Image Processing

Waleed A. Yousef
Faculty of Computers and Information,
Helwan University.

March 7, 2010

Chapter 1

Digital Image Fundamentals

Objectives:

1. Introductory concepts.
2. Human perception and visual system.
3. Image acquisition and sensing.
4. Sampling and Gray levels.
5. What is a linear operation?

2.1 Elements of Visual Perception

Human visual perception system has limitations; we have to understand these to deal with images and their processing correctly. So, we will give a very little introduction on that.

2.1.1 Structure of the Human Eye

Eye is a camera with lens and detector (like a film); processing is done in the brain not in eye.

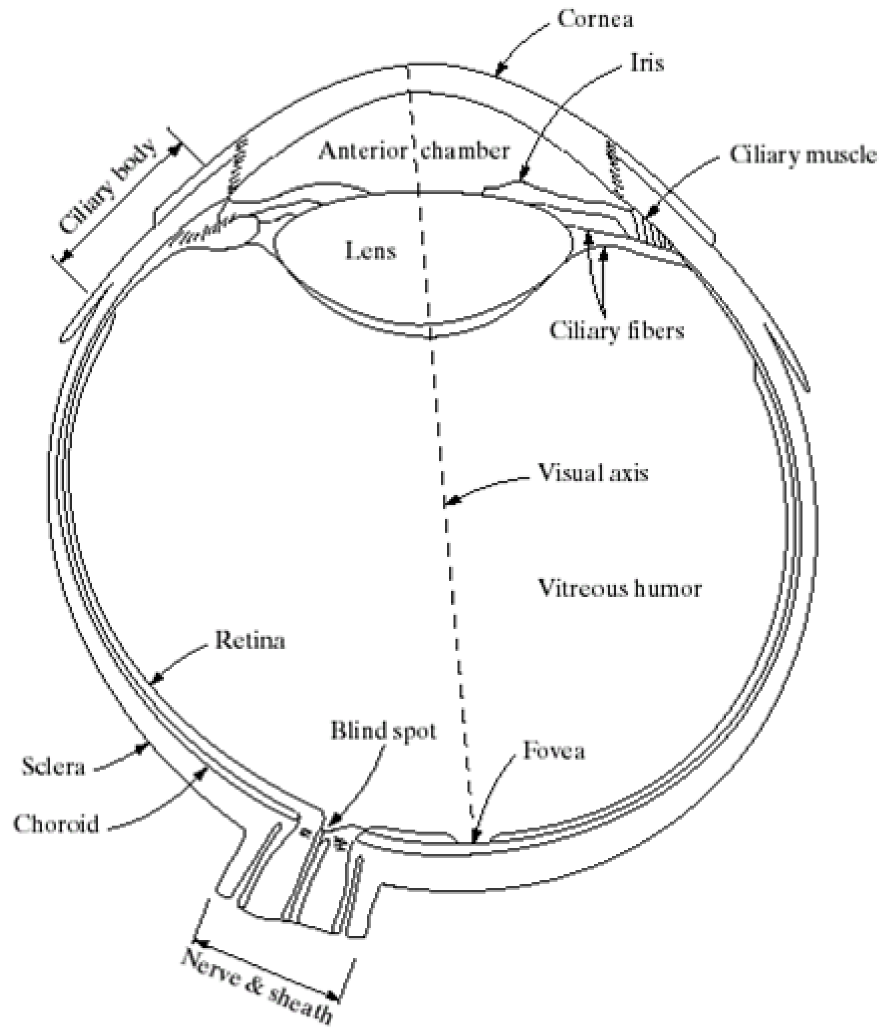


FIGURE 2.1 Simplified diagram of a cross section of the human eye.

Two kinds of receptors in eyes: cones and rods.

6 to 7 millions cones in each eye, almost in the center of the retina, sensitive to colors, each one is connected to its own nerve (*Photopic Vision*)

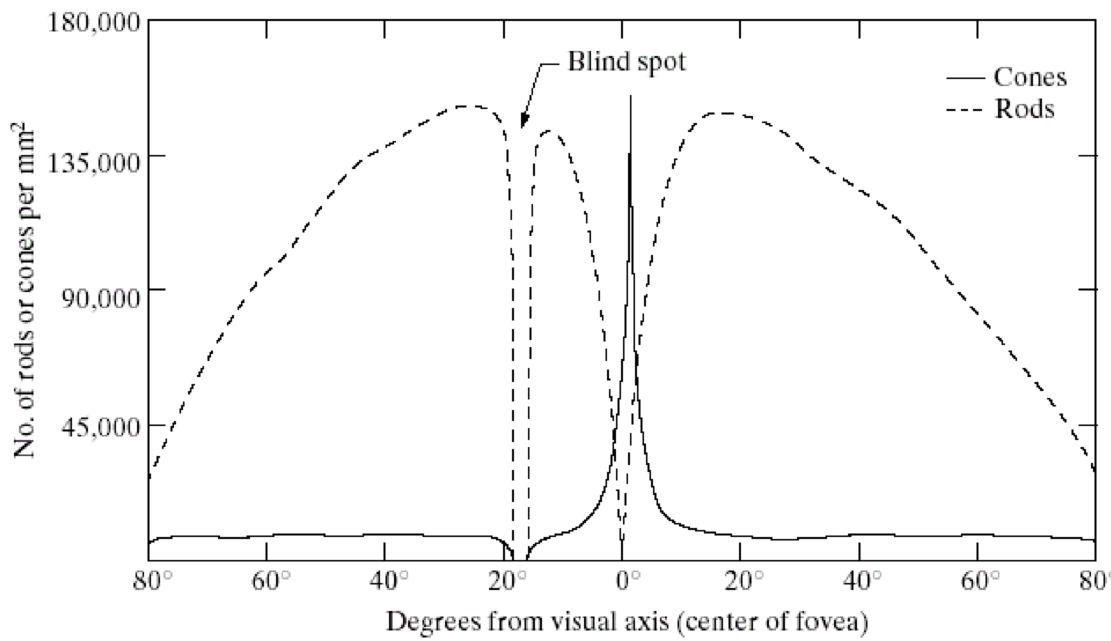


FIGURE 2.2 Distribution of rods and cones in the retina.

75 to 150 million rods per eye, distributed over the retina, all are connected to one nerve.
In moonlight only rods will be stimulated and you can see only shapes with no color (*Scotopic Vision*)

2.1.2 Image Formation in the Eye

The principal difference between the lens of the eye and an ordinary optical lens is that the ormer is flexible.

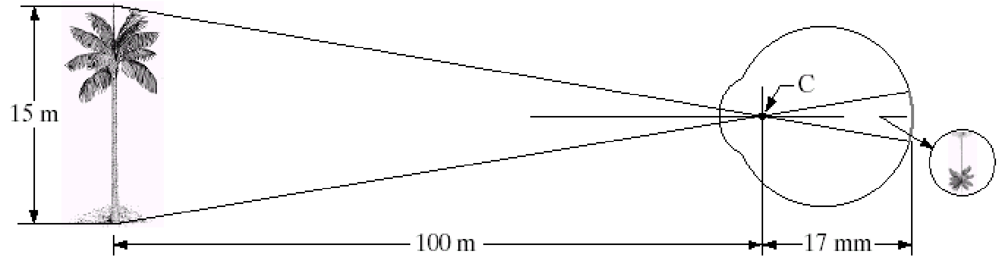


FIGURE 2.3 Graphical representation of the eye looking at a palm tree. Point *C* is the optical center of the lens.

2.1.3 Brightness Adaptation and Discrimination

Eye can discriminate 10^{10} difference intensity levels, ranging from scotopic threshold to the glare limit. It does not adapt but to portion of that range at a time.

Experiments show that the *subjective brightness* (intensity as *perceived* by human visual system) is logarithmic function of actual intensity (see the figure).

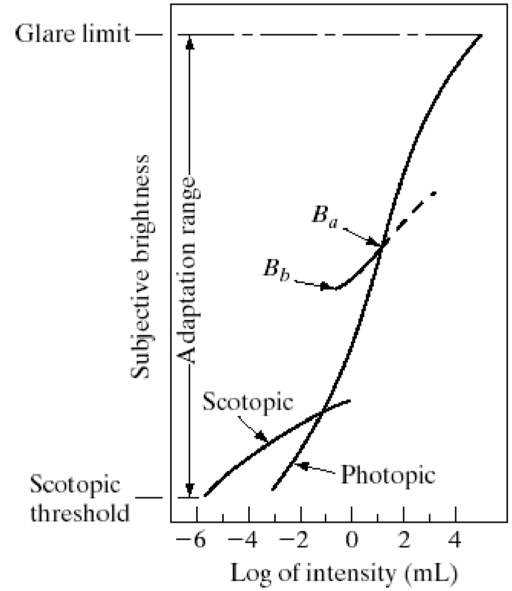
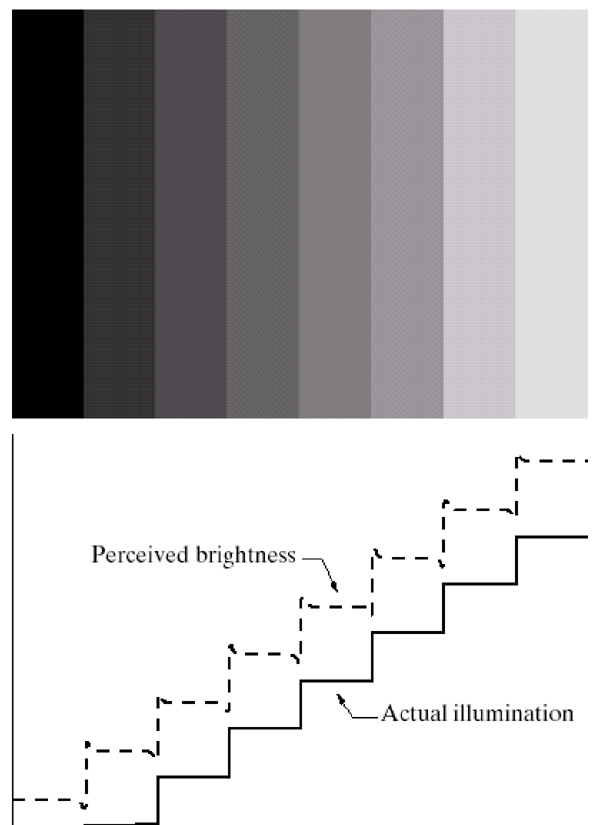


FIGURE 2.4 Range of subjective brightness sensations showing a particular adaptation level.

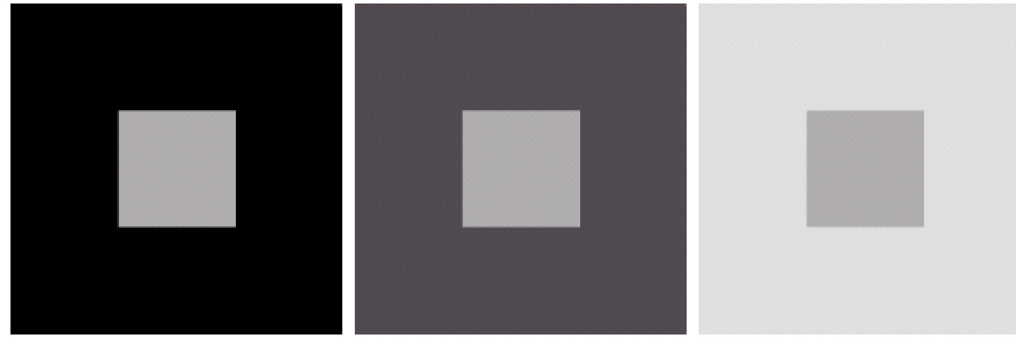
Subjective brightness vary, as well, according to context. The stripes in the figure have constant intensity; however we perceive them as if there is more intensity at the edges! This is known as *Mach bands*, named after Ernst Mach (circa 1865).



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.

Another example is the *Simultaneous contrast*. All the inner squares have the same intensity and gray level; however we perceive them differently because of they have different surroundings.



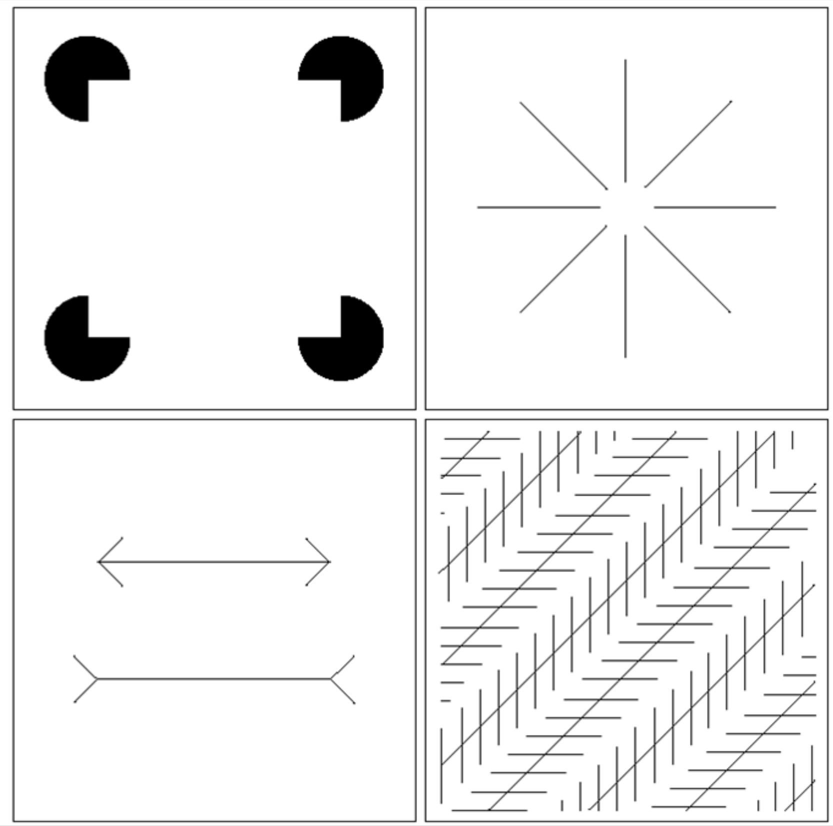
a b c

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 illusion is another example of contextual perception.

a b
c d

FIGURE 2.9 Some well-known optical illusions.



2.3 Light and the Electromagnetic Spectrum

$$\lambda = \frac{c}{\nu},$$
$$c = 2.998 \times 10^8 m/s,$$
$$E = h\nu$$

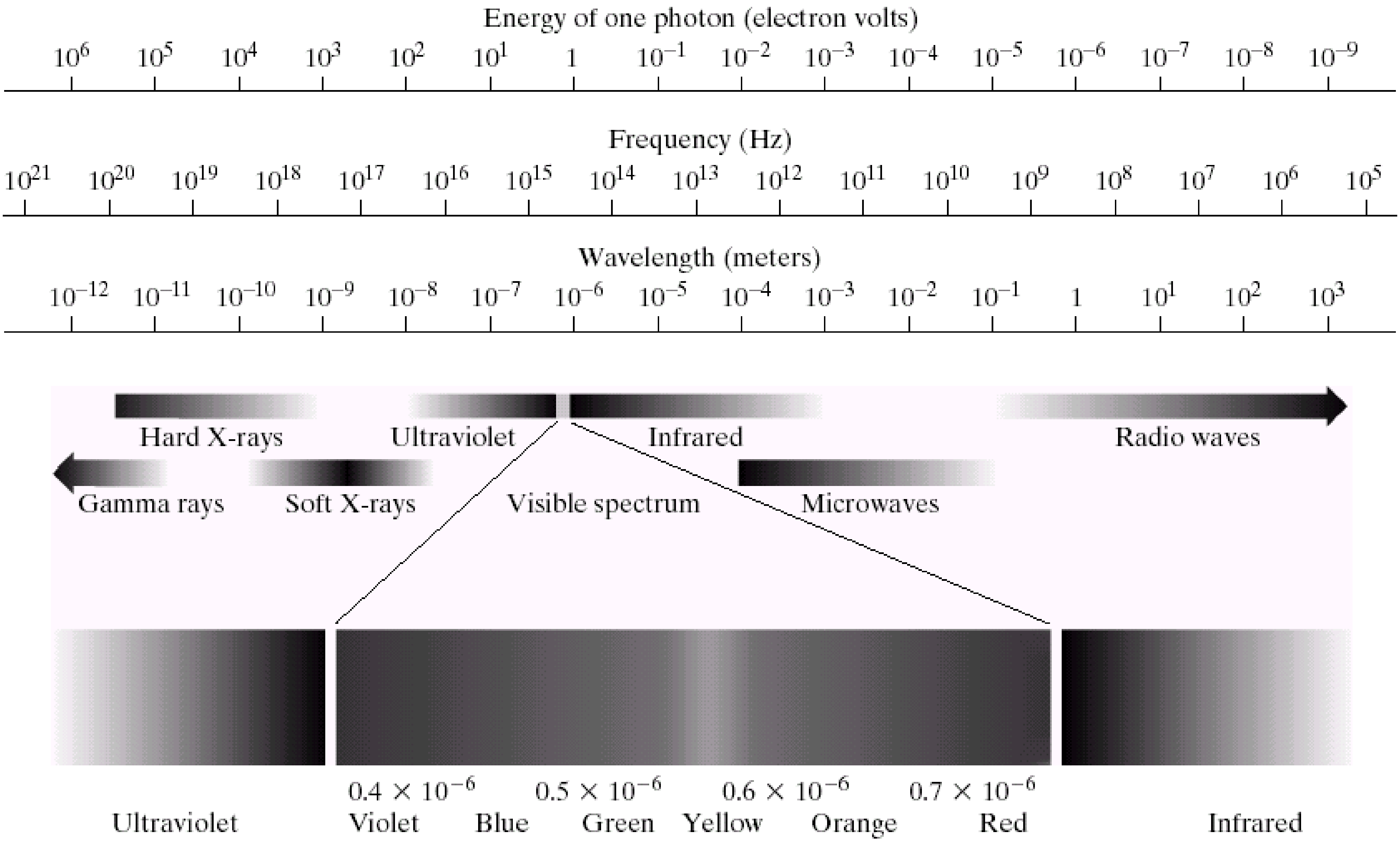


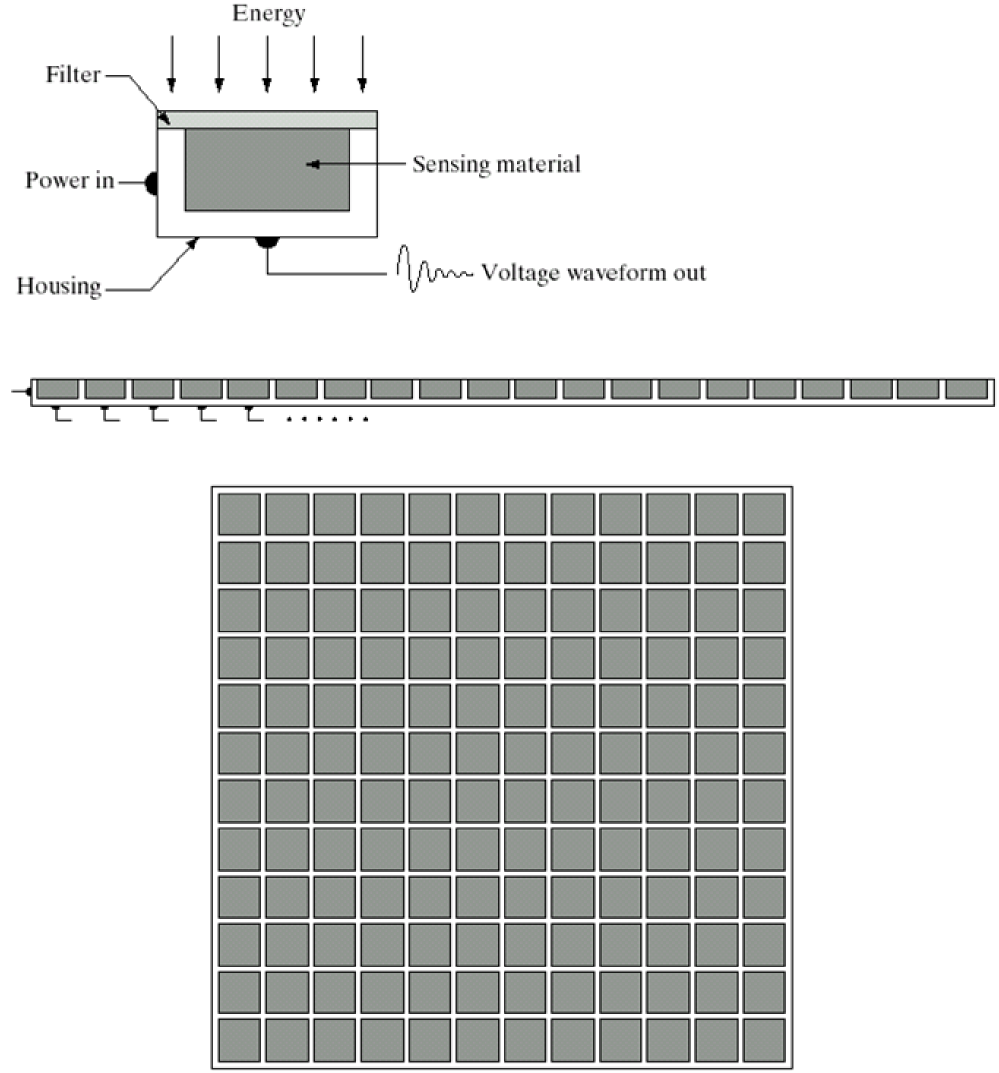
FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

If light is colorless, it is called *achromatic* or *monochromatic* light. The only attribute will be intensity or gray level.

Three basic quantities used to describe the quality of chromatic light:
Radiance: total amount of energy flows from the light source, measured in watts (W).
Luminance: amount of energy an observer *perceives* from the light source, measured in lumens (lm). Infrared has significant radiance but hardly perceived by humans; therefore it has low luminance.
Brightness: this is subjective measure of the light.

2.3 Image Sensing and Acquisition

The energy of incident EM wave is absorbed by photo-converter, then converted to electrical waveform carrying the information.



a
b
c

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

2.3.1 Image Acquisition Using a Single Sensor

A single sensor (as in Figure 2.12 (a)) can generate 2-D image by rotation (Figure 2.13). This method is inexpensive but slow

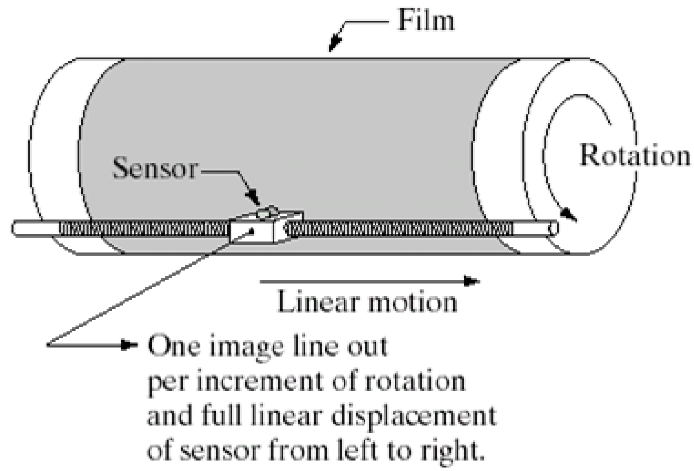


FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

2.3.2 Image Acquisition Using a Sensor Strips

Using sensor strip (as in Figure 2.12 (b)), to acquire 2-D image, we need then motion in one direction (Figure 2.14 (a)). This is used in all flatbed scanners.

To acquire 3-D image, e.g., in CT scan or MRI, we need two motion directions (Figure 2.14 (b)): (1) a rotating X-ray source provides illumination and the portion of the sensors opposite to the source collects the X-ray. (2) a linear motion of the 3-D object, which is the patient.

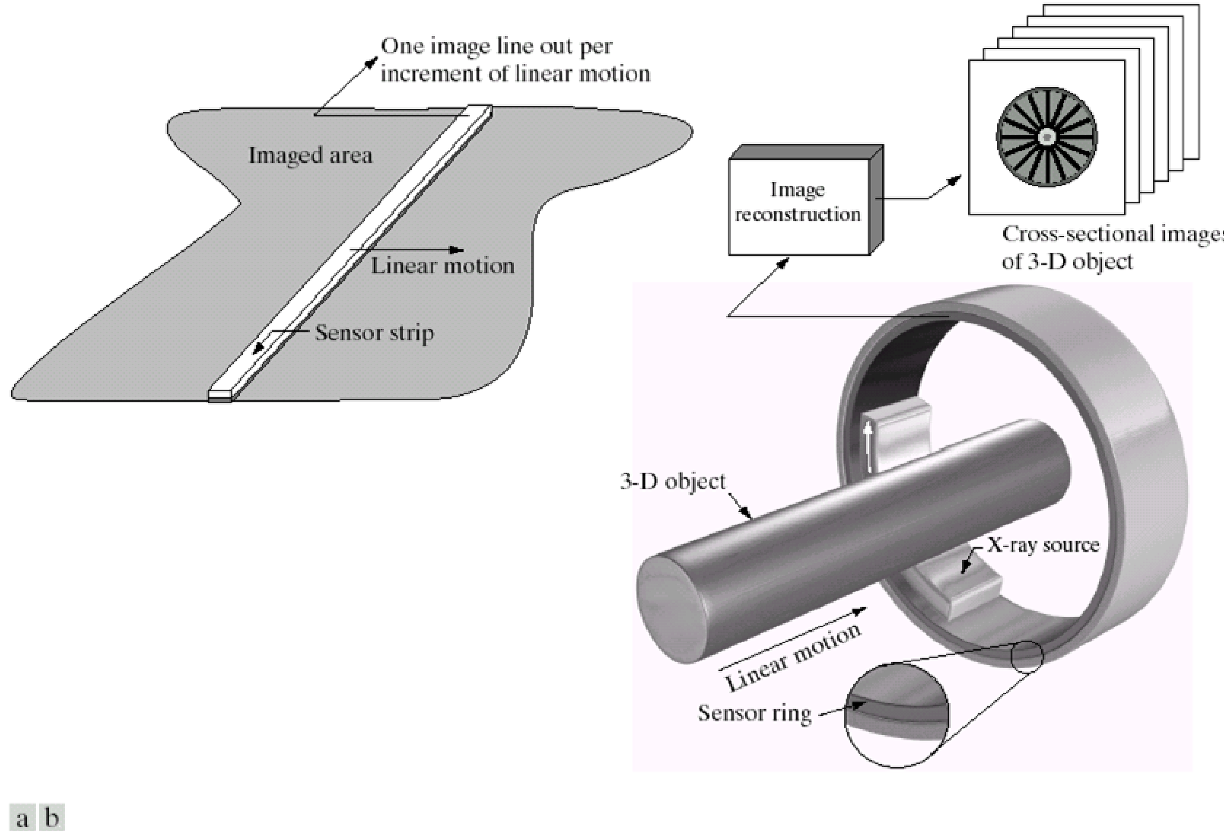


FIGURE 2.14 (a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.

2.3.3 Image Acquisition Using Sensor Arrays

This is almost the technology used, e.g., in digital cameras. Typically, 4000×4000 elements are arrayed (See figure). An advantage is that no motion is required.

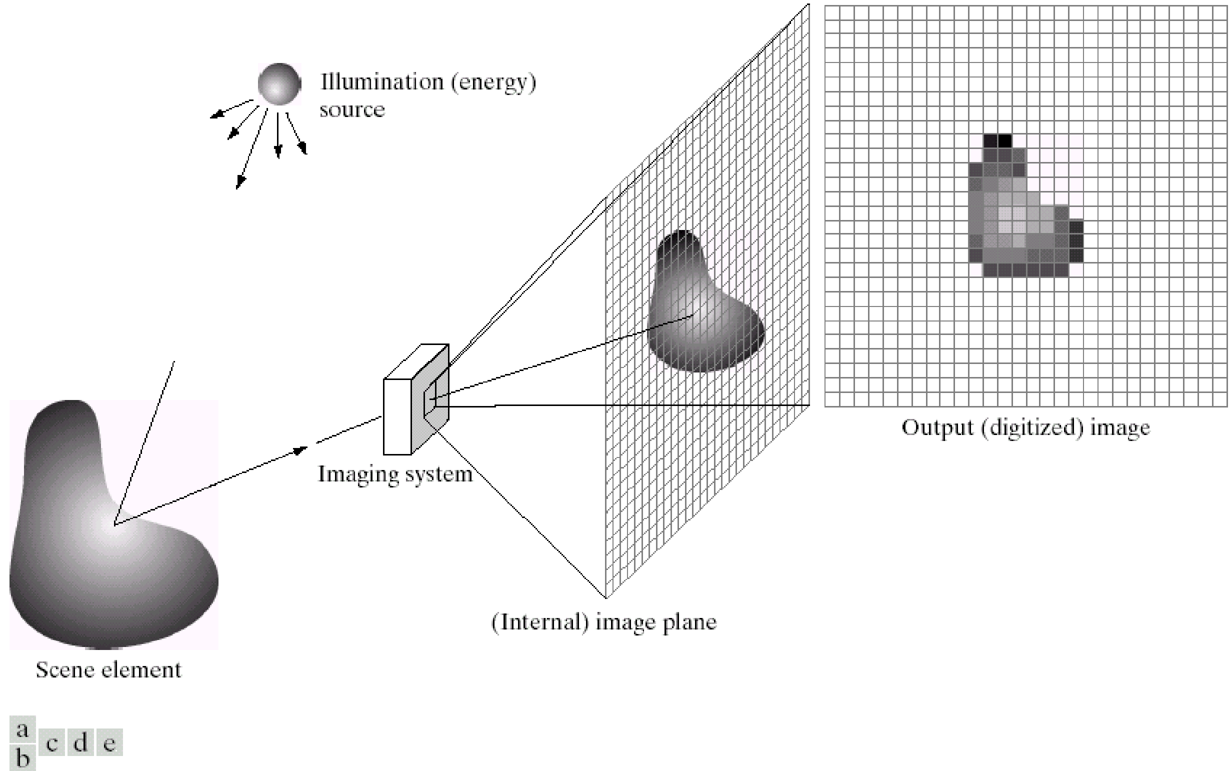


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

2.3.4 A Simple Image Formation Model

The value or amplitude of a monochromatic image at the coordinates (x, y) is:

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

where the illumination of the source is $i(x,y)$ and the reflectance from the object is $r(x,y)$. Therefore

$$f(x,y) = i(x,y)\, r(x,y),$$
$$0 \leq r(x,y) \leq 1.$$

Common practice is to consider $f(x,y)$ be on the scale $[0, L-1]$, where 0 indicates black and $L-1$ indicates white.

2.4.1 Basic Concepts in Sampling and Quantization

Digitization of $f(x, y)$ is to be able to process digitally on a computer. Digitizing coordinates (the x and y) is called *sampling*; digitizing amplitude (f itself) is called *quantization.* Both are in the figure.

Sampling can be applied to analogue signal, e.g., an analogue signal on a film. Or, from the beginning, by the sensor arrangement:

Single Sensor: the step size in both directions determines sampling; the step size should consider the limitations of optical sensor.

Line Sensor: the number of sensors on the strip determines the sampling in that directions; the step increment determines the sampling in the other direction.

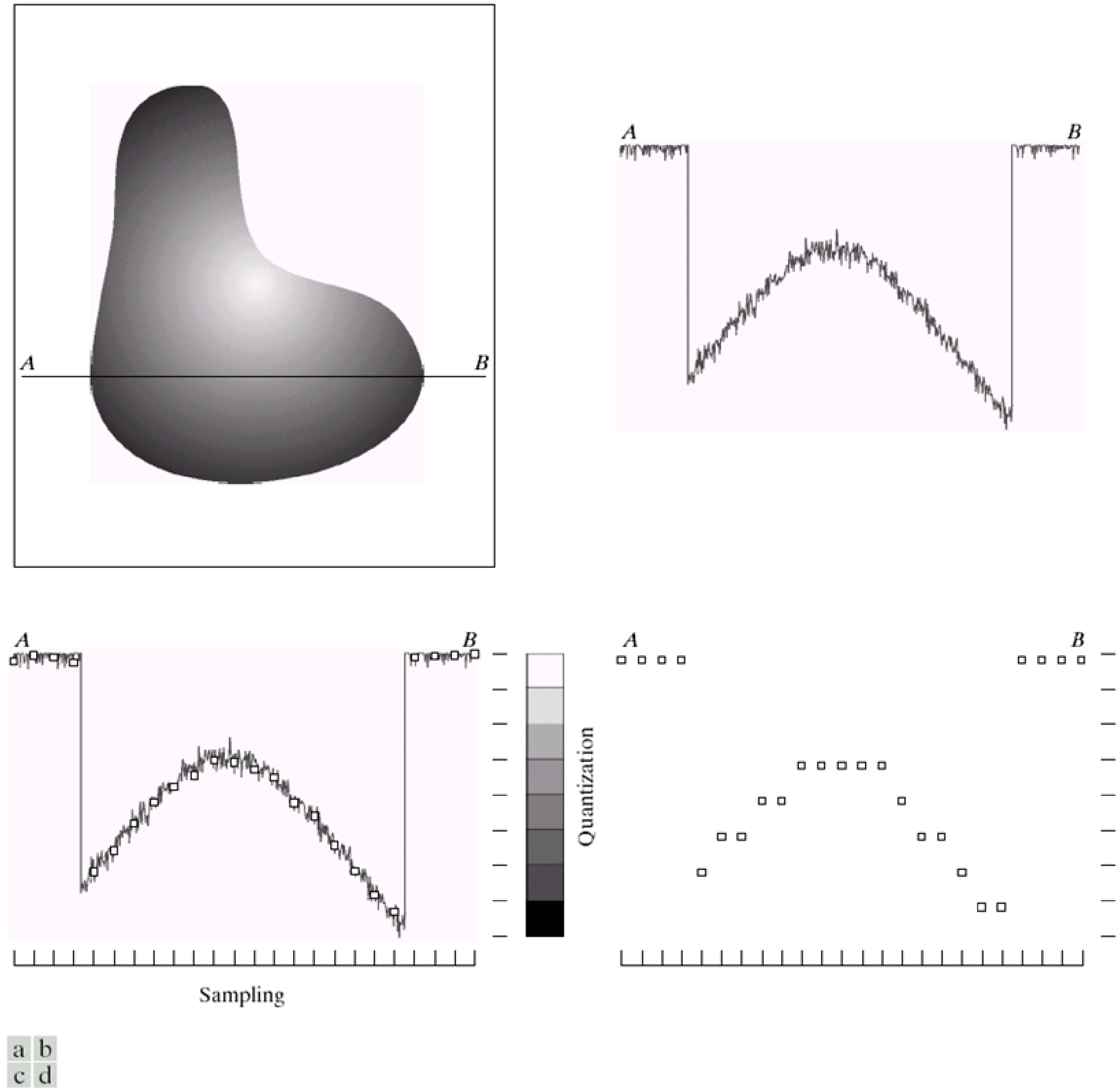
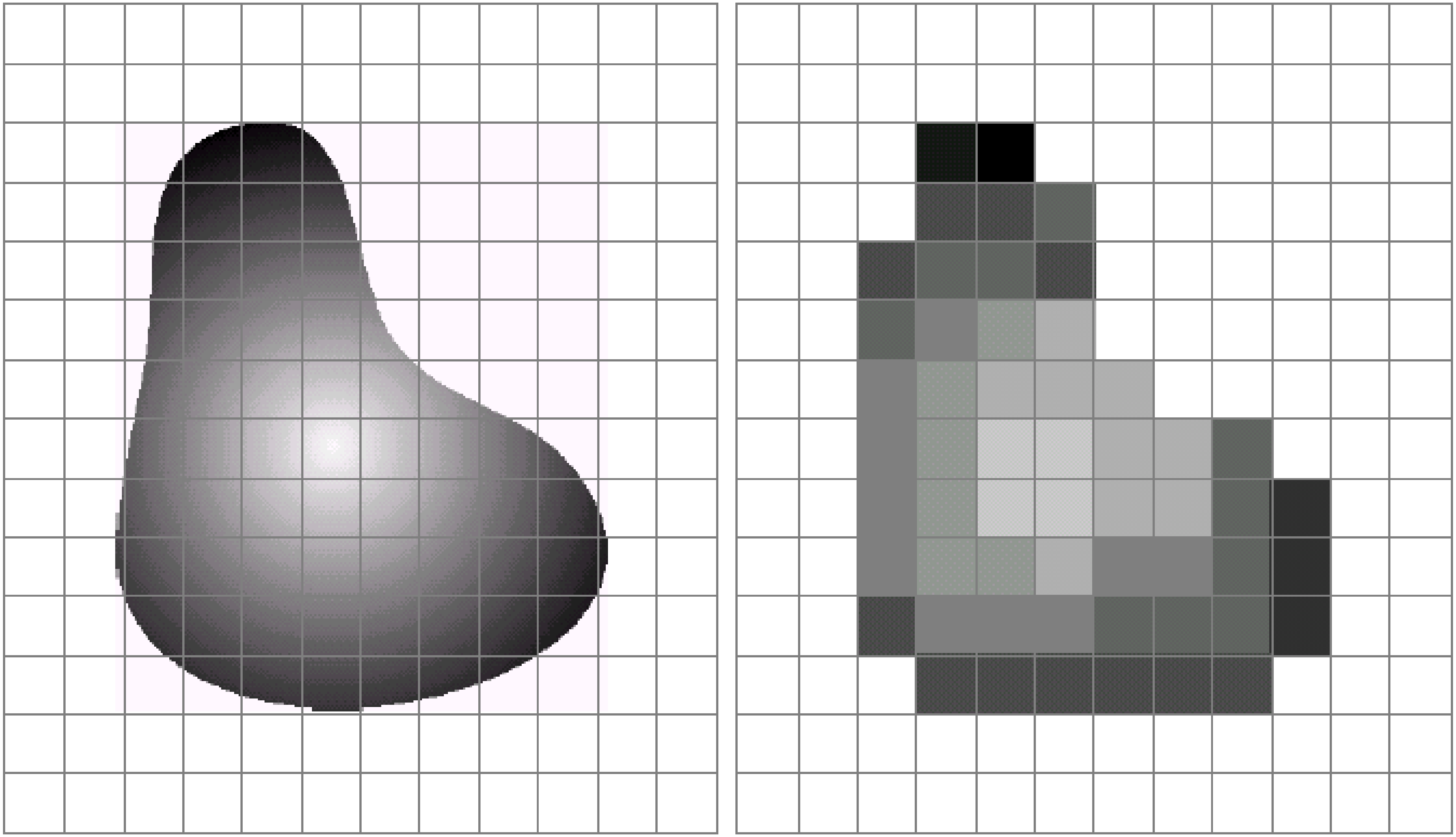


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

Array Sensor: the number of sensors determine sampling in both directions (Fig. 2.17)



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

Notation: $f(x,y)$ has M rows and N columns. The element $(0,1)$ is the second element in the first row; 0 and 1 are not the actual distances on the image.

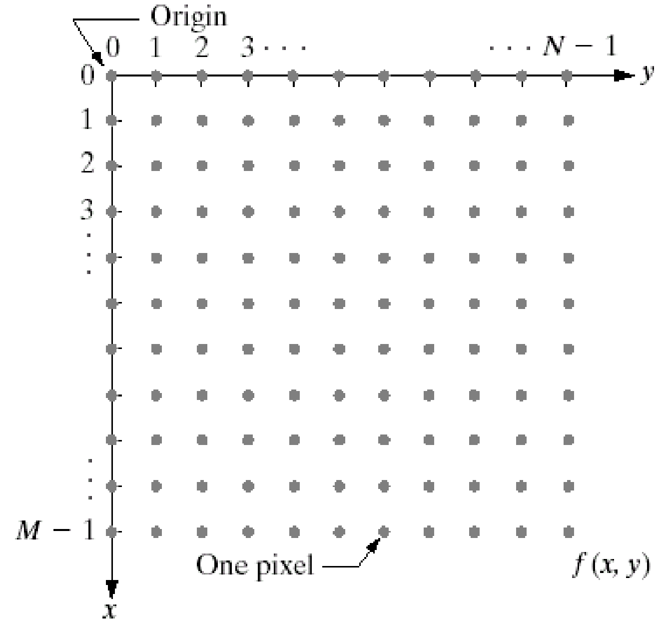


FIGURE 2.18 Coordinate convention used in this book to represent digital images.

When indexing starts with 0, this is consistent with languages like C. In Matlab indexing starts with 1.

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

The values of f are discrete, usually 2^k levels (L) represented by k bits (called k-bit image)

$$L = 2^k$$

Therefore, mathematically

$$f : \mathcal{Z}^2 \rightarrow \mathcal{Z}.$$

Total number of bits required to represent an image is

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

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 determines the spatial resolution, and quantization determines gray-level resolution

Spatial Resolution: is the smallest discernible details in an image. If we have vertical lines, each has a width W followed by a space W, then a line pair is a line along with its space. Resolution is measured in lines pairs per unit length.

Gray-Level Resolution: is the number of gray levels, i.e., 2^k. Typically $k = 8$ or 16.

Examples:

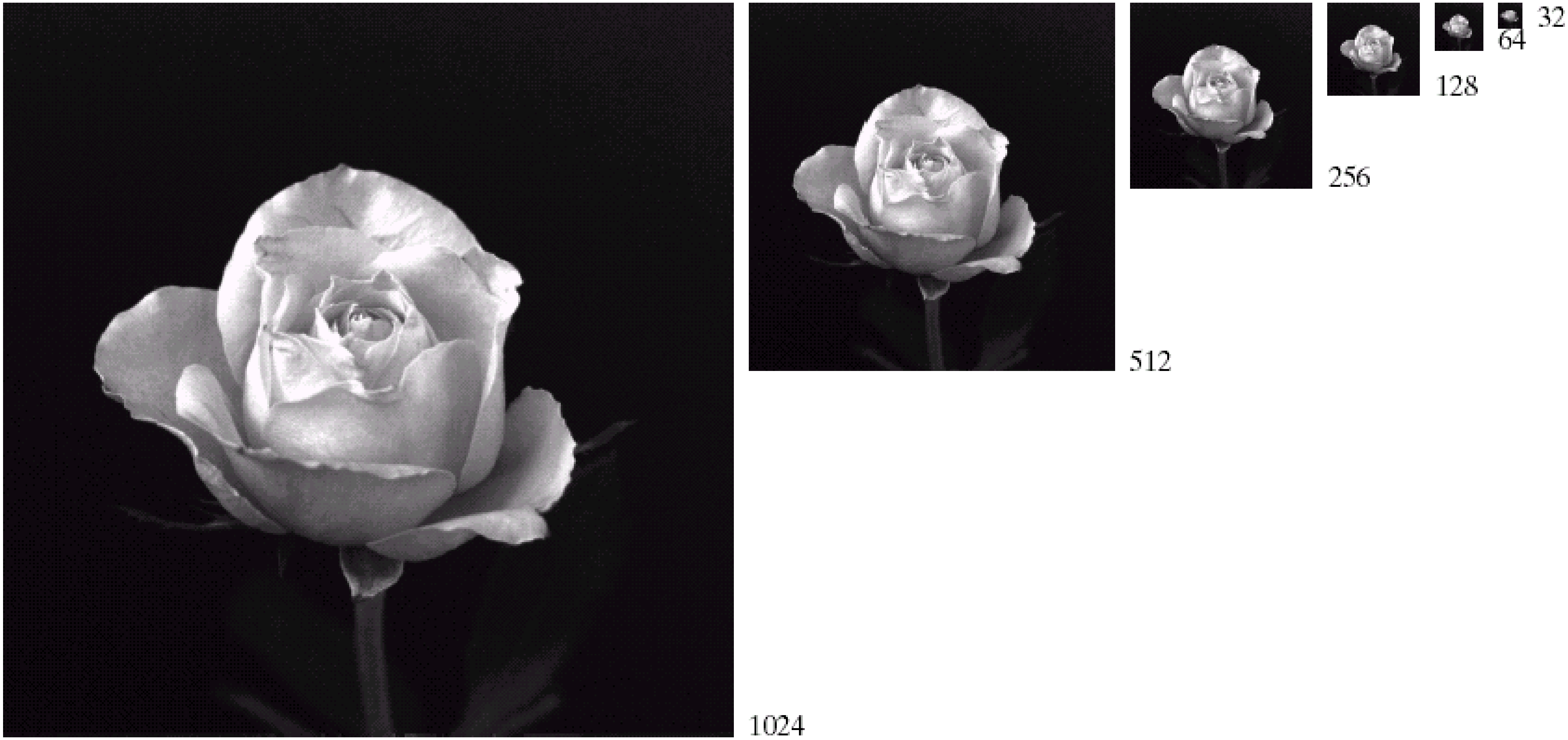
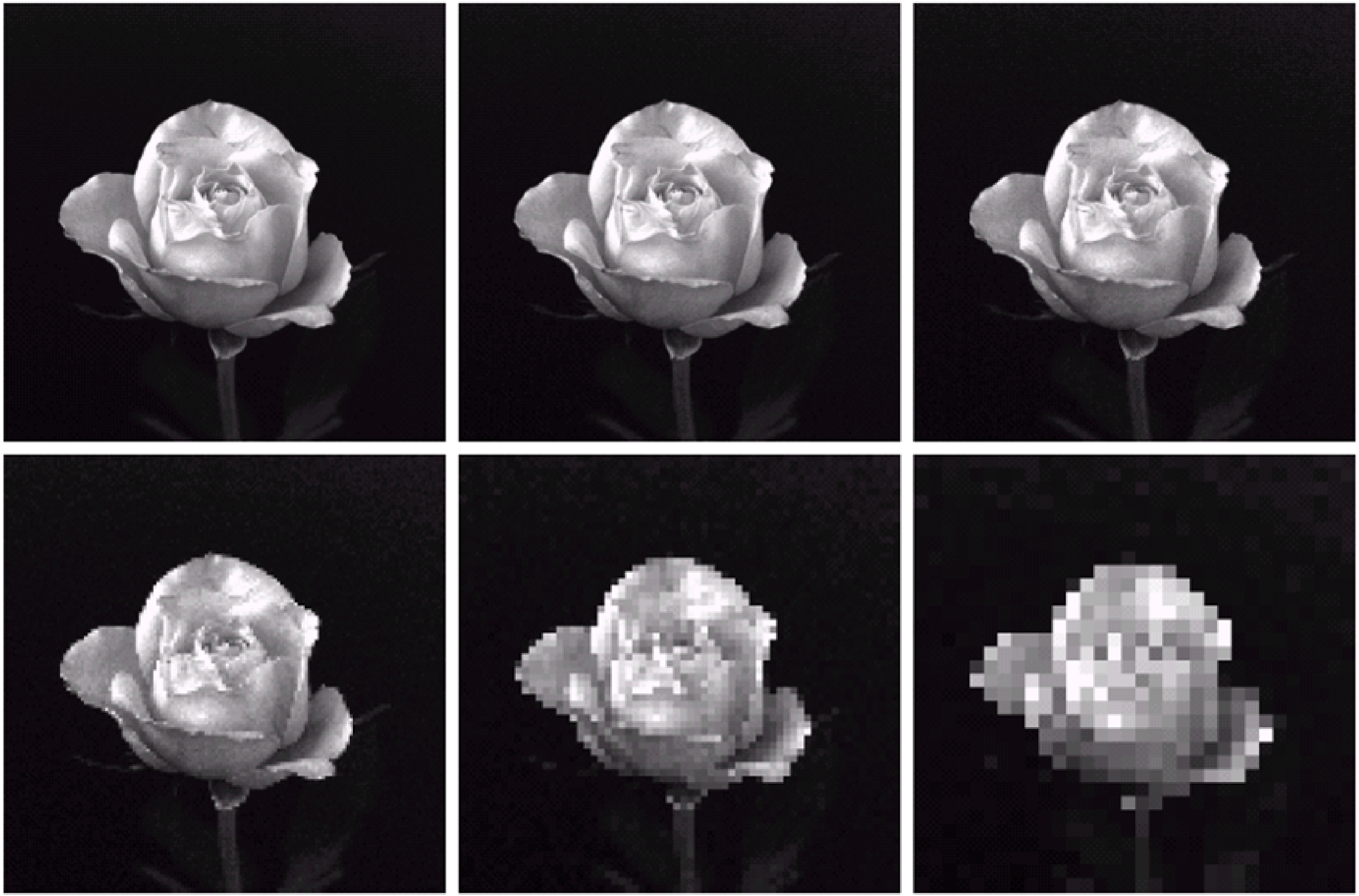


FIGURE 2.19 A 1024×1024, 8-bit image subsampled down to size 32×32 pixels. The number of allowable gray levels was kept at 256.

In this example, we keep the gray-level resolution fixed. Take out rows and columns to sample the picture down, e.g.:

```
fs = f(1:4:end, 1:4:end)
```

This is lossy of course. Let's reconstruct the sampled pictures by resampling up:

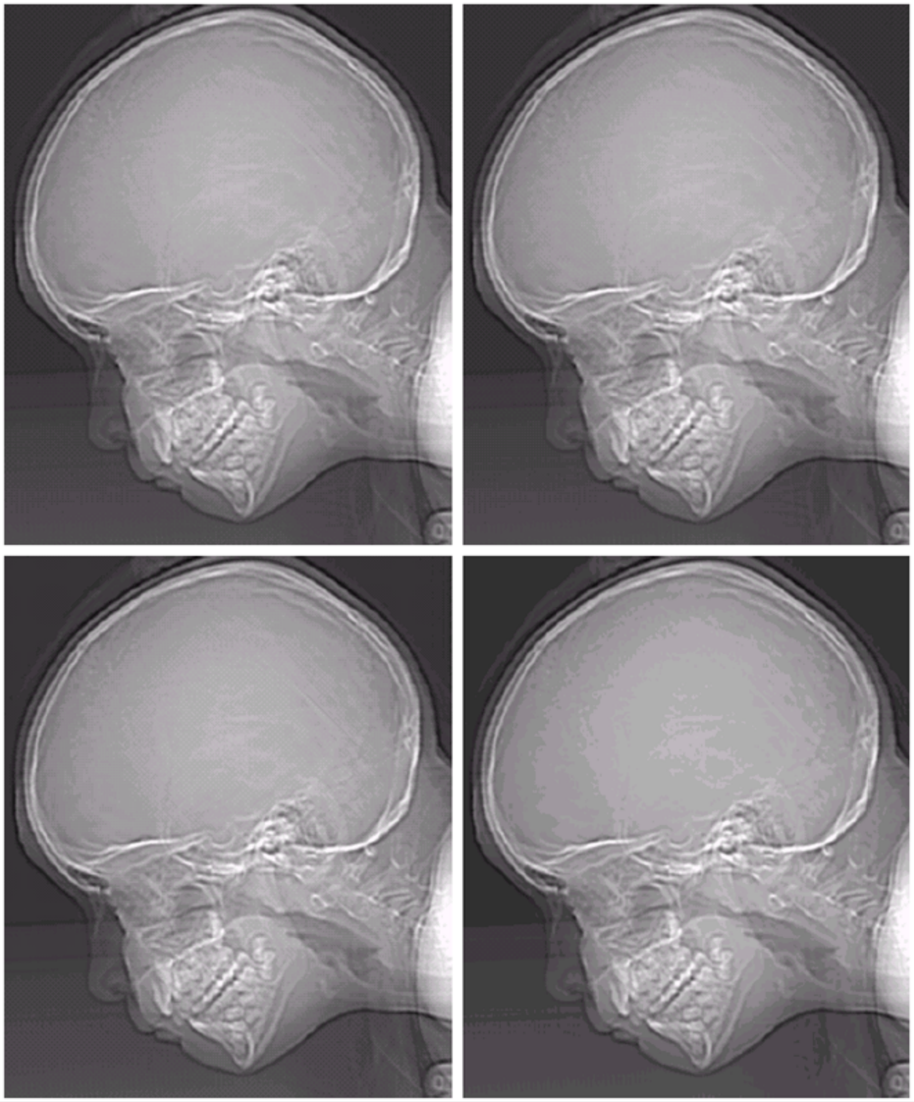


a b c
d e f

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.

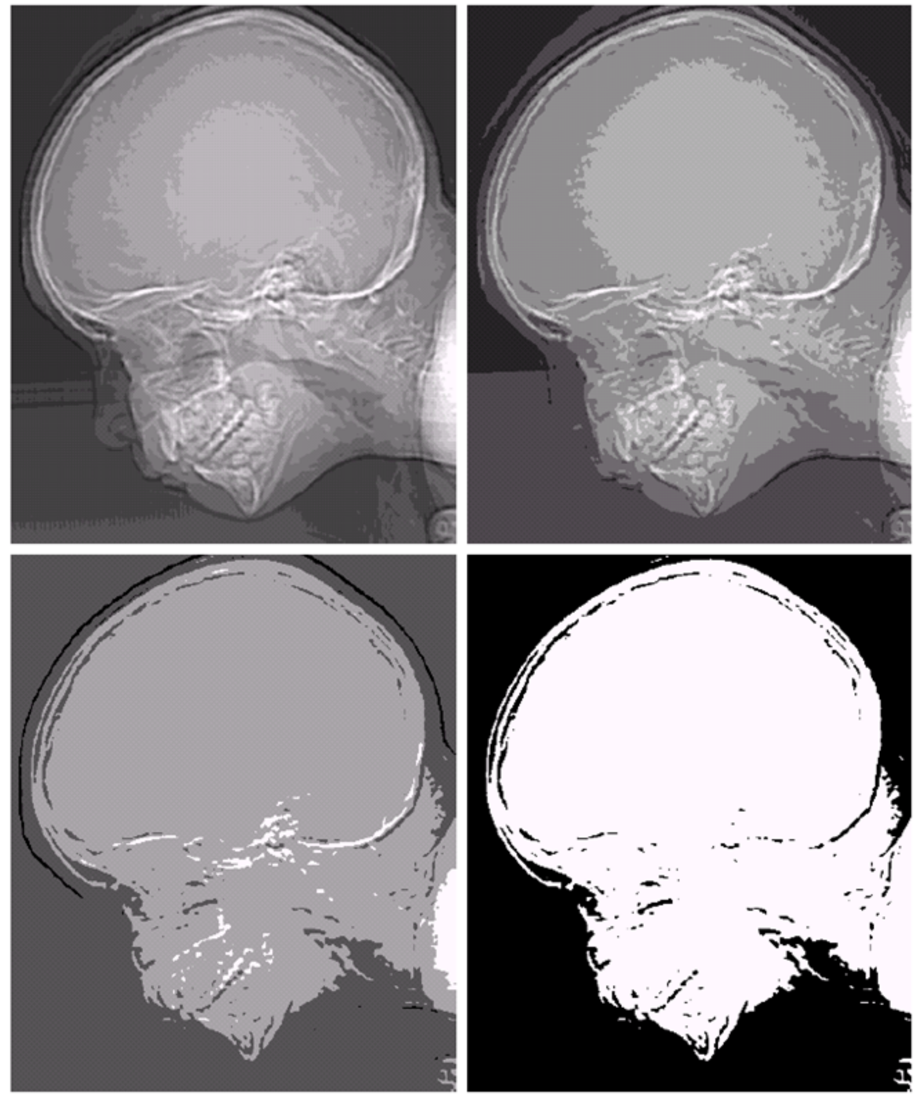
We can re-sample the picture up by replicating the rows and columns. This brings the picture back to the same spatial size but, of course, with less resolution (checkerboard effect). Write a code for it.

In this example we keep the spatial resolution fixed and change the gray-level (the number of bits). This will result in false contouring.



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.



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.)

2.4.5 Zooming and Shrinking Digital Imagies

Original image: $g(x', y')$ with M' rows and N' columns. Transformed image: $f(x, y)$ with M rows and N columns. The new pixel (x, y) is correspondent to the old pixel:

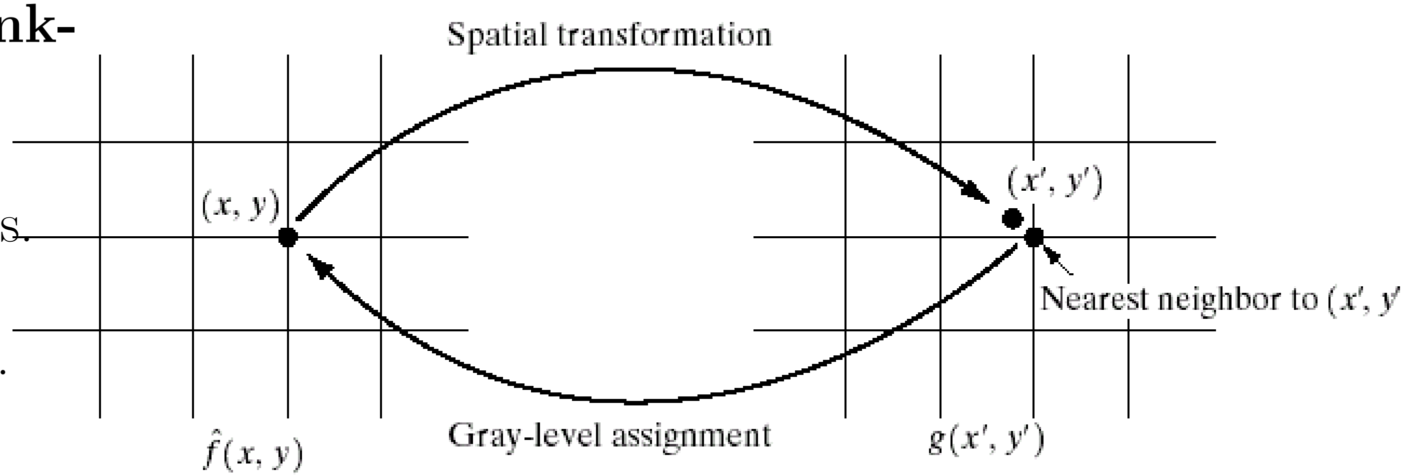


FIGURE 5.33 Gray-level interpolation based on the nearest neighbor concept.

$$(x', y') = \left(x\frac{M'}{M}, y\frac{N'}{N}\right)$$

For nearest neighbor interpolation

$$f(x, y) = g([(x', y')]),$$

where $[\cdot]$ is the rounding function.

Nearest neighbor is equivalent to row deletion (or replication) when $M = \frac{M'}{2^c}$ (or $M = M'2^c$) respectively.

For bilinear interpolation, we assume that

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

where (x', y') takes the four vlaues $(\lfloor x' \rfloor, \lfloor y' \rfloor), (\lfloor x' \rfloor, \lceil y' \rceil), (\lceil x' \rceil, \lfloor y' \rfloor)$, and $(\lceil x' \rceil, \lceil y' \rceil)$. Then we solve for a, b, c, and d. Then

$$f(x, y) = g(x', y')$$

The improvements in overall appearance in case of bilinear is clear for figures d and e. Figure f is blurry, but the image was zoomed 32 times from 1024×1024 to 32×32.

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 Some Basic Relationships Between Pixels

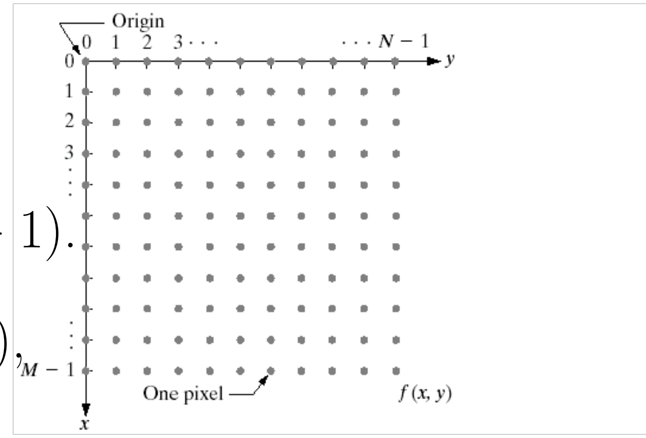
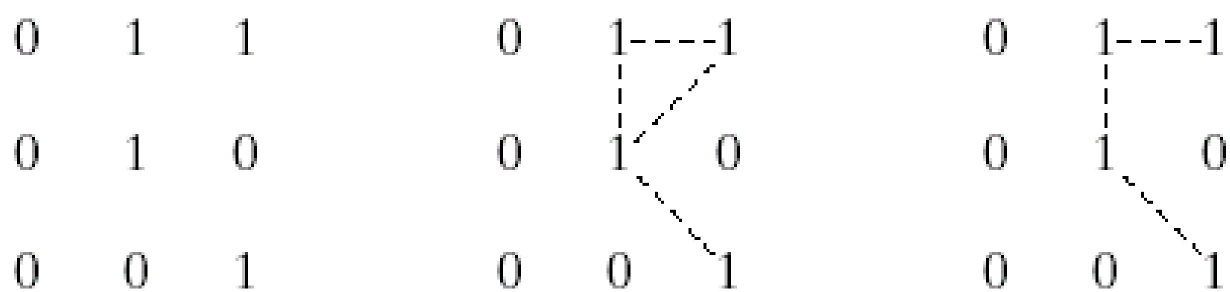


FIGURE 2.18 Coordinate convention used in this book to represent digital images.

Neighbors of a Pixel

The pixel $p = (x, y)$ has:
neighbors $N_{4(p)}$: $(x-1, y)$, $(x+1, y)$, $(x, y+1)$, $(x, y-1)$.

Diagonal neighbors $N_D(p)$: $(x+1, y+1)$, $(x+1, y-1)$, $(x, -1y+1)$, $(x-1, y-1)$.
8-neighbors $N_8(p) = N_4(p) \cup N_D(p)$.



a b c

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

Adjacency

For a set V of gray level values, the pixels $p \in V$ and $q \in V$ are said to be
4-adjacent: if $q \in N_4(p)$.
8-adjacent: if $q \in N_8(p)$.
m-adjacent: if $q \in N_4(p)$ or $(q \in N_D(p)$ and $N_4(p) \cap N_4(q) \notin V)$.
The 1 and 0 values in the figure indicates points belonging (or not belonging) to the set V.

Path (or curve)

is a sequence of pixels $(x_1, y_1), \ldots, (x_n, y_n)$, where the pixels (x_i, y_i) and (x_{i-1}, y_{i-1}) are adjacent (we have to define which adjacency); and is called closed if $(x_1, y_1) = (x_n, y_n)$.

Connectivity

For a set S of pixels, two pixels p and q are said to be connected in S if they are connected through a path belonging entirely to S

Boundary (or border or contour) of a region R

is the set of pixels belonging to R but with at least one neighbor pixel that does not belong to R.

2.5.3 Distance Measures

For any points p, q, and z (pixels in our context) belonging to a set S, the distance function D is called a metric and the tuple (S, D) is called a metric space if

1. $D(p,q) \geq 0$ with equality iff $p = q$,
2. $D(p,q) = D(q,p)$, and
3. $D(p,z) \leq D(p,q) + D(q,z)$

Prove that the the following distance function are metrics over $\mathcal{R}$

Euclidean distance

$$D_e(p,q) = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2}$$

City-block distance

$$D_4(p,q) = |x_p - x_s| + |y_p - y_s|$$

$$\begin{array}{ccccc} & & 2 & & \\ & 2 & 1 & 2 & \\ 2 & 1 & 0 & 1 & 2 \\ & 2 & 1 & 2 & \\ & & 2 & & \end{array}$$

Chessboard distance

$$D_8(p,q) = \max\left(|x_p - x_s|, |y_p - y_s|\right)$$

$$\begin{array}{ccccc} 2 & 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{array}$$

2.5.4 Image Operations on a Pixel Basis

If f and g are two images then for any binary operation $\odot$ defined on $\mathcal{R}$,we define $f \odot g$ as an image h given by

$$h\left(x, y\right) = f\left(x, y\right) \odot g\left(x, y\right) \;\; \forall x, y,$$

2.6 Linear and Nonlinear Operations

By definition, the operator $\mathcal{H}$ is a linear operator if

$$\mathcal{H}\left(af + bg\right) = a\mathcal{H}\left(f\right) + b\mathcal{H}\left(g\right).$$

In our context f and g are images.