Chapter 2: Digital Image Fundamentals

Chapter Overview:

- ✓ To briefly introduce the concept of human vision
- ✓ To discuss lights, EM spectrum imaging characteristics
- ✓ To introduce the fundamental concept of DIP such as sampling, gray level quantization, digital image representation, zooming and shrinking
- ✓ To introduce some basic operations on digital image

Human Eye

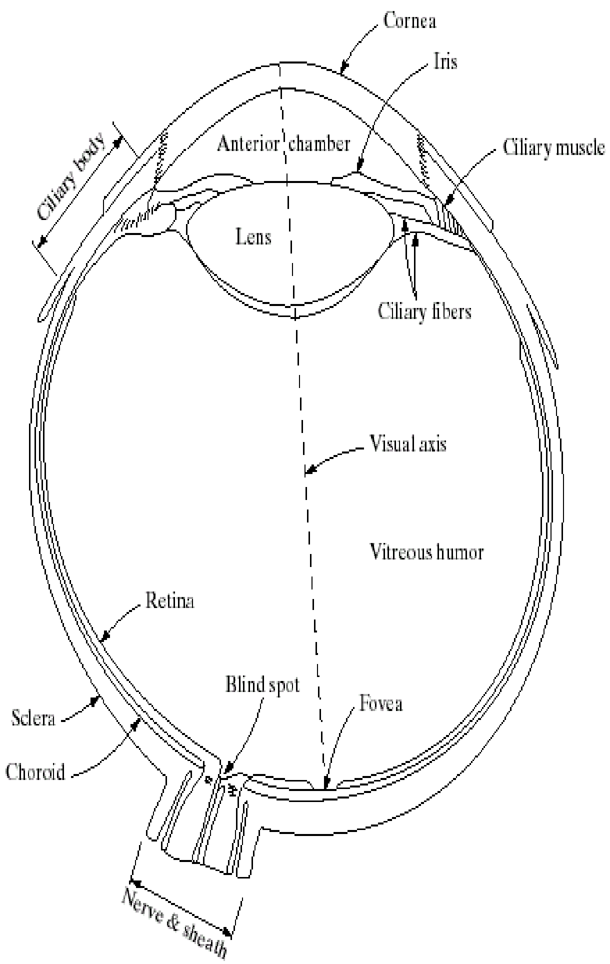


FIGURE 2.1 Simplified diagram of a cross section of the human eye.

1. Shape is nearly a sphere.
2. Average diameter = 20 mm.
3. 3 membranes:
   - Cornea and Sclera - outer cover
   - Choroid
   - Retina- enclose the eye.

Cornea :

- Tough, transparent tissue, covers, the anterior surface of the eye.

Sclera :

- Opaque membrane, enclose the remainder of the optic globe.

Structure of the Human Eye

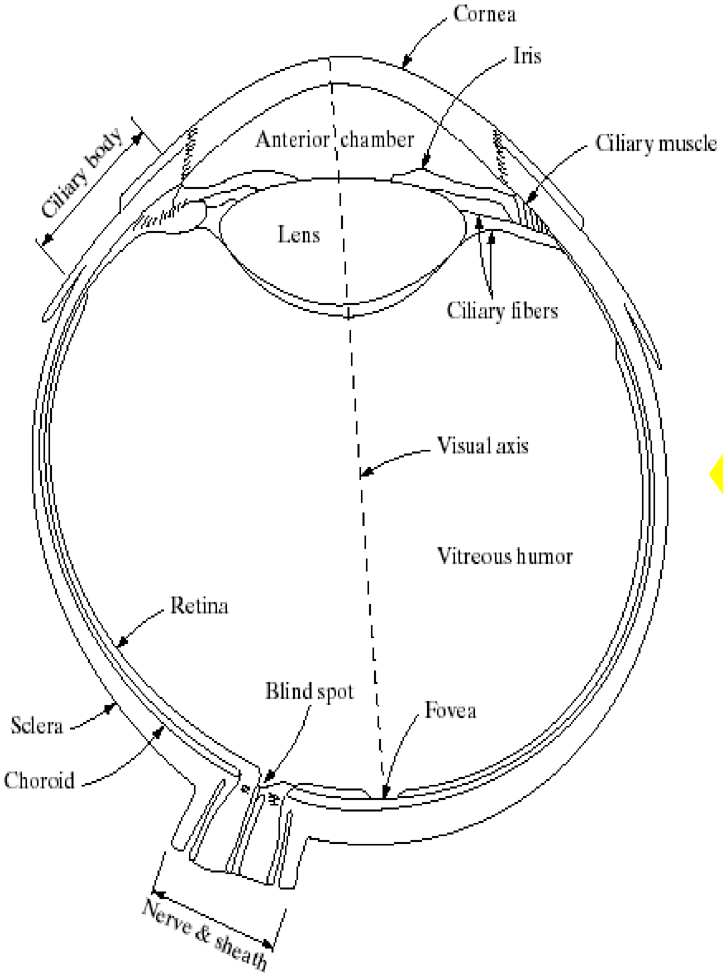


FIGURE 2.1 Simplified diagram of a cro section of the human eye.

Choroid :

- Lies below the sclera, contain network of blood vessels that serve as the major source of nutrition to the eye.

- Superficial injury to choroid could lead severe eye damage as a result of inflammation that restrict blood flow

Lens & Retina

- Lens
  - Contains 60-70% water, around 6% of fat and protiens
  - use for focusing object by the eye
  - Both infrared and ultraviolet light are absorbed appreciably by proteins within the lens structure and, in excessive amounts, can cause damage to the eye.
- Retina
  - Innermost membrane of the eye which lines inside of the wall's entire posterior portion. When the eye is properly focused, light from an object outside the eye is imaged on the retina.

Receptors

- Pattern vision is afforded by the distribution of discrete light receptors over the surface of the retina.
- Receptors are divided into 2 classes:
  - Cones
  - Rods

Cones

- 6-7 million, located primarily in the central portion of the retina called the fovea
- muscles controlling the eye rotate the eyeball until the image of the objects falls on the fovea.
- Highly sensitive to color.
- Each one is connected to its own nerve end, therefore human can resolve fine details.
- Cone vision is called photopic or bright-light vision.

Rods

- 75-150 million, distributed over the retina surface.
- Several rods are connected to a single nerve end, reduce the amount of the detail discernible by these receptors.
- Serve to give a general, overall picture of the field of view.
- Not involved in color vision
- Sensitive to low levels of illumination.
- Rod vision is called scotopic or dim-light vision.

Distribution of rods and cones

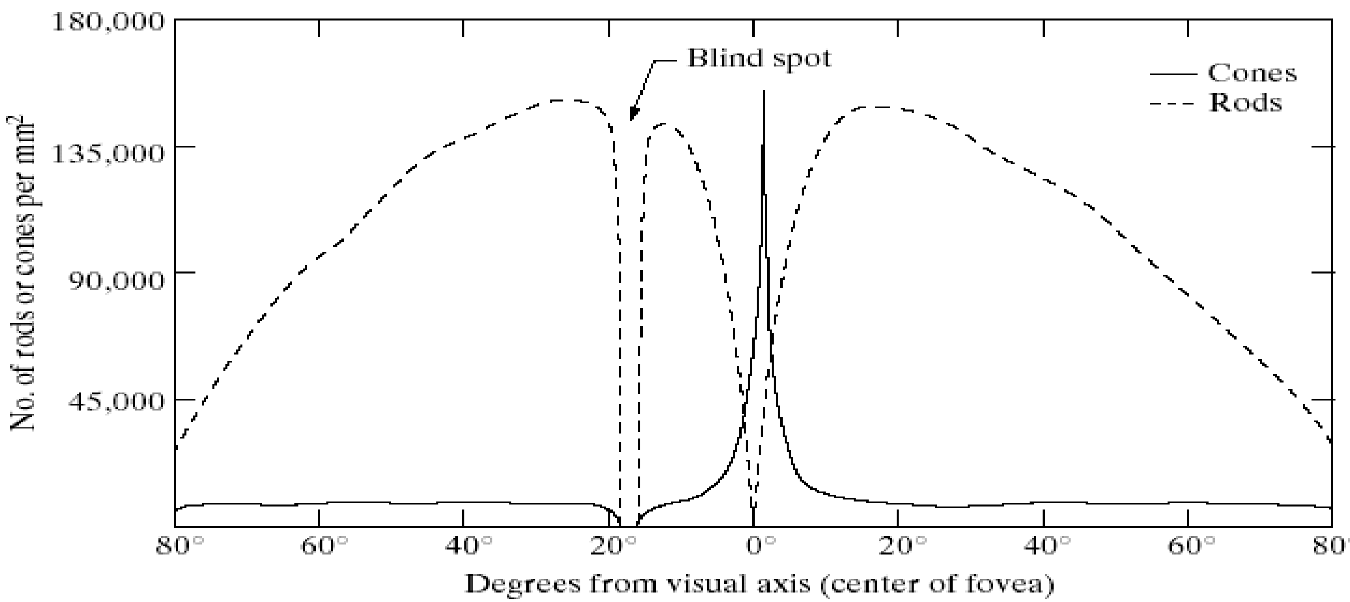


FIGURE 2.2 Distribution of rods and cones in the retina.

- Blind spot – the absence of receptors area.
- Receptor density is measured in degrees from the fovea.
- Cones are most dense in the center of the retina (in the area of the fovea).
- Rods increase in density from the center out to approx. 20° off axis and then decrease in density out to the extreme periphery of the retina

Image Formation in the Eye

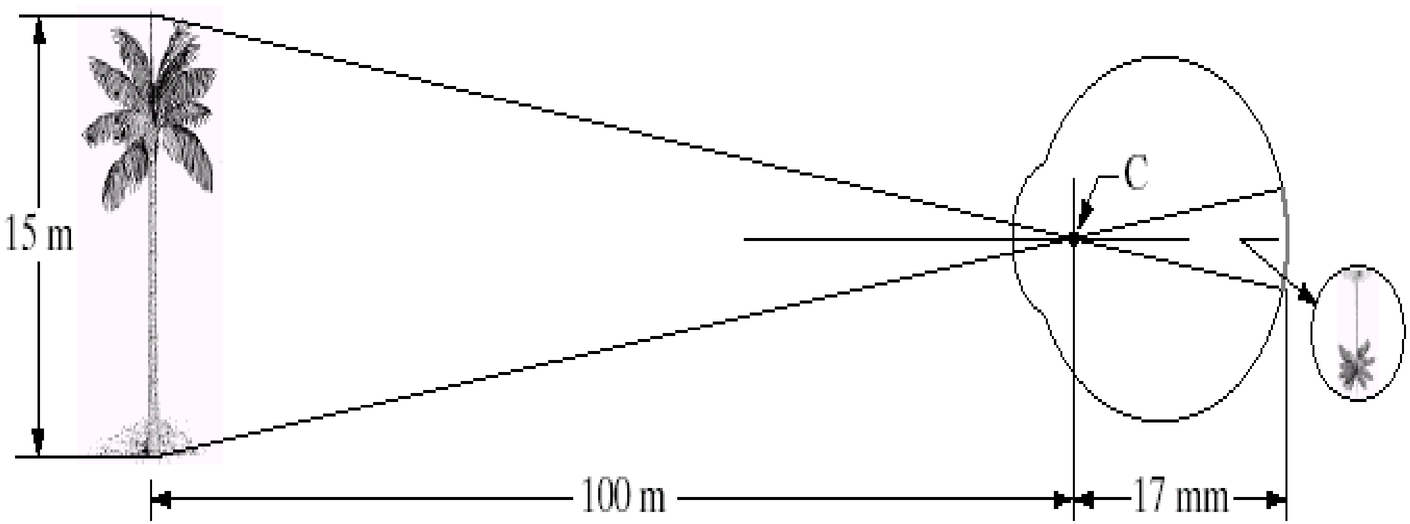


FIGURE 2.3 Graphical representation of the eye looking at a palm tree. Point *C* is the optical center of the lens.

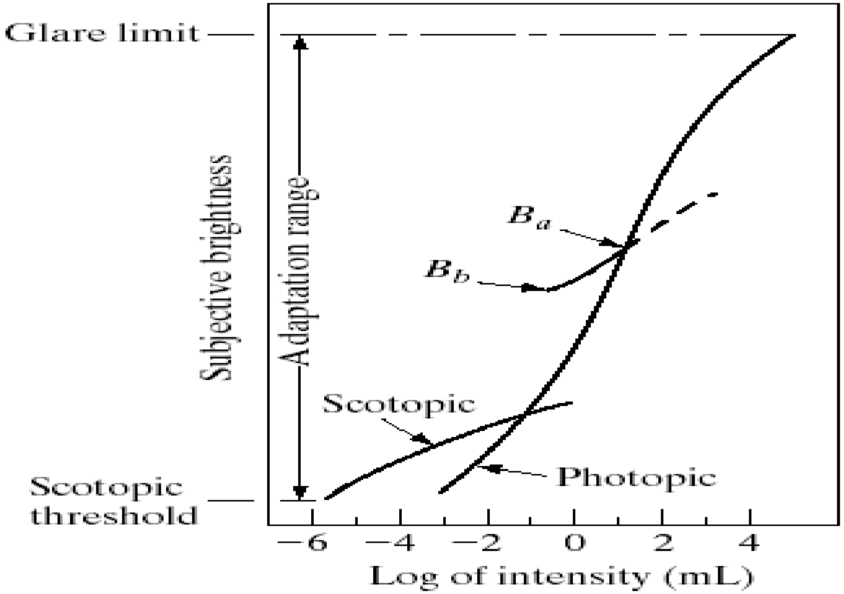
Eye's focal length – the distance between the center of the lens and the retina varies – 17 mm to about 14 mm, as the refractive power of the lens increases from its minimum to its maximum

e.g., Height of the object in the retina, $h/17 = 15/100$.

Thus, $h = 2.55$ mm

Brightness Adaptation & Discrimination

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



•B_a is brightness adaptation level. The short intersecting curve represents the range of subjective brightness that the eye can perceive when adapted to this level.

Contrast sensitivity

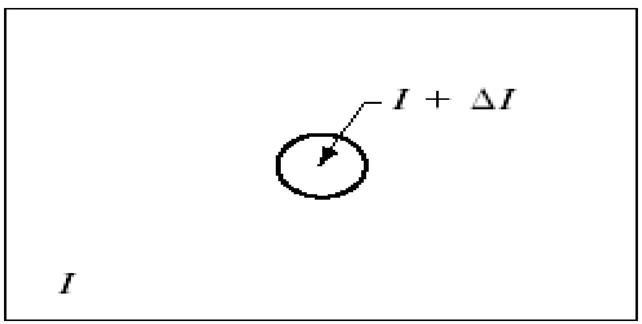


FIGURE 2.5 Basic experimental setup used to characterize brightness discrimination.

Weber's ratio: ΔI_c / I

Good brightness discrimination – small value

=ΔI$_{c/}$ I

Bad brightness discrimination

=**ΔI$_c$ / I** is large

•The ability of the eye to discrimination b/w changes in brightness at any specific adaptation level is of considerable interest

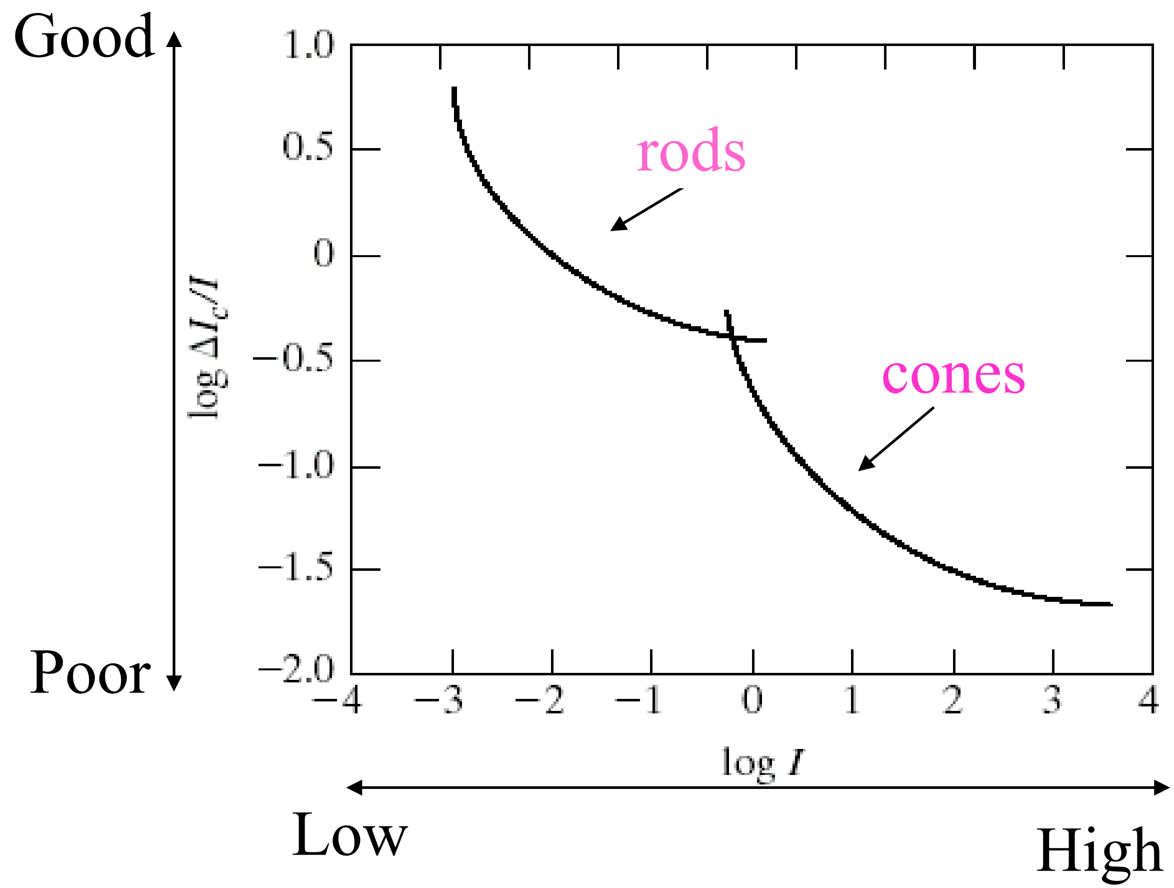
• I is uniformly illumination on a flat area large enough to occupy the entire field of view.

• ΔI_c is the change in the object brightness required to just distinguish the object from the background.

Web ratio

FIGURE 2.6 Typical Weber ratio as a function of intensity.

Discrimination



Level illumination

Weber Ratio

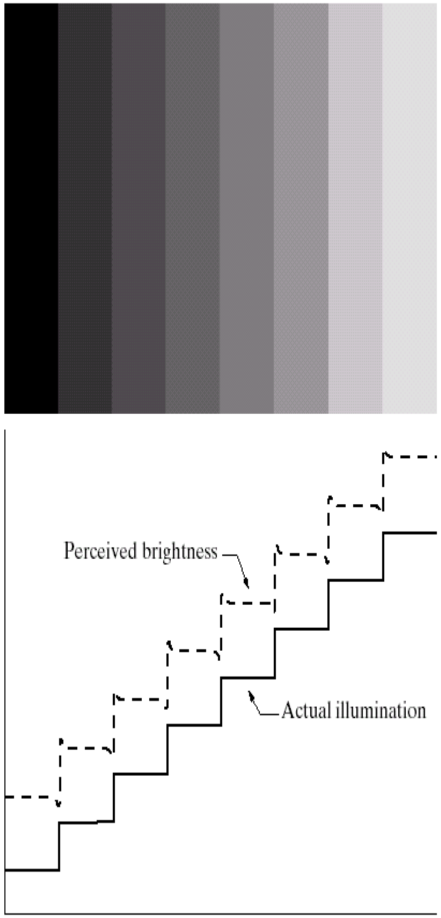
Brightness discrimination is poor

(the weber ratio is large)

at low levels of illumination and improves significantly
(the ratio decreases)
as background illumination increases.

Hard to distinguish the discrimination when it is bright area
but easier when the discrimination is on a dark area

Brighteness vs. Function of intensity



a
b

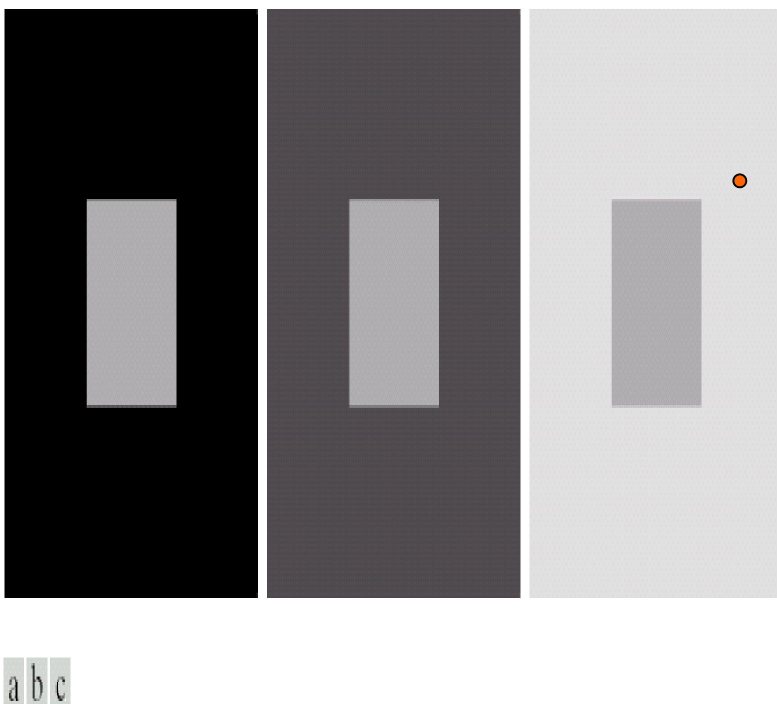
FIGURE 2.7 (a) An example showing that perceived brightness is not a simple function of intensity. The relative vertical positions between the two profiles in (b) have no special significance; they were chosen for clarity.

Brightness is not a simple function of Intensity.

Visual system tends to undershoot or overshoot around the boundary of region of different intensities.

The intensity of the stripes is constant but we actually perceive a brightness pattern is strongly scalloped near the boundaries.

Simultaneous Contrast

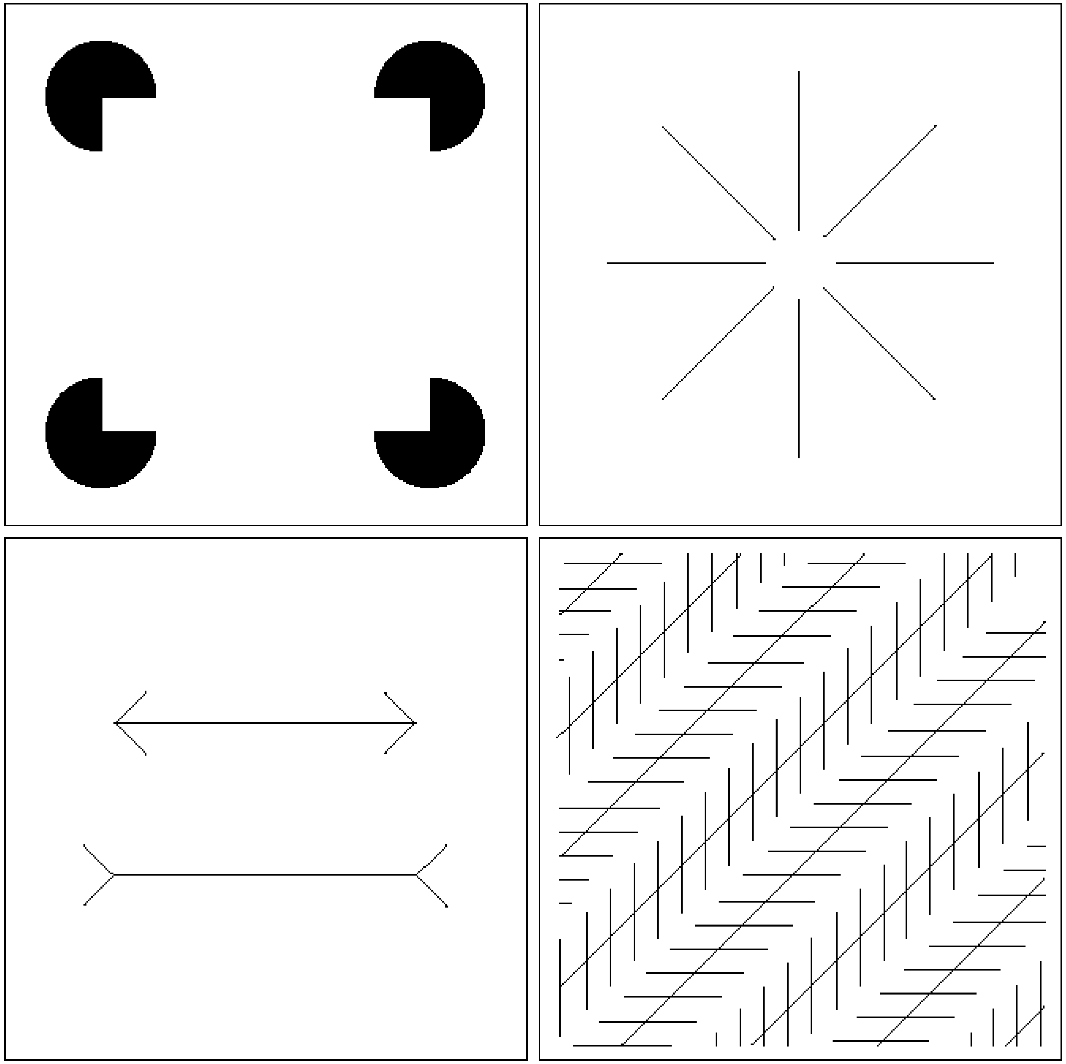


Which small square is the darkest one?

- All the small squares have exactly the same intensity, but they appear to the eye progressively darker as the background becomes brighter.

- Region's perceived brightness does not depend simply on its intensity.

Human Perception Phenomena



a b
c d

FIGURE 2.9 Some well-known optical illusions.

Light & the Electromagnetic Spectrum (EM)

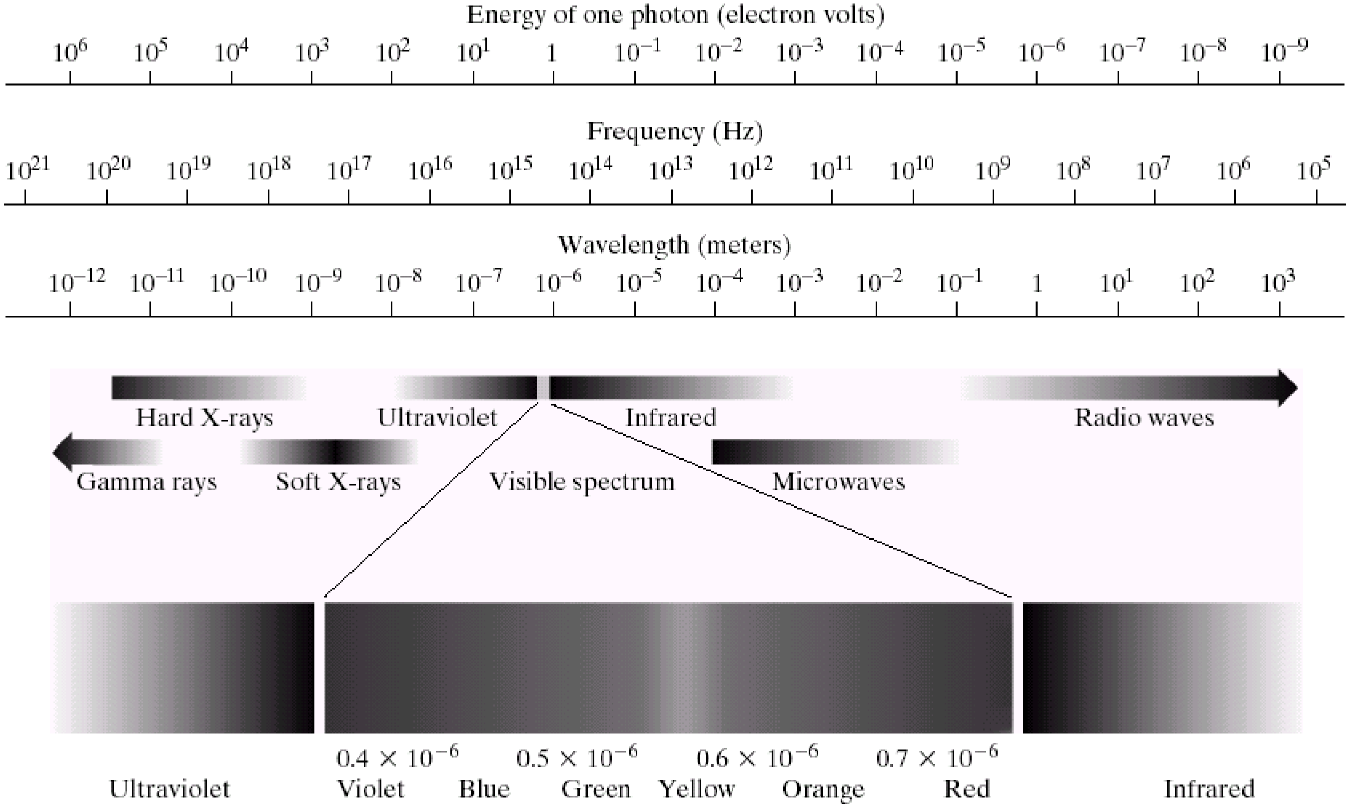


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

Monochromatic vs Chromatic

Void of color (achromatic)

- only attribute is intensity
- generaly used term is gray level

Visible band (color)

- 0.43 micron (violet) to 0.7 micron (red)
-The quality of light is described using 3 basic quantitied
  1. Radiance – total amount of energy that flow from light source (Watt)
  2. Luminance – a measure of energy an observer perceives from a light source (lumens (lm)
  3. Brightness – subjective descriptor of light perception that is practically impossible to measure

Sampling

- Sampling = the spacing of discrete values in the domain of a signal.
- Sampling-rate = how many samples are taken per unit of each dimensions. e.g. samples per second, frames per second, etc.

Quantization

- Quantization = spacing of discrete values in the range of a signal.
- Usually thought of as the number of bits per sample of the signal. e.g. 1 bit per pixel (b/w images), 16-bit audio, 24-bit color images, etc.

Digital Image Representation

- A digital image is an image $f(x,y)$ that has been digitized both in spatial coordinates and brightness.
- The value of f at any point (x,y) is proportional to the brightness (or gray level) of the image at that point.

Digital Image Representation

A digital image can be considered a matrix whose row and column indices identify a point in the image and the corresponding matrix element value identifies the gray level at that point.

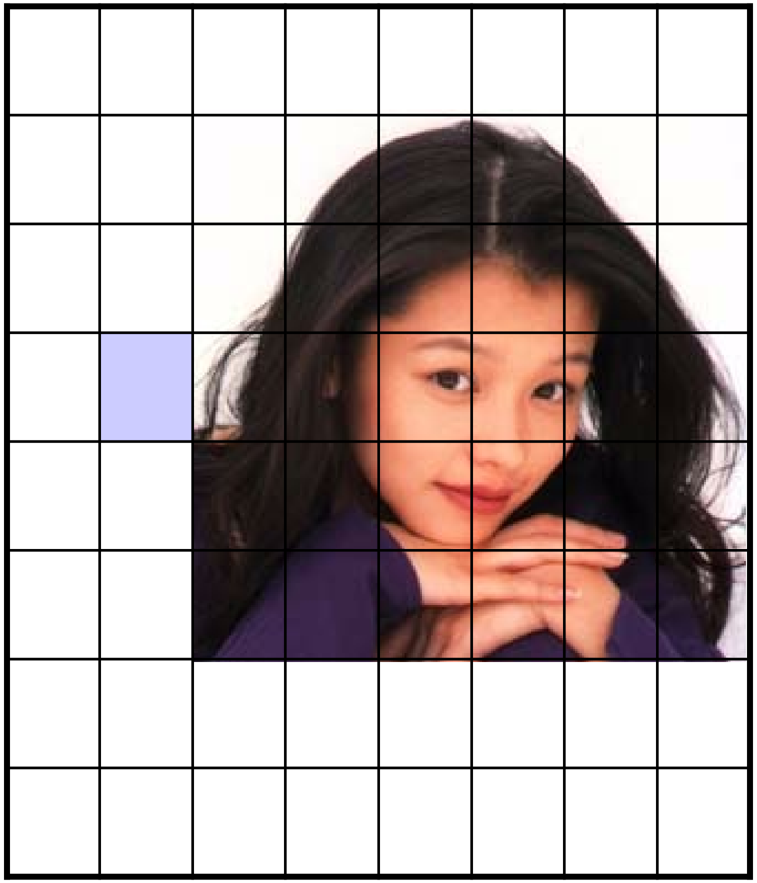


Image Sensing & Acquisition

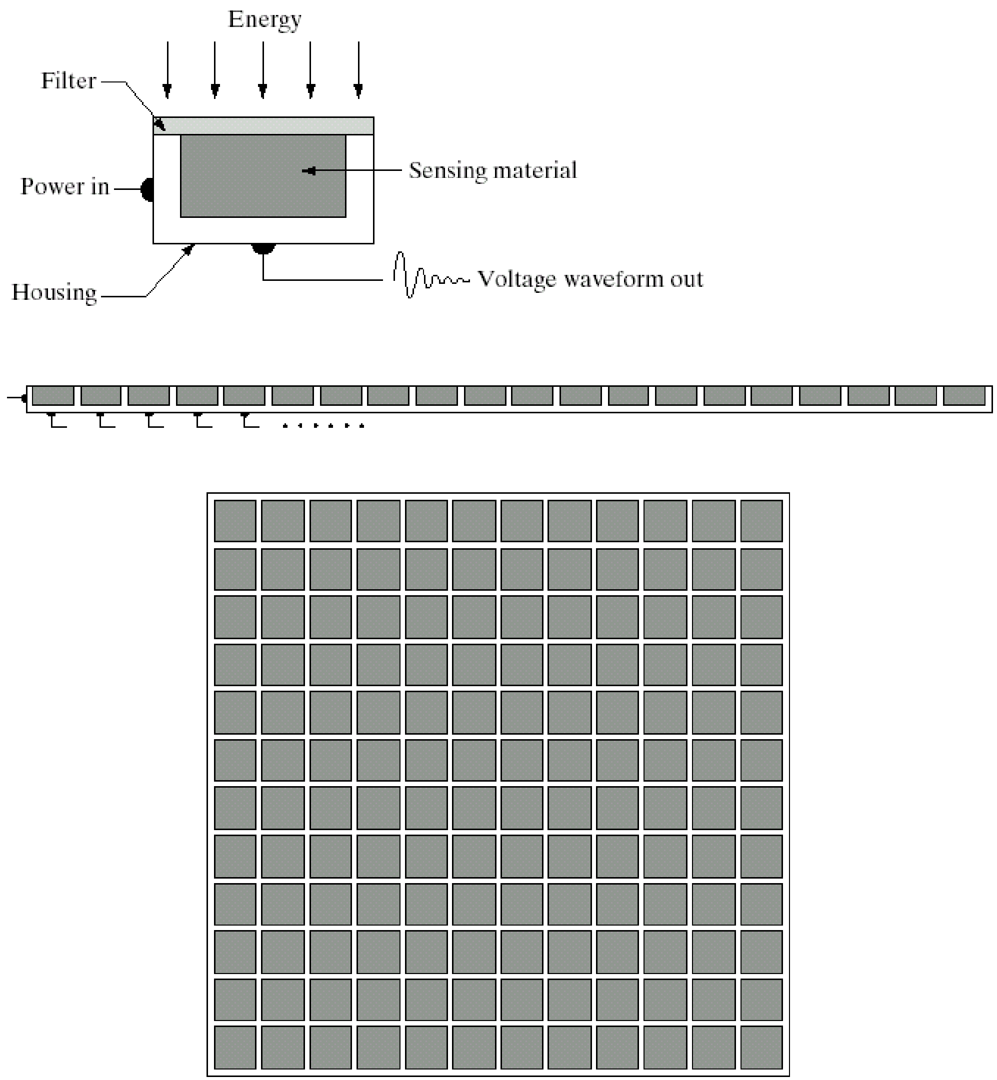


FIGURE 2.12
(a) Single imaging sensor.
(b) Line sensor.
(c) Array sensor.

Image Sensing & Acquisition

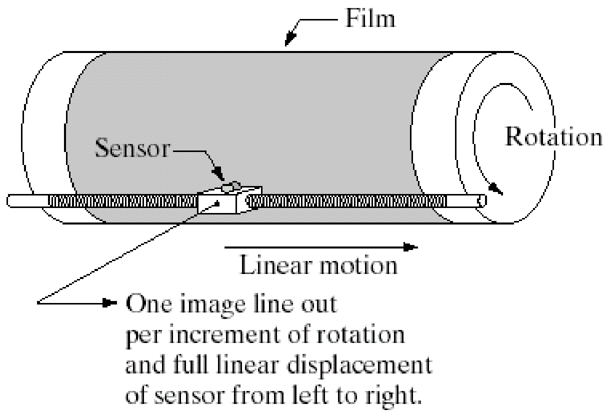
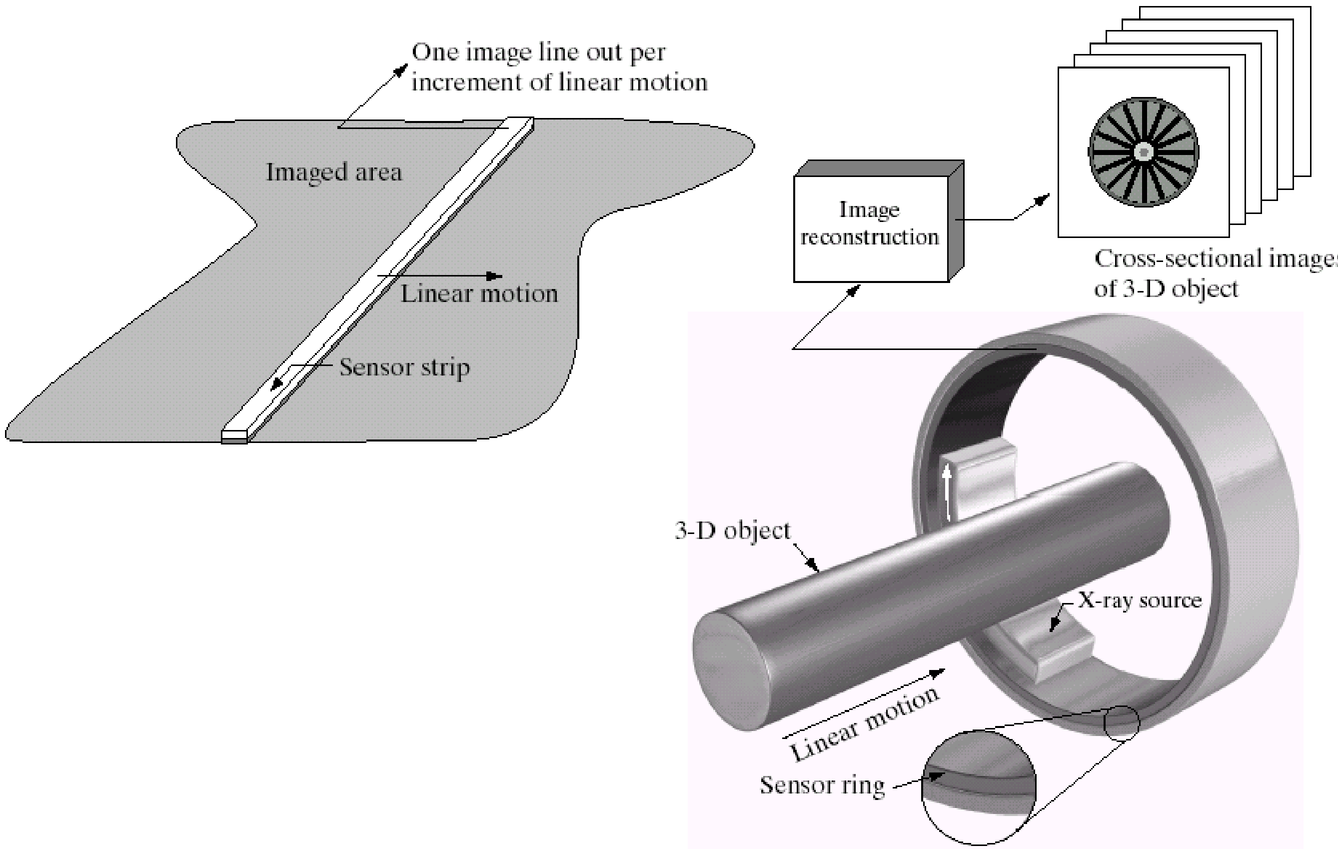


FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

Image Sensing & Acquisition



a b

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

Image Acquisition using Sensor Array

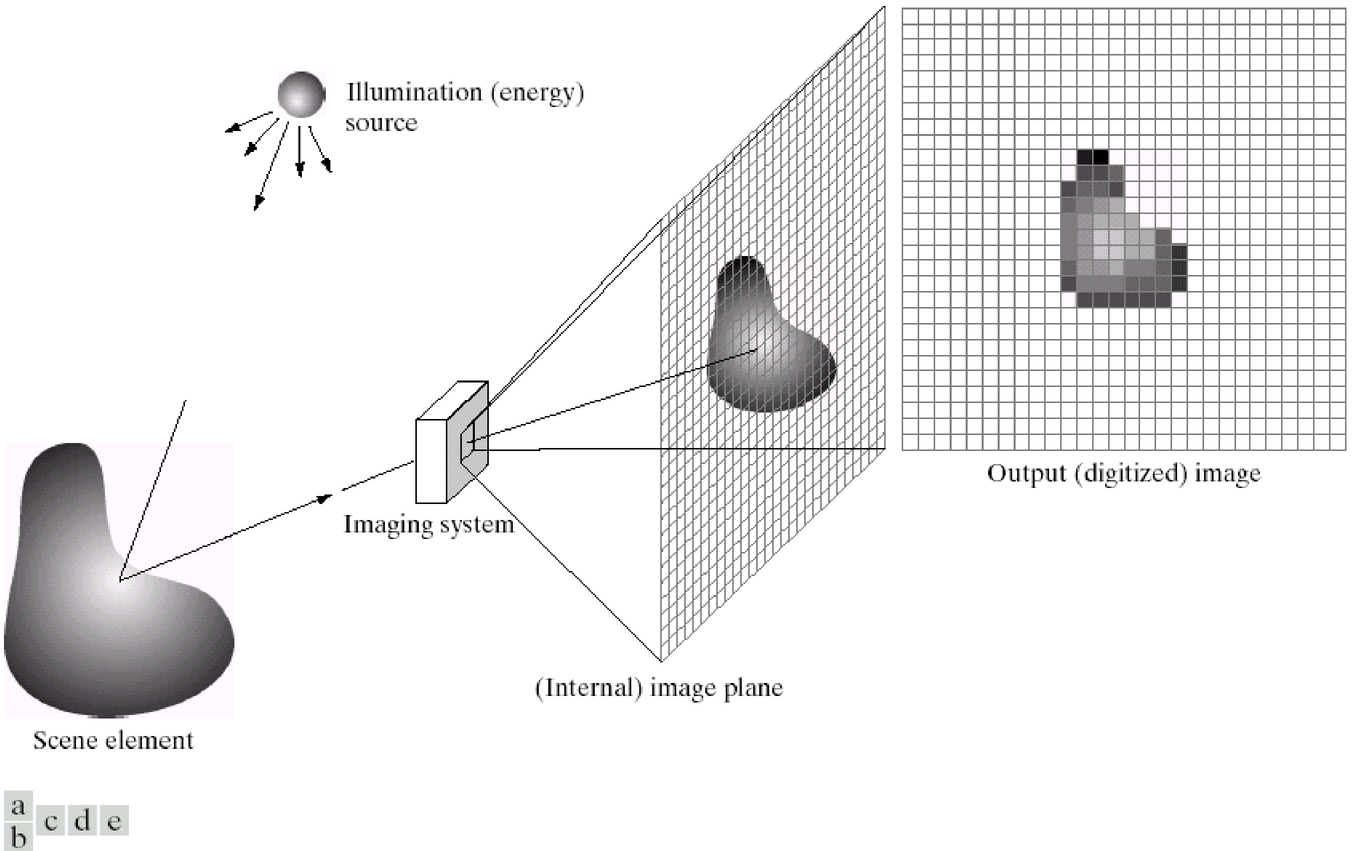


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

Generating Digital Image

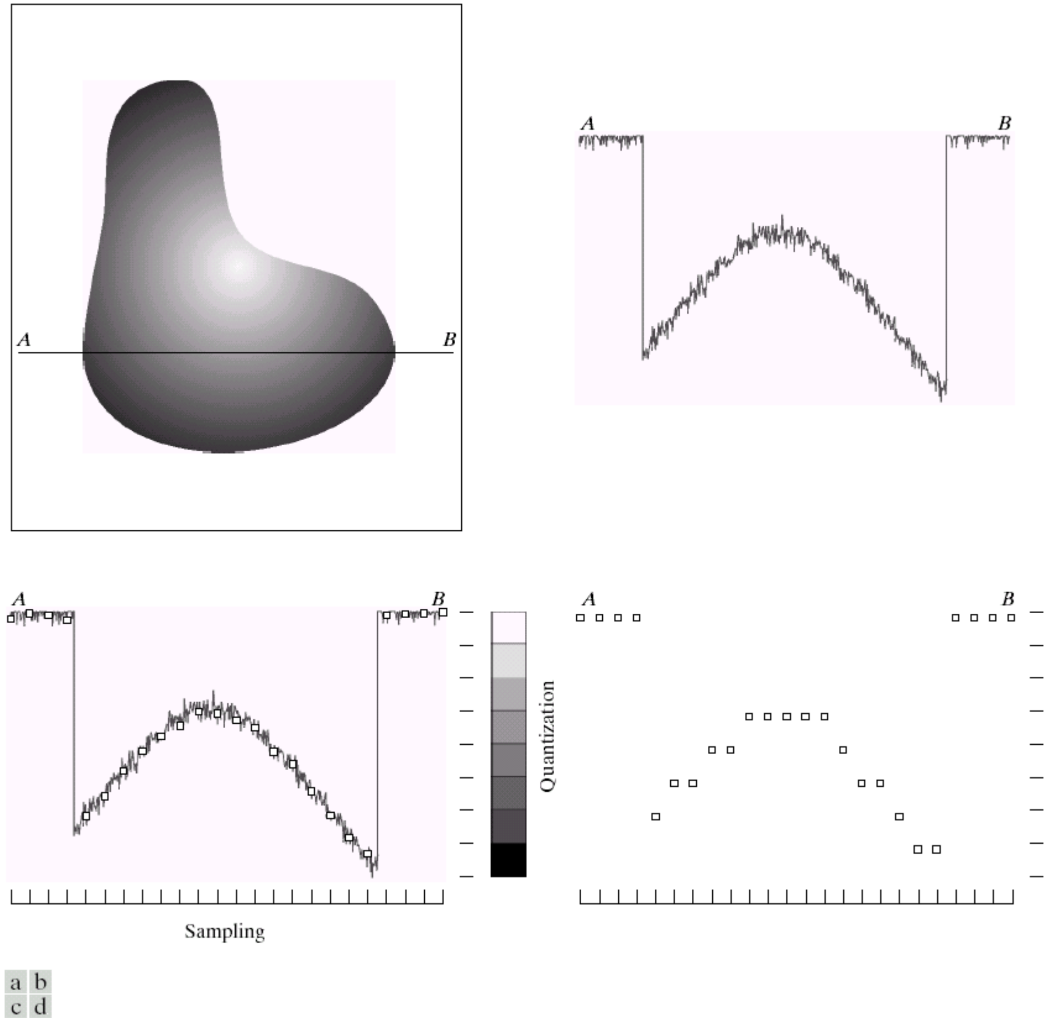
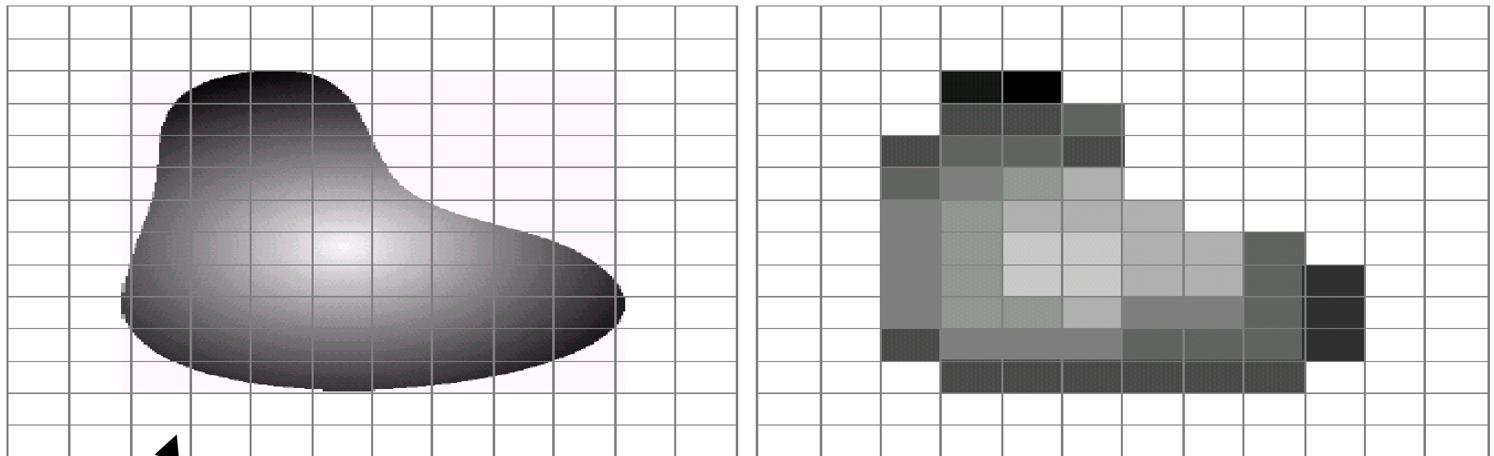


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

Example of Digital Image



a b

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

Continuous image projected onto a sensor array

Result of image sampling and quantization

Light-intensity function

- Image refers to a 2D light-intensity function, $f(x,y)$
- The amplitude of f at spatial coordinates (x,y) gives the intensity (brightness) of the image at that point.
- Light is a form of energy thus $f(x,y)$ must be nonzero and finite

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

Illumination & Reflectance

- The basic nature of $f(x,y)$ may be characterized by 2 components:
  - The amount of source light (illumination) incident on the scene being viewed called Illumination, $\boldsymbol{i}(\mathbf{x},\mathbf{y})$
  - The amount of light (illumination) reflected by the objects in the scene called Reflectance, $\mathbf{r}(\mathbf{x},\mathbf{y})$

Illumination & Reflectance

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

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

Determined by the nature of the light source

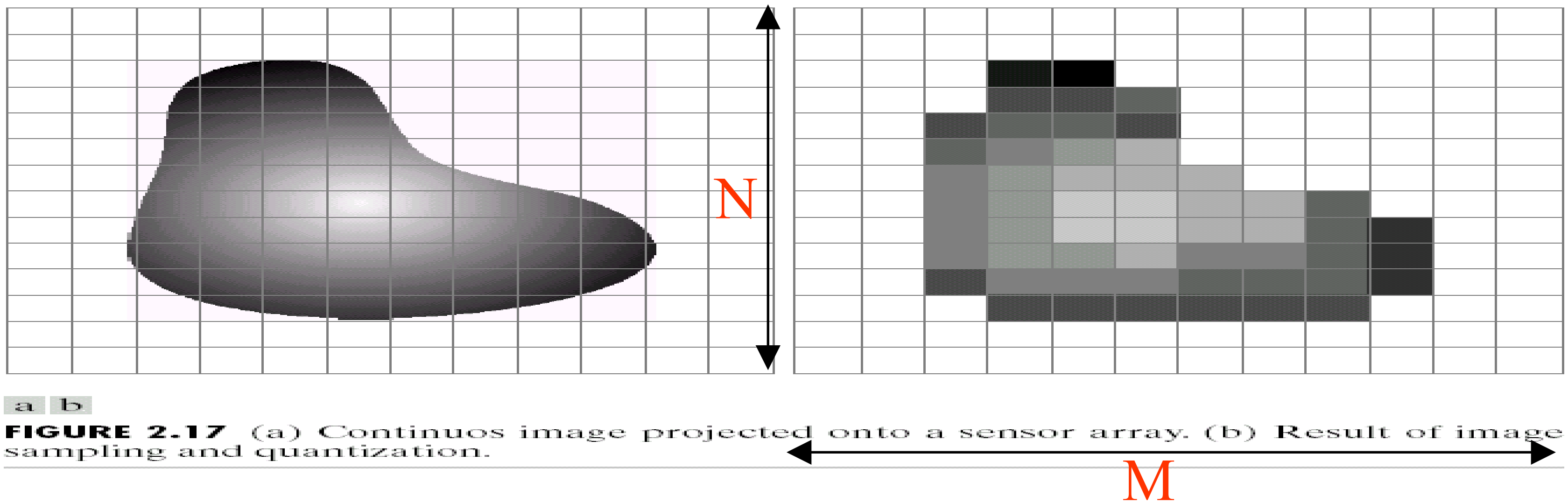
$$0 < r(x,y) < 1$$

Determined by the nature of the objects in scene, bounded from total absorption ($r(x,y) = 0$) to total reflectance ($r(x,y) = 1$) .

Gray Level

- We call the intensity of a monochrome image *f* at coordinate (x,y) the gray level (*l*) of the image at that point.
- Thus, *l* lies in the range $L_{min} \leq l \leq L_{max}$
- L_{min} is positive and L_{max} is finite.
- Gray scale = $[L_{min}, L_{max}]$
- Comman practice, shift the interval to [0 , L]
- 0=black, L=white

Number of Bits



a b

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

- The number of gray levels typically is an integer power of 2

$$L = 2^k$$

- Number of bits required to store a digitized image

$$\mathbf{b = M \times N \times k}$$

Resolution

- Resolution (how much you can see the detail of the image) depends on sampling and gray levels.
- The bigger the sampling rate (n) and the gray scale(g), the better the approximation of the digitized image from the original.
- The more the quantization scale becomes, the bigger the size of the digitized image.

Coordinate Convention in Digital Images

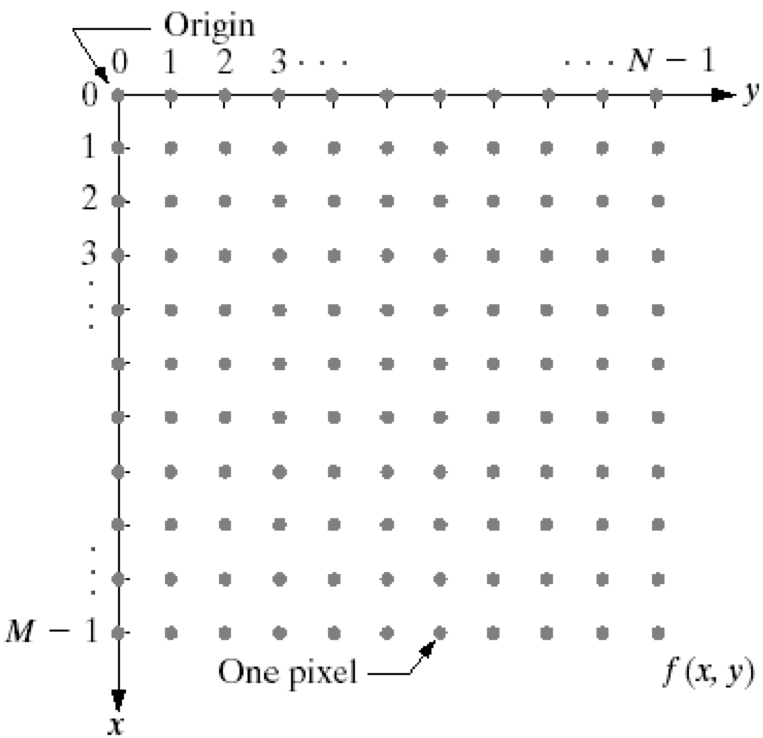


FIGURE 2.18 Coordinate convention used in this book to represent digital images.

Image Size vs No. of Bits

TABLE 2.1
Number of storage bits for various values of N and k.

N/k	1 ($L = 2$)	2 ($L = 4$)	3 ($L = 8$)	4 ($L = 16$)	5 ($L = 32$)	6 ($L = 64$)	7 ($L = 128$)	8 ($L = 256$)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,369,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912

Spatial & Gray Level Resolution

Effect of Varying Number of Samples in Image

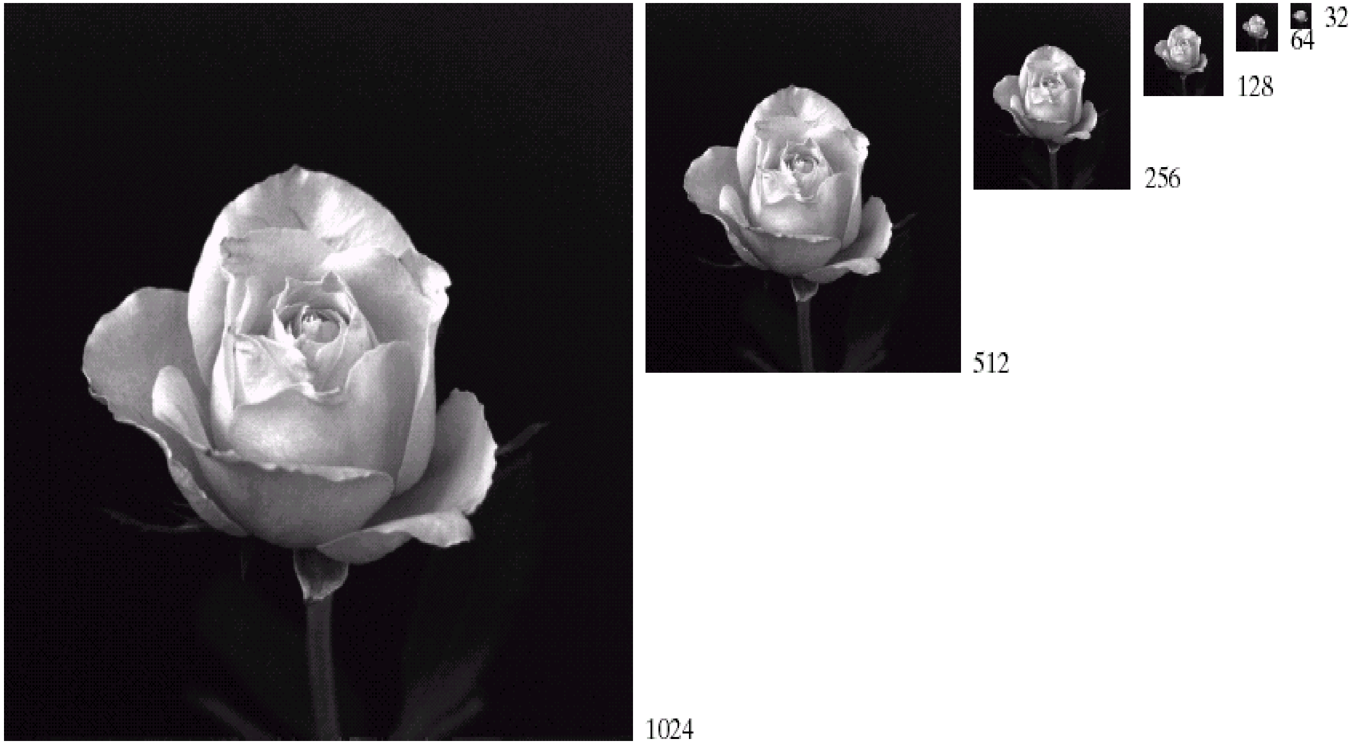
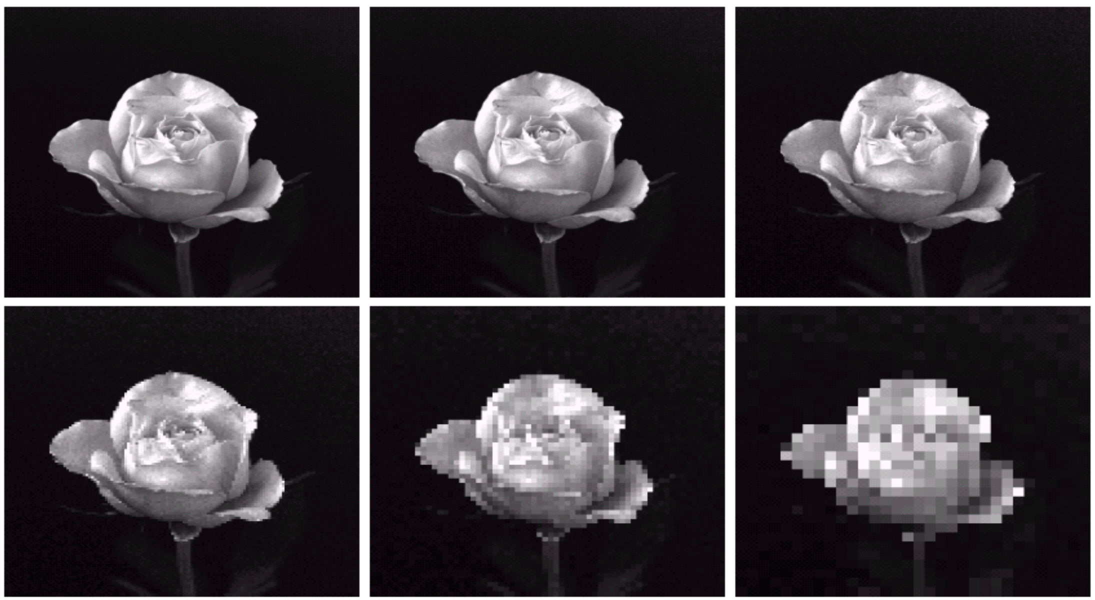


FIGURE 2.19 A 1024 × 1024, 8-bit image subsampled down to size 32 × 32 pixels. The number of allowable gray levels was kept at 256.

Effect of Varying Number of Samples in Image



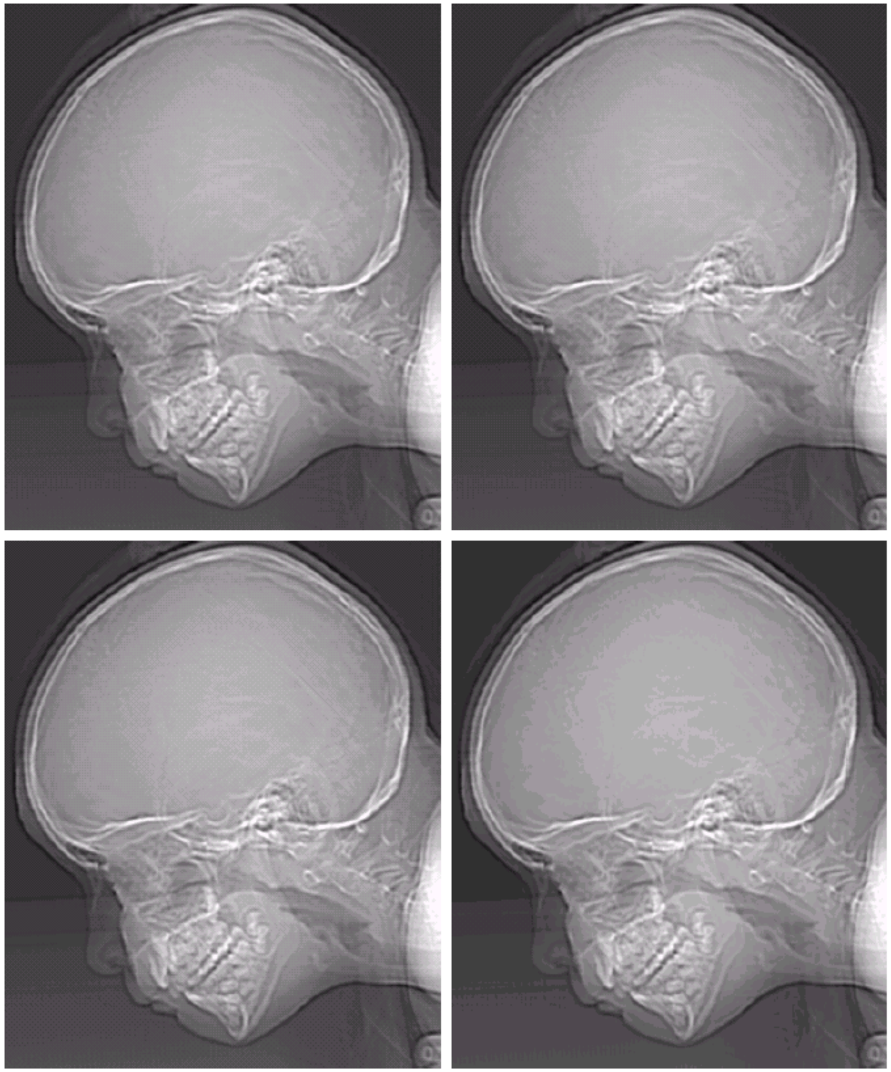
a	b	c
d	e	f

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.

(a)	1024x 1024
(b)	512 x 512
(c)	256 x 256
(d)	128 x 128
(e)	64 x 64
(f)	32 x 32

If the resolution is decreased too much, the checkerboard effect can occur.

Typical Effect of Varying of Gray Level



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.

Typical Effect of Varying of Gray Level

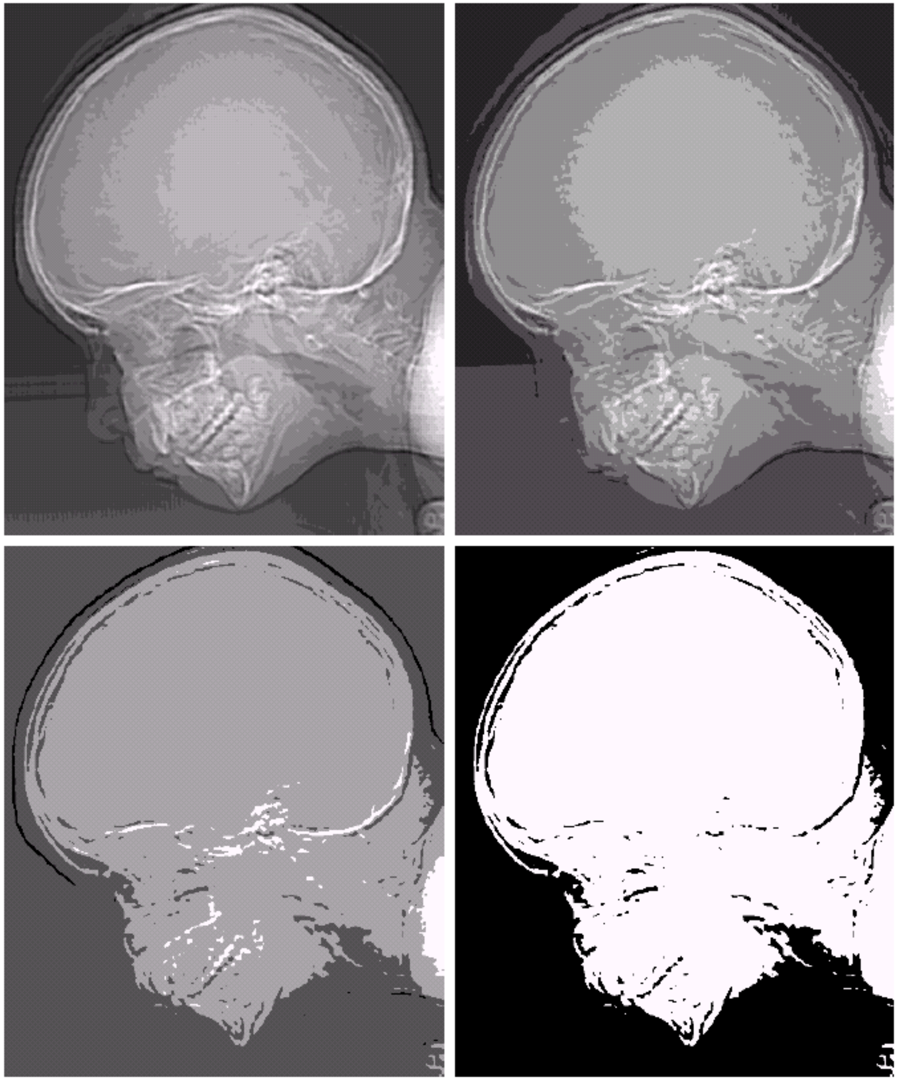


FIGURE 2.21 *(Continued)* (e)–(h) Image displayed in 16, 8, 4, and 2 gray levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)

Zooming of Digital Images

Zooming – 2 steps operation:

1- creating new pixel locations

2- assignment of gray levels to those new locations

e.g.

Image size – 500 x 500 → enlarge to 1.5 times. → 750 x 750

Using nearest neighbor interpolation, bilinear interpolation, etc.

Crude method – row/column replication.

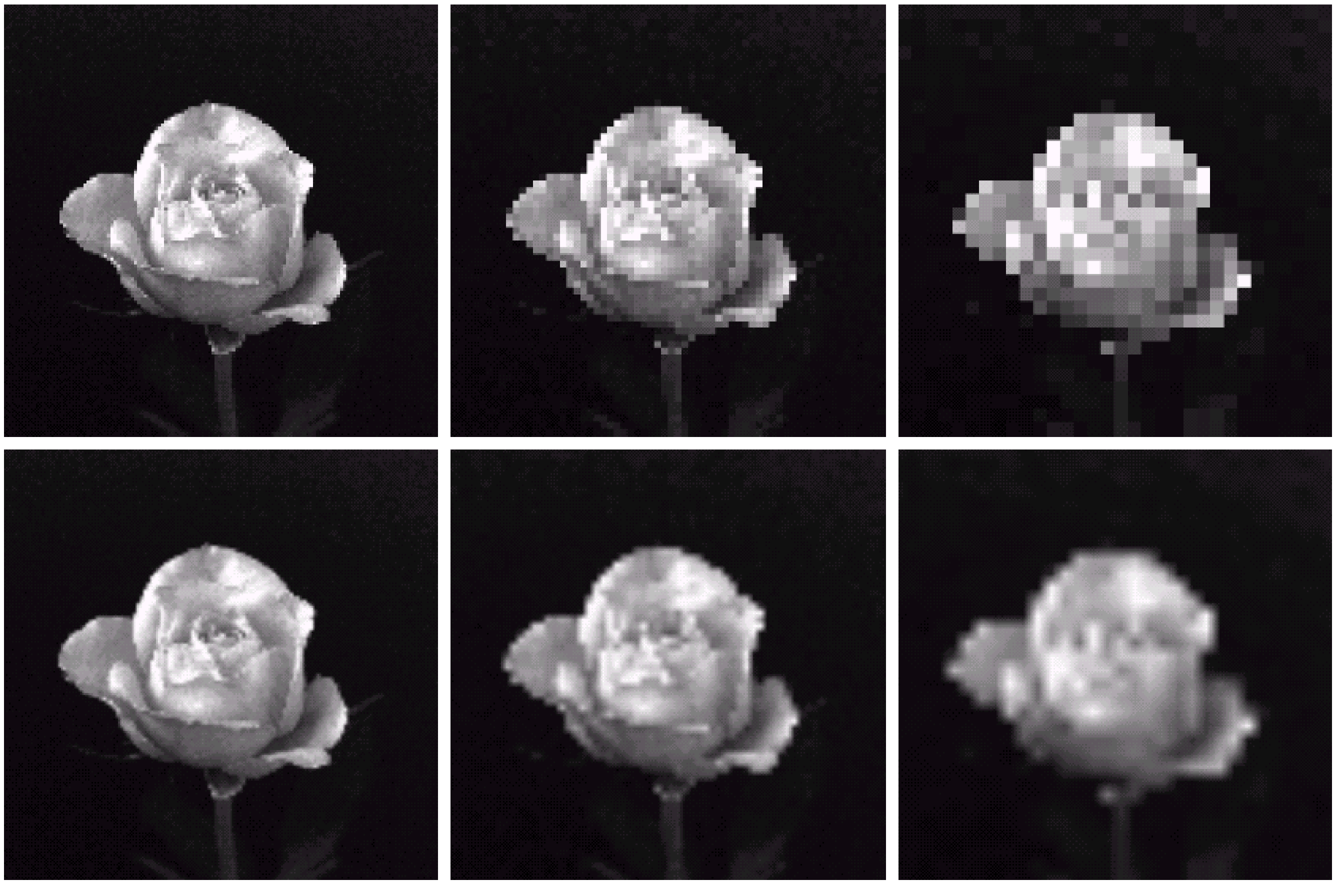
e.g. bilinear interpolation.

$$v(x', y') = ax' + by' + cx'y' + d$$

Where the four coefficient are determined from 4 equations in 4 unknown that can be written using the 4 nearest neighbor of point (x',y')

Shringking of Digital Images

Similar as zooming but by deleting row/column.



a b c
d e f

FIGURE 2.25 Top row: images zoomed from 128 × 128, 64 × 64, and 32 × 32 pixels to 1024 × 1024 pixels, using nearest neighbor gray-level interpolation. Bottom row: same sequence, but using bilinear interpolation.

Basic Relationship b/w pixels

- Neighbors of a pixel
- Connectivity
- Labeling of connected components
- Relations, Equivalences, and Transitive Closure
- Distance measures
- Arithmetic/ Logic Operations

Neighbors of a pixel

- A pixel *p* at coordinate (x,y) has
  - N_4(p): 4-neighbors of *p*
    (x+1,y),(x-1,y),(x,y+1),(x,y-1)
  - N_D (p): 4-diagonal neighbors of p
    (x+1,y+1),(x+1,y-1),(x-1,y+1),(x-1,y-1)
  - N_8(p): 8-neighbors of *p*:
    a combination of N_4(p) and N_D (p):

	X	
X	P	X
	X	

X		X
	P	
X		X

X	X	X
X	P	X
X	X	X

Connectivity

- Let *V* be the set of gray-level values used to defined connectivity
  - 4-connectivity :
    - 2 pixels p and q with values from *V* are 4-connected if q is in the set N_4 (p)
  - 8-connectivity :
    - 2 pixels p and q with values from *V* are 8-connected if q is in the set N_8 (p)
  - M-connectivity (mixed connectivity):
    - 2 pixels p and q with values from *V* are m-connected if
      - q is in the set N_4 (p) or
      - q is in the set N_D (p) and the set $N_4 (p) \cap N_4 (q)$ is empty
      - (the set of pixels that are 4-neighbors of both p and q whose values are from *V*)

Adjacent

- A pixel *p* is adjacent to a pixel *q* if they are connected.
- Two image area subsets S1 and S2 are adjacent if some pixel in S1 is adjacent to some pixel S2.

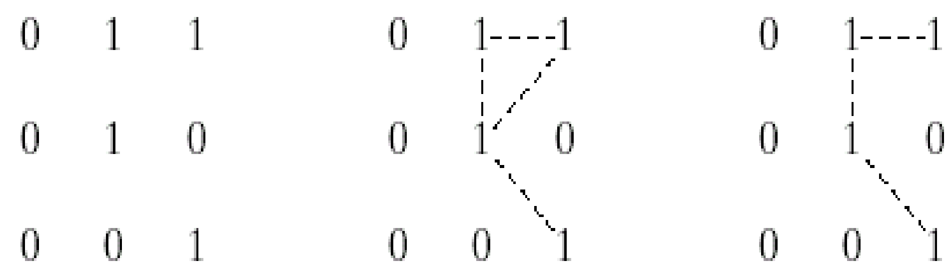
Path (Curve)

- A path (curve) from pixel *p* with coordinates (x,y) to pixel *q* with coordinates (s,t) is a sequence of distinct pixels with coordinates

  $(x_0,y_0),(x_1,y_1),\ldots.(x_n,y_n)$

  where $(x_0,y_0) = (x,y)$, $(x_n,y_n) = (s,t)$ and pixels (x_i,y_i) and (x_{i-1},y_{i-1}) are adjacent for $1 <= i <= n$
- n is the length of the path.
- We can define 4-,8-,or *m*-paths depending on type of adjacency specified.

Connectivity Concept Example



a b c

FIGURE 2.26 (a) Arrangement of pixels; (b) pixels that are 8-adjacent (shown dashed) to the center pixel; (c) *m*-adjacency.

m-connectivity eliminates the multiple path connections that arise in 8-connectivity.

Labeling of Connected Components

- Scan the image from left to right
- Let p denote the pixel at any step in the scanning process.
- Let r denote the upper neighbor of p.
- Let t denote the left-hand neighbor of p, respectively.
- When we get to p, points r and t have already been encountered and labeled if they were 1 's.

Labeling of Connected Components

- If the value of p=0,move on.
- If the value of p=1,examine *r* and *t*.
  - If they are both 0, assign a new label to p.
  - If only one of them is 1,aasign its label to p.
  - If they are both 1
    - If they have the same label,assign that label to *p*.
    - If not, assign one of the labels to *p* and make a note that the two labels are equivalent.(*r* and *t* are connected through p).
- **At the end of the scan, all points with value 1 have been labeled.**
- **Do a second scan, assign a new label for each equivalent labels.**

What shall we do with 8-connected components?

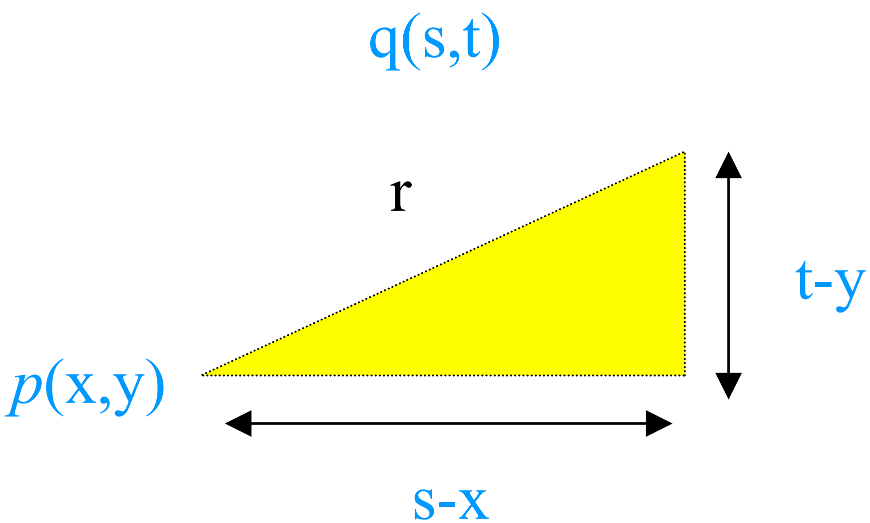
- Do the same way but examine also the upper diagonal neighbors of p.
  - If p is 0, move on.
  - If p is 1
    - If all four neighbors are 0,assign a new label to p.
    - If only one of the neighbors is 1,assign its label to p.
    - If two or more neighbors are 1, assign one of the label to p and make a note of equivalent classes.
    - After complete the scan, do the second round and introduce a unique label to each equivalent class.

Distance Measures

- For pixel *p*,*q* and *z* with coordinates (x,y), (s,t) and (u,v) respectively,
- D is a distance function or metric if
  - (a) $D(p,q) \geq 0$; $D(p,q) = 0$ iff $D=q$
  - (b) $D(p,q) = D(q,p)$
  - (c) $D(p,z) \leq D(p,q) + D(q,z)$

Euclidean distance between *p* and *q*

- $D_e(p,q)=[(x-s)^2+(y-t)^2]^{1/2}$



Radius (r) centered at (x,y)

City-block distance D_4 distance

- $D_4(p,q)=|x-s|+|y-t|$
- Diamond centered at (x,y)
- $D_4=1$ are 4-neighbors of (x,y).

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

Chessboard distance : D_8 distance

- $D_8(p,q) = \max(|x-s|,|y-t|)$

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

Square centered at (x,y)

D_4 and D_8 distance

- D_4 distance and D_8 distance between pixel *p* and *q* = length 4- and 8- path between those 2 points, respectively.
- We can consider both D_4 and D_8 distances b/w *p* and *q* regardless of whether a connected path exists between them because the definitions of these distances involve only the coordinates.

M-connectivity's distance

- Distance of m-connectivity of the path between 2 pixels depends on values of pixels along the path.
- e.g, if only connectivity of pixels valued 1 is allowed, find the m-distance b/w p and p_4

	P3	P4
P1	P2	
P		

	0	1
0	1	
1		

Distance = 2

	1	1
0	1	
1		

Distance =3

	1	1
1	1	
1		

Distance = 4

Arithmetic Operators

- Used extensively in most branches of image processing.
- Arithmetic operations b/w 2 pixels *p* and *q* :
  - Addition : p+q used in image average to reduce noise.
  - Subtraction: p-q basic tool in medical imaging.
  - Multiplication: p x q
    - To correct gray-level shading result from non-uniformities in illumination or in the sensor used to acquire the image.
  - Division : p ÷ q
- Arithmetic operation entire images are carried out pixel by pixel.

Logic operations

- AND : p AND q (p.q)
- OR : p OR q (p+q)
- COMPLEMENT : NOT q (q)
- Logic operations apply only to binary images.
- Arithmetic operations apply to multi-valued pixels
- Logic operations used for tasks such as masking, feature detection, and shape analysis.
- Logic operations perform pixel by pixel.

Mask Operation

- Besides pixel-by-pixel processing on entire images, arithmetic and logical operations are used in neighborhood oriented operations.

		⋮		
	Z_1	Z_2	Z_3	
…	Z_4	Z_5	Z_6	…
	Z_7	Z_8	Z_9	
		⋮		

Mask Operation

- Let the value assigned to a pixel be a function of its gray level and the gray level of its neighbors.
- e.g., replace the gray value of pixel Z_5 with the average gray values of it's neighborhood within a 3 x 3 mask.

$$Z=1/9(Z_1+Z_2+Z_3+\ldots+Z_9)$$

Mask operator

- In general term:

$$Z = 1/9\ Z_1 + 1/9\ Z_2 + 1/9\ Z_3 + \ldots.. + 1/9\ Z_9$$

$$= w_1Z_1 + w_2Z_2 + w_3Z_3 + \ldots.. + w_9Z_9$$

$$= \sum_{i=1}^{9} w_i z_i$$

W1	W2	W3
W4	W5	W6
W7	W8	w9

$\longrightarrow$

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Mask coefficient

- Proper selection of the coefficients and application of the mask at each pixel position in an image makes possible a variety of useful image operations
  - Noise reduction
  - Region thinning
  - Edge detection
- Applying a mask at each pixel location in an image is a computationally expensive task.

Image Geometry

- Basic transformations : expressed in 3D Cartesian coordinate system (x,y,z)
  - Translation
  - Scaling
  - Rotation
  - Concatenation and inverse transformation

Geometry General Form

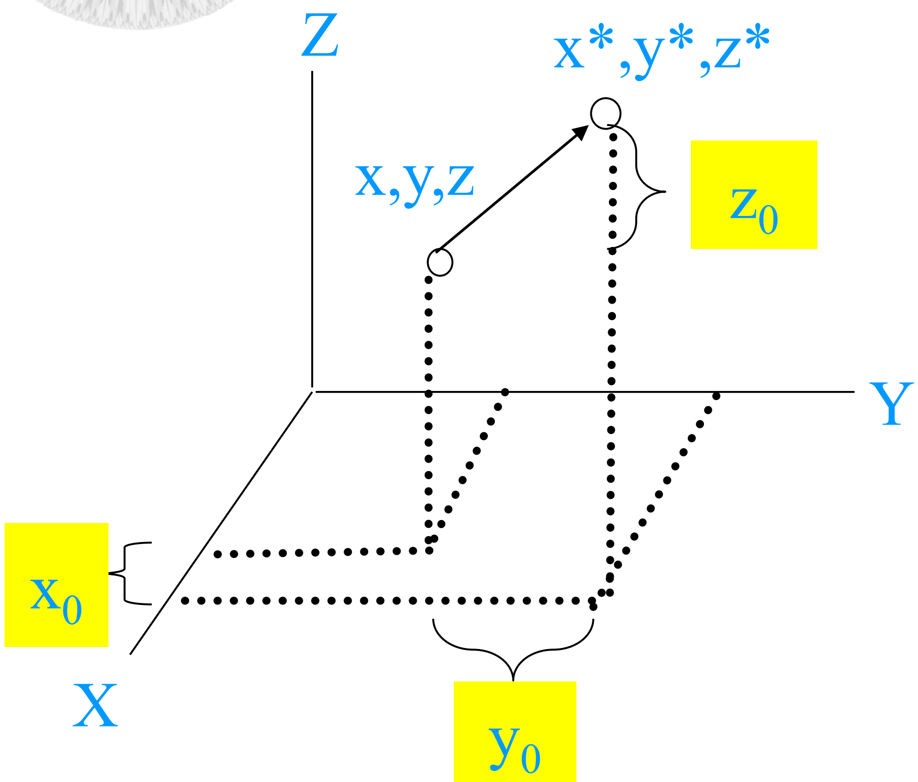
$$V^* = AV$$

A : 4x4 transformation matrix

V : column vector containing the original coordinates

V* : column vector whose components are the transformed coordinates

Translation



$$X^* = X + X_0$$
$$Y^* = Y + Y_0$$
$$Z^* = Z + Z_0$$

$$\begin{bmatrix} X^* \\ Y^* \\ Z^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{x} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Translation matrix

$$T = \begin{bmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$V^* = TV$$

Scaling

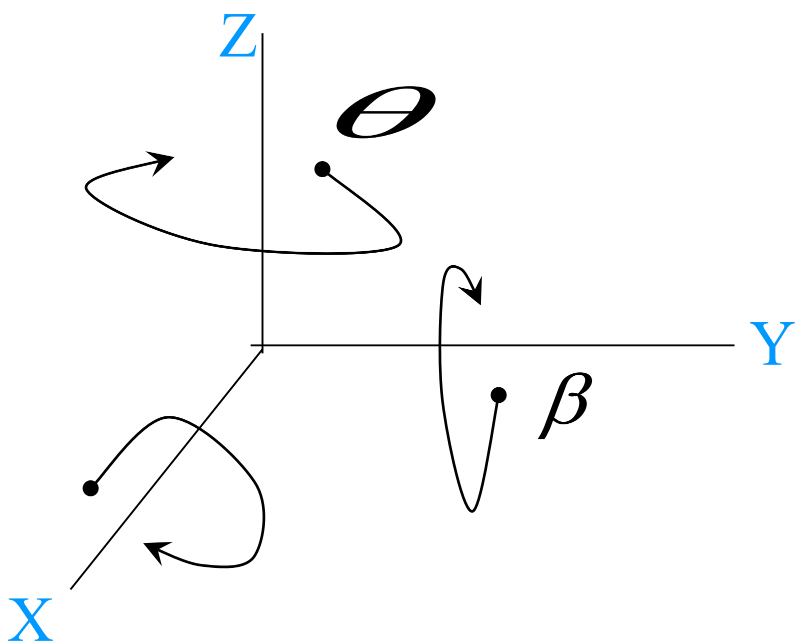
- Scaling by factors S_x, S_y and S_z along the X,Y,Z axes

$$s = \begin{bmatrix} Sx & 0 & 0 & 0 \\ 0 & Sy & 0 & 0 \\ 0 & 0 & Sz & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$X^* = S_x X$
$Y^* = S_y Y$
$Z^* = S_z Z$

Rotation in Z axis

- Rotation of a point about Z axis by an angle θ (clockwise).



$$x^* = x\cos\theta + y\sin\theta$$

$$y^* = -x\sin\theta + y\cos\theta$$

$$Z^* = z$$

$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Rotation in X axis

- Rotation of a point about X axis by an angle α

$$R_{\alpha} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Rotation in Y Axis

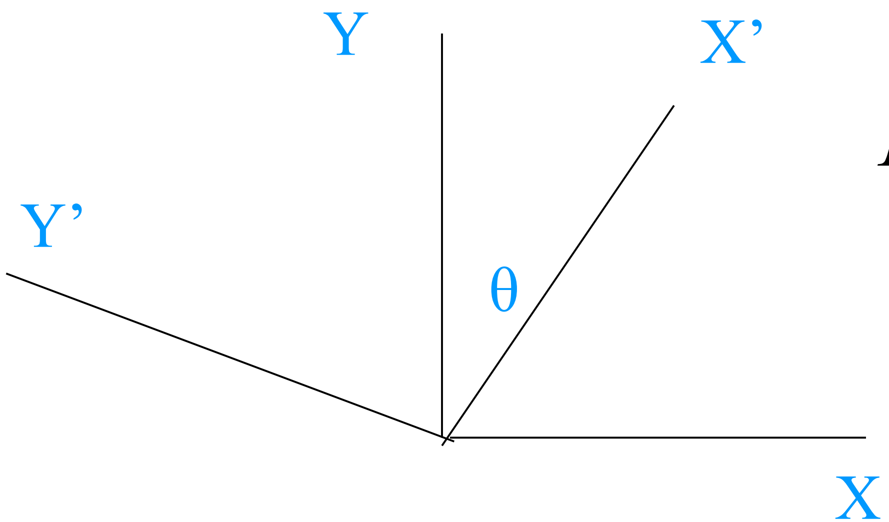
- Rotation of a point about Y axis by an angle β

$$R_{\beta} = \begin{bmatrix} \cos\beta & 0 & -\sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

If we rotate of a point about one axis, it will affect on the coordination of the other 2 axes.

Exercise

- Rotate a point about Z axis.



$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Concatenation & Inverse Transformation

- Several transformation can be represented by a single 4x4 transformation matrix.
- e.g., translation, scaling, and rotation about the Z axis of a point V is given by

$$
\begin{aligned}
V^* &= \mathrm{R}_\theta(\mathrm{S}(\mathrm{TV})) \\
&= \mathrm{AV}
\end{aligned}
$$