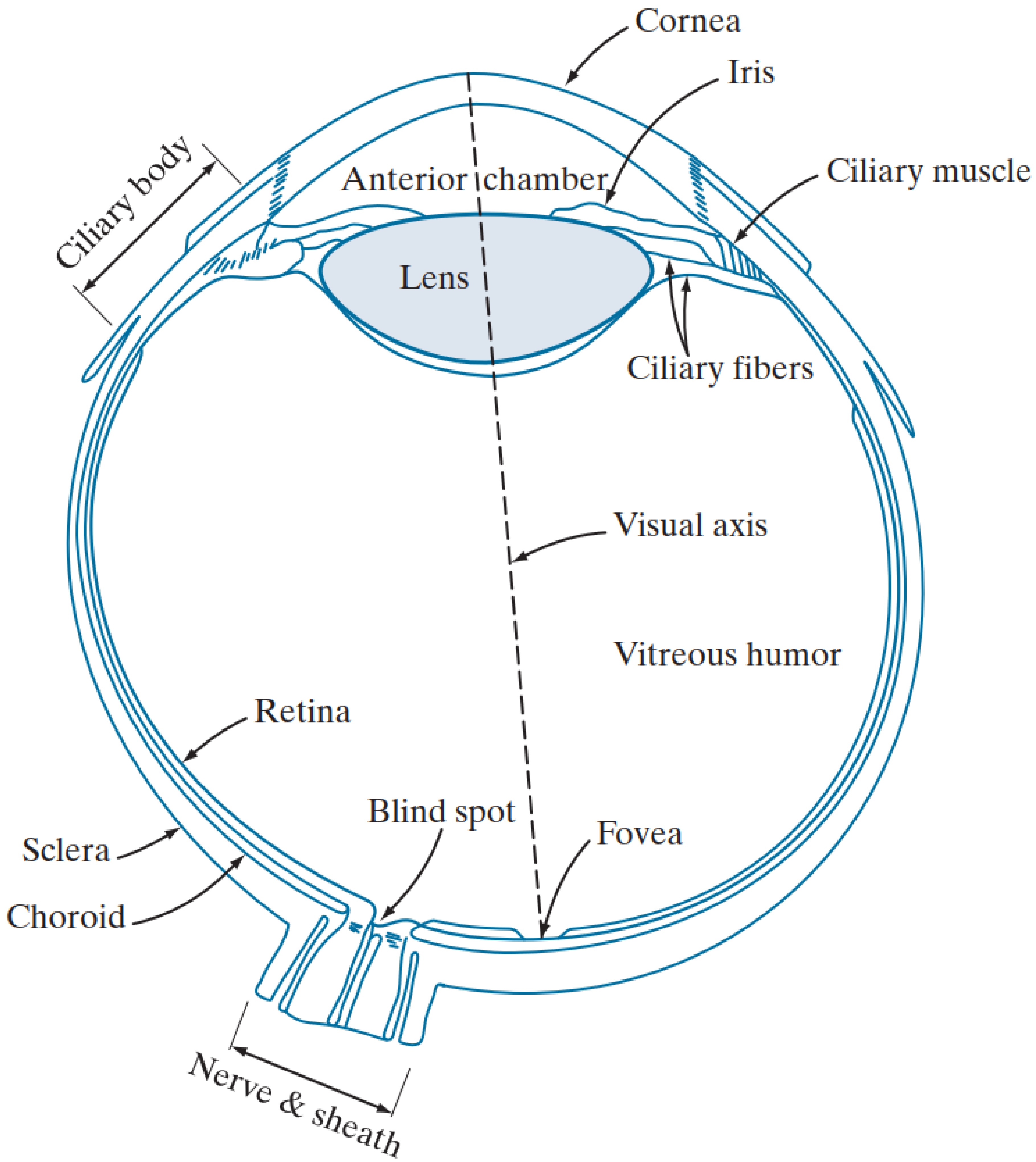
Digital image fundamentals

Alberto Carini

Structure of the human eye

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



Structure of the human eye

FIGURE 2.2
Distribution of rods and cones in the retina.

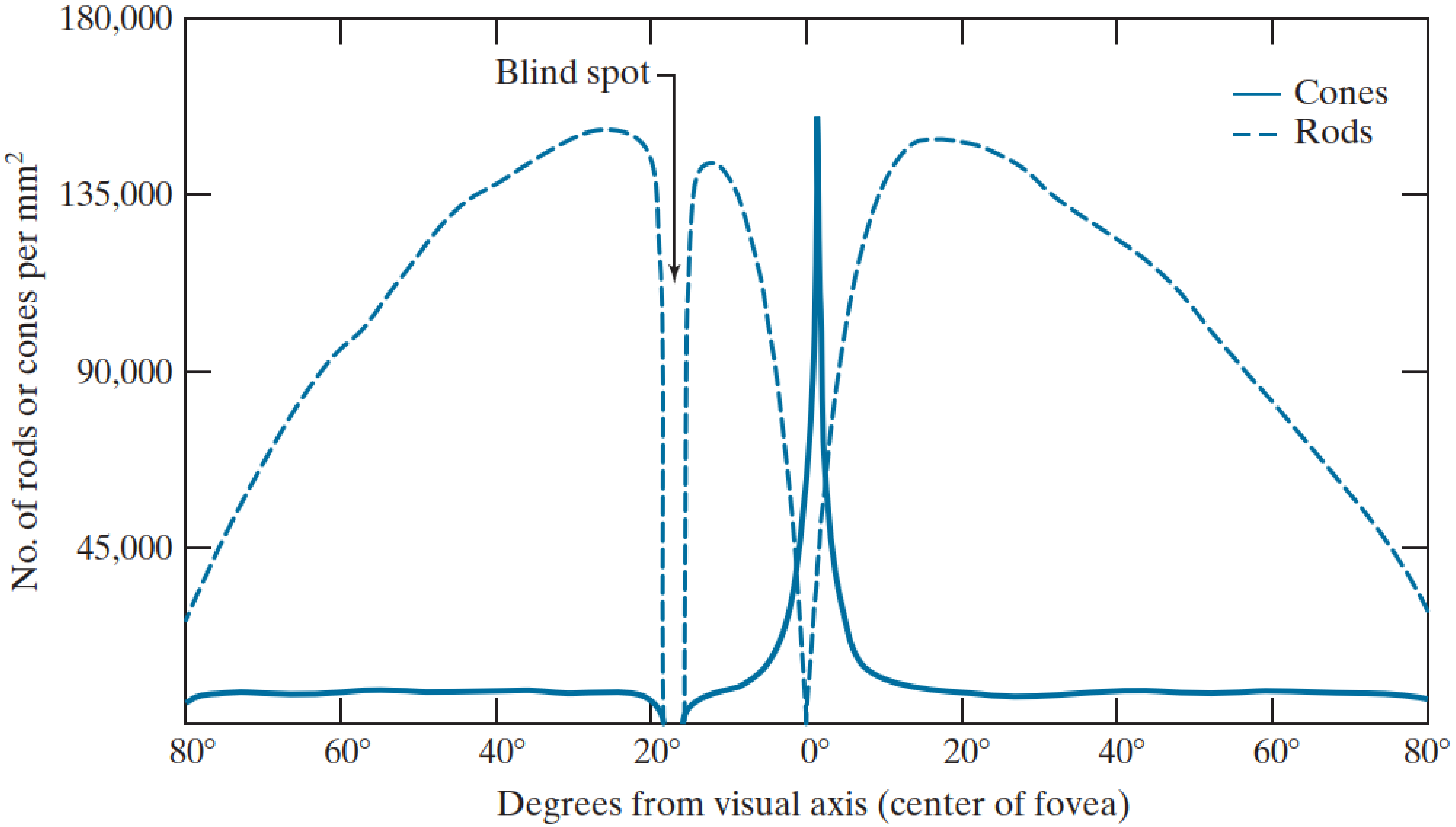
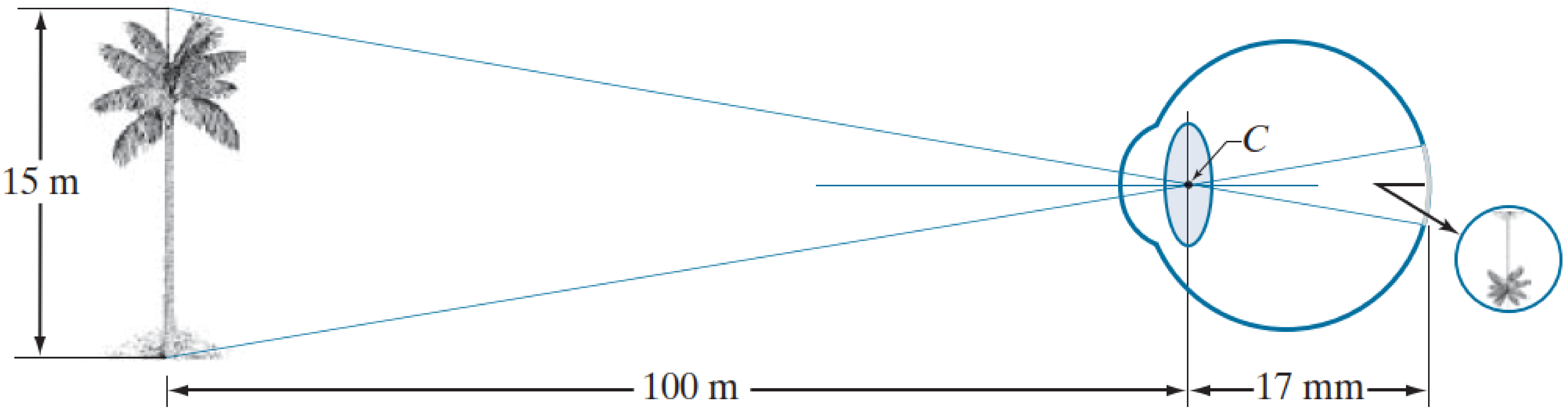


Image formation in the eye

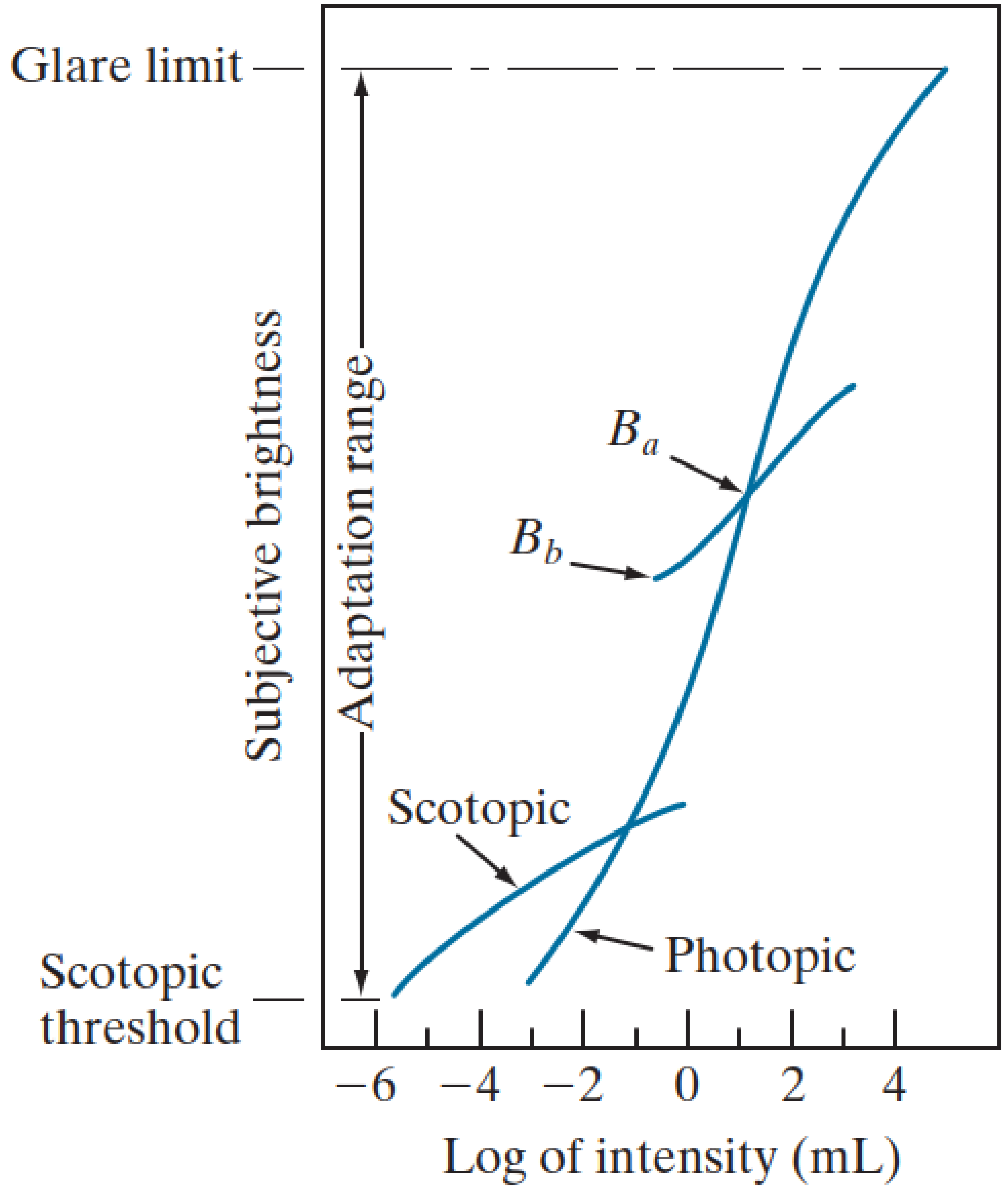
FIGURE 2.3 Graphical representation of the eye looking at a palm tree. Point *C* is the focal center of the lens.



- Suppose that a person is looking at a tree 15 m high at a distance of 100 m.
- Letting h denote the height of that object in the retinal image, the geometry of the figure yields 15/100 = h/17 or h = 2.5 mm.

Brightness adaptation or discrimination

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



Brightness adaptation or discrimination

FIGURE 2.5
Basic experimental setup used to characterize brightness discrimination.

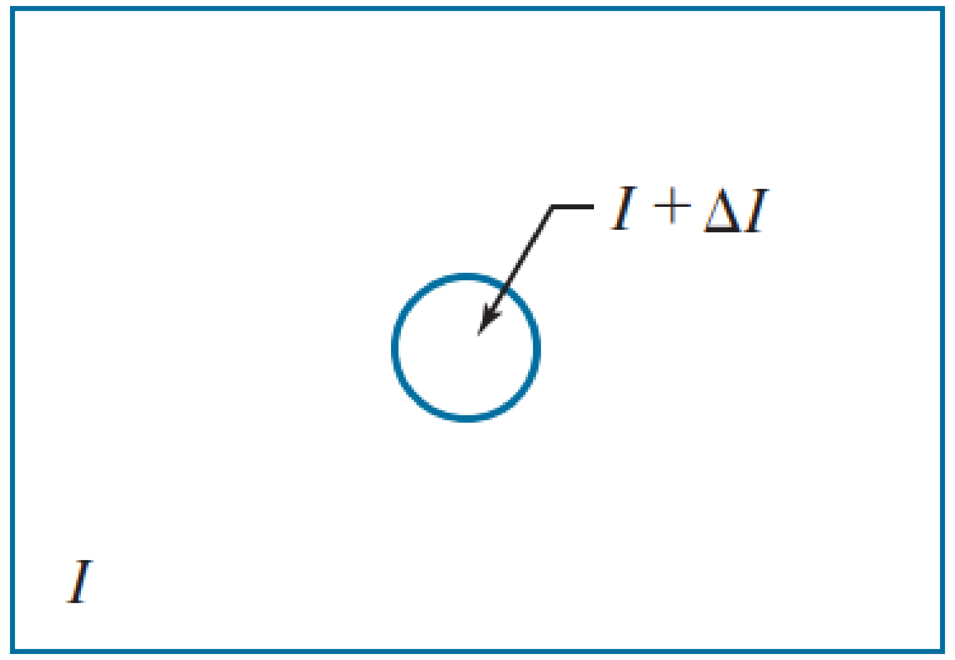
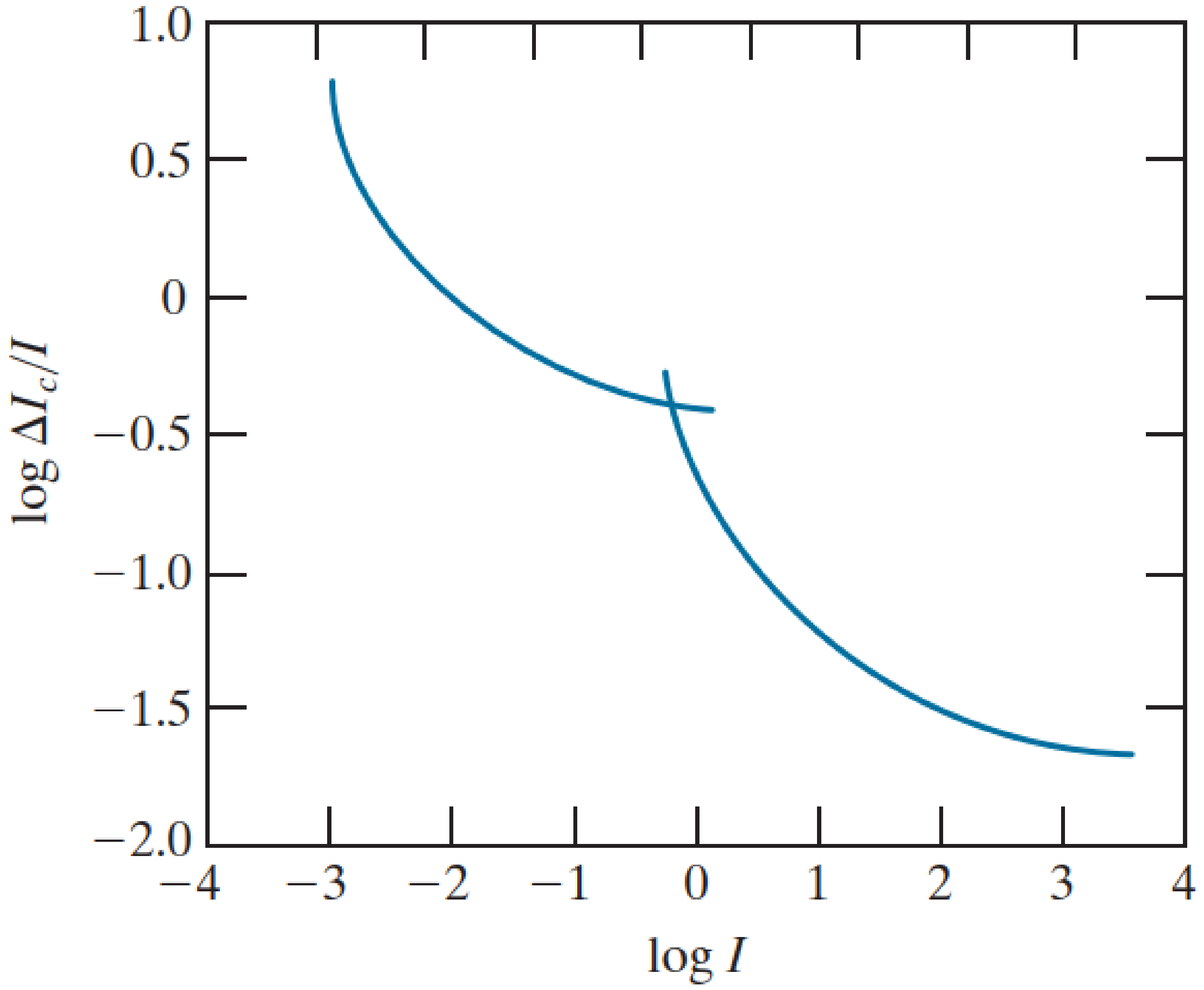


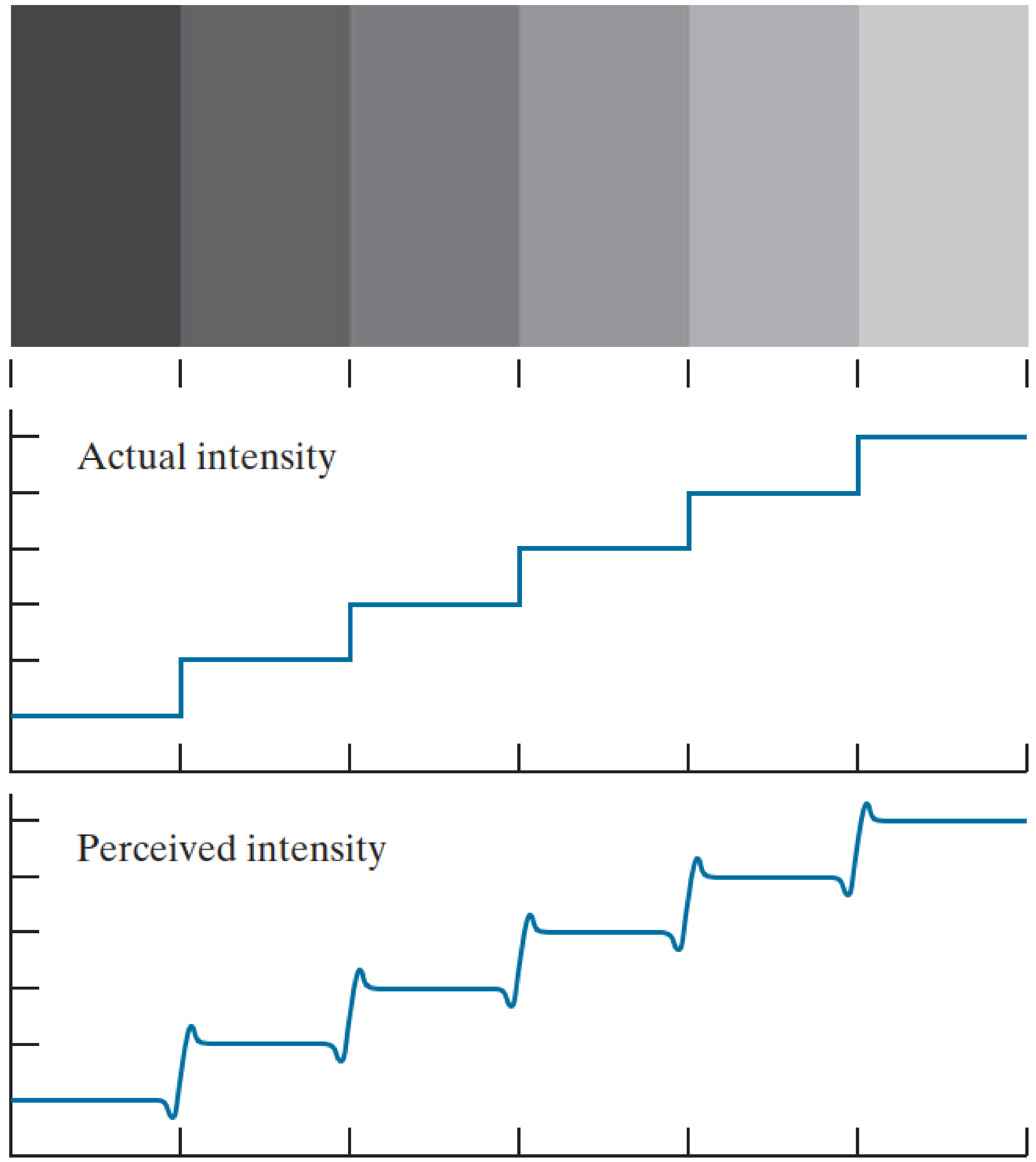
FIGURE 2.6
A typical plot of the Weber ratio as a function of intensity.



Brightness adaptation or discrimination

a
b
c

FIGURE 2.7 Illustration of the Mach band effect. Perceived intensity is not a simple function of actual intensity.

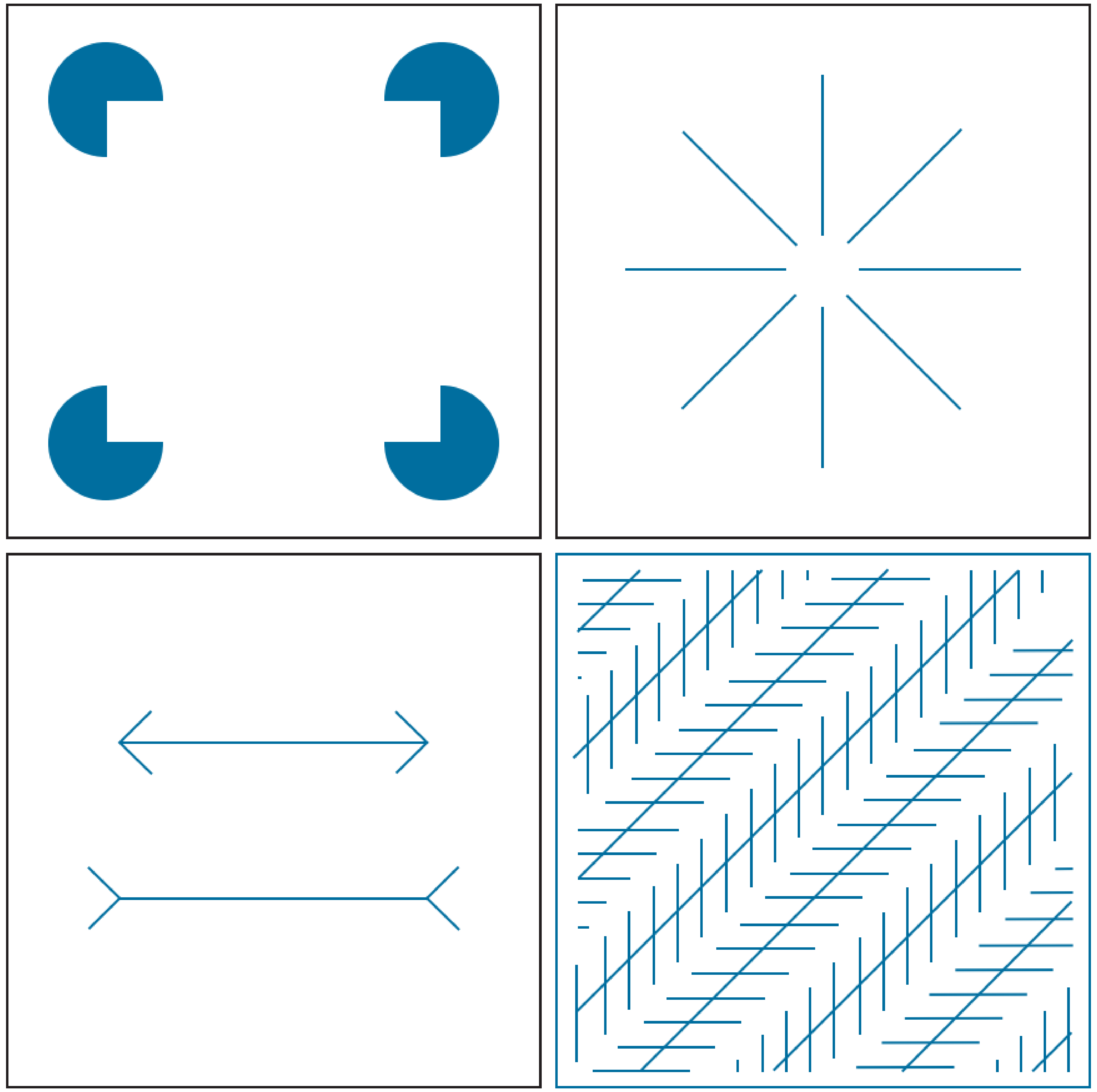


Brightness adaptation or discrimination

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.

Brightness adaptation or discrimination



a b
c d

FIGURE 2.9 Some well-known optical illusions.

Light and the electromagnetic spectrum

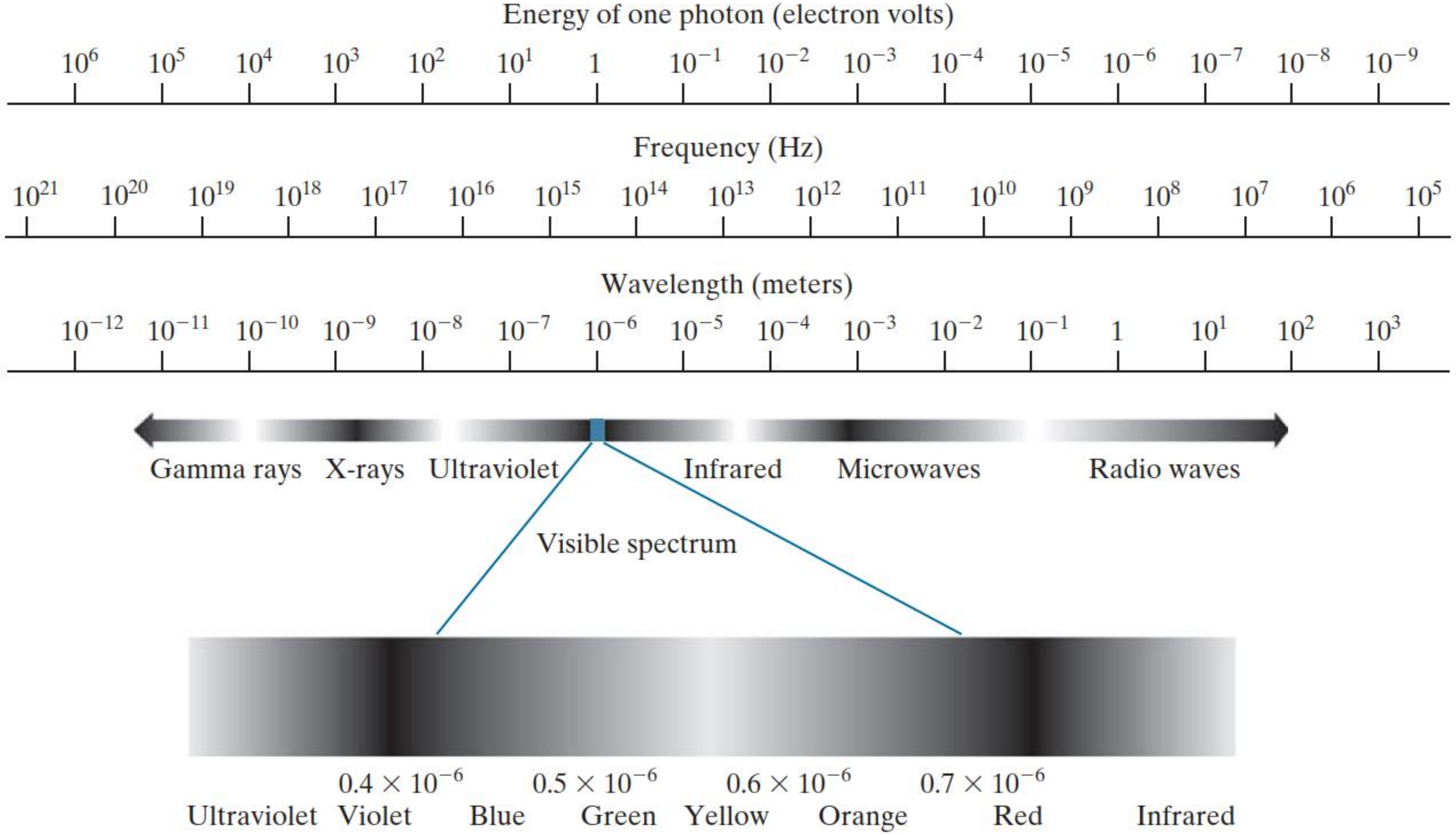


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanations, but note that it encompasses a very narrow range of the total EM spectrum.

$$\lambda = \frac{c}{\nu}$$

$$E = h\nu$$

Light and the electromagnetic spectrum

- Light that is void of color is called monochromatic (or achromatic) light.
- The only attribute of monochromatic light is its *intensity*.
- Because the intensity of monochromatic light is perceived to vary from black to grays and finally to white, the term gray level is used commonly to denote monochromatic intensity.
- The range of values of monochromatic light from black to white is usually called the *gray scale*, and monochromatic images are frequently referred to as *grayscale images*.
- *Chromatic* light spans the electromagnetic energy spectrum from approximately 0.43 to 0.79 um. In addition to frequency, three other quantities are used to describe a chromatic light source: radiance, luminance, and brightness.
- *Radiance* is the total amount of energy that flows from the light source, measured in watts (W).
- *Luminance*, measured in lumens (lm), gives a measure of the amount of energy an observer perceives from a light source.
- *Brightness* is a subjective descriptor of light perception that is practically impossible to measure. It embodies the achromatic notion of intensity and is one of the key factors in describing color sensation.

Some definitions

Lumen:

Photometrically, it is the luminous flux emitted within a unit solid angle (one steradian) by a point source having a uniform luminous intensity of one candela.

Candela:

The candela is the luminous intensity in a given direction of a source that emits monochromatic radiation of frequency 540 x 10^{12} hertz and that has a radiant intensity in that direction of 1/683 watt per steradian.

Image sensing and acquisition

a
b
c

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

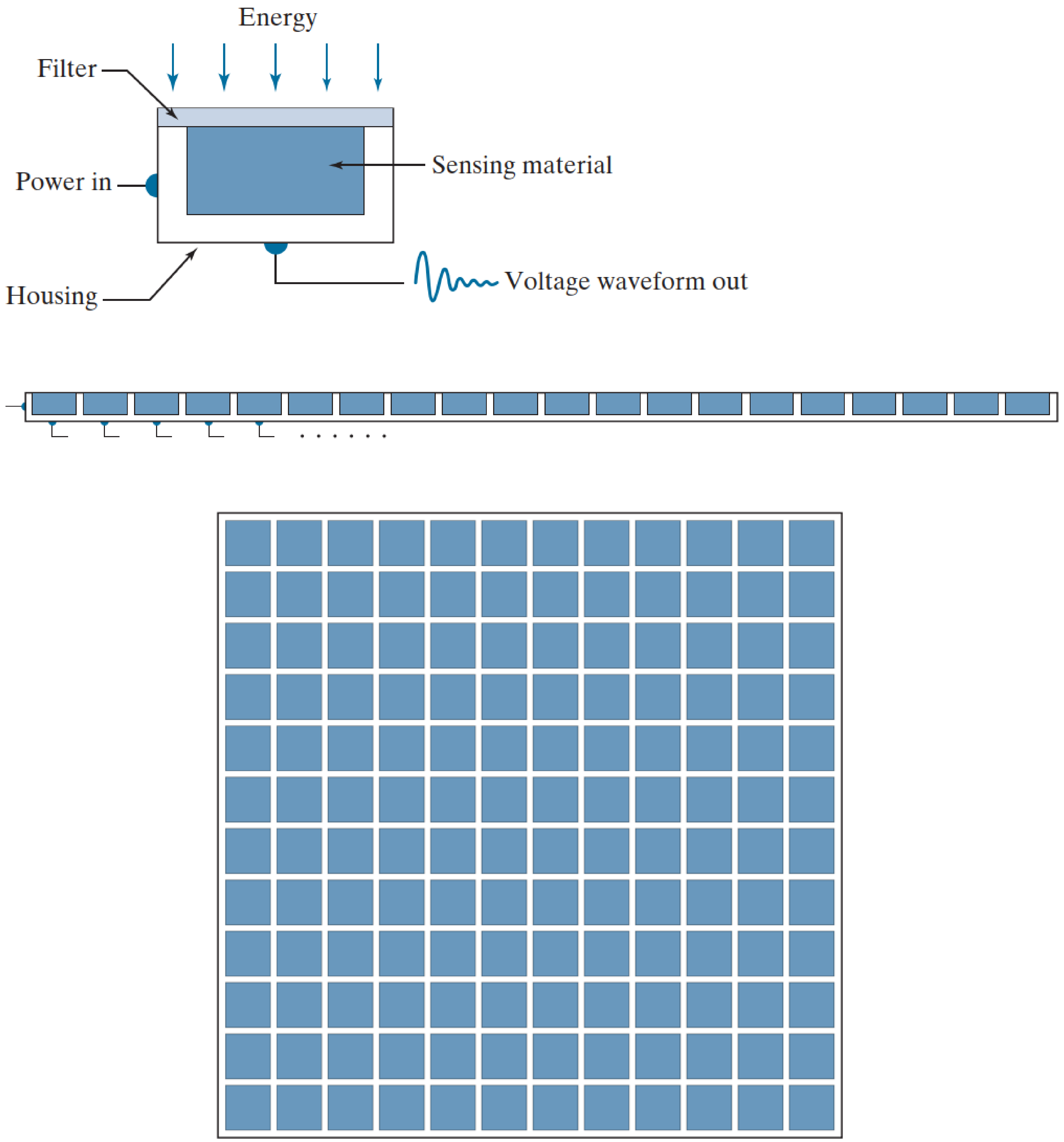


Image sensing and acquisition

FIGURE 2.13
Combining a single sensing element with mechanical motion to generate a 2-D image.

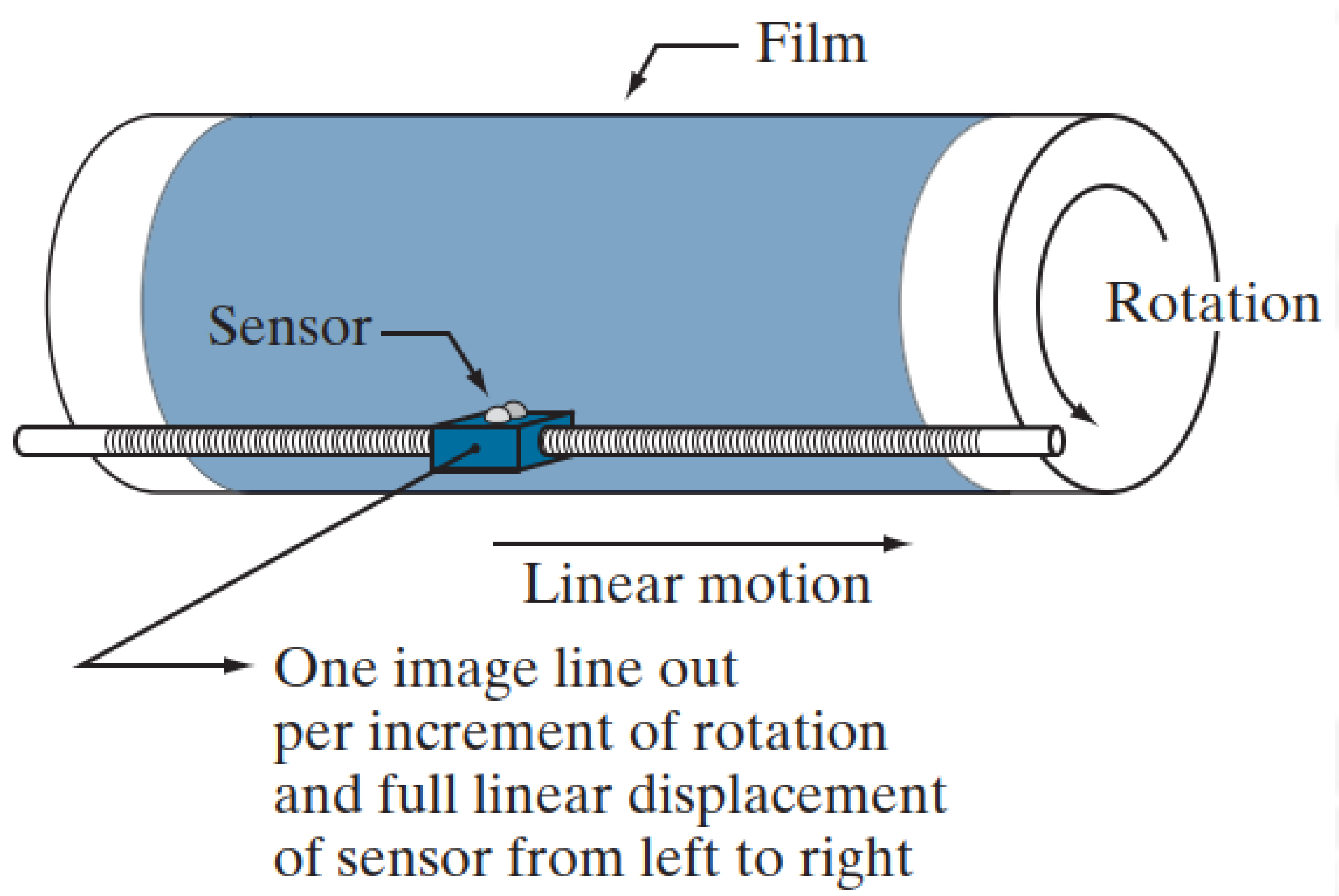


Image sensing and acquisition

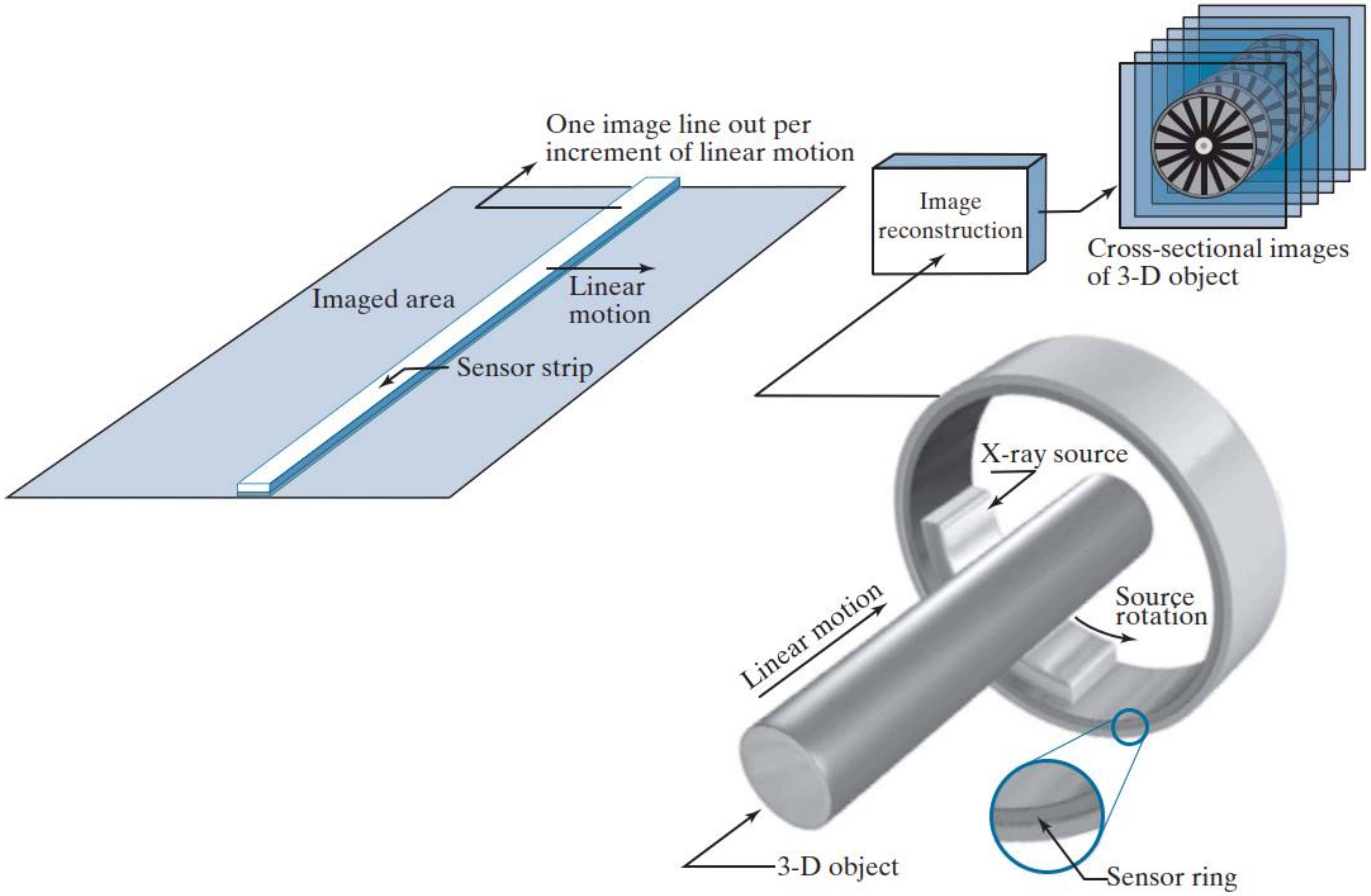
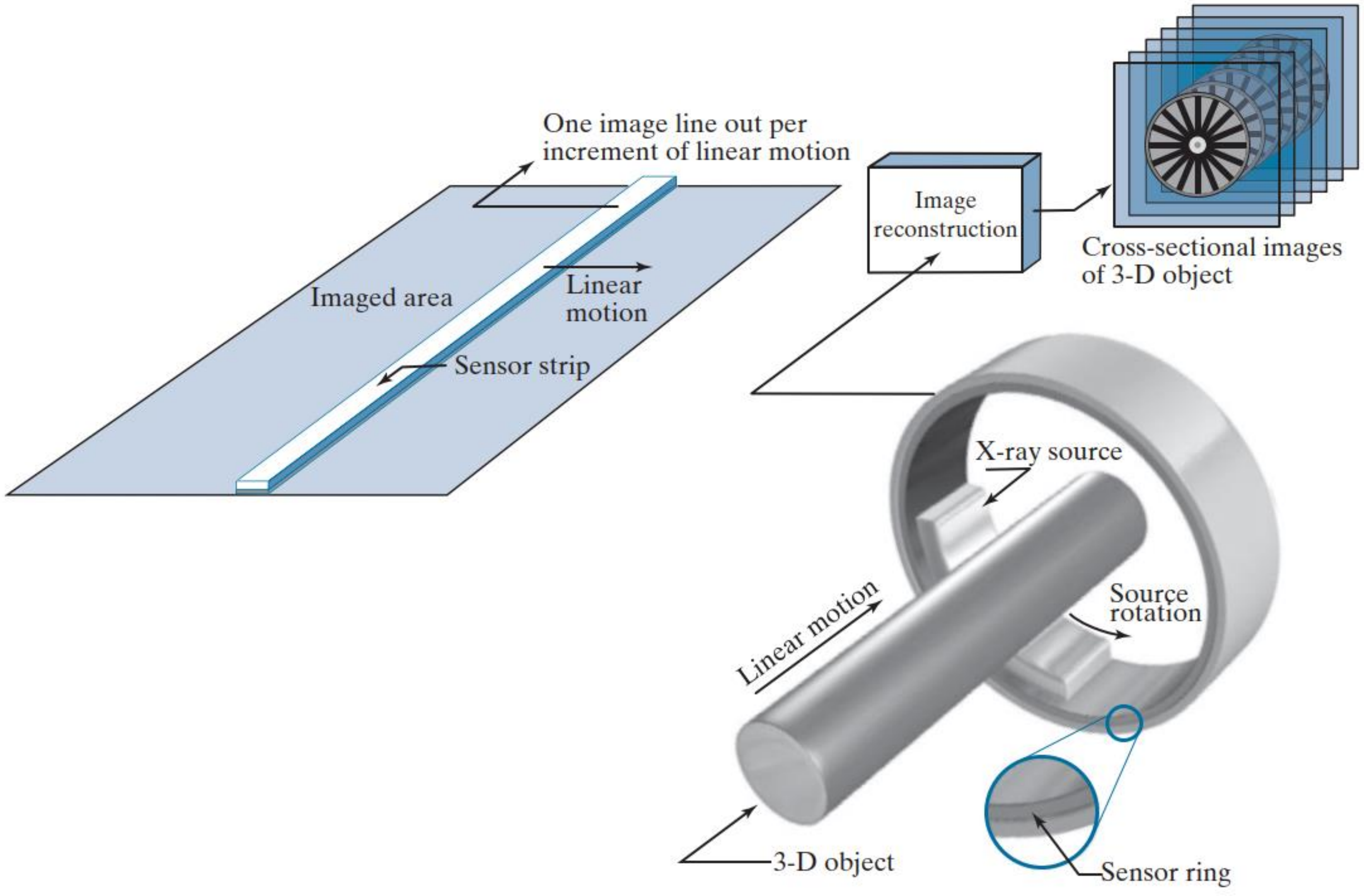


FIGURE 2.14 (a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.

Image sensing and acquisition

a b

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



A simple image formation model

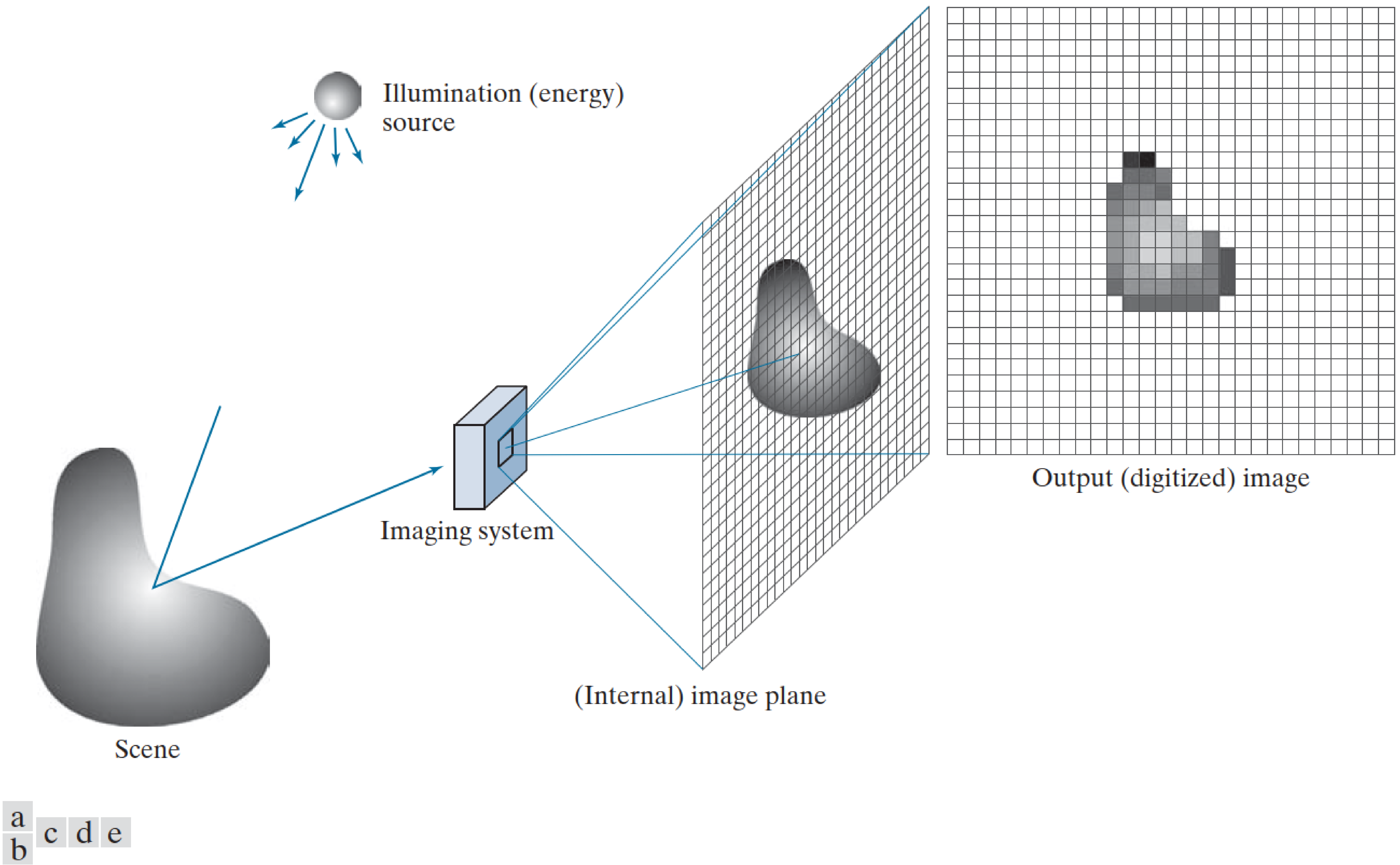


FIGURE 2.15 An example of digital image acquisition. (a) Illumination (energy) source. (b) A scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

A simple image formation model

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

- Function f (x, y) is characterized by two components:
  (1) the amount of source illumination incident on the scene being viewed, and
  (2) the amount of illumination reflected by the objects in the scene.
- These are called the *illumination* and *reflectance* components, and are denoted by i(x, y) and r(x, y), respectively:

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

where

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

and

$$0 \leq r(x,y) \leq 1$$

- Are applicable also to images formed via transmission of the illumination through a medium.: *transmissivity* instead of a *reflectivity* but same limits.

A simple image formation model

- On a clear day, the sun may produce in excess of 90,000 lm/m^2 of illumination on the surface of the earth.
- This value decreases to less than 10,000 lm/m^2 on a cloudy day.
- On a clear evening, a full moon yields about 0.1 lm/m^2 of illumination.
- The typical illumination level in a commercial office is about 1,000 lm/m^2.
- Similarly, the following are typical values of $r(x, y)$:
  - 0.01 for black velvet,
  - 0.65 for stainless steel,
  - 0.80 for flat-white wall paint,
  - 0.90 for silver-plated metal, and
  - 0.93 for snow.

A simple image formation model

- Let the intensity (gray level) of a monochrome image at any coordinates (x, y) be denoted by

$$\ell = f(x, y)$$

- it is evident that ℓ lies in the range

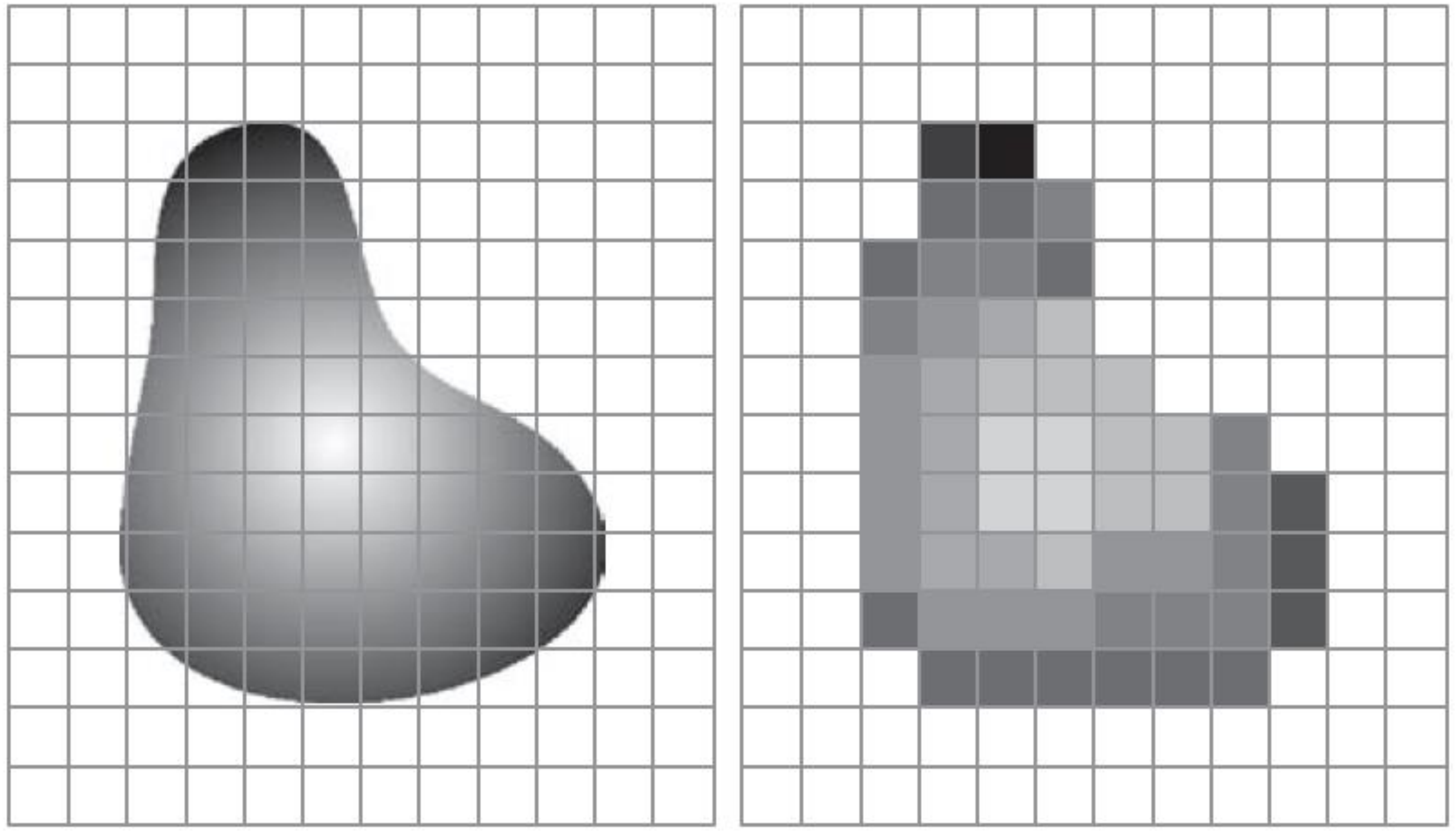
$$L_{\min} \leq \ell \leq L_{\max}$$

- The interval $[L_{min}, L_{max}]$ is called the intensity (or gray) scale.
- Common practice is to shift this interval numerically to the interval [0, 1], or [0,C], where $\ell = 0$ is considered black and $\ell = 1$ (or C) is considered white on the scale. All intermediate values are shades of gray varying from black to white.

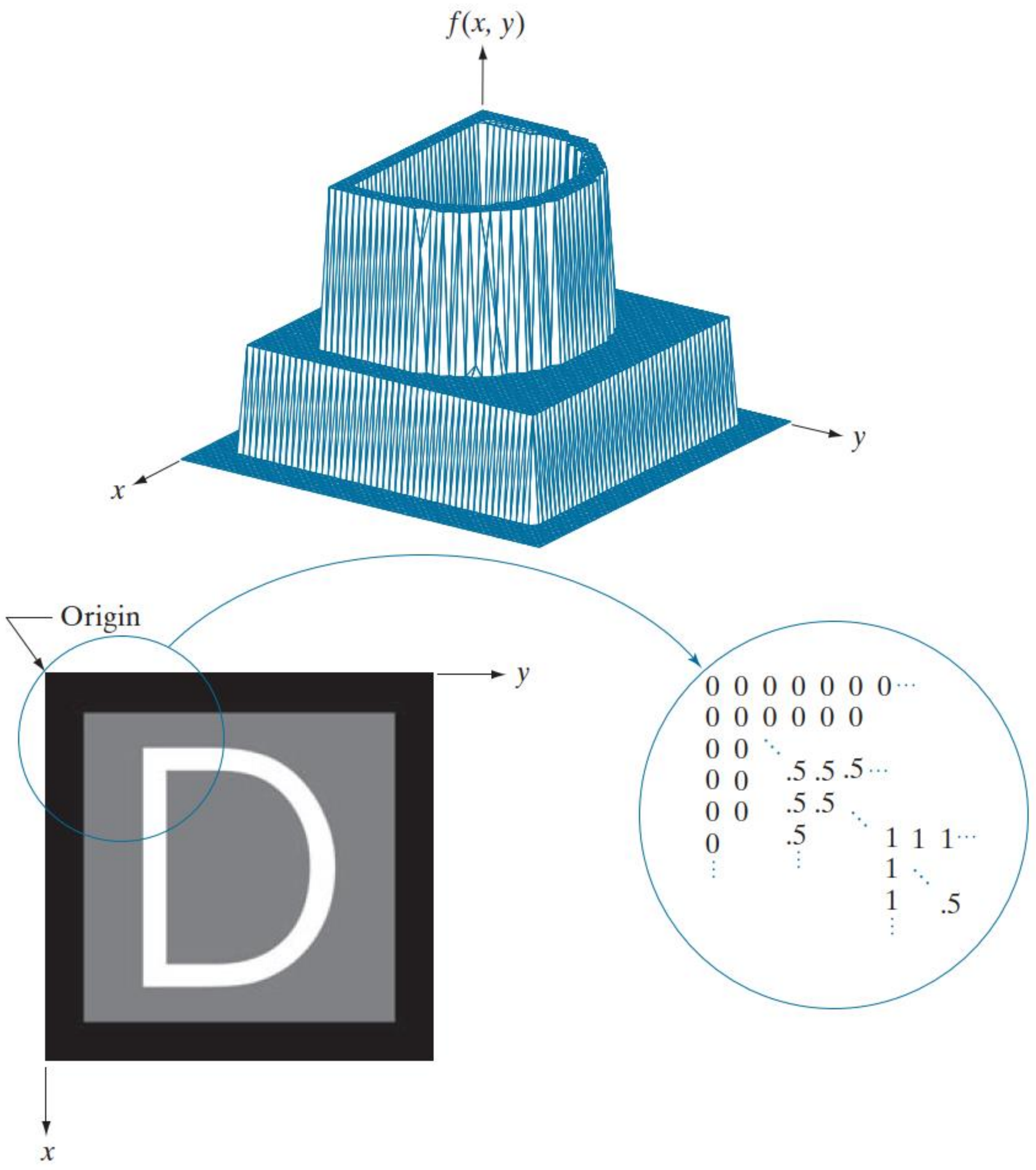
Image sampling and quantization

a b

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



Representation of a digital image

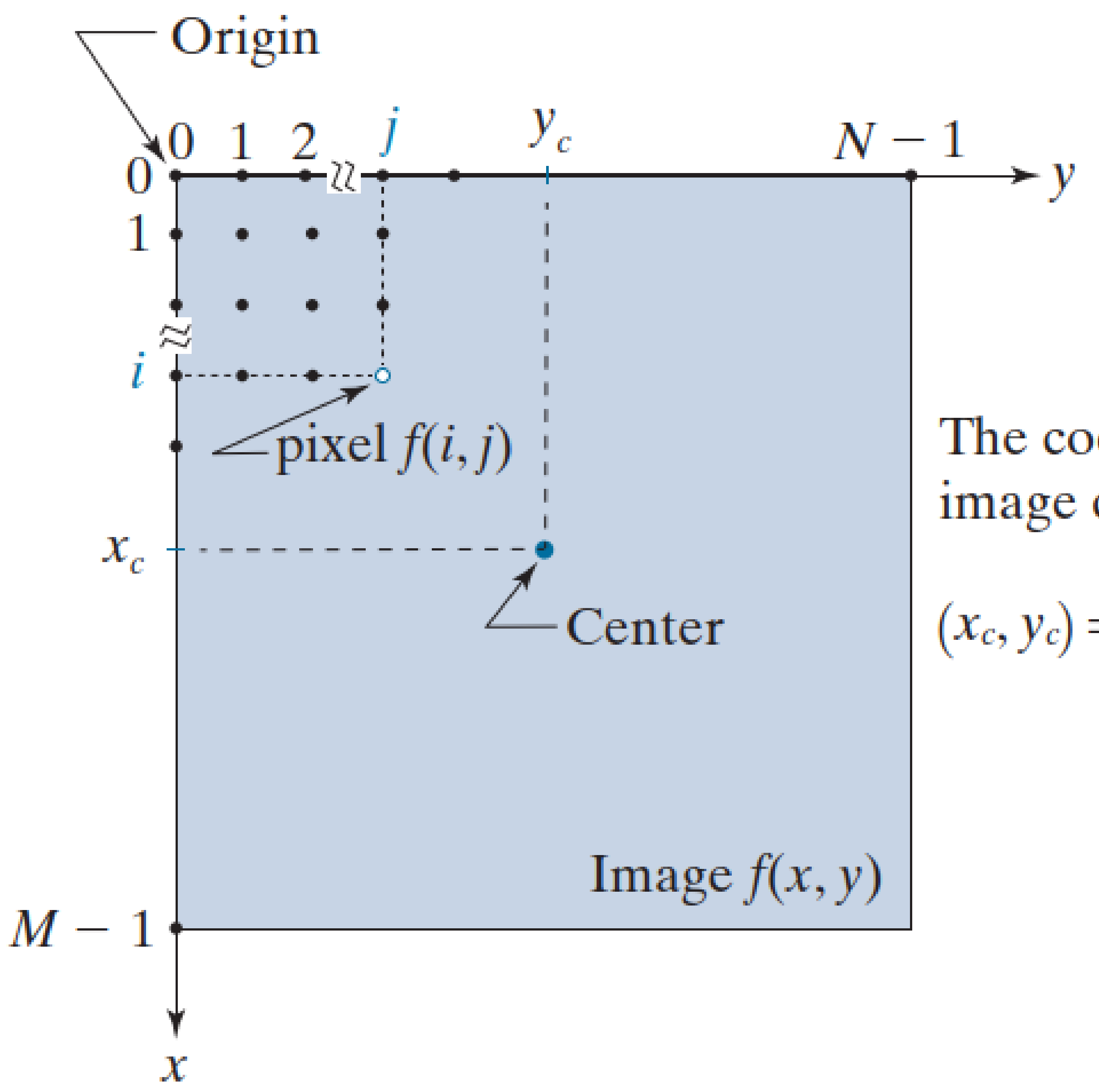


a
b c

FIGURE 2.18
(a) Image plotted as a surface. (b) Image displayed as a visual intensity array. (c) Image shown as a 2-D numerical array. (The numbers 0, .5, and 1 represent black, gray, and white, respectively.)

Representation of a digital image

FIGURE 2.19 Coordinate convention used to represent digital images. Because coordinate values are integers, there is a one-to-one correspondence between x and y and the rows (r) and columns (c) of a matrix.



The coordinates of the image center are

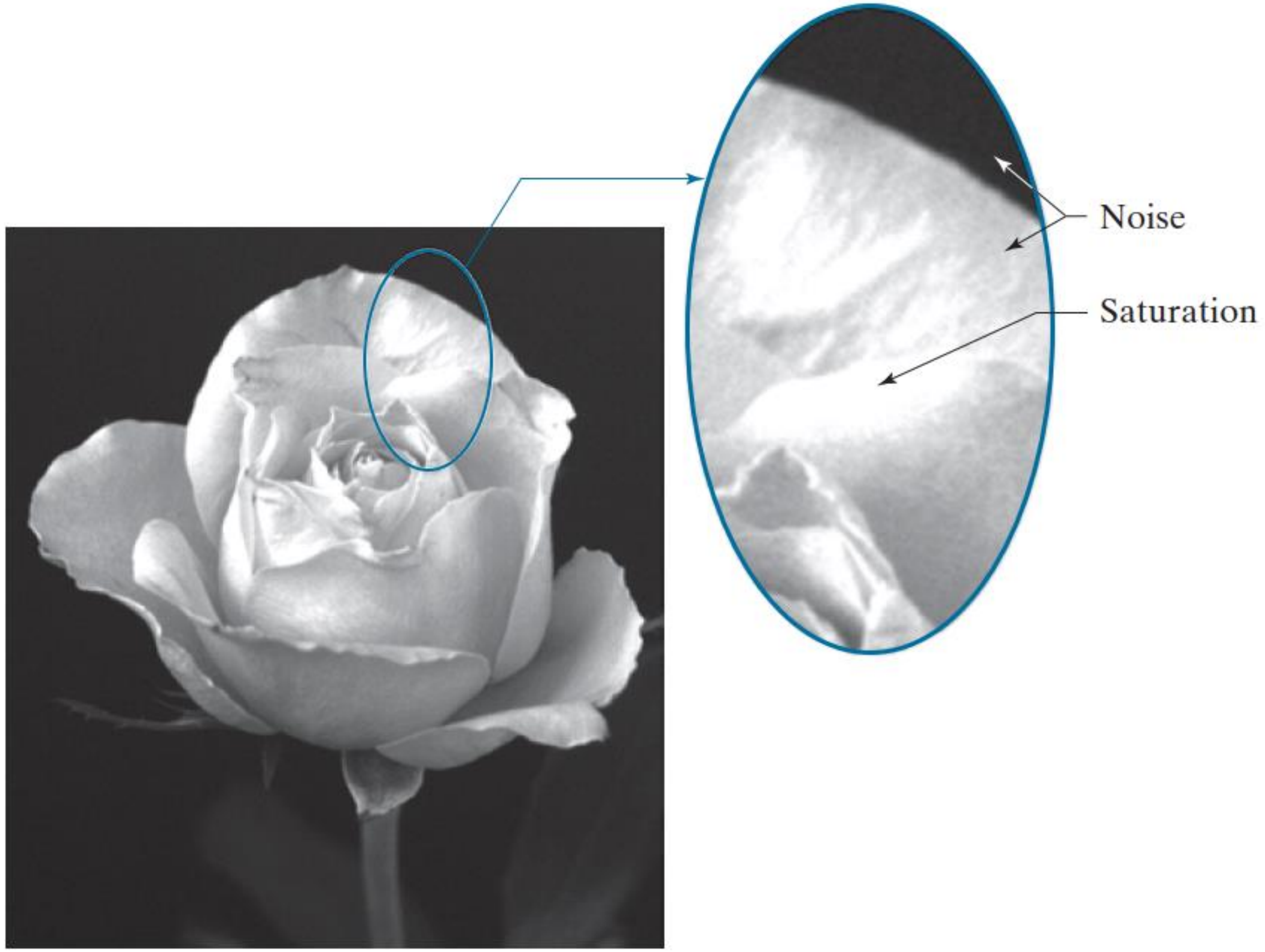
$$(x_c, y_c) = \left(\text{floor}\left(\frac{M}{2}\right), \text{floor}\left(\frac{N}{2}\right)\right)$$

Representation of a digital image

- Sometimes, the range of values spanned by the gray scale is referred to as the *dynamic range*.
- Here, we define the *dynamic range of an imaging system* to be the ratio of the maximum measurable intensity to the minimum detectable intensity level in the system. As a rule, the upper limit is determined by *saturation* and the lower limit by *noise*.
- The dynamic range establishes the lowest and highest intensity levels that a system can represent and, consequently, that an image can have.
- Closely associated with this concept is *image contrast*, which we define as the difference in intensity between the highest and lowest intensity levels in an image. The *contrast ratio* is the ratio of these two quantities.

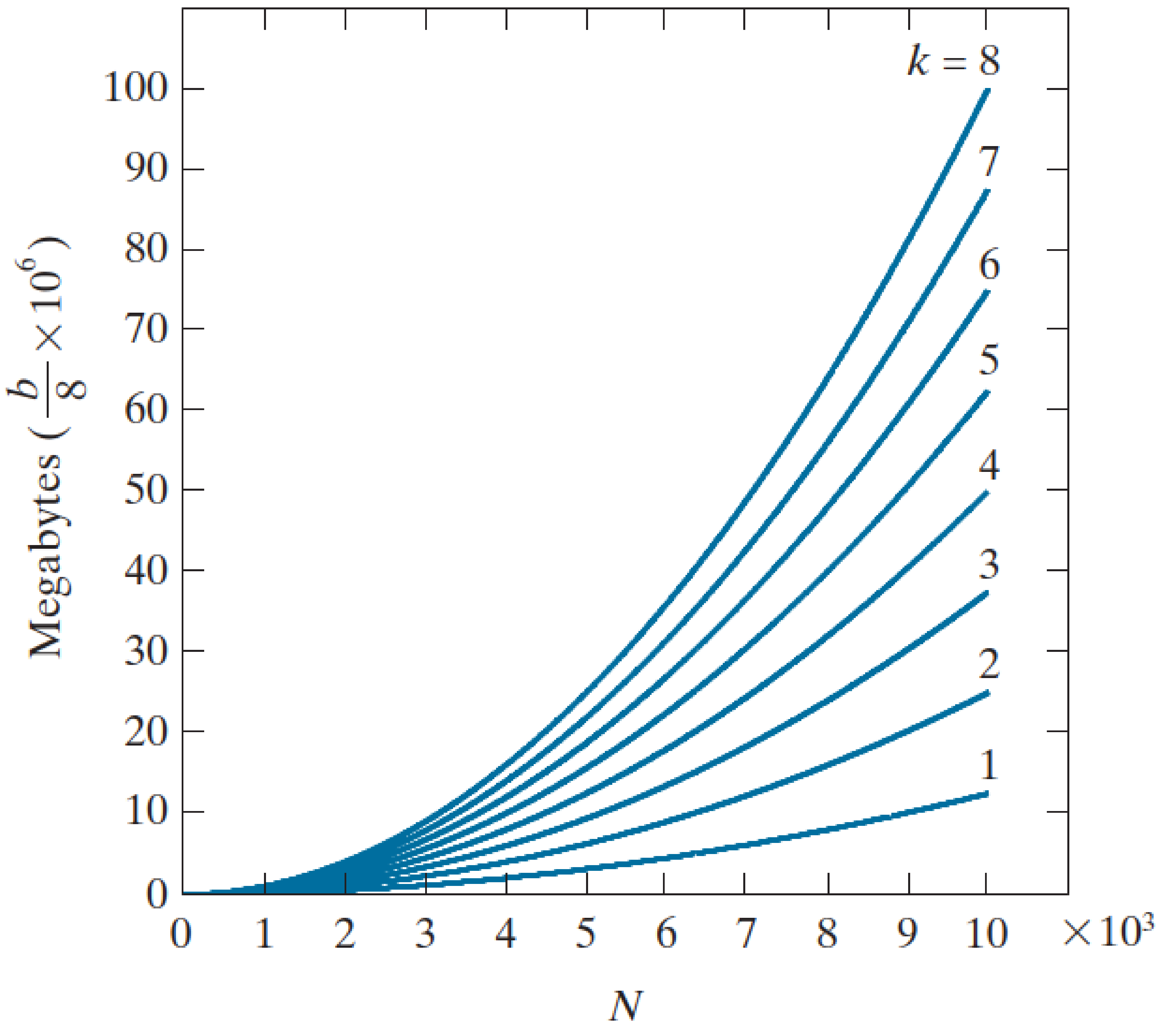
Representation of a digital image

FIGURE 2.20
An image exhibiting saturation and noise. Saturation is the highest value beyond which all intensity values are clipped (note how the entire saturated area has a high, constant intensity level). Visible noise in this case appears as a grainy texture pattern. The dark background is noisier, but the noise is difficult to see.



Representation of a digital image

FIGURE 2.21
Number of megabytes required to store images for various values of N and k.



The number, b, of bits required to store a digital image is

$$b = M \times N \times k$$

When $M = N$, this equation becomes

$$b = N^2 k$$

Spatial intensity and resolution

- Intuitively, spatial resolution is a measure of the smallest discernible detail in an image.
- Quantitatively, spatial resolution can be stated in several ways, with *line pairs per unit* distance, and *dots (pixels) per unit distance* being common measures.
- Suppose that we construct a chart with alternating black and white vertical lines, each of width W units. The width of a line pair is thus 2W, and there are W/2 line pairs per unit distance.
- A widely used definition of image resolution is the largest number of discernible line pairs per unit distance (e.g., 100 line pairs per mm).
- *Dots per unit distance* is a measure of image resolution used in the printing and publishing industry. In the U.S., this measure usually is expressed as *dots per inch* (dpi).
- Newspapers are printed with a resolution of 75 dpi, magazines at 133 dpi, glossy brochures at 175 dpi, books also with 2400 dpi.
- **To be meaningful, measures of spatial resolution must be stated with respect to spatial units.**
- *Intensity resolution* similarly refers to the smallest discernible change in intensity level.
- Based on hardware considerations, the number of intensity levels usually is an integer power of two.

Spatial intensity and resolution

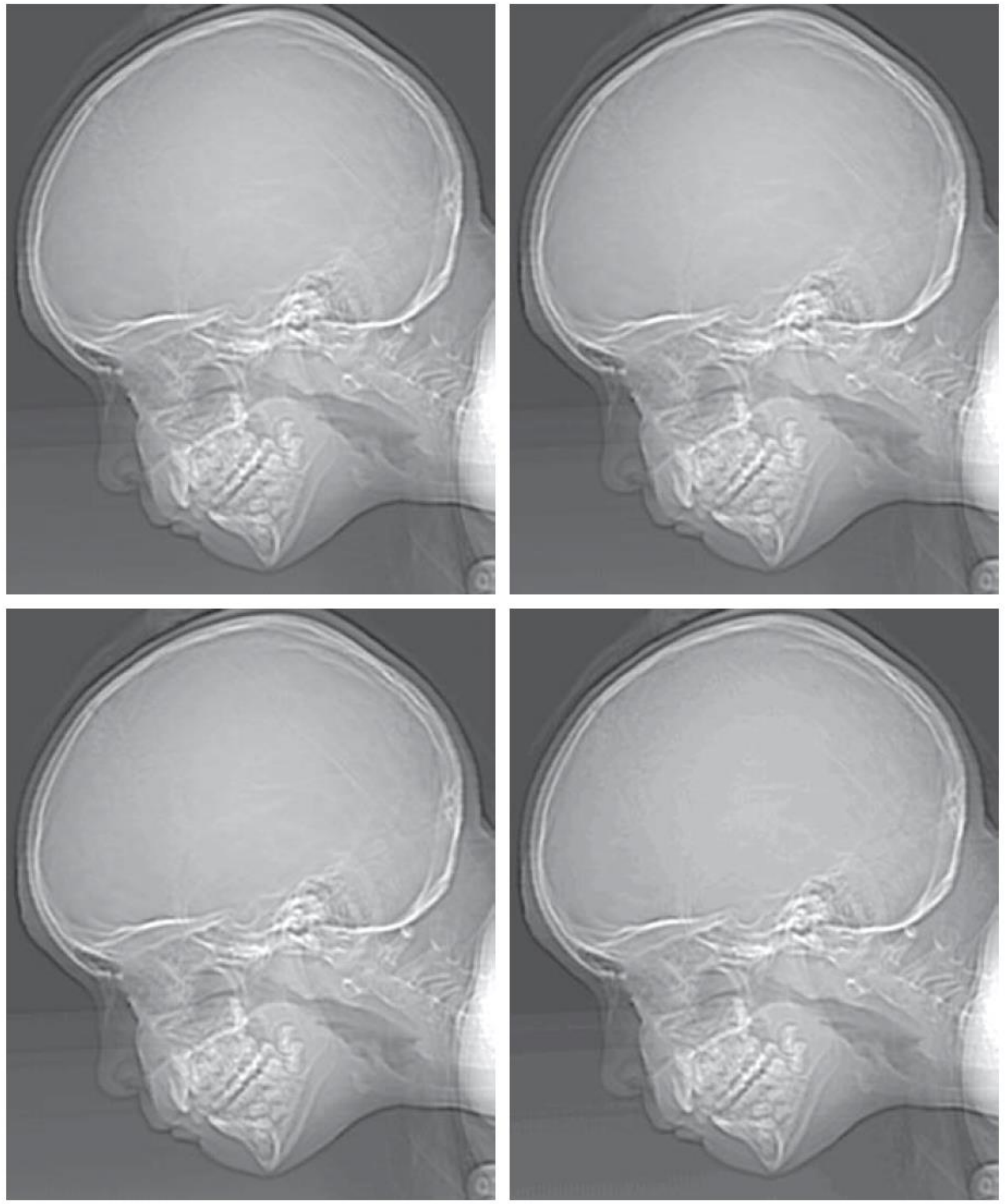
a b
c d

FIGURE 2.23 Effects of reducing spatial resolution. The images shown are at: (a) 930 dpi, (b) 300 dpi, (c) 150 dpi, and (d) 72 dpi.

Spatial intensity and resolution

a b
c d

FIGURE 2.24
(a) 774 × 640, 256-level image. (b)-(d) Image displayed in 128, 64, and 32 intensity levels, while keeping the spatial resolution constant. (Original image courtesy of the Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)



e f
g h

FIGURE 2.24
(*Continued*)
(e)-(h) Image displayed in 16, 8, 4, and 2 intensity levels.

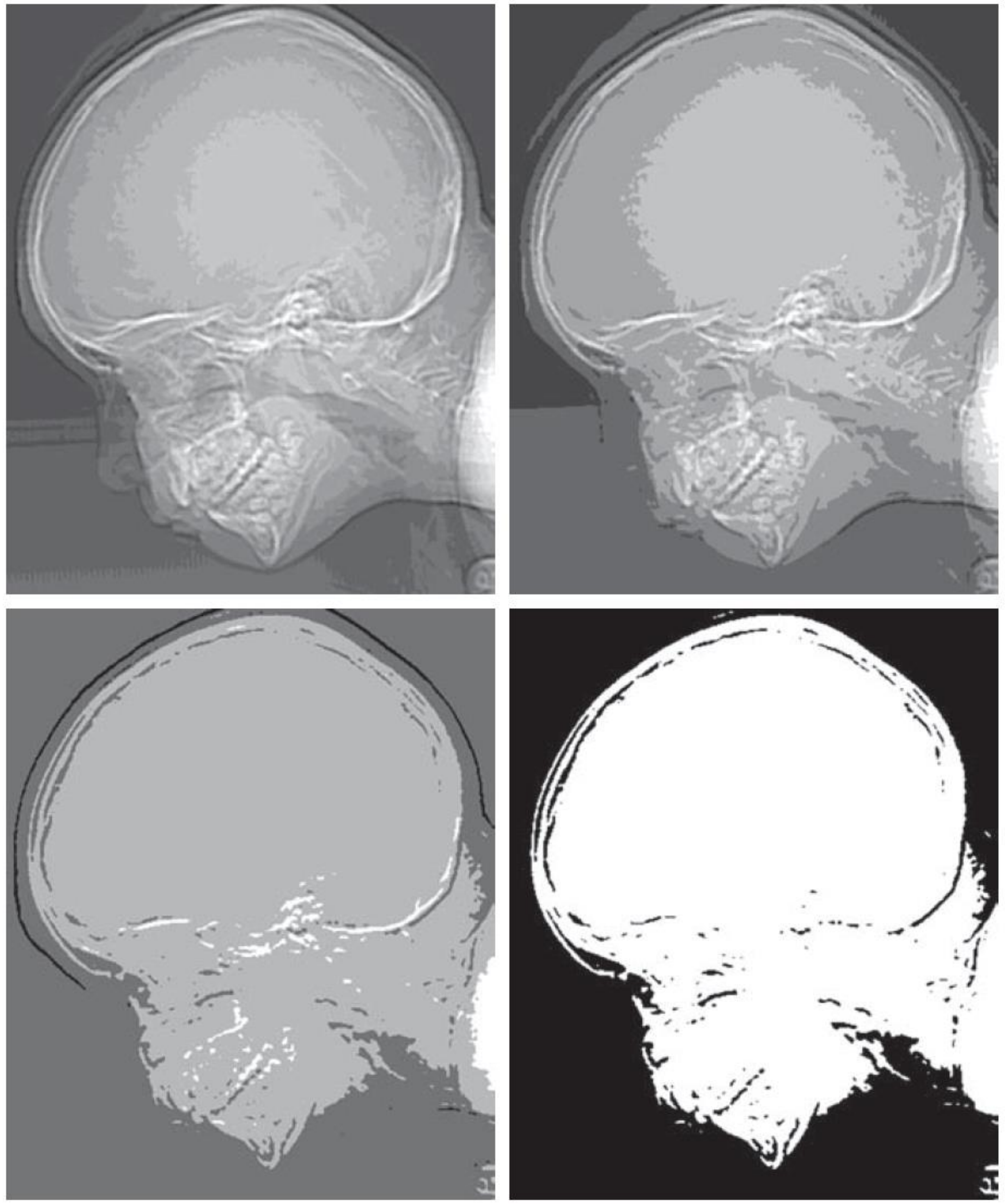


Image interpolation

- Interpolation is used in tasks such as zooming, shrinking, rotating, and geometrically correcting digital images.
- *Interpolation* is the process of using known data to estimate values at unknown locations.
- Suppose that an image of size 500 × 500 pixels has to be enlarged 1.5 times to 750 × 750 pixels.
- A simple way to visualize zooming is to create an imaginary 750 × 750 grid with the same pixel spacing as the original image, then shrink it so that it exactly overlays the original image.
- Obviously, the pixel spacing in the shrunken 750 × 750 grid will be less than the pixel spacing in the original image.
- To assign an intensity value to any point in the overlay…

Image interpolation

- *nearest neighbor interpolation* assigns to each new location the intensity of its nearest neighbor in the original image
- In *bilinear interpolation* the assigned value is obtained using the equation

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

  where the four coefficients are determined from the four equations in four unknowns that can be written using the four nearest neighbors of point (x, y).
- *Bicubic interpolation* involves the sixteen nearest neighbors of a point. The intensity value assigned to point (x, y) is obtained using the equation

$$v(x, y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^i y^j$$

- The sixteen coefficients are determined from the sixteen equations with sixteen unknowns that can be written using the sixteen nearest neighbors of point (x, y).
- Bicubic interpolation is the standard used in commercial image editing applications,

Image interpolation

a b c

FIGURE 2.27 (a) Image reduced to 72 dpi and zoomed back to its original 930 dpi using nearest neighbor interpolation. This figure is the same as Fig. 2.23(d). (b) Image reduced to 72 dpi and zoomed using bilinear interpolation. (c) Same as (b) but using bicubic interpolation.

Geometric transformations

- We use geometric transformations modify the spatial arrangement of pixels in an image. These transformations are called rubber-sheet transformations because they may be viewed as analogous to "printing" an image on a rubber sheet, then stretching or shrinking the sheet according to a predefined set of rules.
- Geometric transformations of digital images consist of two basic operations:

1. Spatial transformation of coordinates.
2. Intensity interpolation that assigns intensity values to the spatially transformed pixels.

- Our interest is in so-called *affine transformations*, which include scaling, translation, rotation, and shearing.
- The key characteristic of an affine transformation in 2-D is that it preserves points, straight lines, and planes.

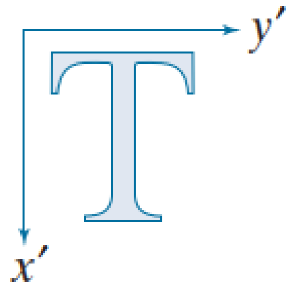
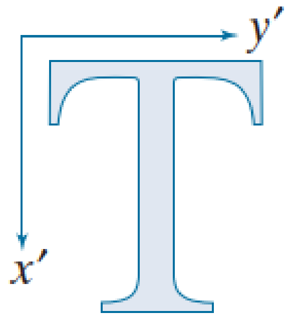
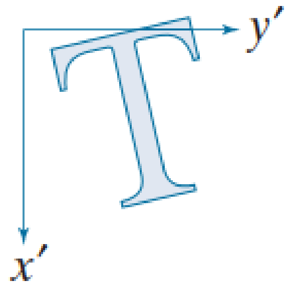
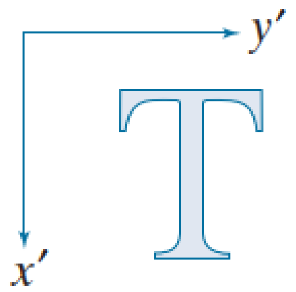
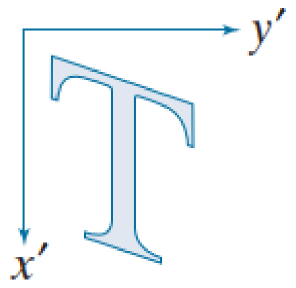
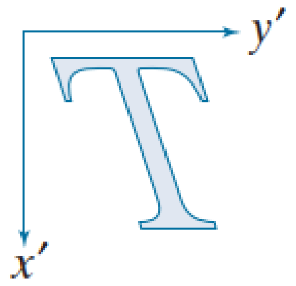
Geometric transformations

- It is possible to express all four affine transformations (scaling, translation, rotation, and shearing) using a single 3 × 3 matrix in the following general form:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

- This transformation can scale, rotate, translate, or sheer an image, depending on the values chosen for the elements of matrix **A**.
- The preceding transformation moves the coordinates of pixels in an image to new locations.
- To complete the process, we have to assign intensity values to those locations.
- This task is accomplished using *intensity interpolation*.

Geometric transformations

Transformation Name	Affine Matrix, A	Coordinate Equations	Example
Identity	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = y$	
Scaling/Reflection (For reflection, set one scaling factor to –1 and the other to 0)	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = c_x x$ $y' = c_y y$	
Rotation (about the origin)	$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x\cos\theta - y\sin\theta$ $y' = x\sin\theta + y\cos\theta$	
Translation	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + t_x$ $y' = y + t_y$	
Shear (vertical)	$\begin{bmatrix} 1 & s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + s_v y$ $y' = y$	
Shear (horizontal)	$\begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = s_h x + y$	

Geometric transformations

We can use the equation in two basic ways:

1. **Forward mapping**, which consists of scanning the pixels of the input image and, at each location (x,y), computing the spatial location (x′, y′) of the corresponding pixel in the output image using directly the equation.
   - A problem with the forward mapping approach is that two or more pixels in the input image can be transformed to the same location in the output image-
2. **Inverse mapping** scans the output pixel locations and, at each location (x′, y′), computes the corresponding location in the input image using $(x,y) = A^{-1}(x', y')$. It then interpolates among the nearest input pixels to determine the intensity of the output pixel value.
   - Inverse mappings are more efficient to implement than forward mappings.

Geometric transformations

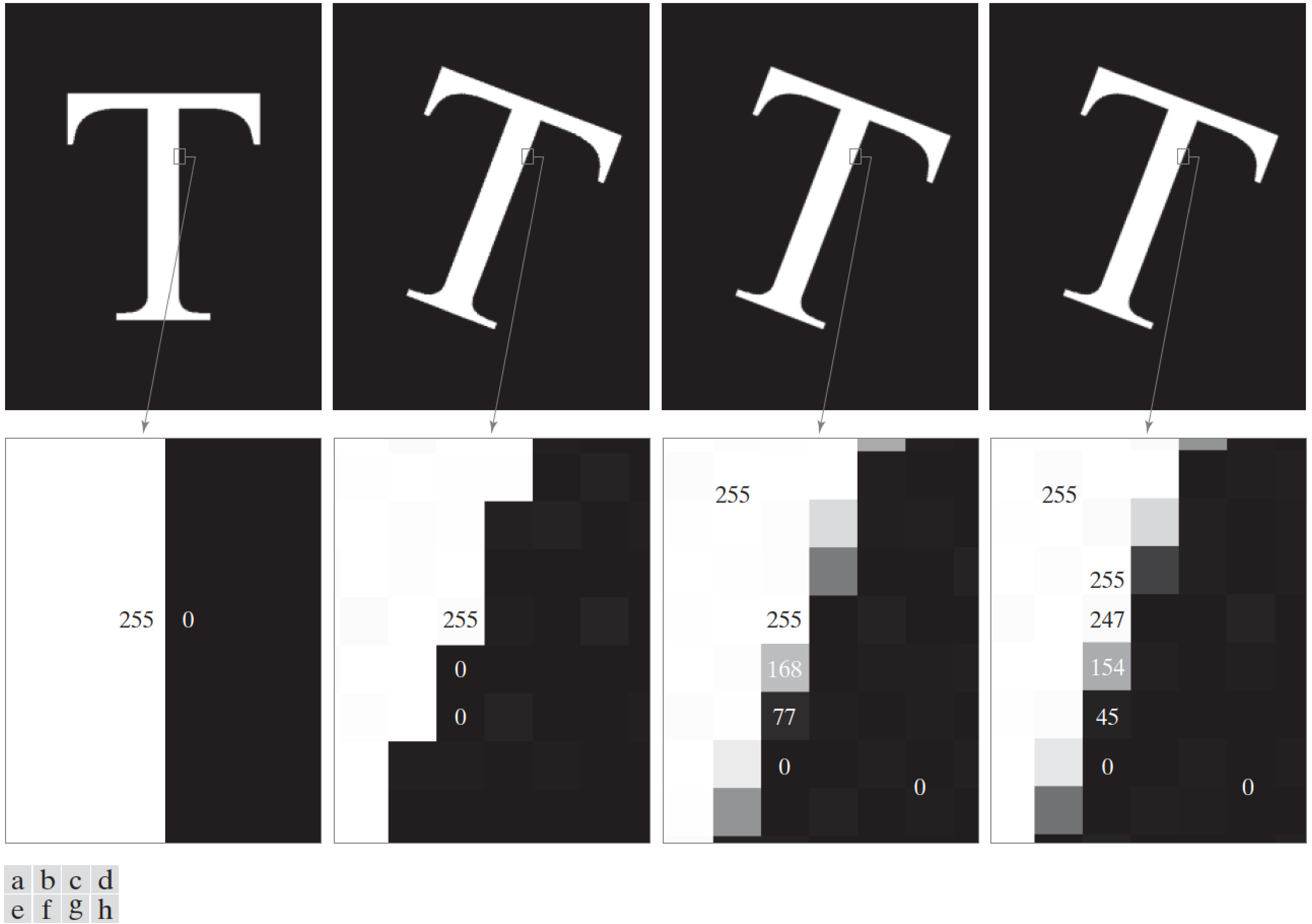


FIGURE 2.40 (a) A 541 × 421 image of the letter T. (b) Image rotated −21° using nearest-neighbor interpolation for intensity assignments. (c) Image rotated −21° using bilinear interpolation. (d) Image rotated −21° using bicubic interpolation. (e)-(h) Zoomed sections (each square is one pixel, and the numbers shown are intensity values).

Geometric transformations

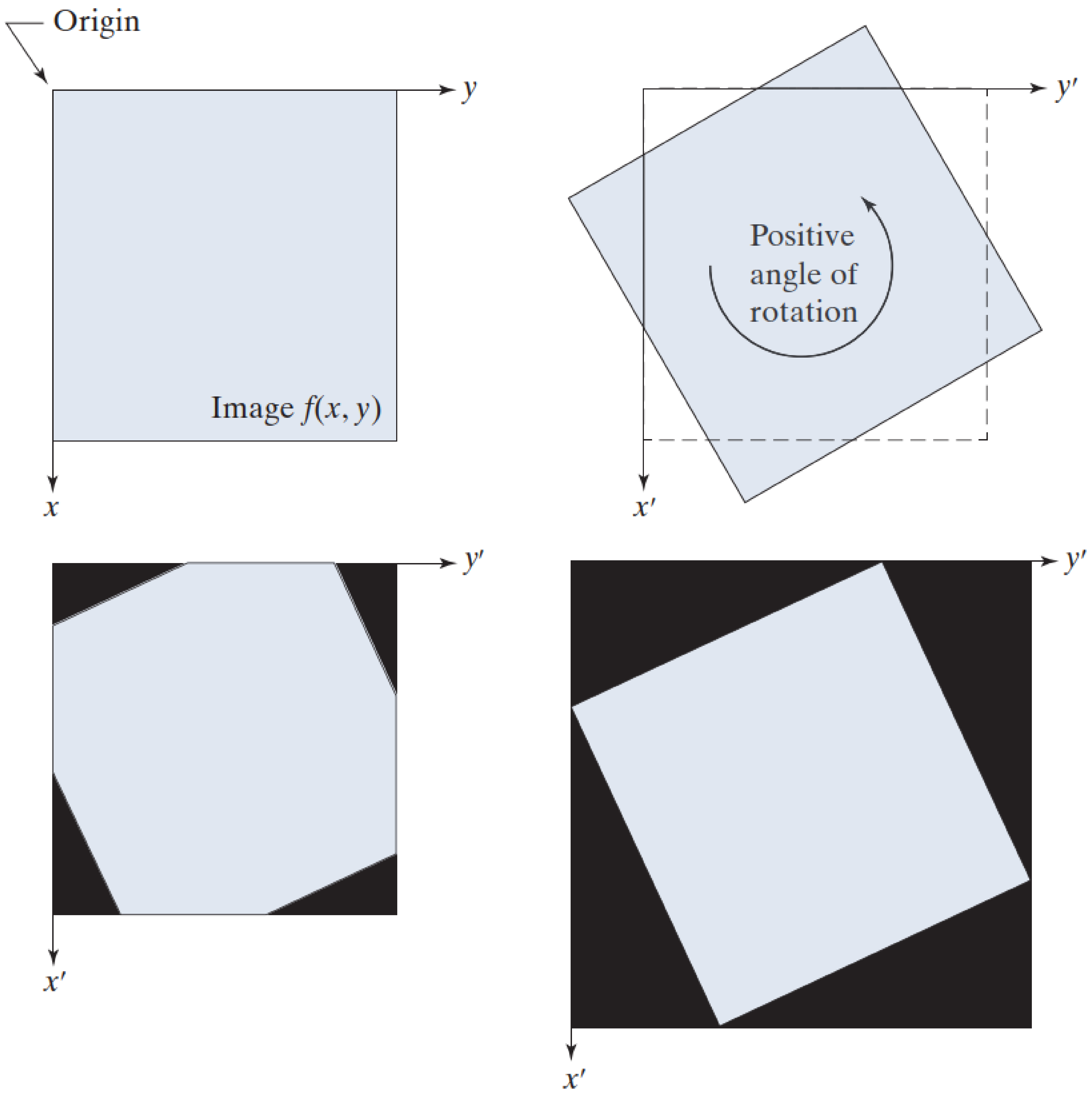


FIGURE 2.41
(a) A digital image.
(b) Rotated image (note the counterclockwise direction for a positive angle of rotation).
(c) Rotated image cropped to fit the same area as the original image.
(d) Image enlarged to accommodate the entire rotated image.

Study:

- Rafael Gonzalez, Richard Woods, “Digital Image Processing”, 4th edition, Pearson, 2018
  - Chapter 2.1, 2.2, 2.3, 2.4, 2.5 (geometric transformations)