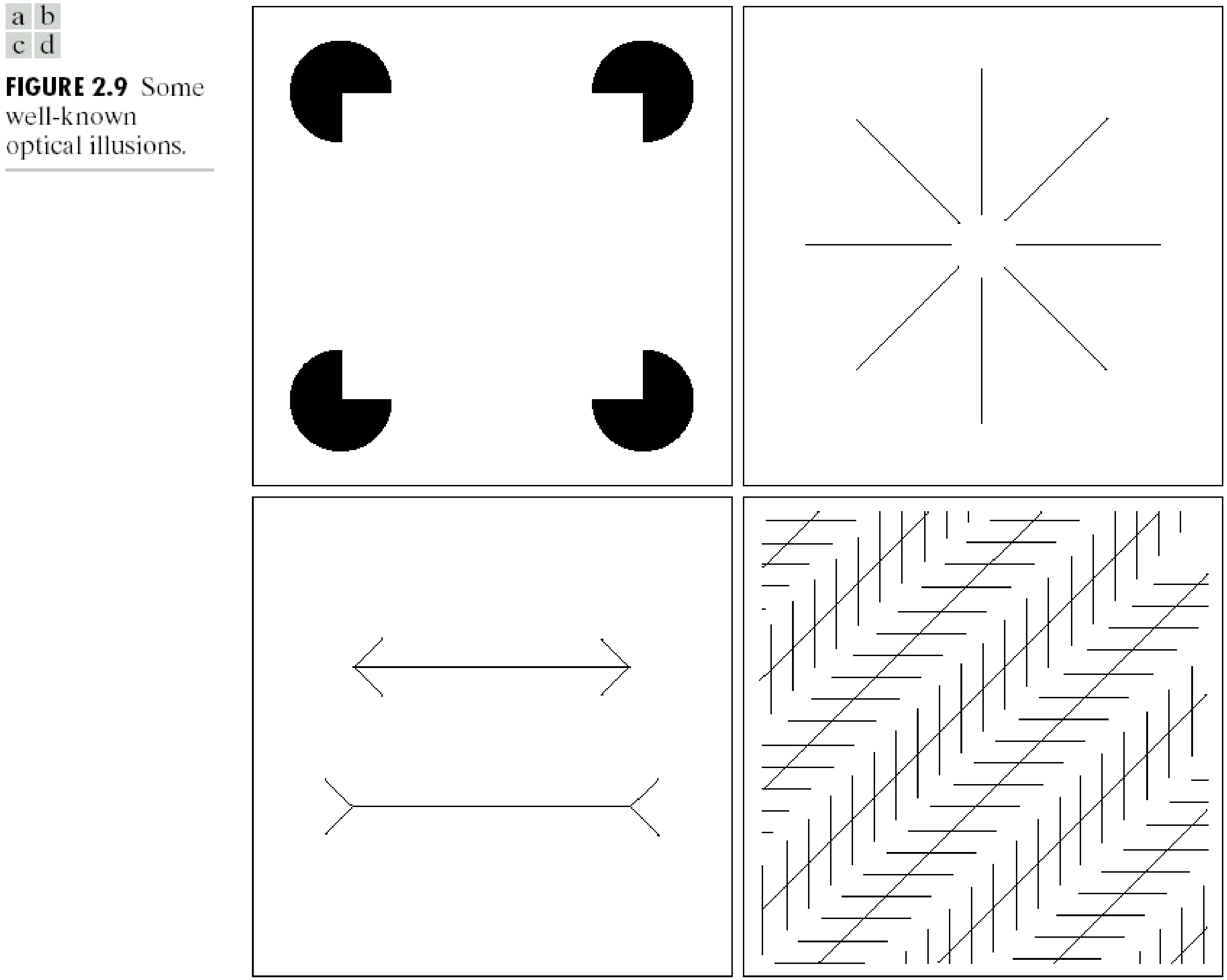


FIGURE 2.9 Some well-known optical illusions.

Introduction to Grayscale Images

- Image acquisition
  - Light and Electromagnetic spectrum
  - Sampling and Quantization
- Image perception
  - Structure of human eyes
  - Image formation in human eyes
  - Human vision system
- Image representation
  - Spatial and bit-depth resolution
  - Local neighborhood

Image Represented by a Matrix

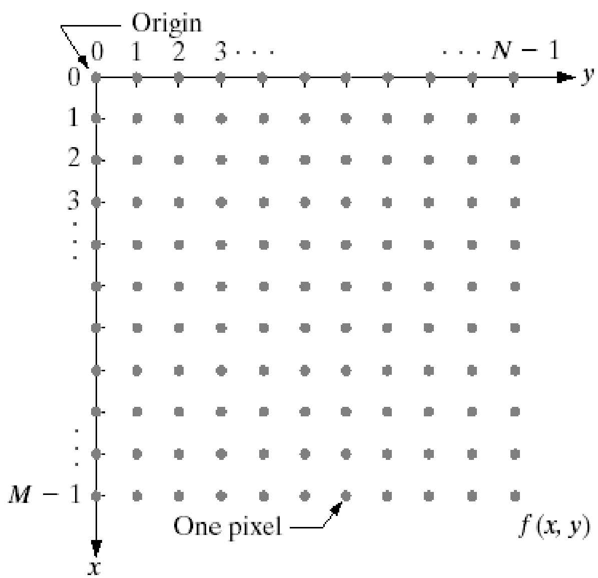
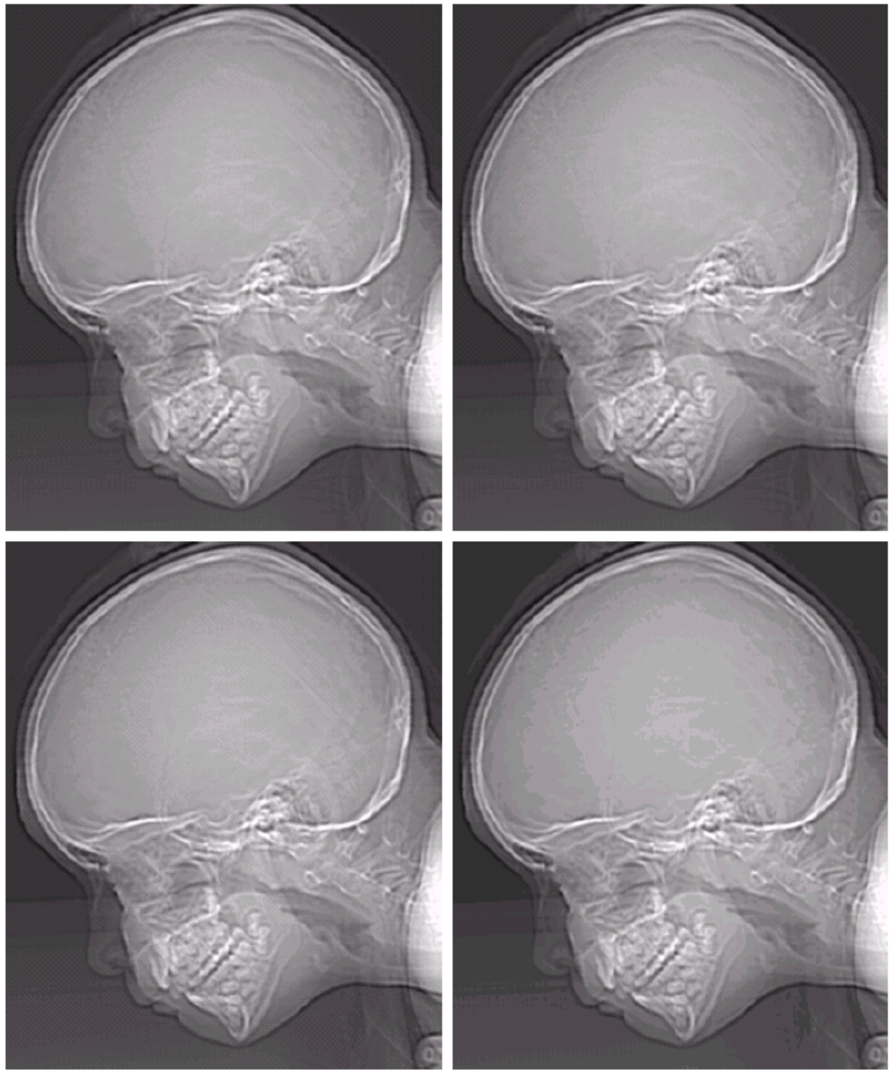


FIGURE 2.18 Coordinate convention used in this book to represent digital images.

Spatial resolution

Bit-depth resolution

Bit-depth Resolution



a b
c d

FIGURE 2.21 (a) 452 × 374, 256-level image. (b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant.

Bit-depth Resolution (Con'd)

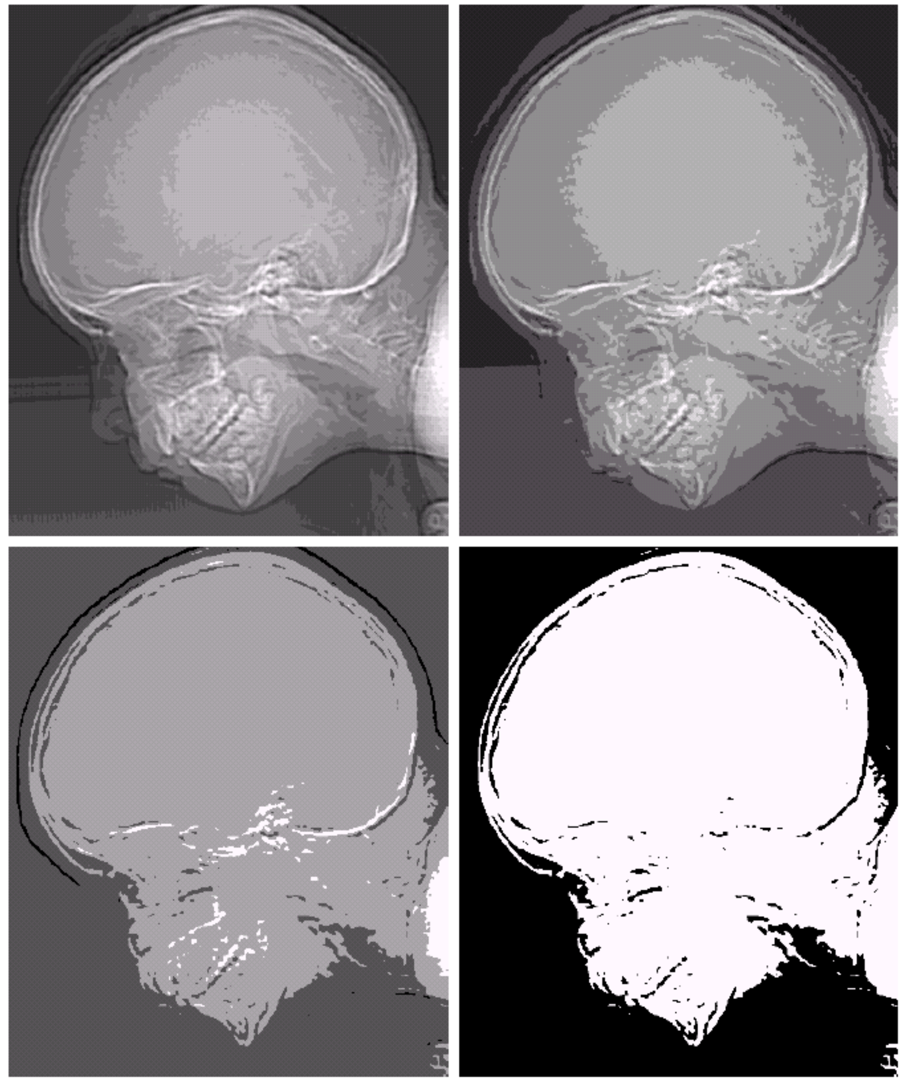
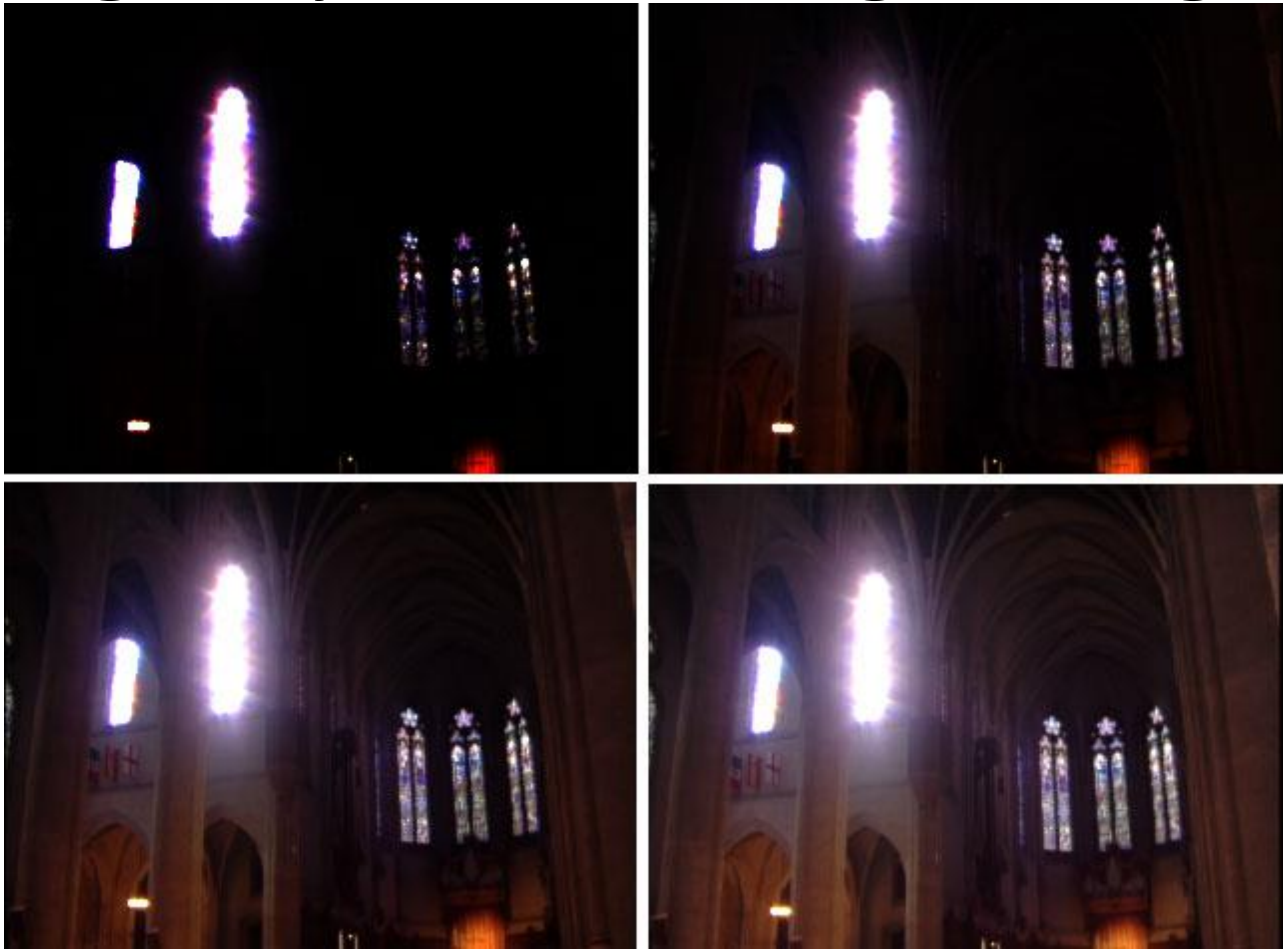


FIGURE 2.21 *(Continued)* (e)–(h) Image displayed in 16, 8, 4, and 2 gray levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)

High Dynamic Range Imaging



Q: Can we generate a HDR image (16bpp) by a standard camera?

A: Yes, adjust the exposure and fuse multiple LDR images together

Spatial Resolution

FIGURE 2.19 A 1024 × 1024, 8-bit image subsampled down to size 32 × 32 pixels. The number of allowable gray levels was kept at 256.

Image Resampling

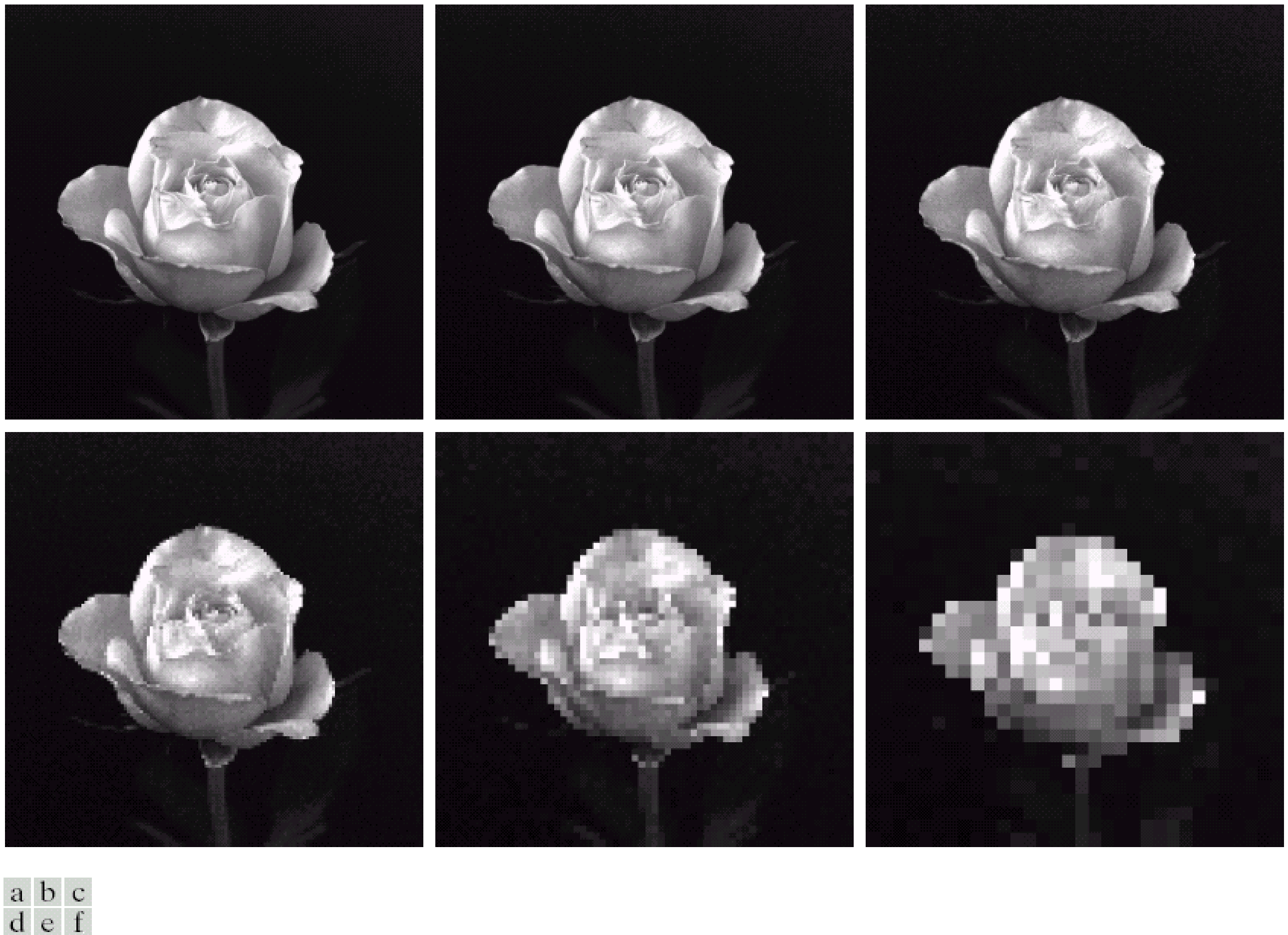
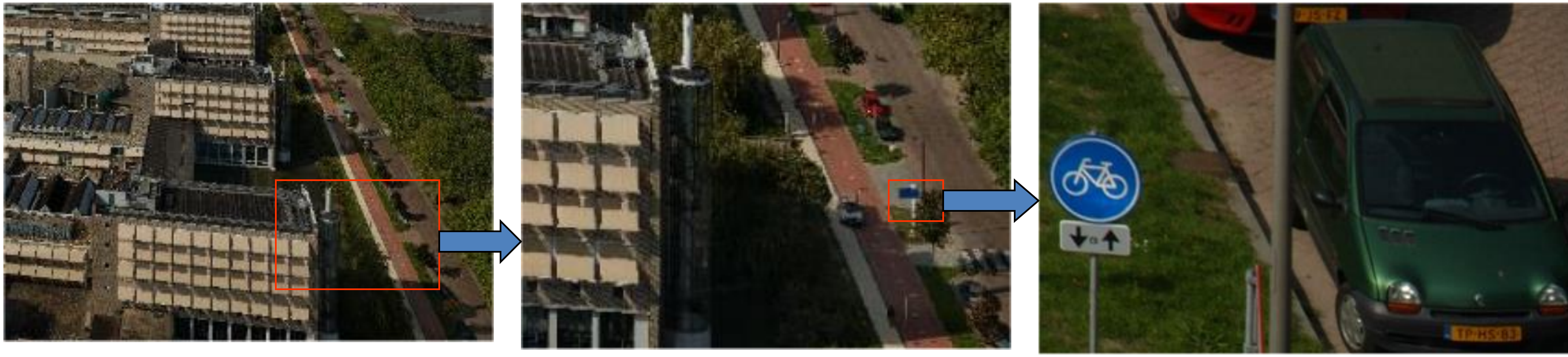


FIGURE 2.20 (a) 1024 × 1024, 8-bit image. (b) 512 × 512 image resampled into 1024 × 1024 pixels by row and column duplication. (c) through (f) 256 × 256, 128 × 128, 64 × 64, and 32 × 32 images resampled into 1024 × 1024 pixels.

Towards Gigapixel



Mega-pel ⟶ Giga-pel

Photographers and artists have manually or semi-automatically stitched hundreds of mega-pel pictures together to demonstrate how a giga-pel picture looks like → **the power of pixels**

http://triton.tpd.tno.nl/gigazoom/Delft2.htm

Block-based Processing

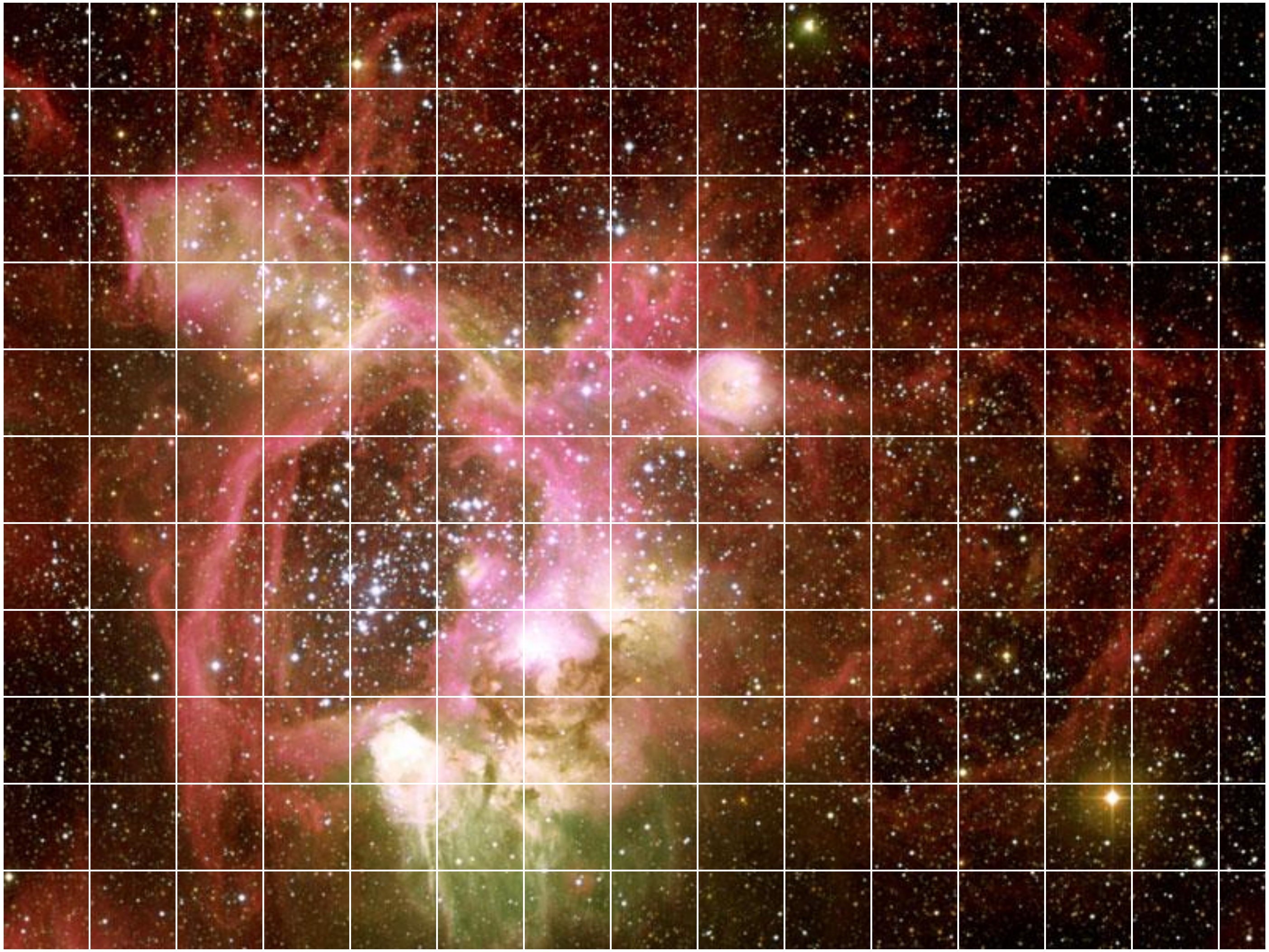
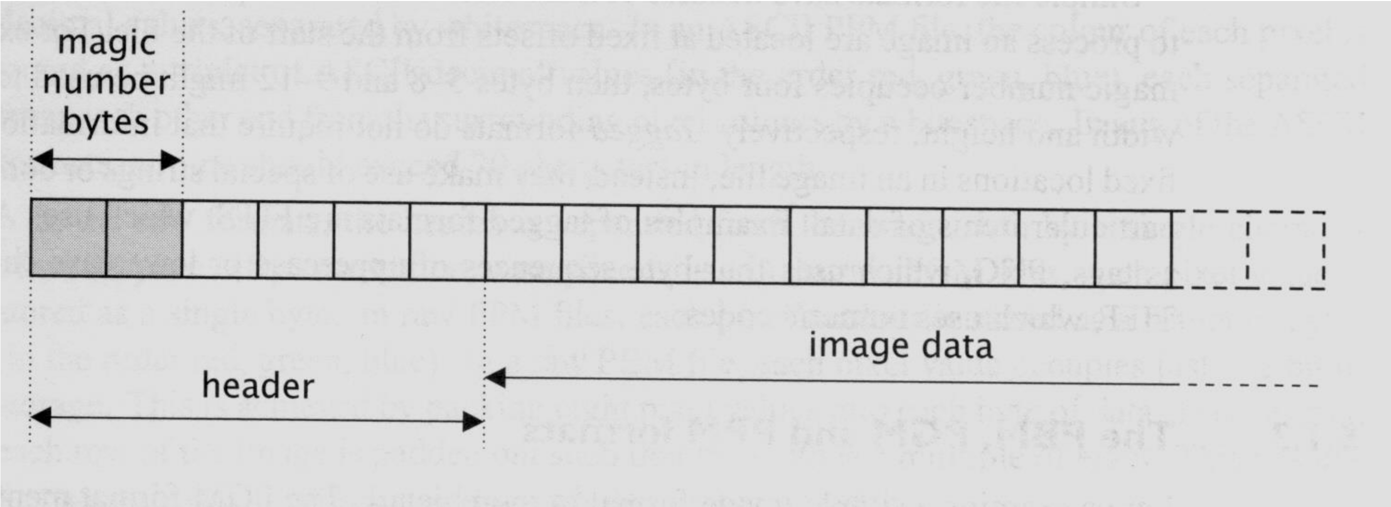


Image file formats

- Many image formats adhere to the simple model shown below (line by line, no breaks between lines).
- The header contains at least the width and height of the image.
- Most headers begin with a **signature** or “magic number” - a short sequence of bytes for identifying the file format.



Comparison of image formats

Image File Format	No. Bytes "Hi"	No. Bytes "Cars"
PGM	595	509,123
GIF	192	138,267
TIF	918	171,430
PS	1591	345,387
HIPS	700	160,783
JPG (lossless)	684	49,160
JPG (lossy)	619	29,500