Chapter 2 Digital Image Fundamentals

The purpose of this chapter is to introduce a number of basic concepts in digital image processing, which will be used throughout this course.

2.1 Elements of Visual Perception

Although the filed of digital image processing is built on a foundation of mathematical formulations, human intuition and analysis play a central role in the choice of one technique versus another, which often is made based on subjective and visual judgments.

We focus our interest in the mechanics and parameters related to how images are formed and perceived by humans.

Structure of the Human Eye

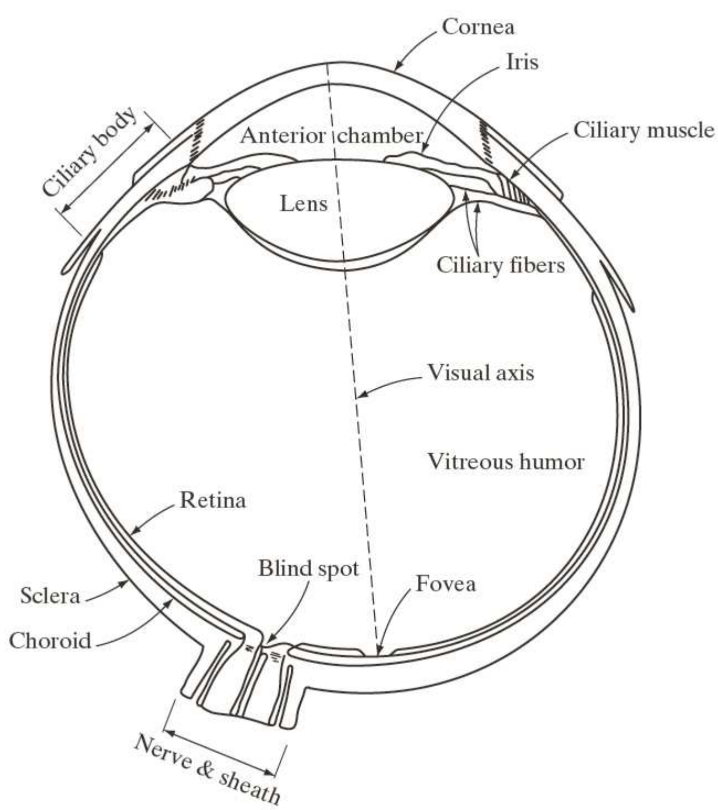


FIGURE 2.1 Simplified diagram of a cross section of the human eye.

Image of Formation in the Eye

In the human eye, the distance between the lens and the imaging region (the retina) is fixed, and the focal length needed to achieve proper focus is obtained by varying the shape of the lens.

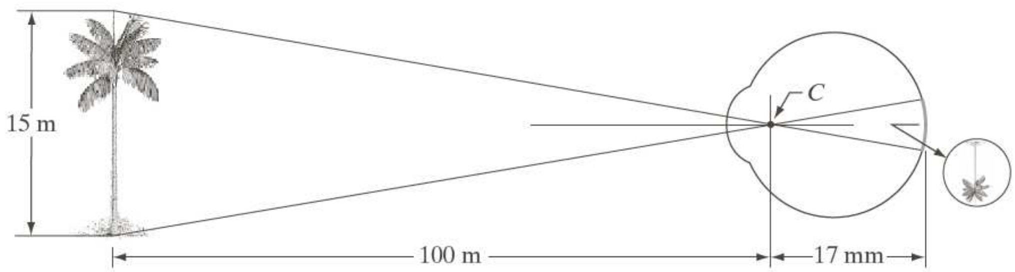


FIGURE 2.3 Graphical representation of the eye looking at a palm tree. Point *C* is the optical center of the lens.

In an ordinary photographic camera, the converse is true. The lens has a fixed focal length, and focusing at various distances is achieved by varying the distance between the lens and the imaging plane.

Brightness Adaptation and Discrimination

Since digital images are displayed as a discrete set of intensities, the eye's ability to discriminate between different intensity levels is an important issue.

The range of light intensity levels adapted by human visual system is enormous, on the order of 10^{10}, from the scotopic threshold to the glare limit.

Experimental evidence indicates that subjective brightness is a logarithmic function of the light intensity incident on the eye.

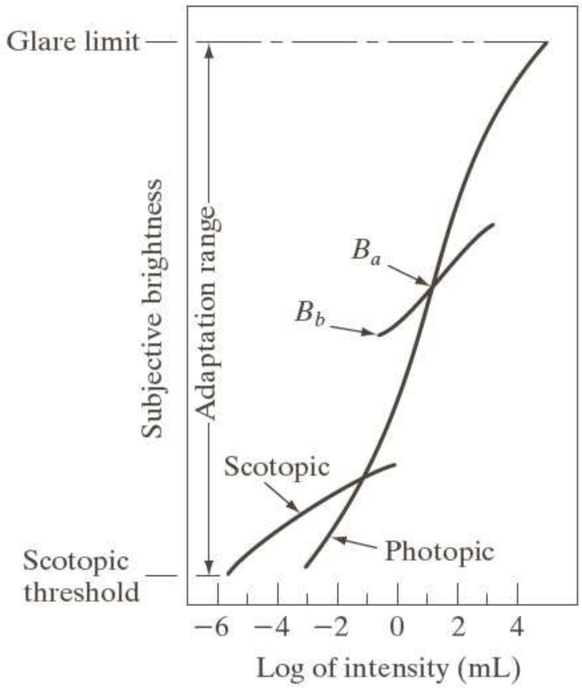


FIGURE 2.4 Range of subjective brightness sensations showing a particular adaptation level.

Figure 2.4 shows that the visual system cannot operate over such a range simultaneously.

The total range of distinct intensity levels the eye can discriminate simultaneously is rather small when compared with the total adaptation range.

For any given set of conditions, the current sensitivity level of the visual system is called the brightness adaptation level, for example, as B_a shown in Figure 2.4. It represents the range of subjective brightness that the eye can perceive when adapted to this level.

Another important issue is the ability of the eye to discriminate between changes in light intensity at any specific adaptation level.

Figure 2.5 shows the idea of a basic experiment to determine the human visual system for brightness discrimination.

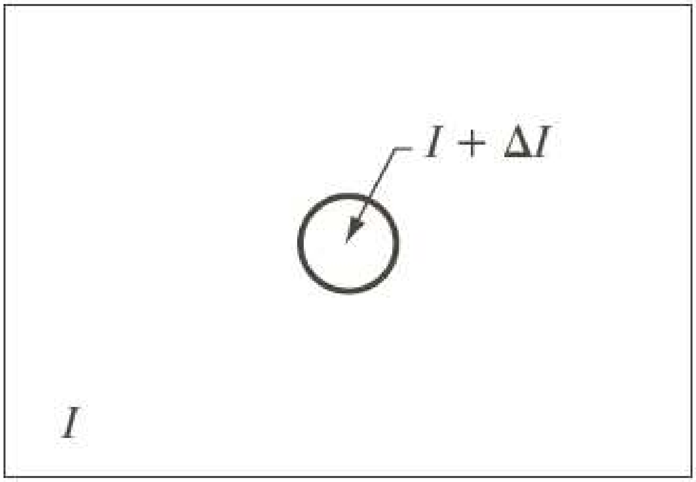


FIGURE 2.5 Basic experimental setup used to characterize brightness discrimination.

An increment of illumination, ΔI, is added to the field in the form of a short-duration flash.

If ΔI is not bright enough, the subject says "no".

As ΔI gets stronger, the subject may give a positive response of "yes"

Finally, when ΔI becomes strong enough, the subject will give a positive response of "yes" all the time.

Let ΔI_c denote the increment of illumination discriminable 50% of the time with background illumination I. It is called the Weber ratio.

A small value of $\Delta I_c / I$ means that a small change in intensity is discriminable. It represents "good" brightness discrimination.

A plot of $\log \Delta I_c / I$ as a function of $\log I$ has the general shape shown in Figure 2.6.

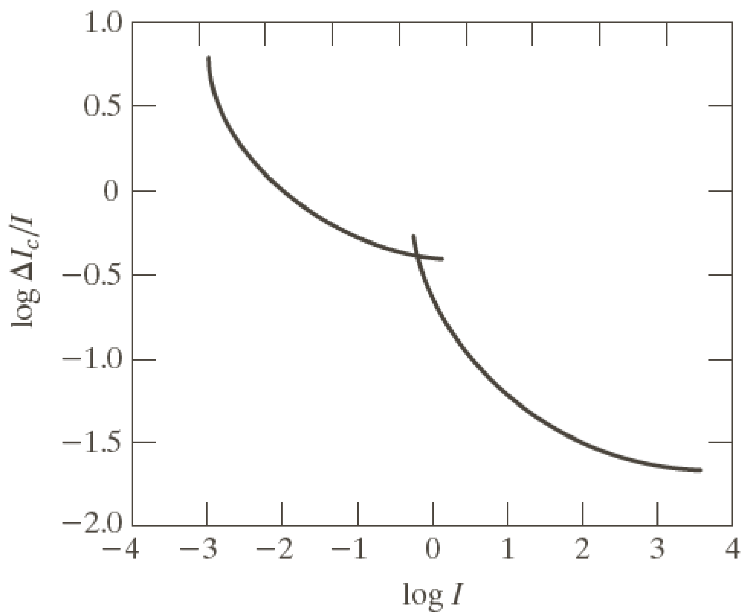
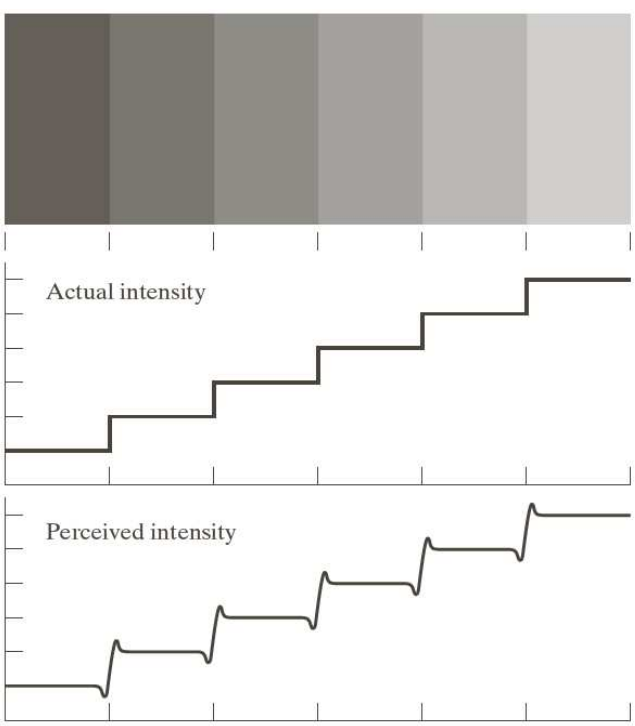


FIGURE 2.6 Typical Weber ratio as a function of intensity.

Figure 2.6 shows that brightness discrimination improves significantly as background illumination increases.

Two phenomena demonstrate that perceived brightness is not a simple function of intensity.



a
b
c

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

Figure 2.7 (a) shows the fact that the visual system tends to undershoot or overshoot around the boundary of regions of different intensities.

As shown in Figure 2.7 (c), although the intensity of the stripes is constant, we actually perceive a brightness pattern that is strongly scalloped near the boundaries.

Another phenomenon, called simultaneous contrast, is related to the fact that a region's perceived brightness does not depend simply on its intensity.

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.

In Figure 2.8, all the center squares have exactly the same intensity, though they appear to the eye to become darker as the background gets lighter.

Other examples of human perception phenomena are optical illusions, in which the eye fills in nonexisting information or wrongly perceives geometrical properties of objects.

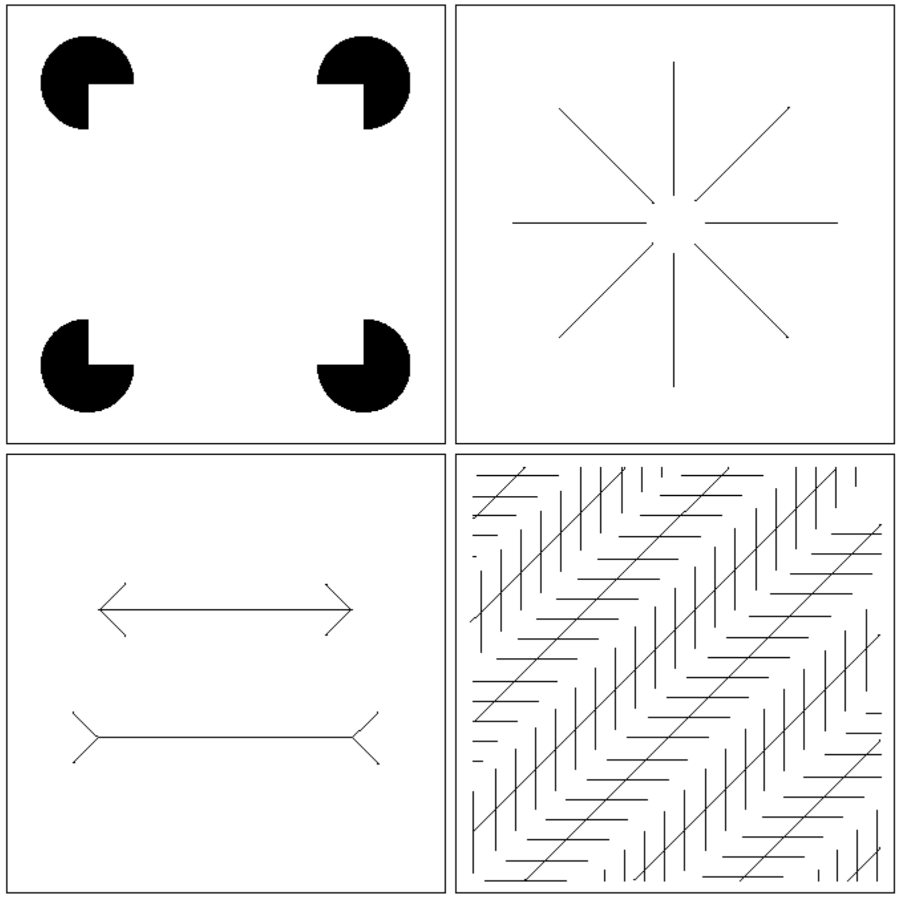
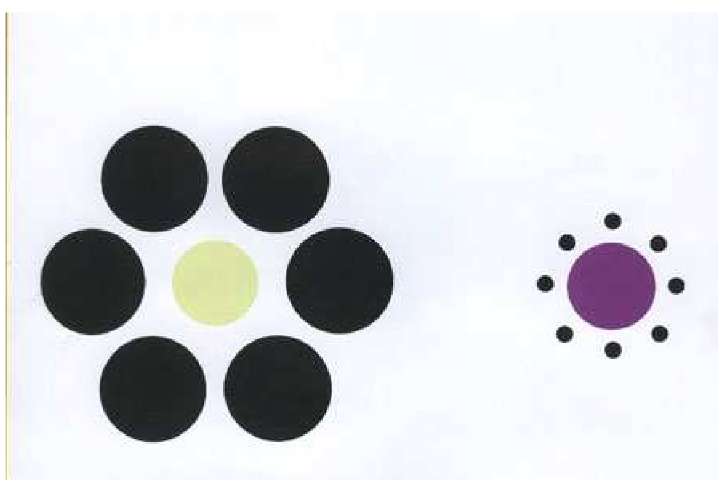
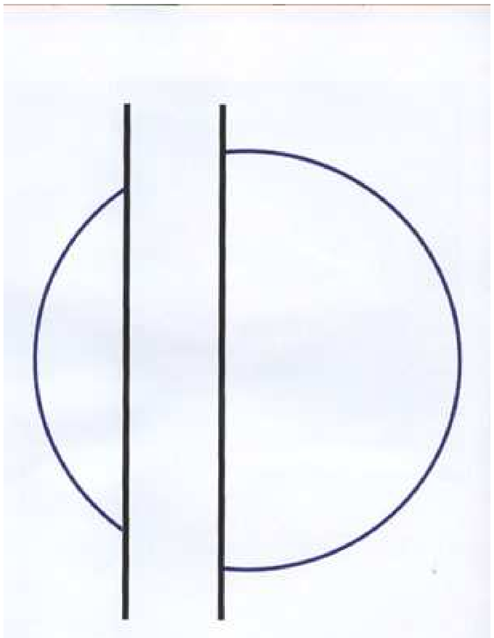
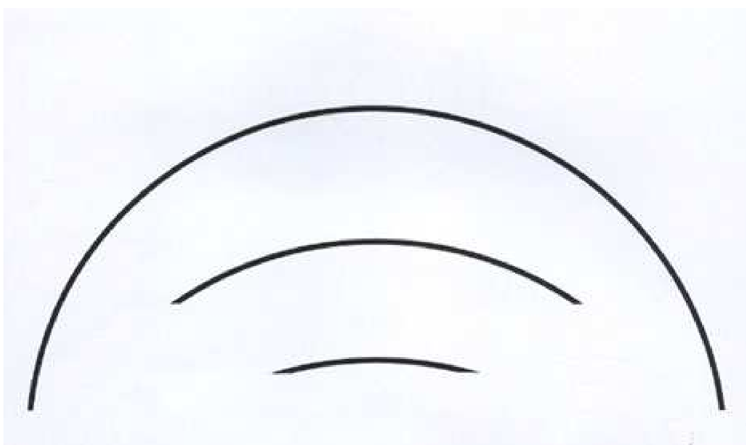


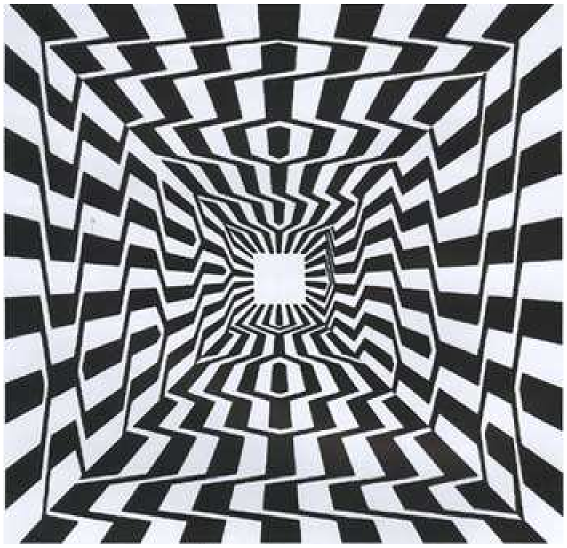
FIGURE 2.9 Some well-known optical illusions.

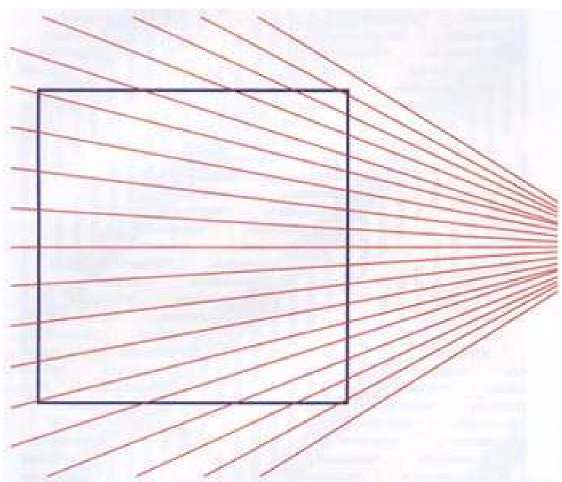
More examples:

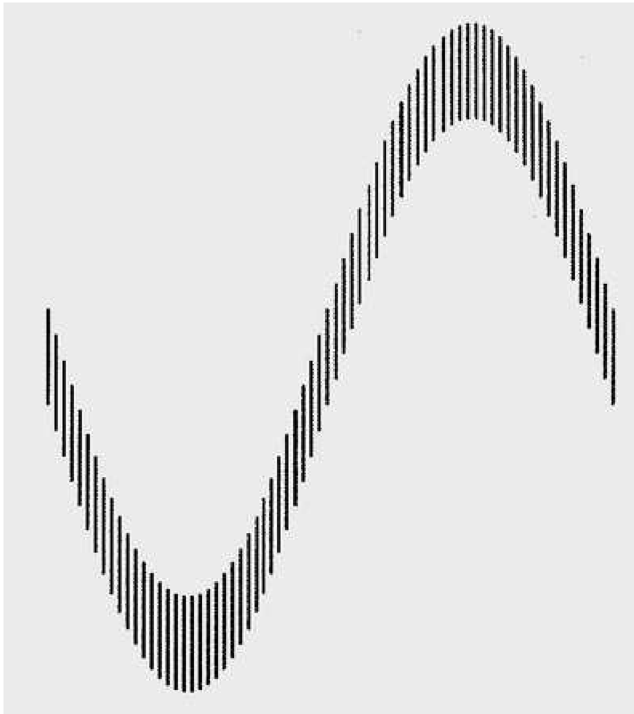


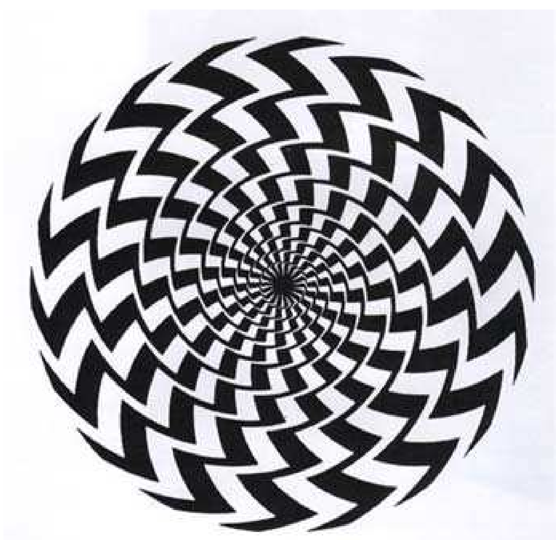












2.2 Light and the Electromagnetic Spectrum

In 1666, Sir Isaac Newton discovered that when a beam of sunlight is passed through a glass prism, the emerging beam of light is consists of a continuous spectrum of colors from violet at one end to red at the other.

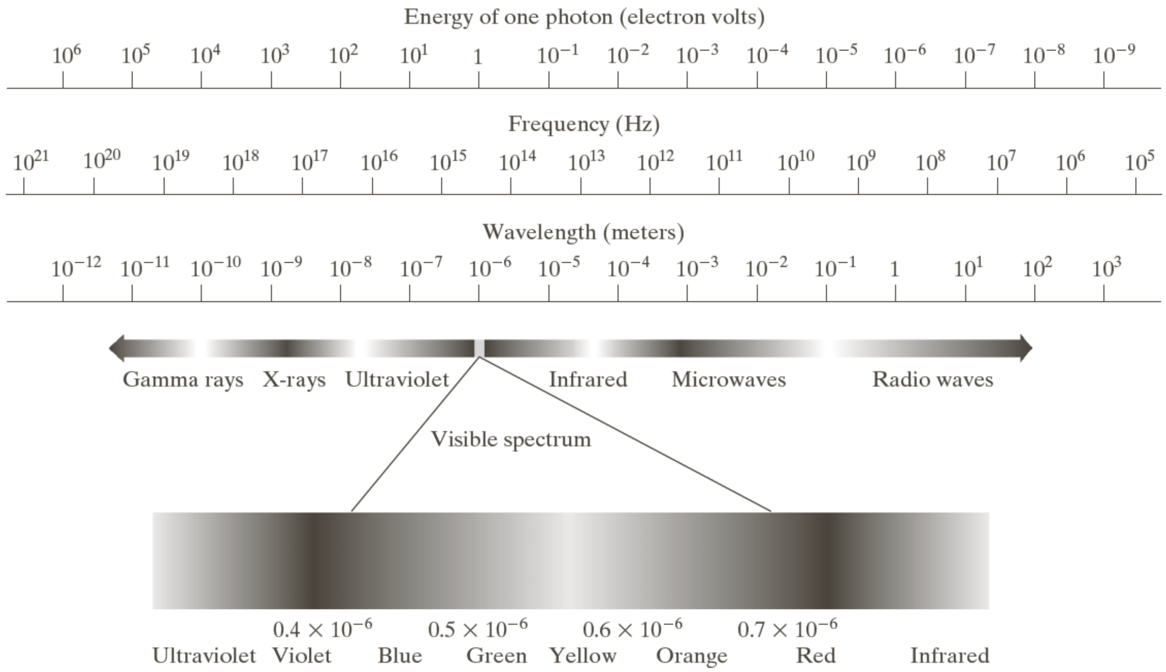


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

As Figure 2.10 shows that the range of colors we perceive in visible light represents a very small portion of the electromagnetic spectrum.

Wavelength λ and frequency ν are related by the expression

$$\lambda = \frac{c}{\nu} \tag{2.2-1}$$

where c is the speed of light ($2.998 \times 10^8 m/s$).

The energy of the various components of the electromagnetic spectrum is given by

$$E = h\nu \tag{2.2-2}$$

where h is Planck's constant.

Electromagnetic wave can be visualized as propagating sinusoidal waves with wavelength λ, or they can be thought of as a stream of massless particles traveling in a wavelike pattern. Each massless particle contains a bundle of energy, which is called a photon.

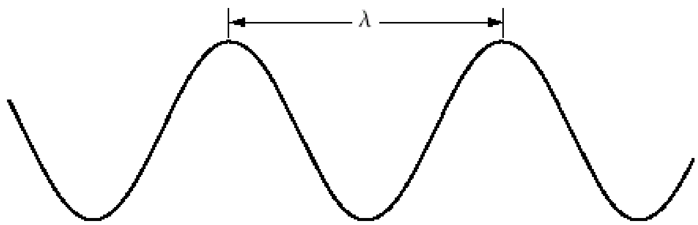


FIGURE 2.11 Graphical representation of one wavelength.

Regarding to (2.2-2), energy is proportional to frequency, so the higher-frequency electromagnetic phenomena carry more energy per photon.

Light is a particular type of electromagnetic radiation that can be sensed by human eye. The visible band of the electromagnetic spectrum spans the range from about $0.43\ \mu m$ (violet) to about $0.79\ \mu m$ (red).

The colors that human perceive in an object are determined by nature of the light reflected from the object.

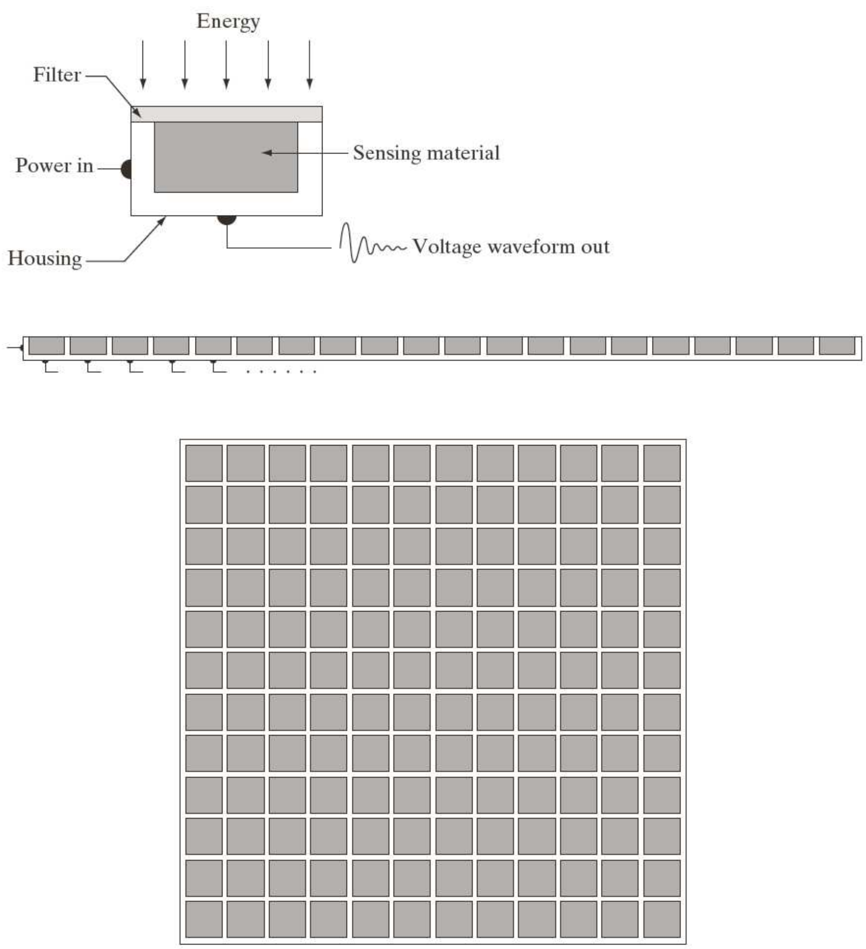
Light that is void of color is called monochromatic light. The only attribute of monochromatic light is its intensity or amount. Because the intensity is perceived to vary from black to white, the term gray level is used to denote monochromatic intensity.

The terms intensity and gray level are used interchangeably in this textbook.

The range of measured values of monochromatic light from black to white is usually called the gray scale, and monochromatic images are frequently referred to as gray-scale images.

2.3 Image Sensing and Acquisition

Figure 2.12 shows the three principal sensor arrangements used to transform illumination energy into digital images.



a
b
c

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

Incoming energy is transformed into a voltage by the combination of input electrical power and sensor material that is responsive to the particular type of energy being detected.

The output voltage waveform is the response of the sensor(s), and a digital quantity is obtained from each sensor by digitizing its response.

Image Acquisition Using a Single Sensor

In order to generate a $2\text{-}D$ image using a single sensor, there has to be relative displacements in both the x- and y- directions between the sensor and the area to be imaged.

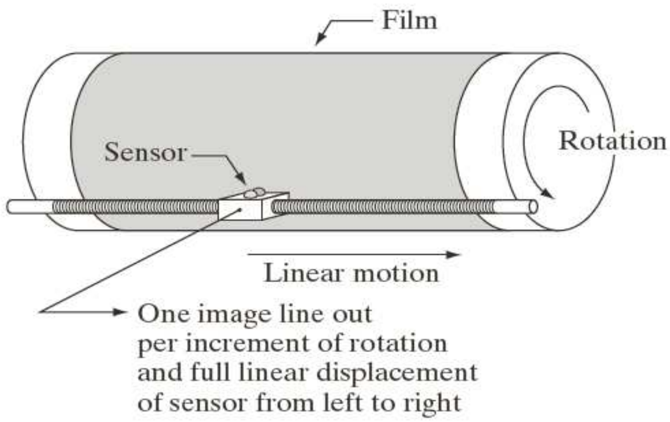


FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

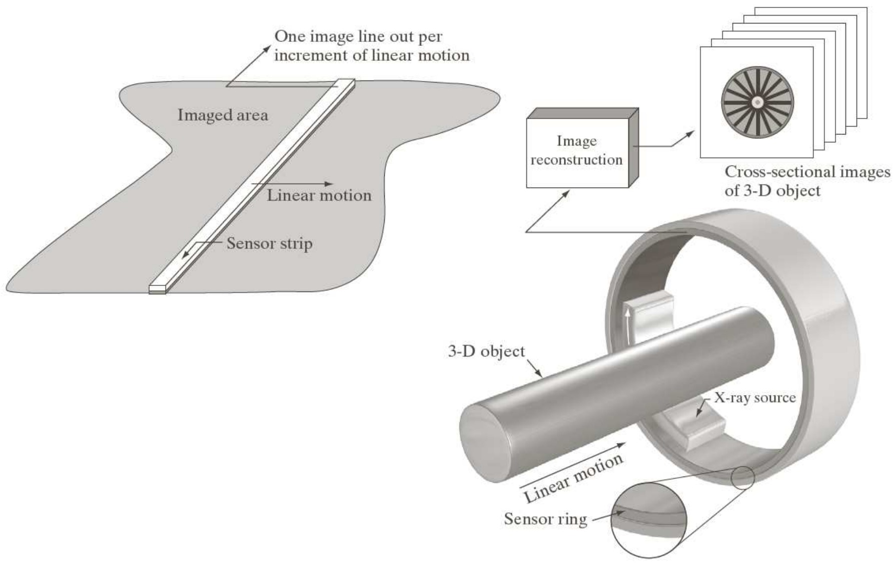
Figure 2.13 shows an arrangement used in high-precision scanning.

Other similar mechanical arrangements use a flat bed, with the sensor moving in two linear directions.

Another example of imaging with a single sensor places a laser source coincident with the sensor. Moving mirrors are used to control the outgoing beam in a scanning pattern and to direct the reflected laser signal onto the sensor.

Image Acquisition Using Sensor Strips

As Figure 2.12 (b) shows, an in-line arrangement of sensors in the form of a sensor strip is used much more than a single sensor.



a b

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

Figure 2.14 (a) shows the type of arrangement used in most flat bed scanners.

Sensor strips mounted in a ring configuration are used in medical and industrial imaging to obtain cross-sectional images of $3\text{-}D$ objects, as Figure 2.14 (b) shows.

Image Acquisition Using Sensor Arrays

Figure 2.12 (c) shows individual sensors arranged in the form of a $2\text{-}D$ array.

This is the predominated arrangement found in digital cameras. Since the sensor array is two-dimensional, its key advantage is that a complete image can be obtained by focusing the energy pattern onto the surface of the array.

The principal manner in which array sensors are used is shown in Figure 2.15.

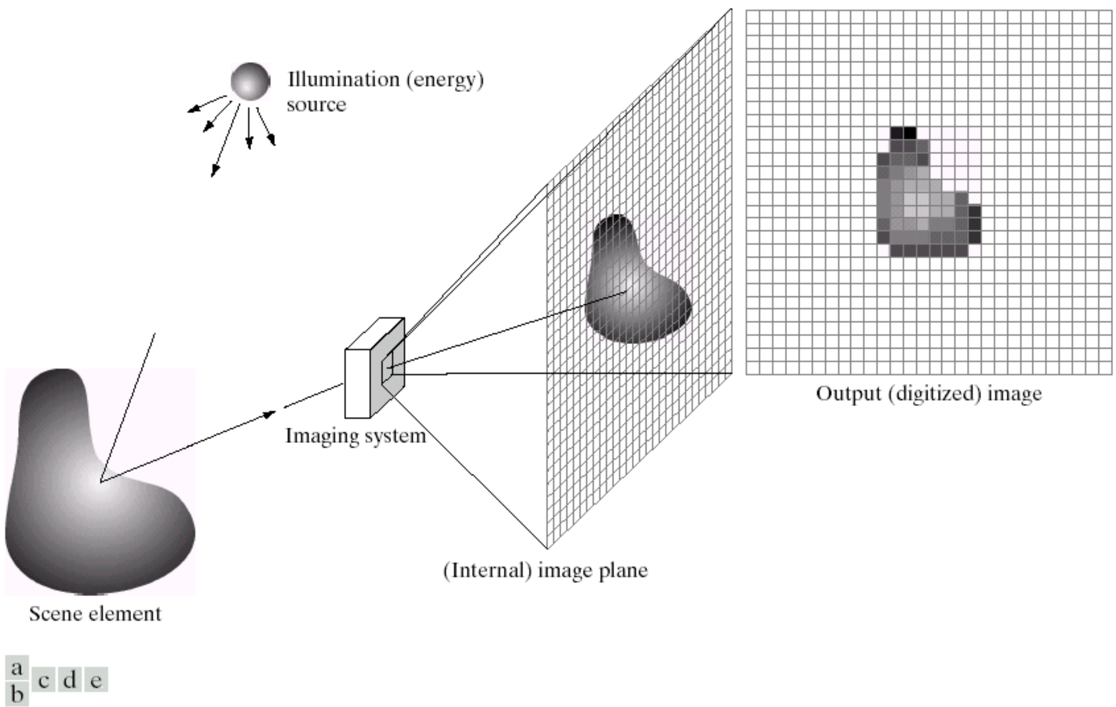


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

A Simple Image Formation Model

We denote images by two-dimensional functions of the form $f(x, y)$. The value of f at coordinates (x, y) is a positive scalar quantity whose physical meaning is determined by the source of the image.

The function $f(x, y)$ must be nonzero and finite

$$0 < f(x, y) < \infty \,. \qquad (2.3\text{-}1)$$

The function $f(x, y)$ may be characterized by two components:

(1) the amount of source illumination incident on the scene being viewed;

(2) the amount of illumination reflected by the objects in the scene.

These are called illumination and reflectance components and are denoted by $i(x, y)$ and $r(x, y)$. These two functions combine to form $f(x, y)$:

$$f(x, y) = i(x, y)r(x, y) \tag{2.3-2}$$

where

$$0 < i(x, y) < \infty \tag{2.3-3}$$

and

$$0 < r(x, y) < 1, \tag{2.3-4}$$

which means that reflectance is bounded by 0 (total absorption) and 1 (total reflectance).

Example 2.1: Some typical values of illumination and reflectance

2.4 Image Sampling and Quantization

To create a digital image, we need to convert the continuous sensed data into digital form. This involves two processes: sampling and quantization.

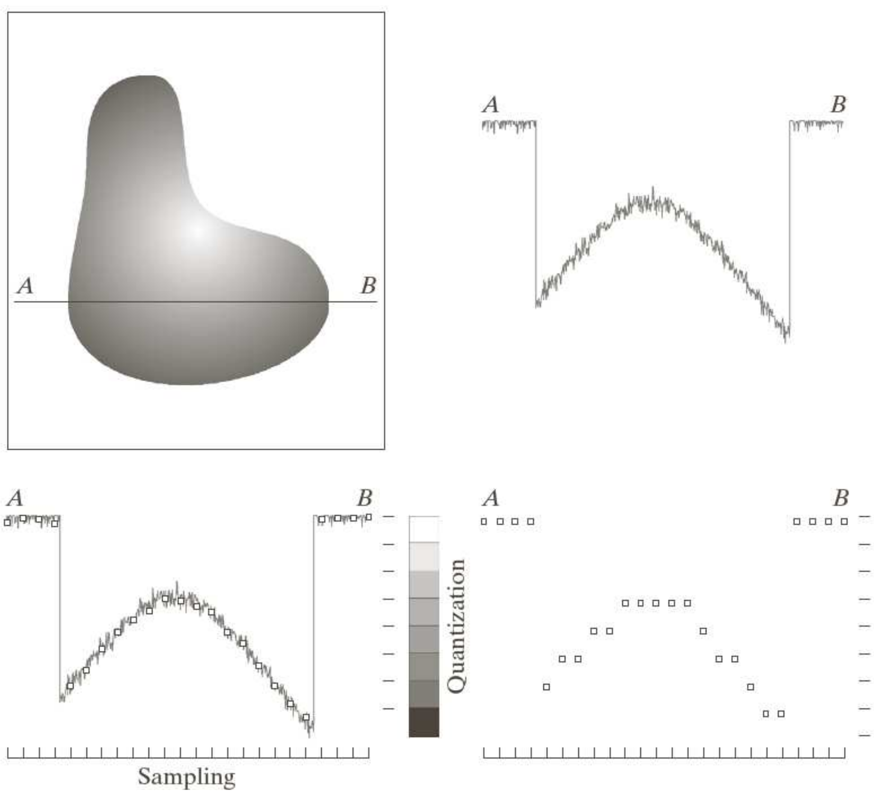
Basic Concepts in Sampling and Quantization

To convert an image to digital form, we have to sample the function in both coordinates and in amplitude.

Digitizing the coordinate values is called sampling.

Digitizing the amplitude values is called quantization.

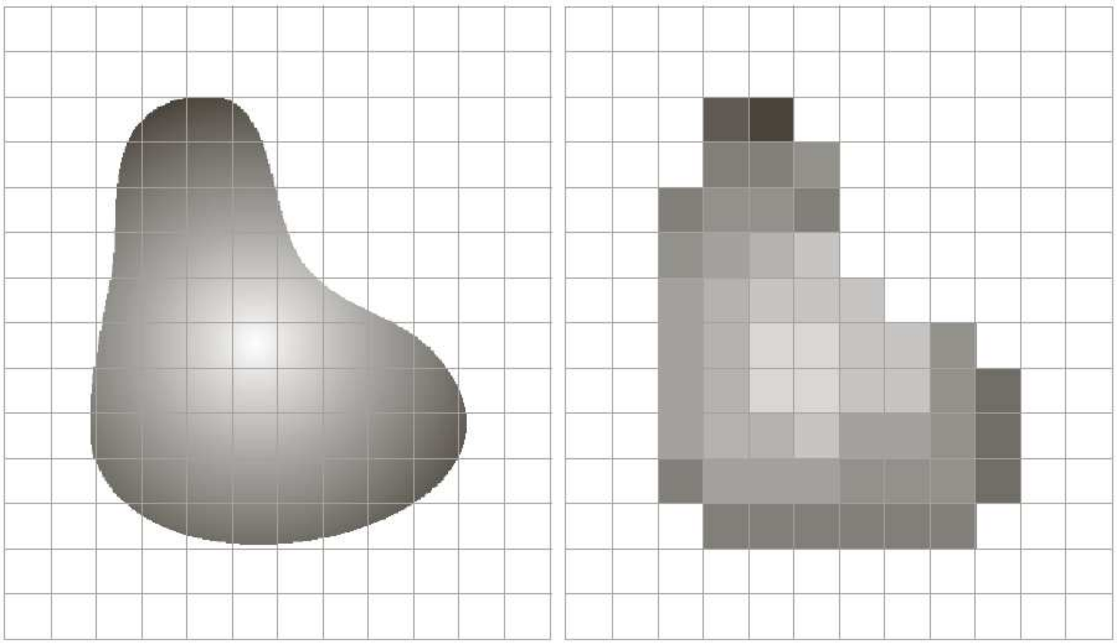
Figure 2.16 shows the basic idea behind sampling and quantization.



a b
c d

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.

When a sensing array is used for image acquisition, there is no motion and the number of sensors in the array establishes the limits of sampling in both directions.



a b

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

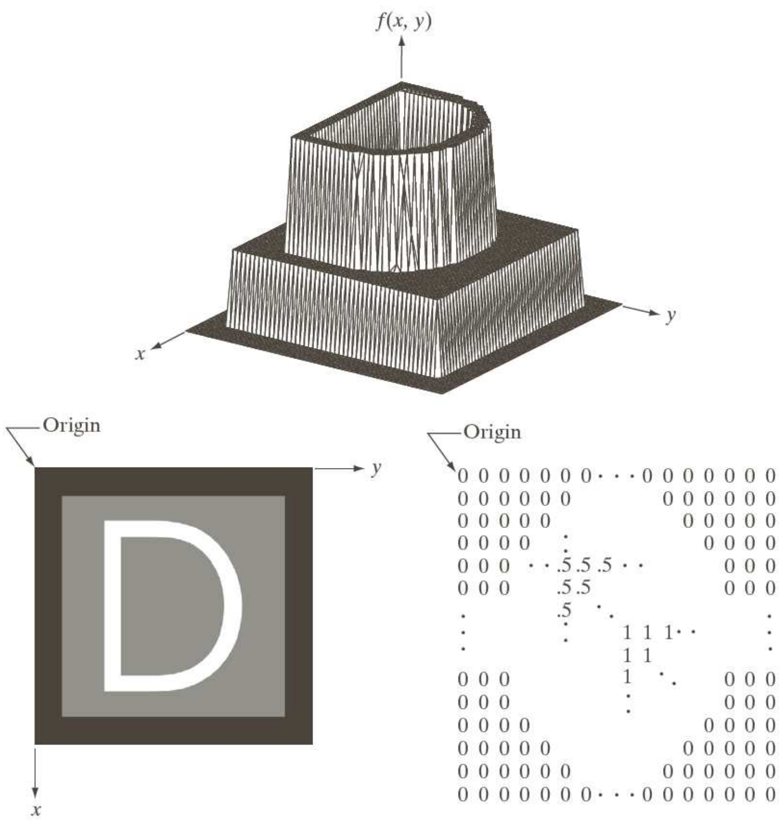
Representing Digital Images

By applying sampling and quantization, we can convert a continuous image function of two continuous variables, $f(s,t)$, into a digital image $f(x,y)$, which contains M rows and N columns. (x,y) are discrete coordinates: $x = 0,1,2,\ldots,M-1$ and $y = 0,1,2,\ldots,N-1$.

In general, the value of the image at any coordinates (x,y) is denoted by $f(x,y)$, where x and y are integers.

The section of the real plane spanned by the coordinates of an image is called the spatial domain.

Figure 2.18 shows three basic ways to represent $f(x,y)$.



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 (0, .5, and 1 represent black, gray, and white, respectively).

The representations in Figure 2.18 (b) and (c) are the most useful. Image displays allow us to view results at a glance, and numerical arrays are used for processing and algorithm development.

In equation form, we write the representation of an $M \times N$ numerical array as

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N-1) \end{bmatrix} \quad (2.4\text{-}1)$$

In some discussions, we use a more traditional matrix notation to denote a digital image as its elements:

$$\mathbf{A} = \begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & \vdots & & \vdots \\ a_{M-1,0} & a_{M-1,1} & \cdots & a_{M-1,N-1} \end{bmatrix} \qquad (2.4\text{-}2)$$

Due to storage and quantizing hardware considerations, the number of intensity levels typically is an integer power of 2:

$$L = 2^k \qquad (2.4\text{-}3)$$

We assume that the discrete levels are equally spaced and that they are integers in the interval $[0, L-1]$.

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.

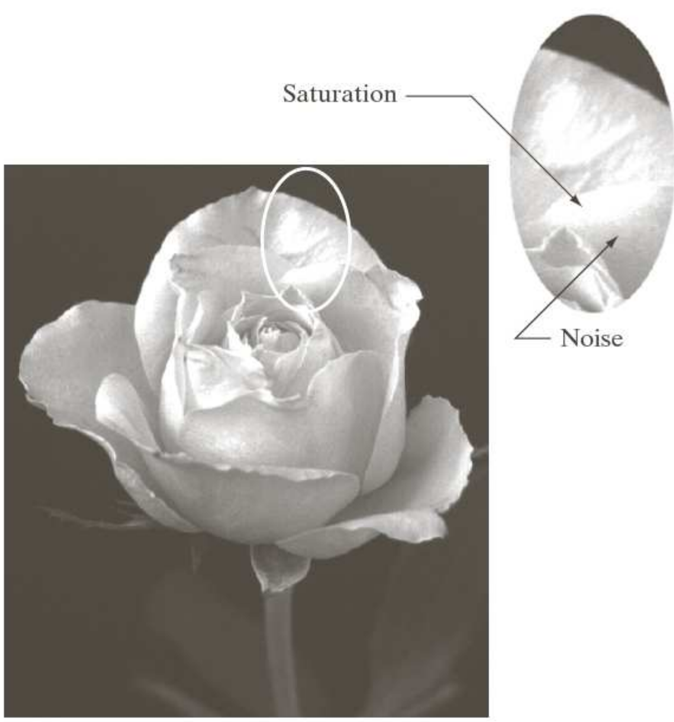


FIGURE 2.19 An image exhibiting saturation and noise. Saturation is the highest value beyond which all intensity levels are clipped (note how the entire saturated area has a high, *constant* intensity level). Noise in this case appears as a grainy texture pattern. Noise, especially in the darker regions of an image (e.g., the stem of the rose) masks the lowest detectable true intensity level.

Closely associated with the concept of dynamic range is image contrast, which is defined as the difference in intensity between the highest and lowest intensity levels in an image.

The number of bits required to store a digitized image is

$$b = M \times N \times k \,. \tag{2.4-4}$$

When $M = N$, (2.4-4) becomes

$$b = N^2 k \,. \tag{2.4-5}$$

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 and Intensity Resolution

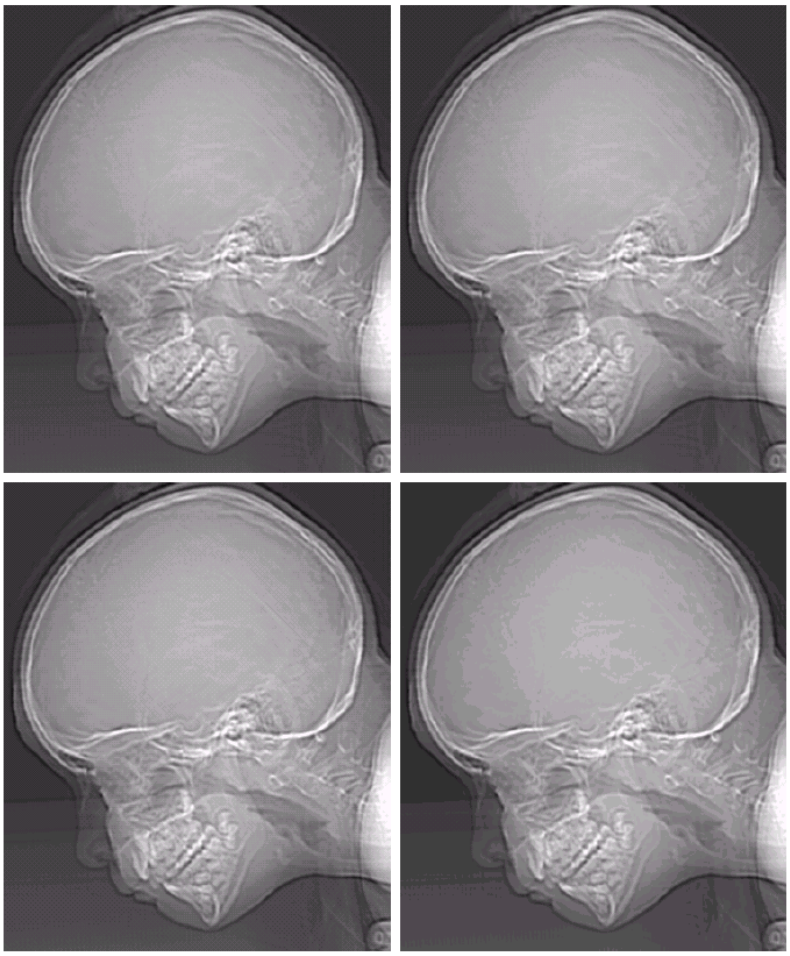
Spatial resolution is a measure of the smallest discernible detail in an image.

Example 2.2: Illustration of the effects of reducing image spatial resolution.

a b
c d

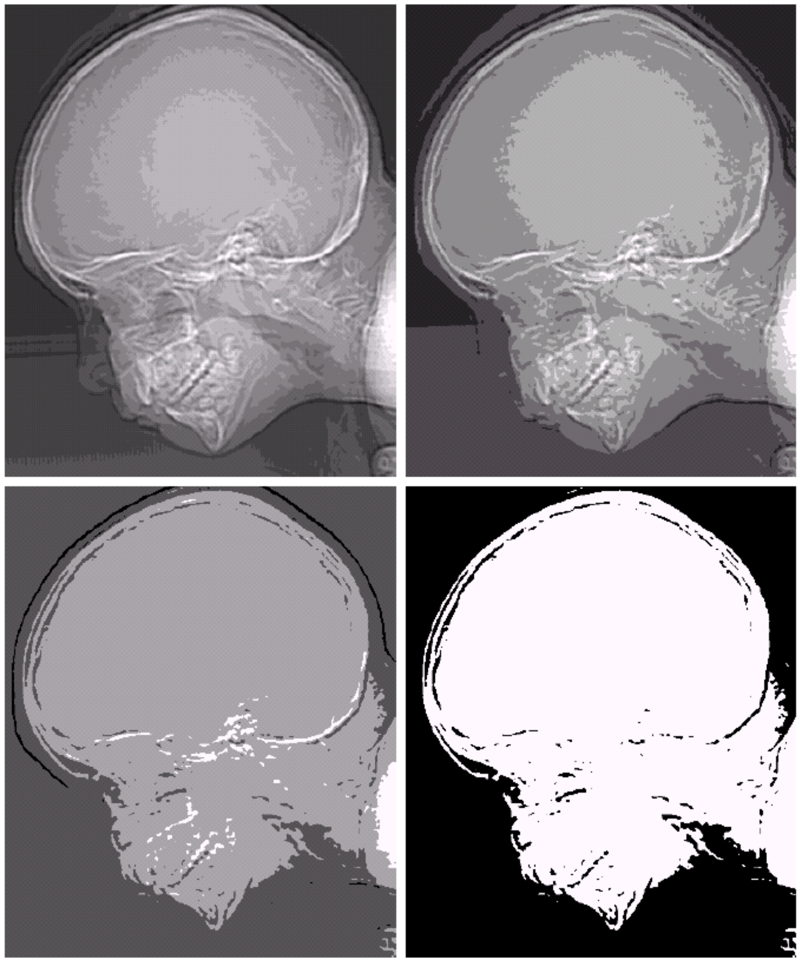
FIGURE 2.20 Typical effects of reducing spatial resolution. Images shown at: (a) 1250 dpi, (b) 300 dpi, (c) 150 dpi, and (d) 72 dpi. The thin black borders were added for clarity. They are not part of the data.

Example 2.3: Typical effects of varying the number of intensity levels in a digital image.



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.)

The results shown in Example 2.2 and 2.3 illustrate the effects produced on image quality by varying N and k independently.

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.)

Figure 2.2 shows the effects on image quality produced by varying *N* and *k* simultaneously.

Sets of these three types of images were generated by varying *N* and *k*, and observers were then asked to rank them according to their subjective quality.

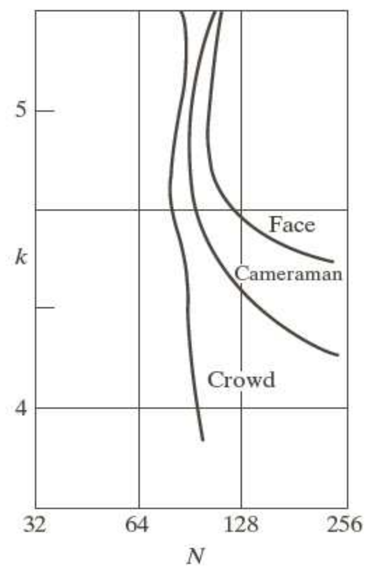


FIGURE 2.23 Typical isopreference curves for the three types of images in Fig. 2.22.

Each point in the *Nk*-plane represents an image having values of *N* and *k* equal to the coordinates of that point.