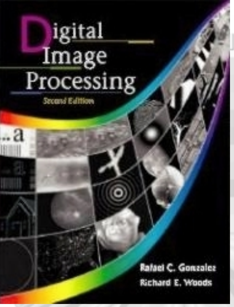


Chapter 2: Digital Image Fundamentals

- Digital image processing is based on
  - Mathematical and probabilistic models
  - Human intuition and analysis
    - Visual perceptions
    - Resolution (judgement)
    - Adaptation

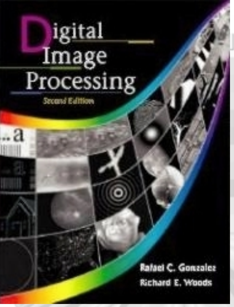
2.1 Visual Perception

- Observing the animal activity → New inventions, *i.e.*, airplane, submarine, etc.
- Image processing/Computer Vision

  How images are formed in the eye ?

  Eye's physical limitation ?

  Human visual interpretation of images ?



2.1.1 Structure of human eyes

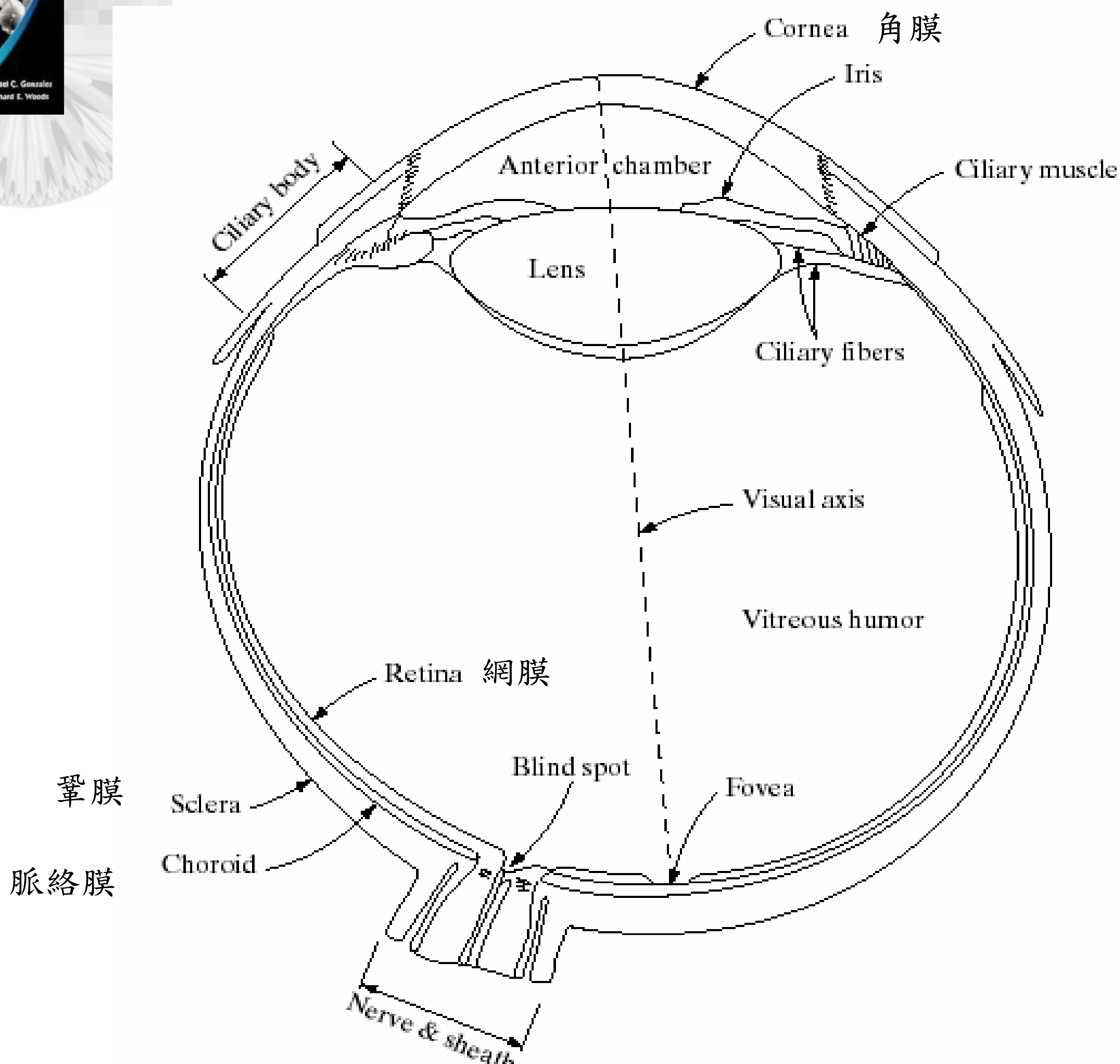
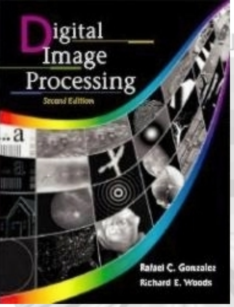


FIGURE 2.1 Simplified diagram of a cross section of the human eye.



2.1.1 Structure of human eyes

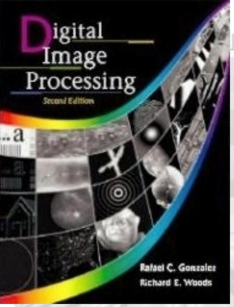
- Three membranes enclose the eye:
  - Cornea (角膜) and sclera(鞏膜)
    - Cornea is a tough, transparent tissue cover the anterior surface of the eye.
    - Sclera is a opaque membrane enclose the remainder of the optic globe.
  - Choroid(脈絡膜)
    - A network of blood vessels for eye nutrition
    - At its anterior extreme, it is divided into the *ciliary body* and *iris diaphragm*.
    - The central opening (the *pupil*) varies in diameter from 2 to 8 mm.
  - Retina (網膜)

2.1.1 Structure of human eyes

- Lens is made of concentric layer of fibrous cells and is suspended by fiber that attached to the ciliary body.
- The lens absorbs approximately 8% of the visible light spectrum.
- The lens contains 60-70% water and 6% fat and protein.

2.1.1 Structure of human eyes

- Retina lines the insides of the wall's interior portion with two classes of receptors:
  - Cones:
    - 6 – 7 millions located primarily in the central portion of the retina
    - Highly sensitive to color
    - Photopic or bright-light vision
  - Rods
    - 75- 150 millions distributed over the retinal surface.
    - Not involved in color vision and sensitive to low-illumination.
    - Scoptopic or dim vision



2.1.1 Structure of human eyes

- Receptor density is measured in degrees from the fovea (fig. 2.2).
- The cones are most dense in the center of retina.
- Fovea : a sensor array of size 1.5mm×1.5mm

  Density of cones in the area of fovea is 150,000 element/mm^2.

  The number of cones in fovea is 337,000 elements.

2.1.1 Structure of human eyes

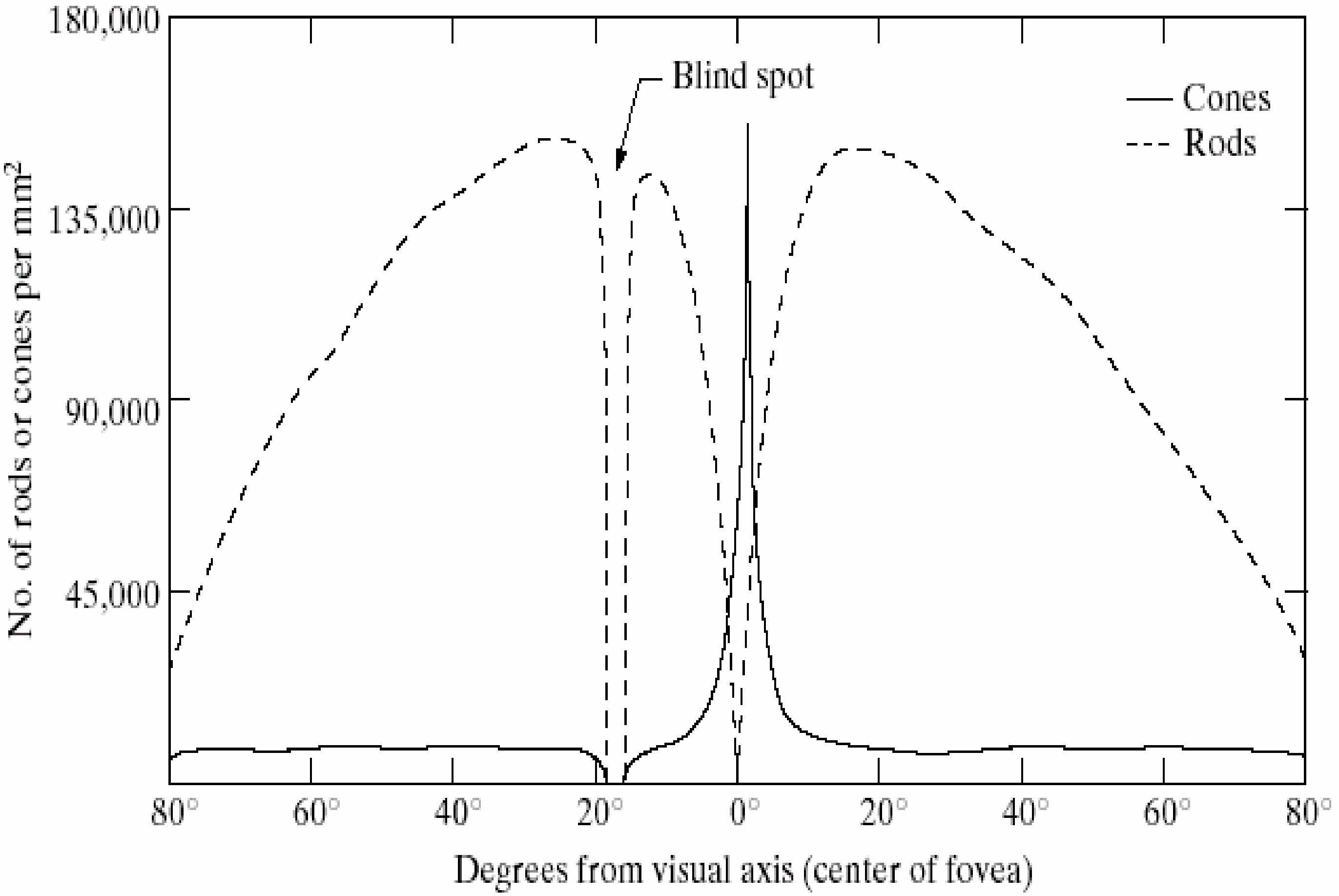
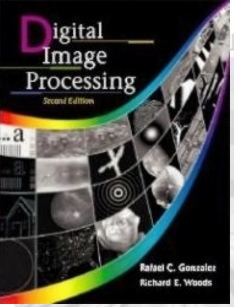


FIGURE 2.2 Distribution of rods and cones in the retina.



2.1.2 Image Formation in the Eyes

- The distance between the center of the lens and the retina (focal length) varies from 17mm to 14mm.
- The shape of lens is controlled by the tension of fibers of the ciliary body.
- The retinal image is reflected primarily in the area of fovea.
- Perception = excitation of light receptors, which transform radiant energy into electrical impulses that are ultimately decoded by the brain.

2.1.3 Brightness adaptation and discrimination

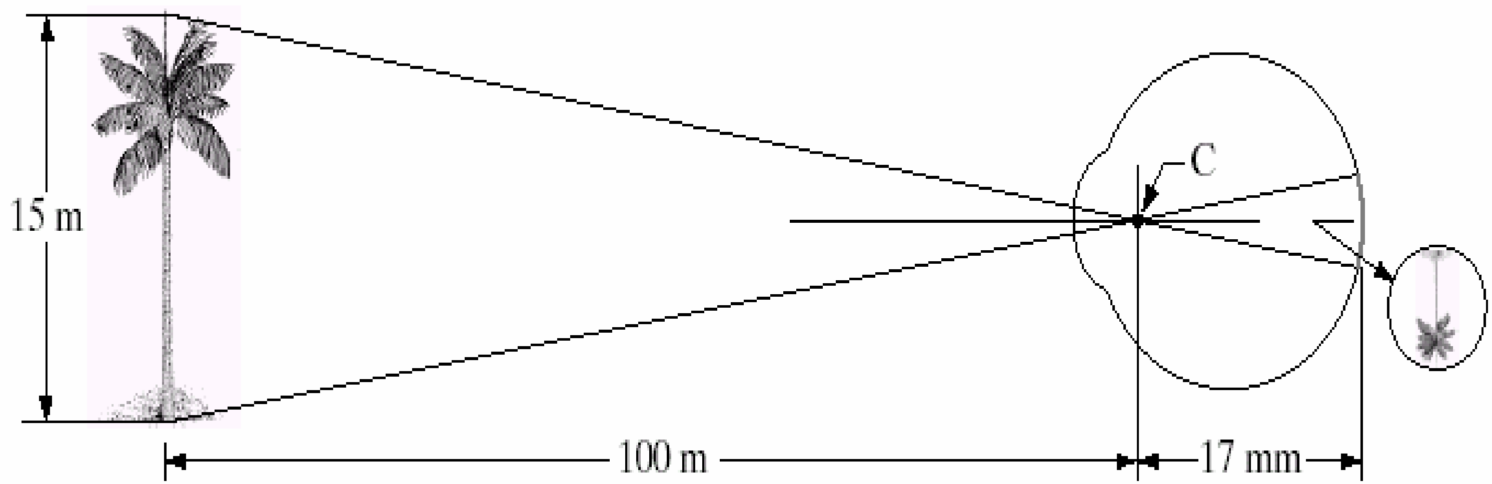
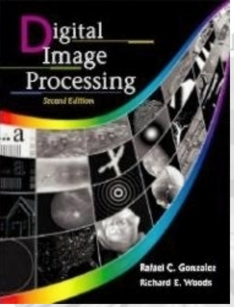


FIGURE 2.3 Graphical representation of the eye looking at a palm tree. Point *C* is the optical center of the lens.

2.1.3 Brightness adaptation and discrimination

- The range of light intensity levels to which the human visual system can adapt is enormous – on the order of 10^{10}.
- The ***subjective brightness*** is a logarithmic function of light intensity incident on the eye.
- In photopic vision, the range is about 10^{6}.
- Brightness adaptation.
  - The current sensitivity level it can discriminate simultaneously is rather small compared with the total adaptation range
  - ***Brightness adaptation level***: the current sensitive level of the visual system.



2.1.3 Brightness adaptation and discrimination

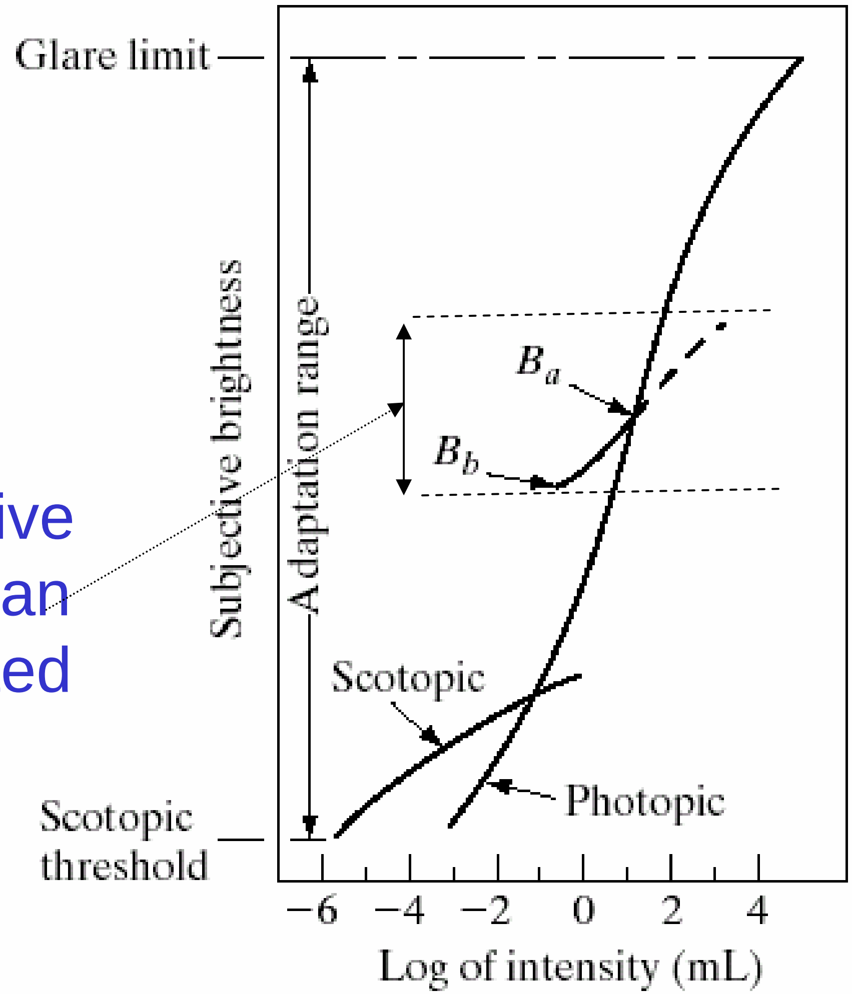


FIGURE 2.4 Range of subjective brightness sensations showing a particular adaptation level.

The range of subjective brightness that eye can perceive when adapted to this level

2.1.3 Brightness adaptation and discrimination

Experiments: Apply a short–duration flash at a circle to see if ΔI is bright enough

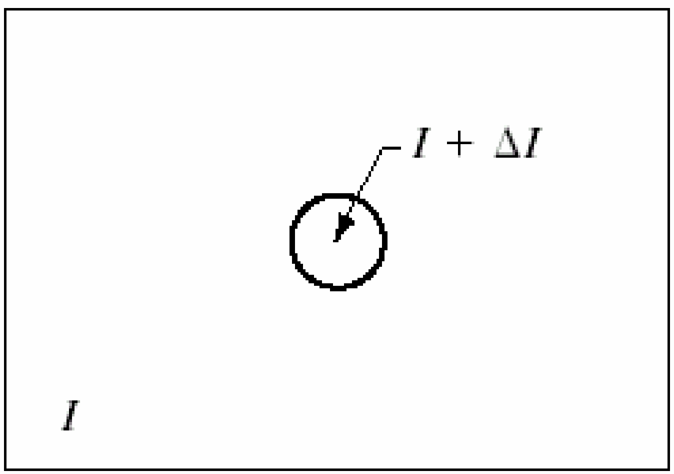
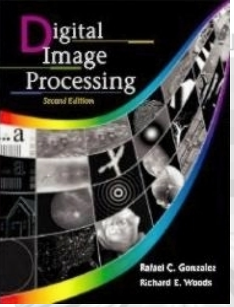


FIGURE 2.5 Basic experimental setup used to characterize brightness discrimination.

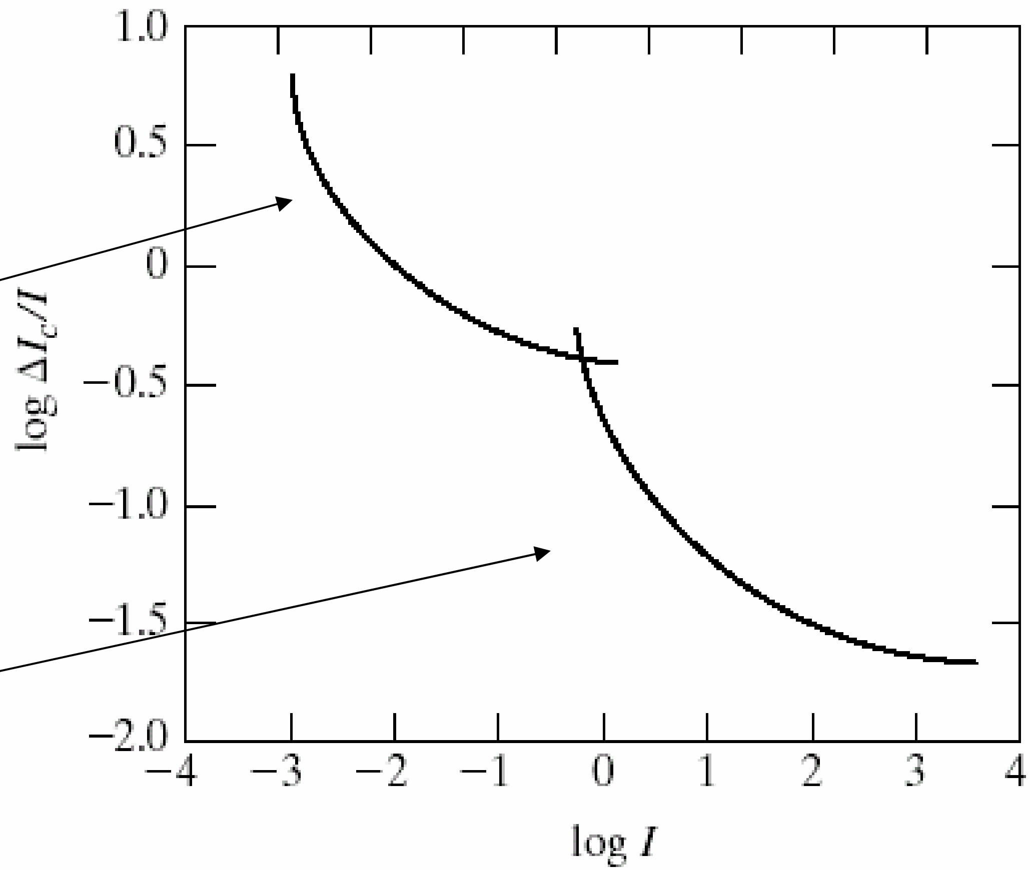


2.1.3 Brightness adaptation and discrimination

- The ΔI_c is the increment of illumination discriminable 50% of the time with the background illumination I.
- The quantity $\Delta I_c/I$ is called the Weber ratio.
- The ***smaller*** $\Delta I_c/I$ means that a small percentage change in intensity is discriminable – ***good brightness discrimination***
- If the background illumination is constant, the intensity of object is allowed to vary incrementally from never perceived to always being perceived.
- Typically the observer can discern a totally from one to two dozens different intensity changes.
  - The number of gray level for digital image
  - *Contouring effect* - not sufficient number of gray level.

2.1.3 Brightness adaptation and discrimination

FIGURE 2.6
Typical Weber ratio as a function of intensity.



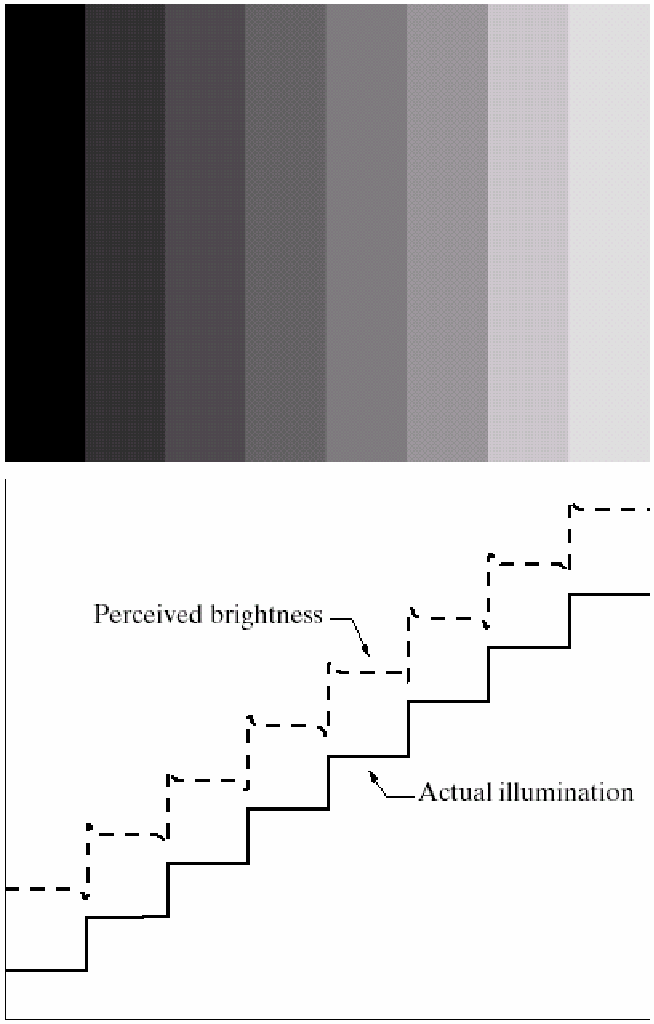
Low-level illumination vision (rod cells)

High-level illumination vision (cone cells) (better discrimination)

2.1.3 Brightness adaptation and discrimination

- ***Perceived brightness*** is not a simple function of ***intensity***
  - Visual system tends to overshoot or undershoot around the boundary of regions of different intensity (Fig. 2.7) – ***March bands effect.***
  - A region's perceived brightness does not simply depend on its intensity *(fig. 2.8)* – ***Simultaneous contrast***.

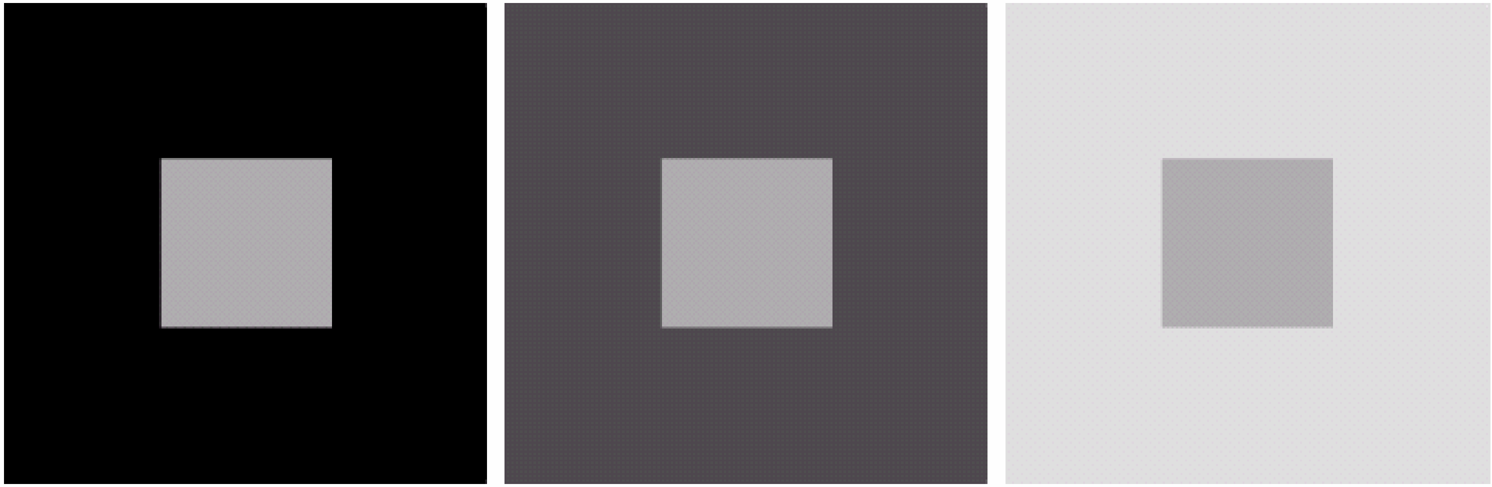
2.1.3 Brightness adaptation and discrimination



a
b

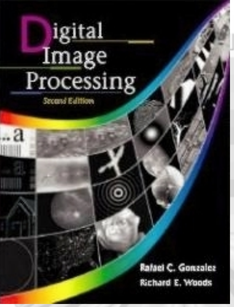
FIGURE 2.7 (a) An example showing that perceived brightness is not a simple function of intensity. The relative vertical positions between the two profiles in (b) have no special significance; they were chosen for clarity.

2.1.3 Brightness adaptation and discrimination

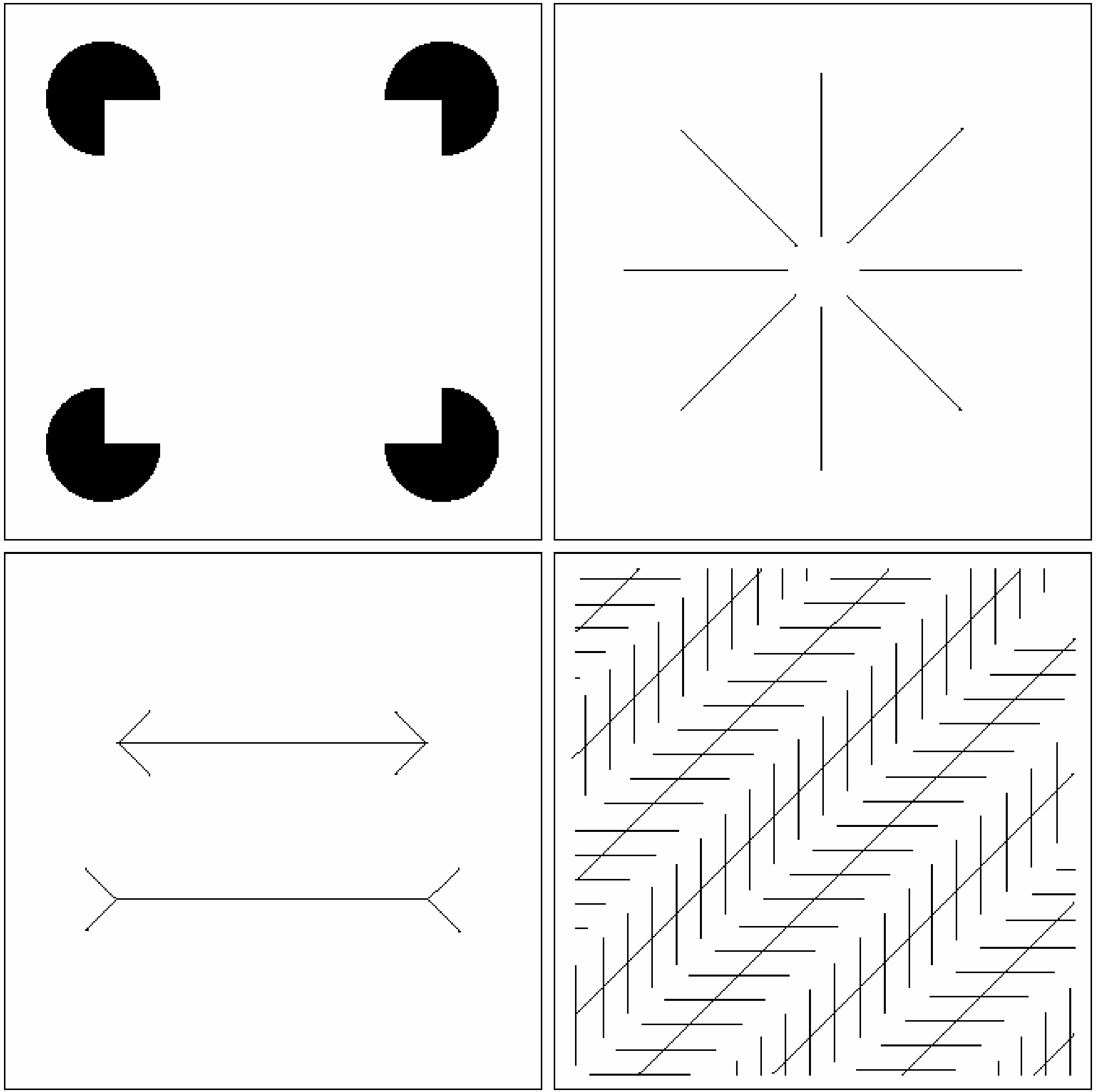


a b c

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.



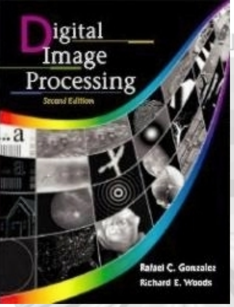
2.1.3 Brightness adaptation and discrimination



a b
c d

FIGURE 2.9 Some well-known optical illusions.

Optical illusion: perceive nonexisting or error information.



2.14 Light and the EM Spectrum

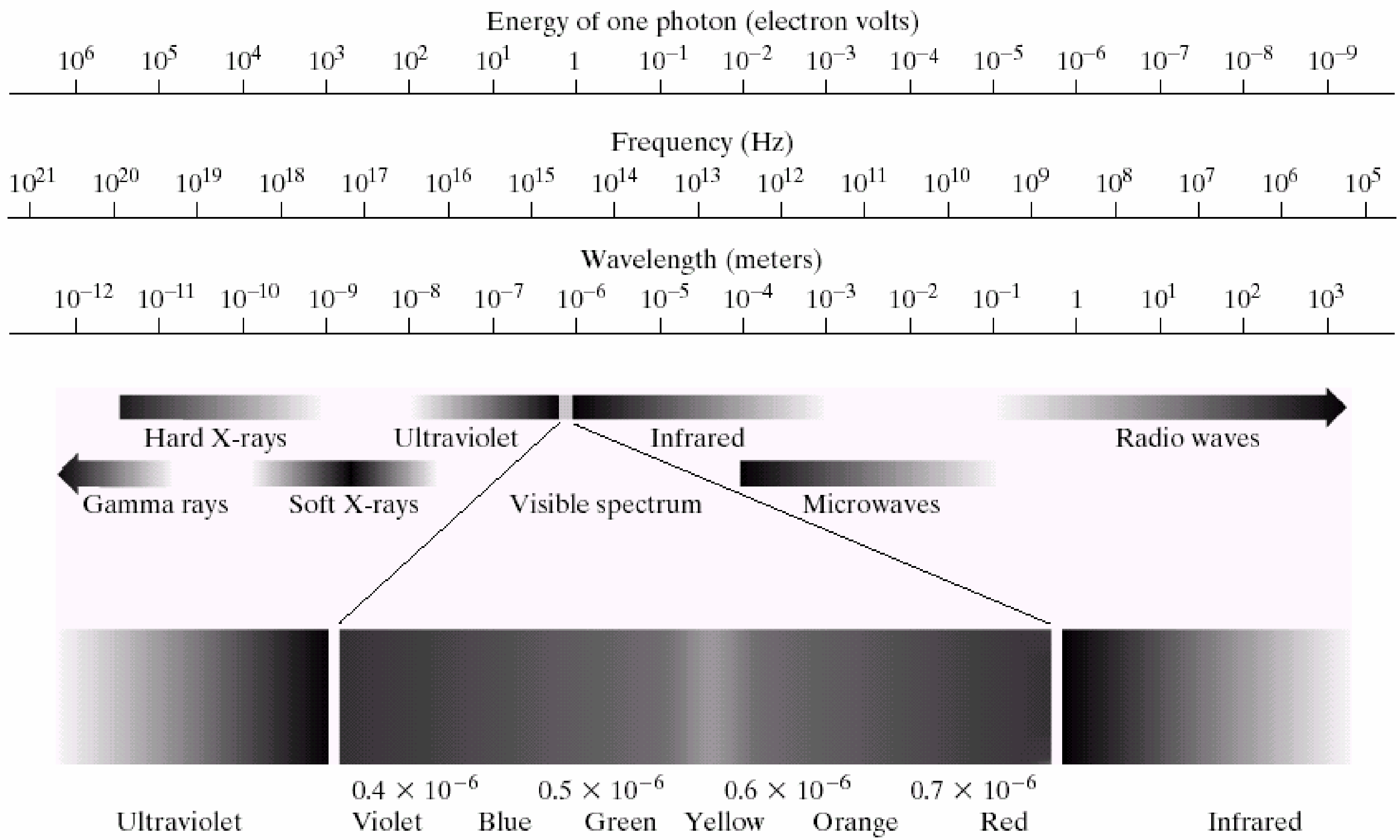
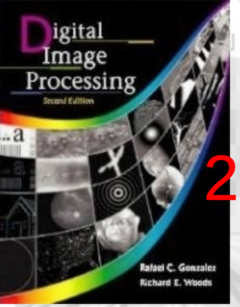


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.



2.1.3 Brightness adaptation and discrimination

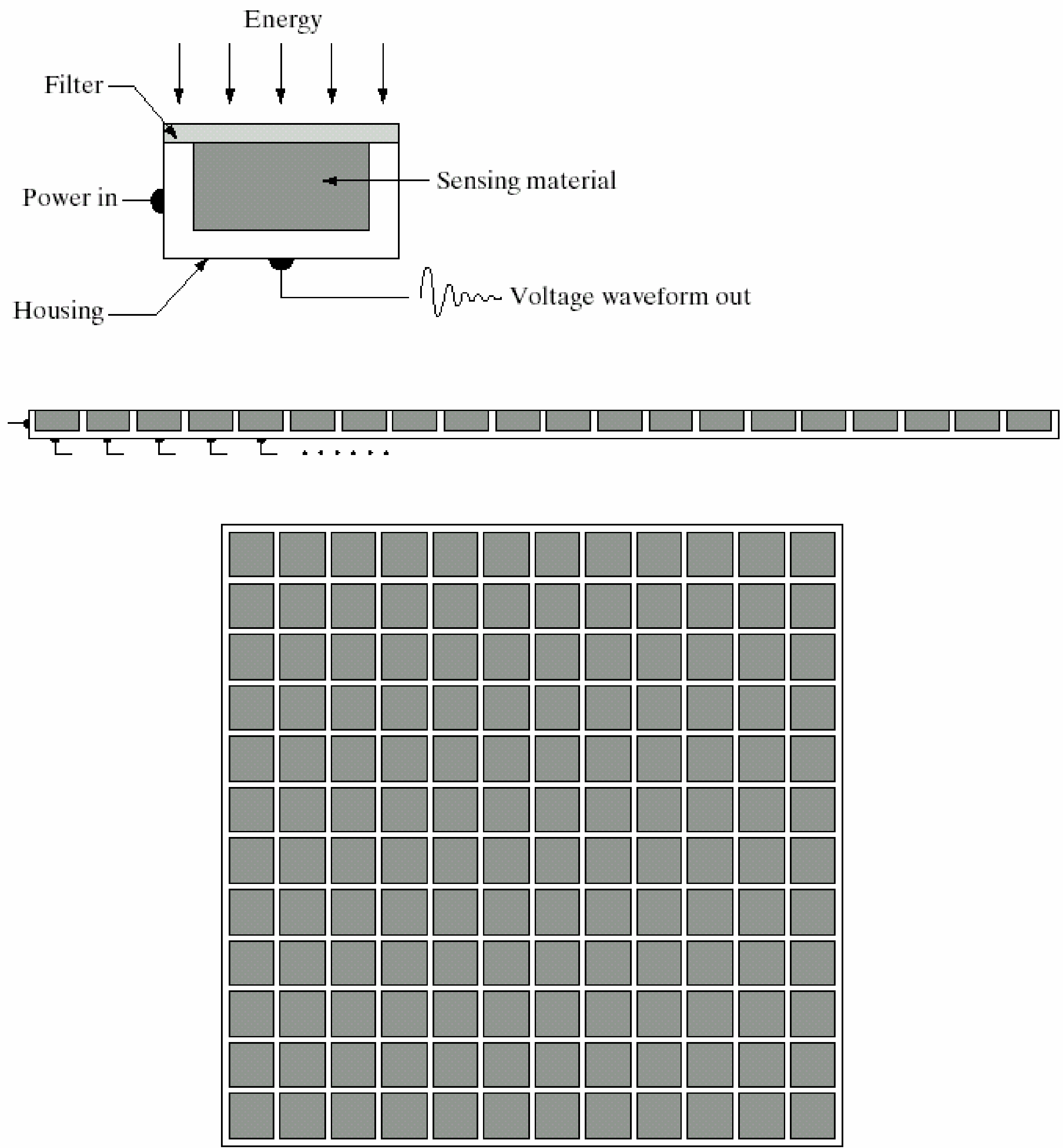
- Light is a particular type of EM radiation that can be seen by human eye.
- Green object reflect light with wavelengths primarily in 500 to 570 nm range.
- ***Chromatic light*** spans EM spectrum from 0.43 μm (violet) to 0.79 μm (red)
  - Radiance: energy in Watt
  - Luminance: in lumens(lm) the amount of energy the observer perceives
  - Brightness: subjective description of light perception.

2.3 Image Sensing and Acquisition

- Image = illumination + scene
  - A visible light source illuminates a 3-D scene.
  - Illumination originate from
    - Conventional EM source, infrared, X-ray, Ultrasound.
    - Computer-generated illumination pattern
- Photo-converter (phosphor screen), convert the EM energy into visible light.

a
b
c

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



2.3.1 A single sensor

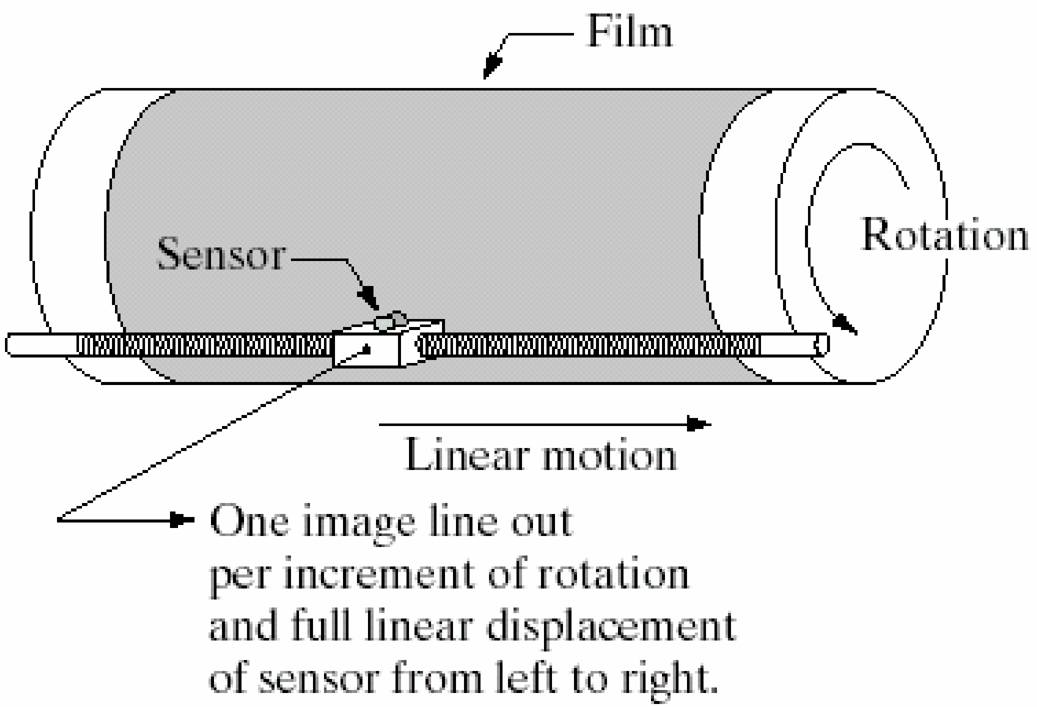
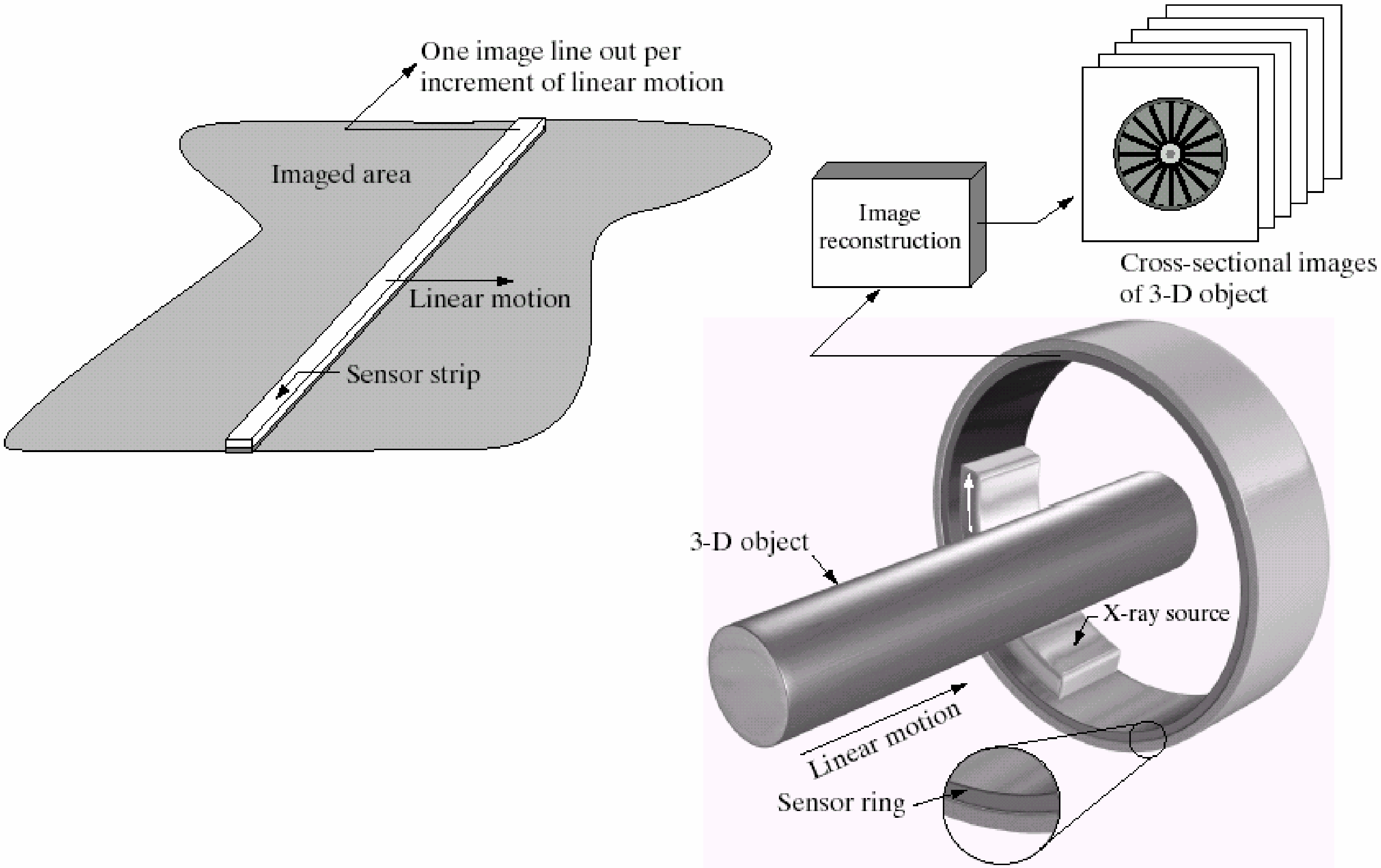


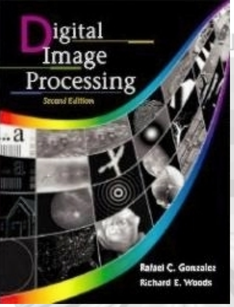
FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

2.3.1 A sensor strip



a b

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



2.3.3 A sensor array

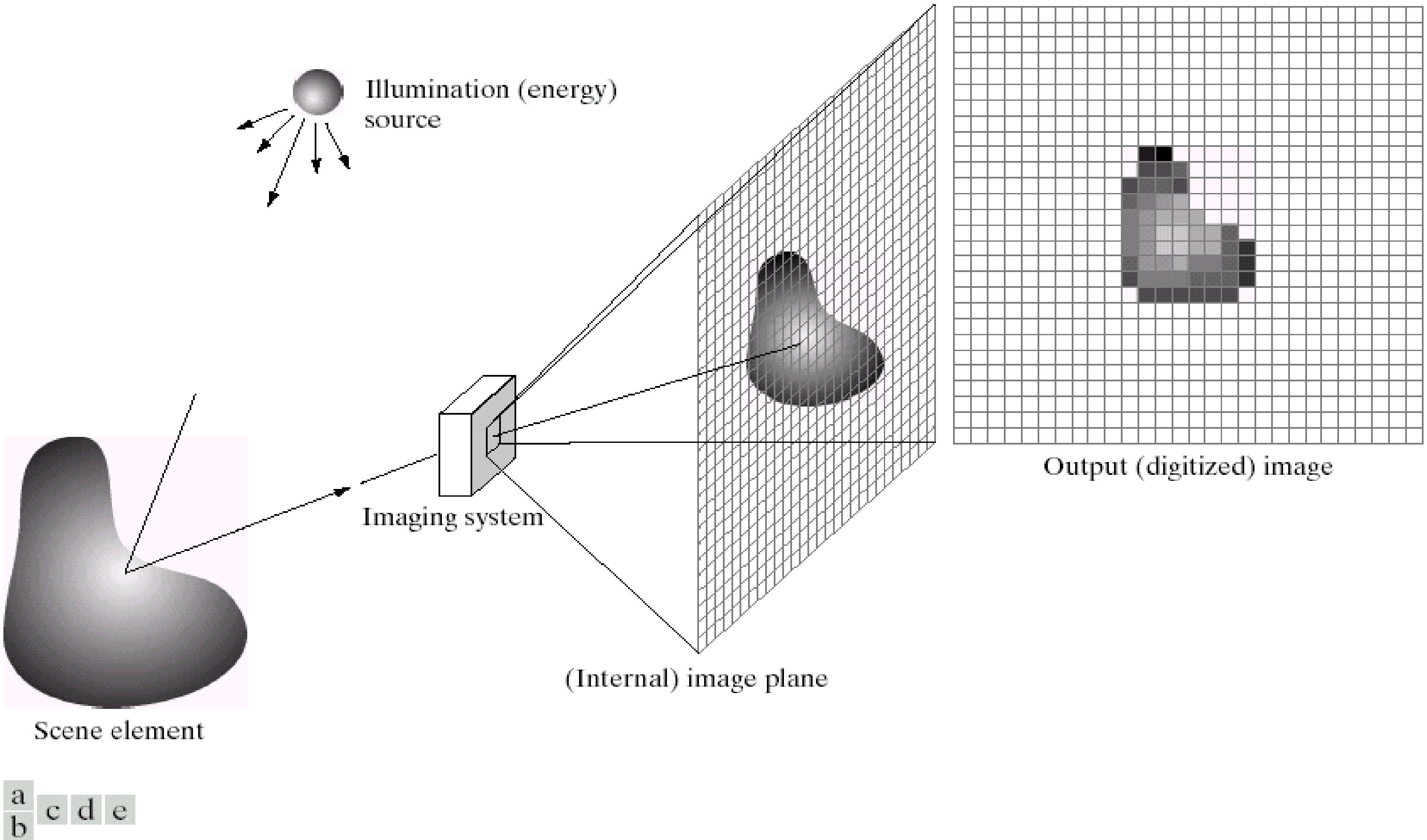
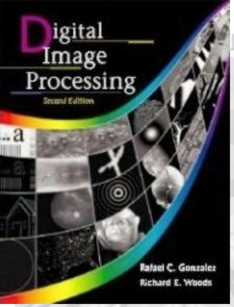


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

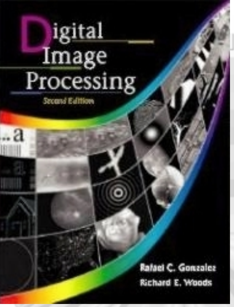


2.3.4 Image formation model

- For monochromatic image 2-D array: $f(x, y)$
- The $f(x, y)$ is characterized by two components:
  - The amount of source illumination incident on the scene, i.e., $i(x,y)$.
  - The amount of illumination reflected by the objects in the scene, *i.e.*, ***reflectivity*** $r(x, y)$.
- $f(x, y)=i(x, y)\, r(x, y)$

  where $0 < i(x, y) < \infty$ and $0 < r(x, y) < 1$
- Reflectivity function: $r(x, y)$
- For X-ray, ***transmissivity*** function
- The ***intensity*** of monochrome image is

  $L_{min} \le f(x,y) \le L_{max}$ $L_{min} = i_{min} r_{min}$ and $L_{max} = i_{max} r_{max}$
- Indoor: $L_{min}=10$ and $L_{max}=1000$



2.4 Image Sampling and Quantization

- To acquire digital image from the continuous sensed data $f(x, y)$:
  - Scanning 2-D image to generate 1-D signal
  - Digitization in coordinate values: Sampling
  - Digitization in amplitude values: Quantization.

2.4 Image Sampling and Quantization

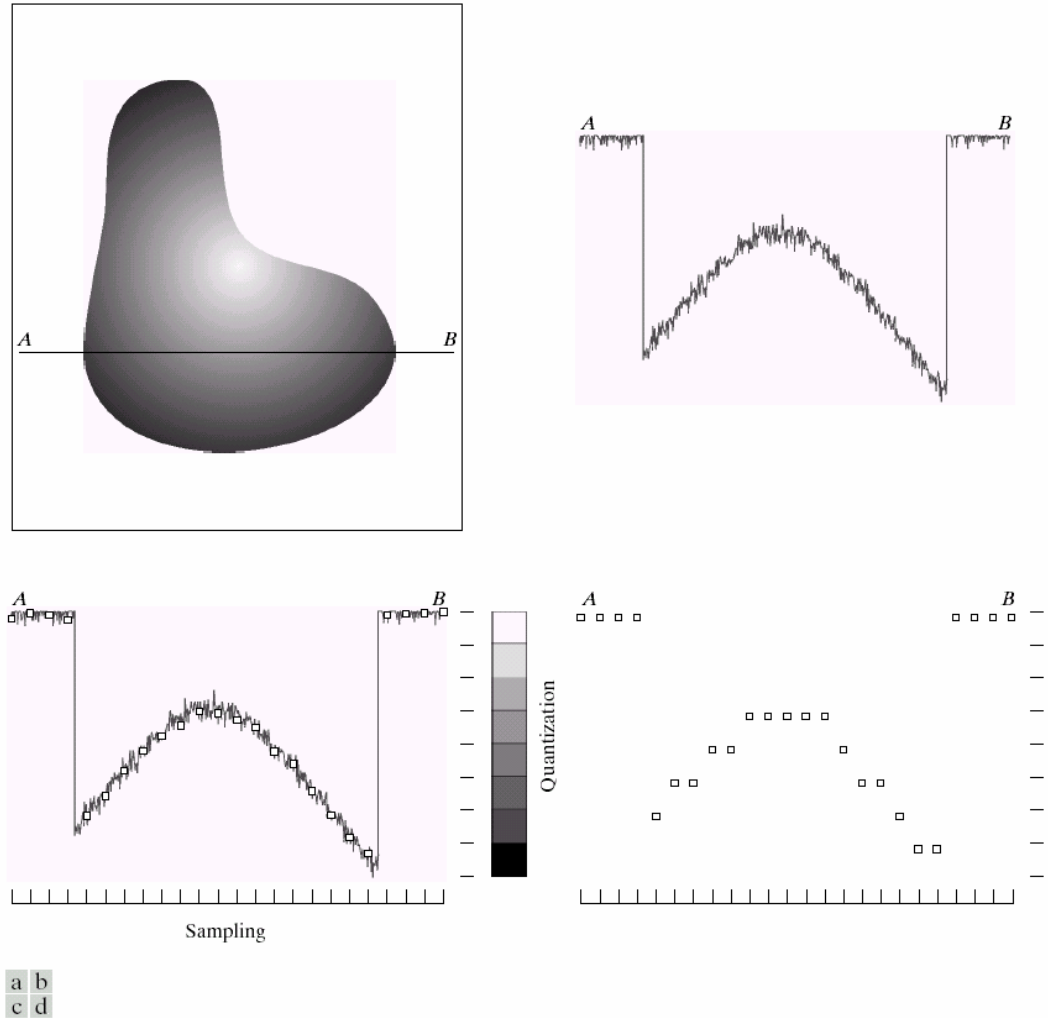
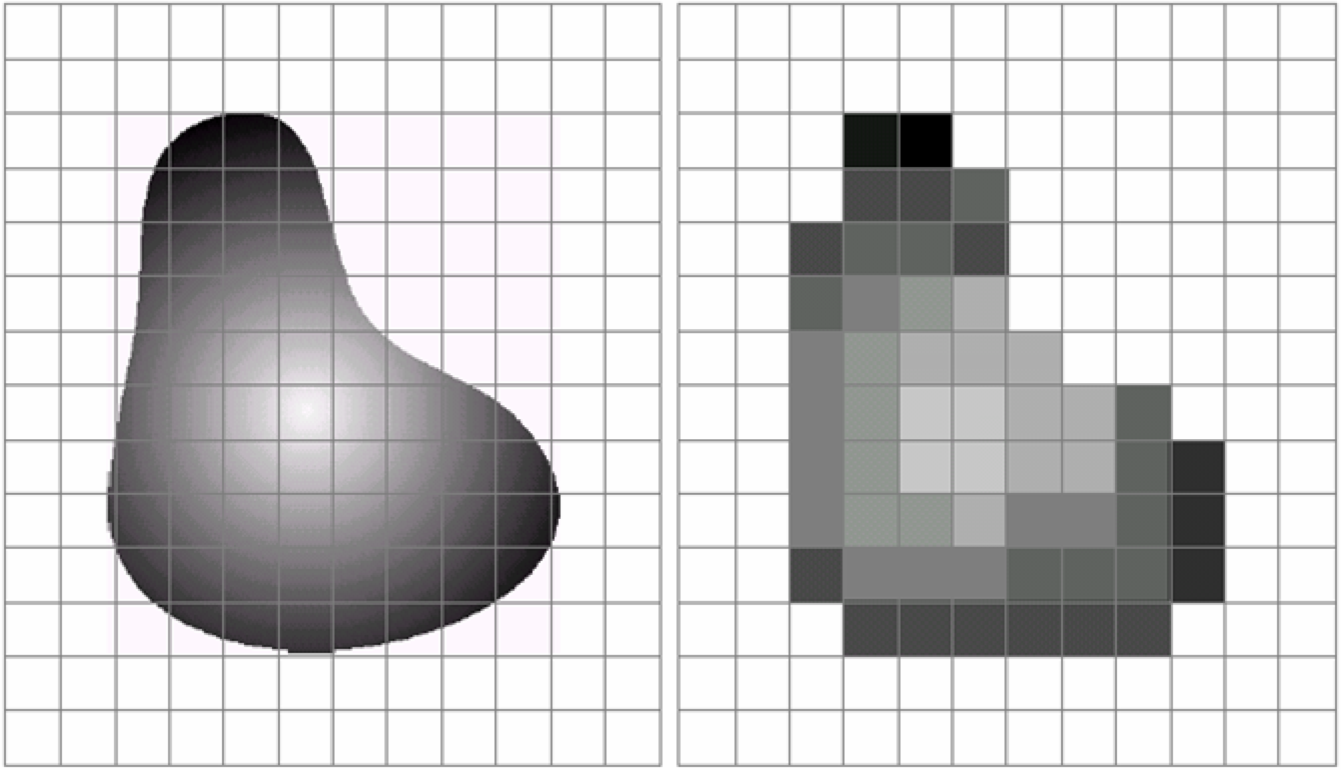


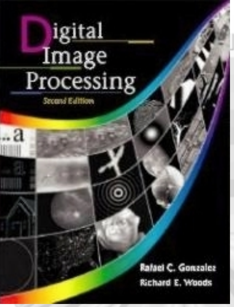
FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

2.4 Image Sampling and Quantization



a b

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



2.4.2 Representing Digital Images

- The resulting image is a 2-D array with M rows and N columns.

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & & f(0,N-1) \\ f(1,0) & f(1,1) & & f(1,N-1) \\ & & & \\ f(M-1,0) & f(M-1,1) & & f(M-1,N-1) \end{bmatrix}$$

- Each element of this matrix is called an image element, picture element, pixel, or pel.

2.4.2 Representing Digital Images

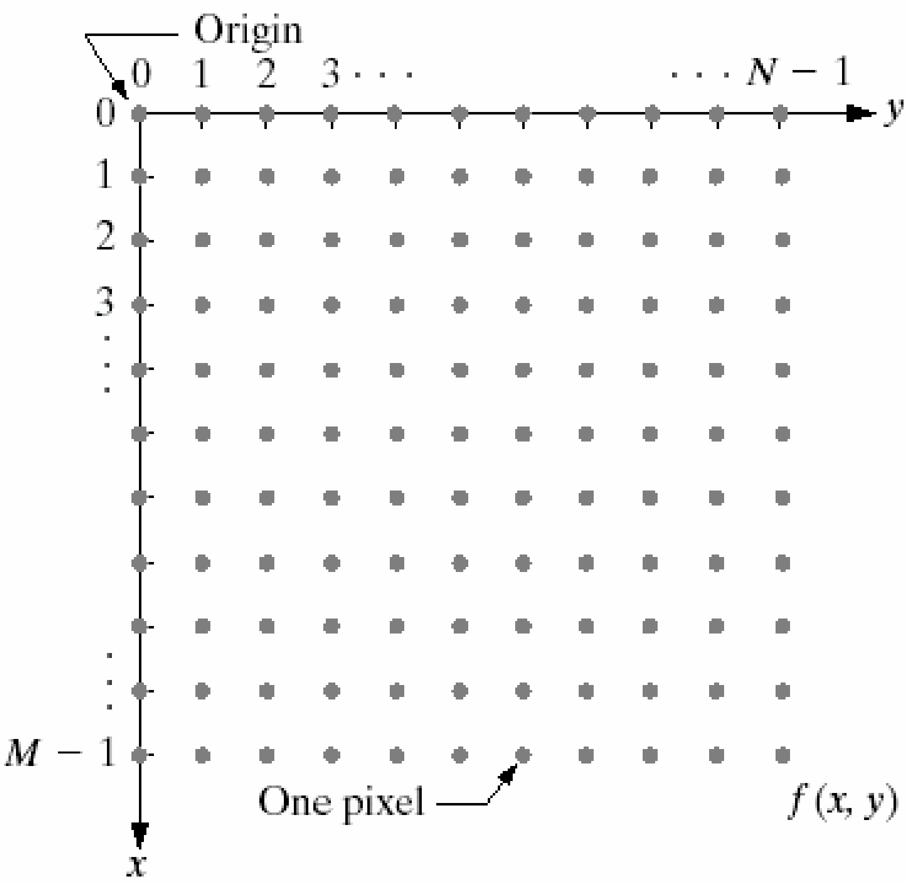
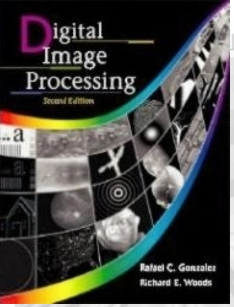


FIGURE 2.18 Coordinate convention used in this book to represent digital images.



2.4 Spatial and Gray-Level resolution

- The digitization process requires to determine the *M*, *N*, and *L*
- *M* and *N* spatial resolution
- *L* gray-level resolution

  $L=2^k$. L=gray-level

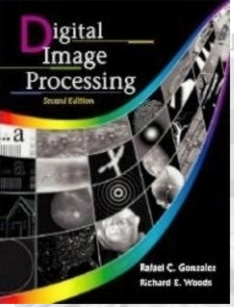
- The number of bits required to store the image

  $b=M\times N\times k$ or $b=N^2\times k$

2.4 Image Sampling and Quantization

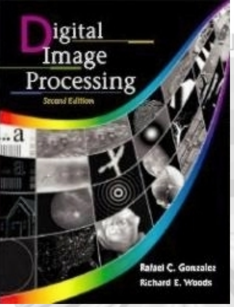
TABLE 2.1
Number of storage bits for various values of N and k.

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



2.4.3 Spatial and Gray-Level resolution

- Sampling → Spatial resolution
- Quantization → Gray-level resolution
- Trade-off between resolution and total bit numbers.
- If total number of bits fixed → how to adjust the trade-off between spatial resolution and gray-level resolution



2.4.3 Spatial and Gray-Level resolution

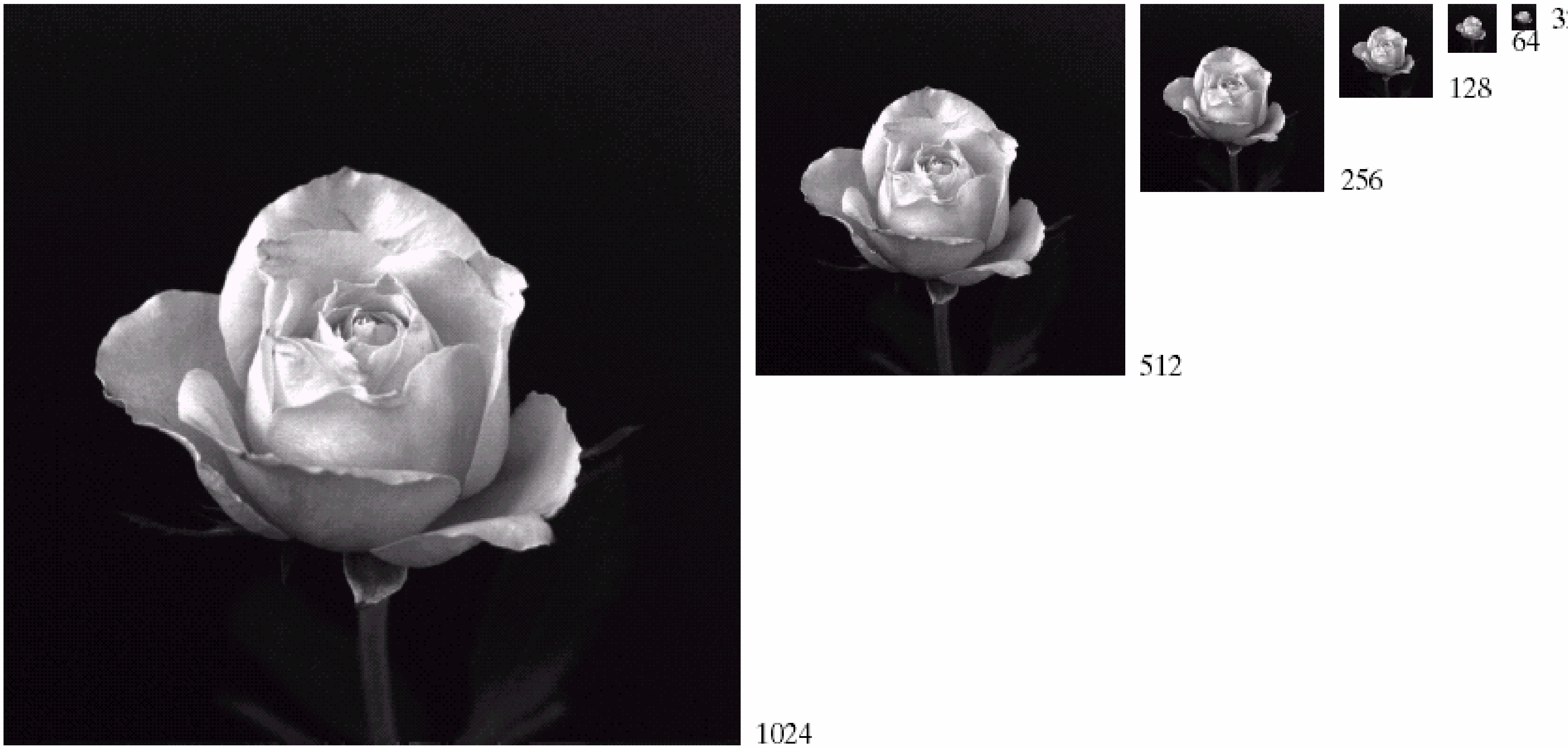
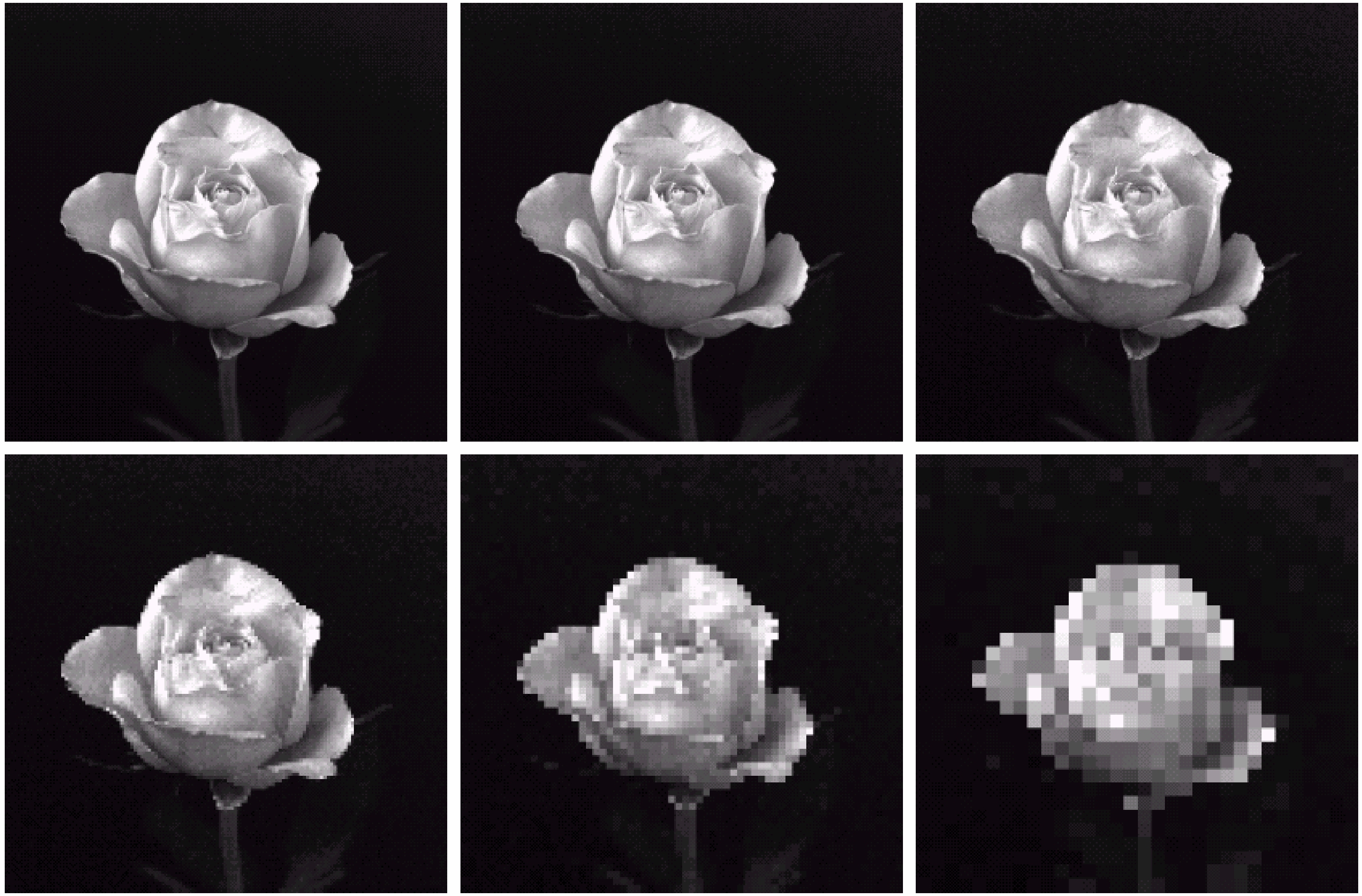


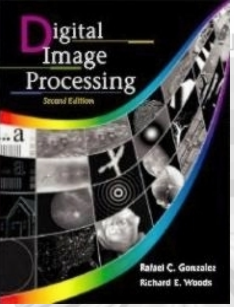
FIGURE 2.19 A 1024 × 1024, 8-bit image subsampled down to size 32 × 32 pixels. The number of allowable gray levels was kept at 256.

2.4.3 Spatial and Gray-Level resolution

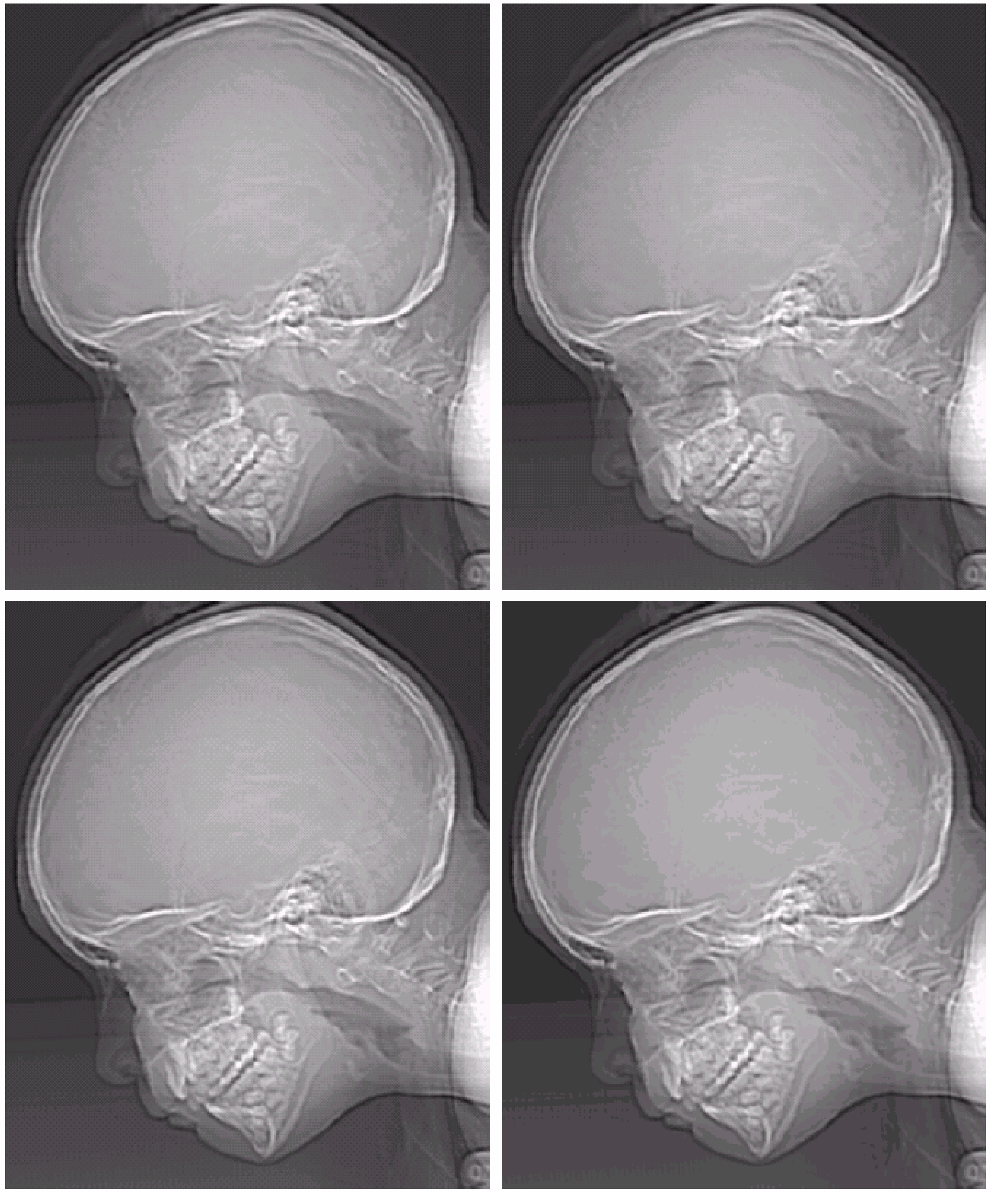


a b c
d e f

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.

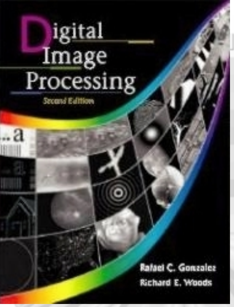


2.4.3 Spatial and Gray-Level resolution

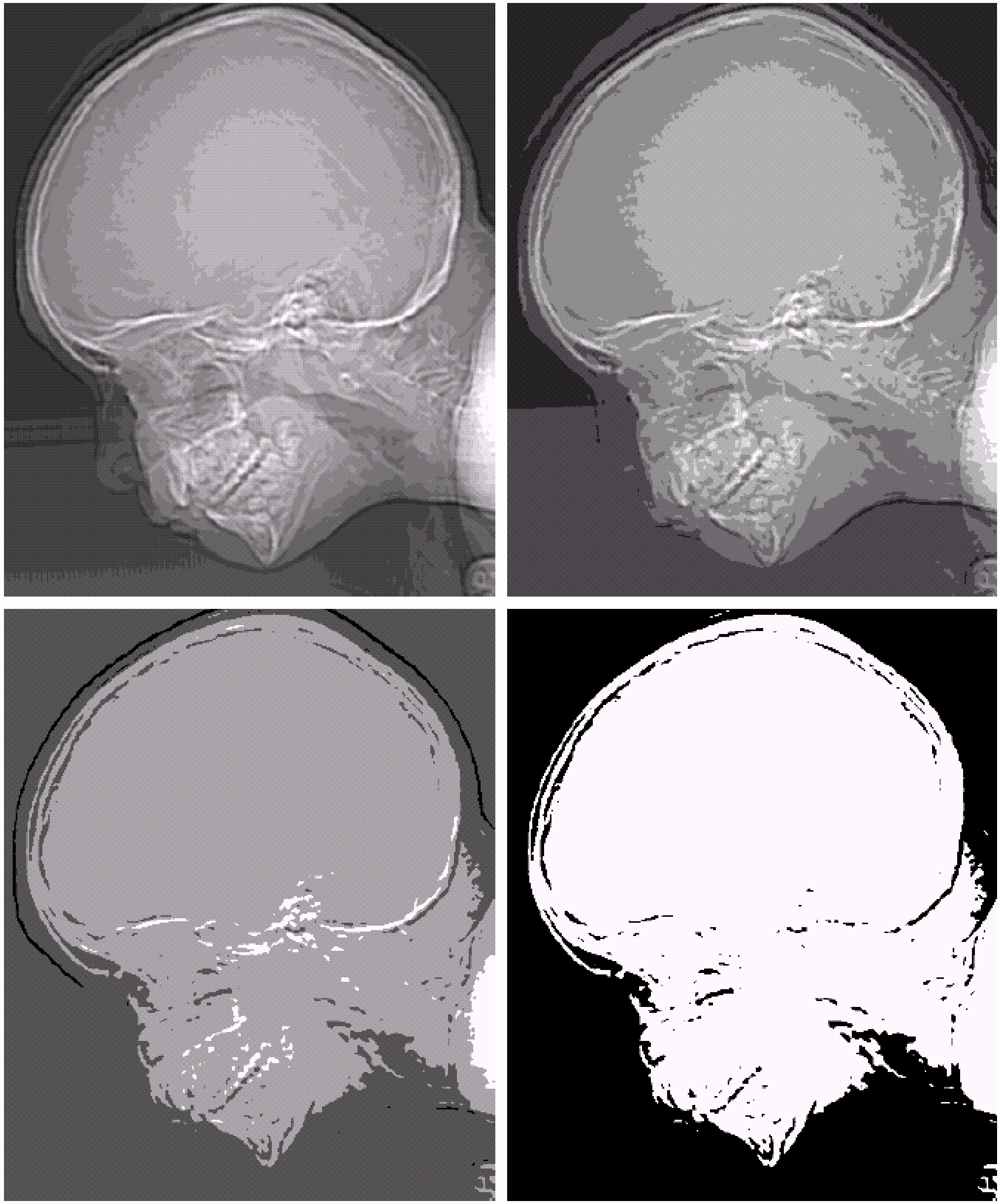


a b
c d

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.



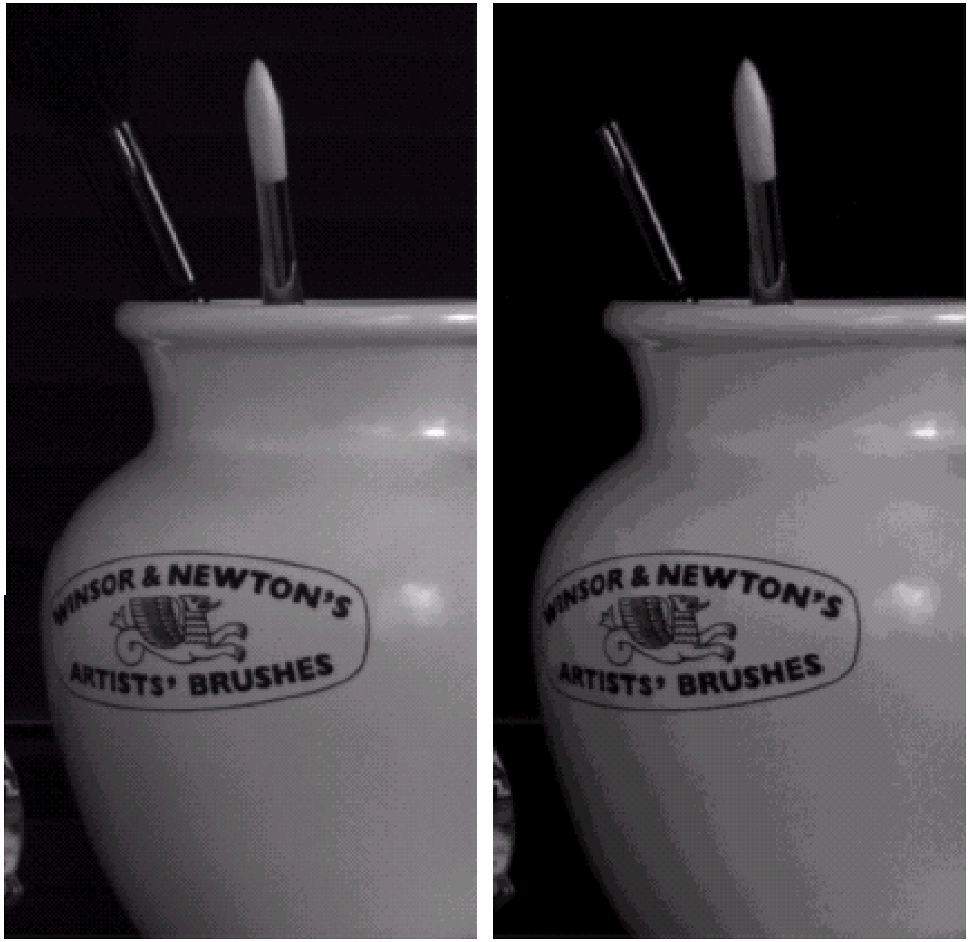
2.4.3 Spatial and Gray-Level resolution



e f
g h

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.)

2.4.3 Spatial and Gray-Level resolution

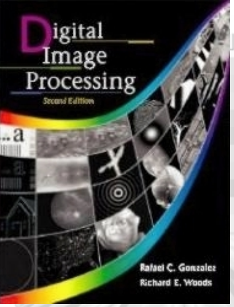


Contouring defect

2.4.3 Spatial and Gray-Level resolution

a b c

FIGURE 2.22 (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail. (Image (b) courtesy of the Massachusetts Institute of Technology.)

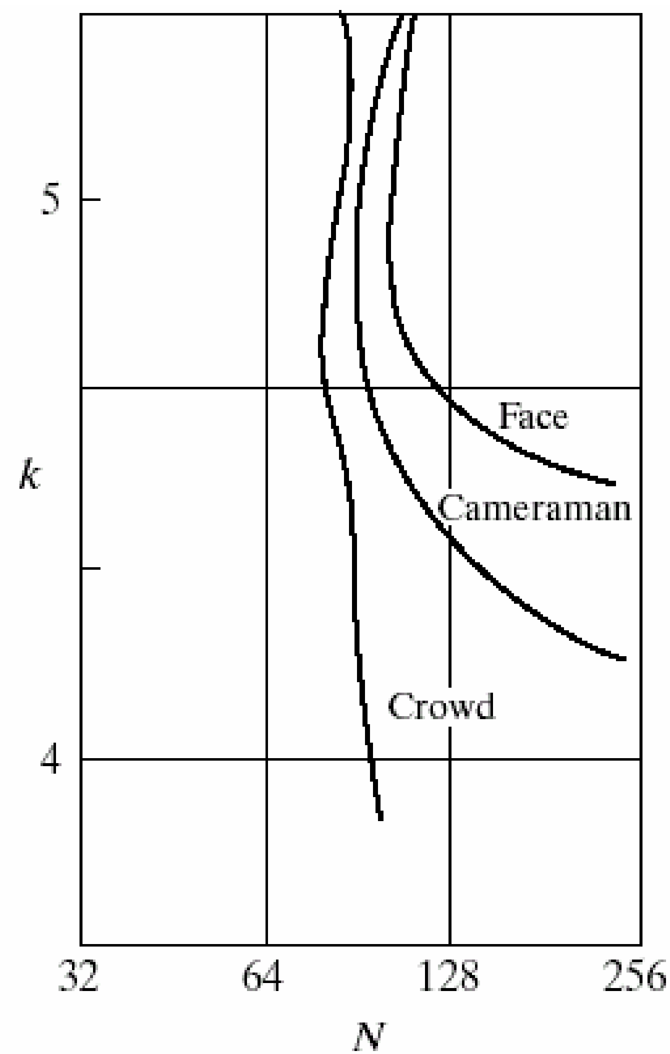


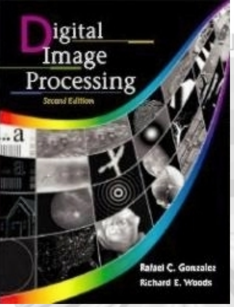
2.4.3 Spatial and Gray-Level resolution

- Varying *N* and *k*, the observers are then asked to rank the image according to their subjective quality.
- The results are summarized in the form of so-called ***isopreference curve***.
- Points on the curve corresponding to the images of equal subjective quality.
- The ***isopreference curve*** tends to become more vertical as the detail of the image increases.

2.4.3 Spatial and Gray-Level resolution

FIGURE 2.23 Representative isopreference curves for the three types of images in Fig. 2.22.





2.4.3 Aliasing and Moire Pattern

- Band-limited function.
- Undersampling – aliasing.
- Aliasing frequencies
- Sampling rate : the number of samples taken per unit distance
- Reduce high frequency component prior to sampling.
- *Moire Pattern* is caused by a break-up of the periodicity, *i.e.*, images are scanned from a printed page, which consists of periodic ink dots.

2.4.4 Aliasing and Moire Pattern

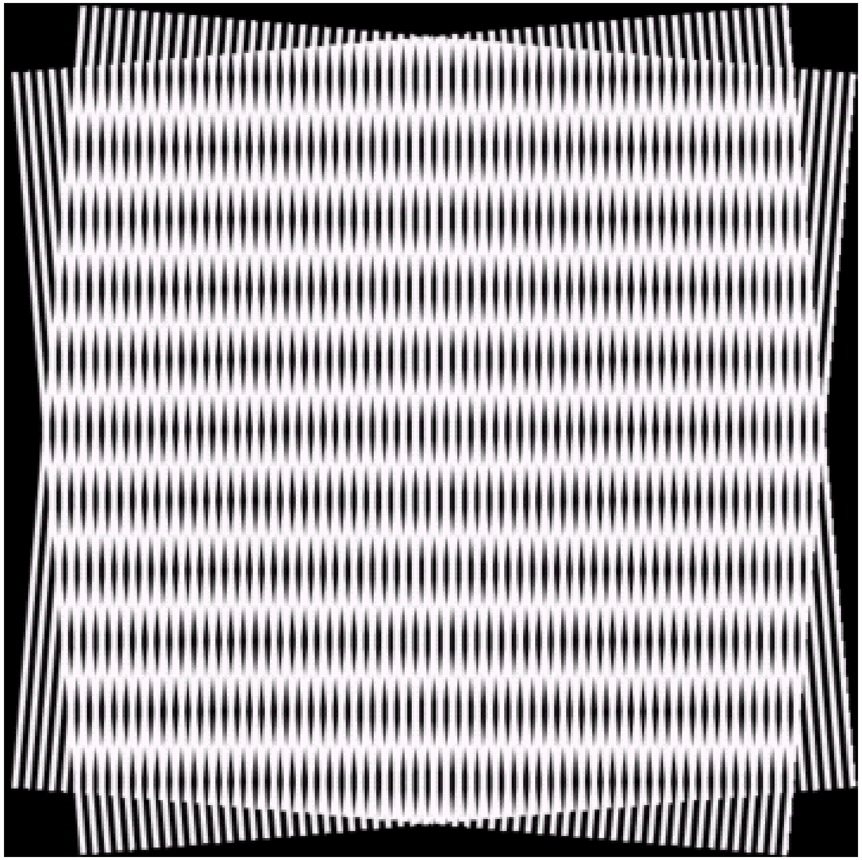
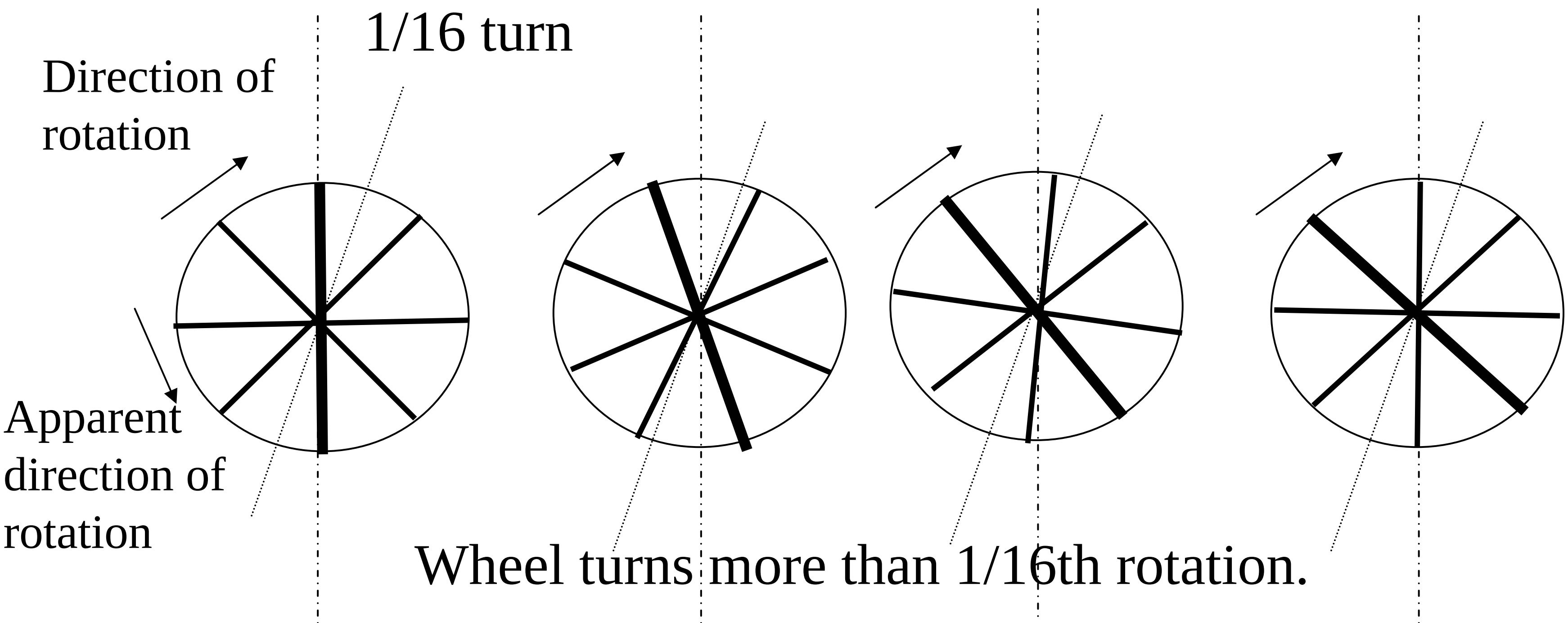
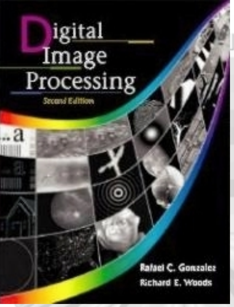


FIGURE 2.24 Illustration of the Moiré pattern effect.

2.4.4 Temporal sampling

- ***Temporal aliasing*** refers to the time-related representation of motion within an image sequence, *i.e.*, wagon wheel in old "Western" movies.





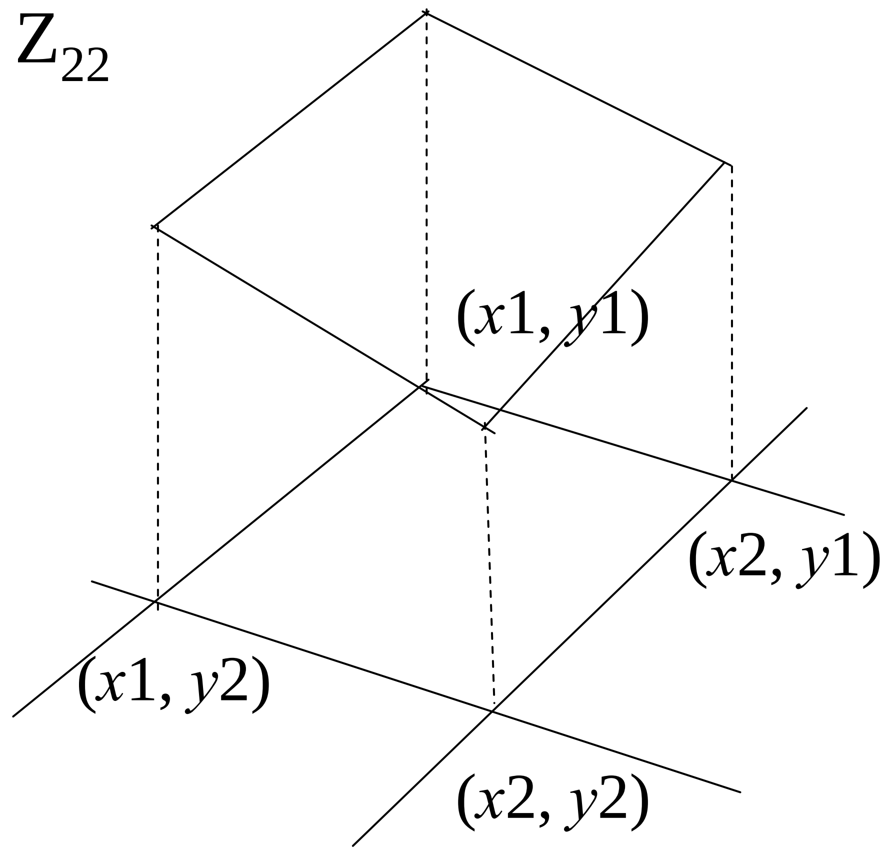
2.4.5 Zooming and Shrinking

- ***Zooming***:
  - Create a new pixel locations
  - Assign a gray-levels to those new locations
- ***Nearest neighbor interpolation***
  - Pixel replication
  - Bilinear interpolation using four nearest neighbors
    - $v(x', y') = ax' + by' + cx'y' + d$
    - where a, b, c, and d are obtained from the gray-level of the four neighbors.
  - Higher-order non-linear interpolation: using more neighbors for interpolation
- ***Shrinking:***
  - Direct shrinking causes aliasing
  - Expansion then Shrinking: blurring the image before shrinking it and reduce aliasing.

2.4.5 Zooming and Shrinking

Nearest neighbor interpolation

- $f(x, y)=a_1+a_2x+a_3y+a_4xy$
- Four corner points: $(x1, y1)$, $(x1, y2)$, $(x2, y1)$, $(x2, y2)$, with intensity Z_{11}, Z_{12}, Z_{21}, Z_{22}
- $Z_{11}=a_1+a_2x1+a_3y1+a_4x1y1$
- $Z_{12}=a_1+a_2x1+a_3y2+a_4x1y2$
- $Z_{13}=a_1+a_2x2+a_3y1+a_4x2y1$
- $Z_{14}=a_1+a_2x2+a_3y2+a_4x2y2$
- Solve for a_1, a_2, a_3, a_4



2.4.5 Zooming and Shrinking

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.

2.5 Basic Relations between pixels

- *Neighbors of a pixel p*
  - Horizontal and vertical neighbors.
    - $(x+1, y), (x-1, y), (x, y+1), (x, y-1)$
  - Four diagonal neighbors.
    - $(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)$
  - 4-neighbors of p: $N_4(p)$.
  - 4-diagonal neighbors of p : $N_D(p)$.
  - 8-neighbors of p: $N_8(p) = N_4(p) \cup N_D(p)$.

2.5 Basic Relations between pixels

- Adjacency

  4-adjacency: p and q are 4-adjacency if $q \in N_4(p)$

  8-adjacency: p and q are 8-adjacency if $q \in N_8(p)$

  m-adjacent (mixed) (*i.e*., Fig. 2.26)

- Path (curve) from $p=(x_0, y_0)$ to $g=(x_n, y_n)$

  consist of a sequence of pixels: $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$

  where pixels (x_i, y_i) and (x_{i-1}, y_{i-1}) are adjacent

- Closed path if $(x_0, y_0)=(x_n, y_n)$

2.5 Basic Relations between pixels

- Connectivity

  S represent a set of pixels in image, Two pixels *p* and *q* are said to connected in *S* if there exists a path between them. For any pixel *p* in *S*, the set of pixels that are connected to it in S is called a ***connected component*** in *S*. If there is ***only one*** connected component, then *S* is called a ***connected set***

- Regions.

  Let *R* be a subset of pixels in image, We call *R* a region if it is a connected set.

- Boundary:

  The set of pixels in a region *R* that have one or more neighbors that are not in *R*.

2.5 Basic Relations between pixels

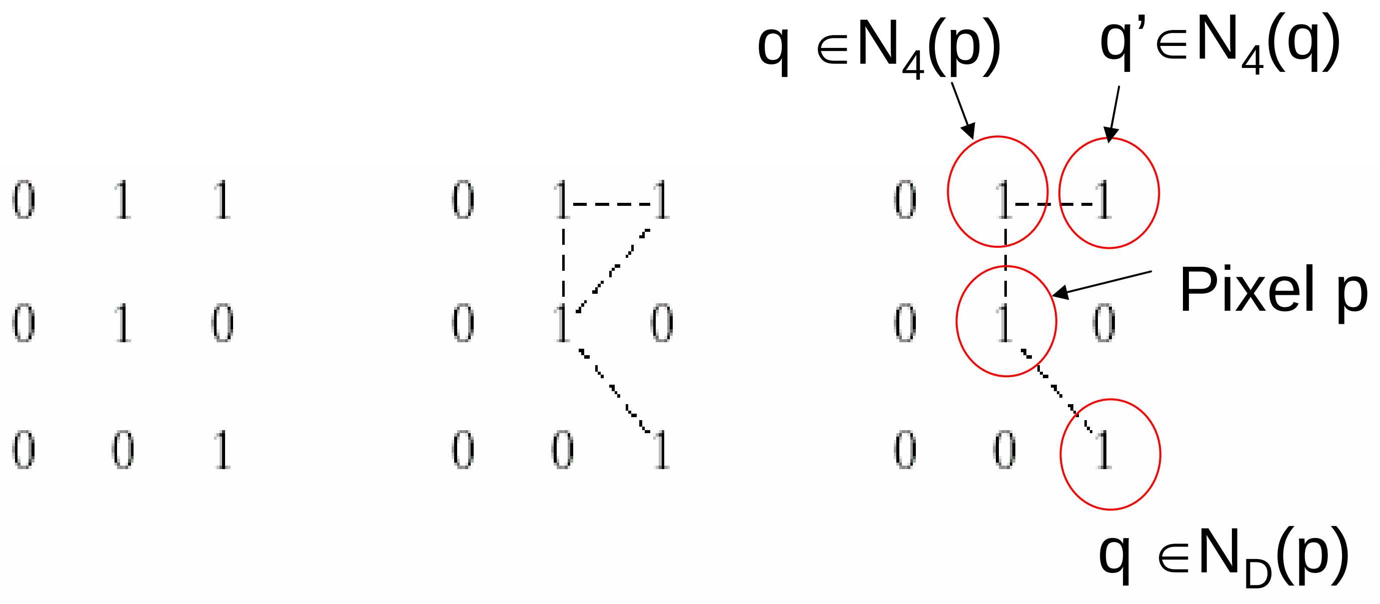


FIGURE 2.26 (a) Arrangement of pixels; (b) pixels that are 8-adjacent (shown dashed) to the center pixel; (c) *m*-adjacency.

2.5 Basic Relations between pixels

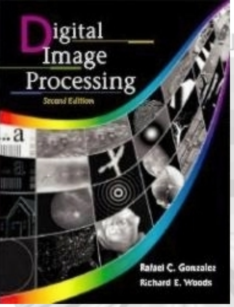
- ***Distance measures***
  - Euclidean distance
  - City-block distance or D_4 distance.

    $D_4(p, q) = | x-s | + | y-t |$
  - D_8 distance or chessboard distance.

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

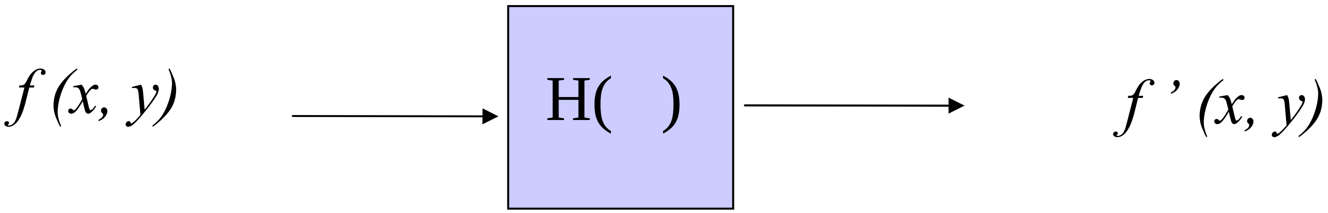
		2		
	2	1	2	
2	1	0	1	2
	2	1	2	
		2		

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



2.5.4 Image Operation on a Pixel basis

- Image processing: different operations applied on the pixels.



- Linear or nonlinear operation
  - H$(af+bg)$=aH(f)+bH(g), H is a linear operator.