to 0.3 enhanced a little more detail in the background, but began to reduce contrast to the point where the image started to have a very slight "washed-out" look, especially in the background. By comparing all results, we see that the best enhancement in terms of contrast and discernable detail was obtained with $\gamma = 0.4$. A value of $\gamma = 0.3$ is an approximate limit below which contrast in this particular image would be reduced to an unacceptable level. ■

EXAMPLE 3.2: Another illustration of power-law transformations.

■ Figure 3.9(a) shows the opposite problem of Fig. 3.8(a). The image to be enhanced now has a washed-out appearance, indicating that a compression of gray levels is desirable. This can be accomplished with Eq. (3.2-3) using values of γ greater than 1. The results of processing Fig. 3.9(a) with $\gamma = 3.0$, 4.0, and 5.0 are shown in Figs. 3.9(b) through (d). Suitable results were obtained with gamma values of 3.0 and 4.0, the latter having a slightly more appealing appearance because it has higher contrast. The result obtained with $\gamma = 5.0$ has areas that are too dark, in which some detail is lost. The dark region to the left of the main road in the upper left quadrant is an example of such an area. ■

a b
c d

FIGURE 3.9 (a) Aerial image. (b)–(d) Results of applying the transformation in Eq. (3.2-3) with $c = 1$ and $\gamma = 3.0, 4.0$, and 5.0, respectively. (Original image for this example courtesy of NASA.)

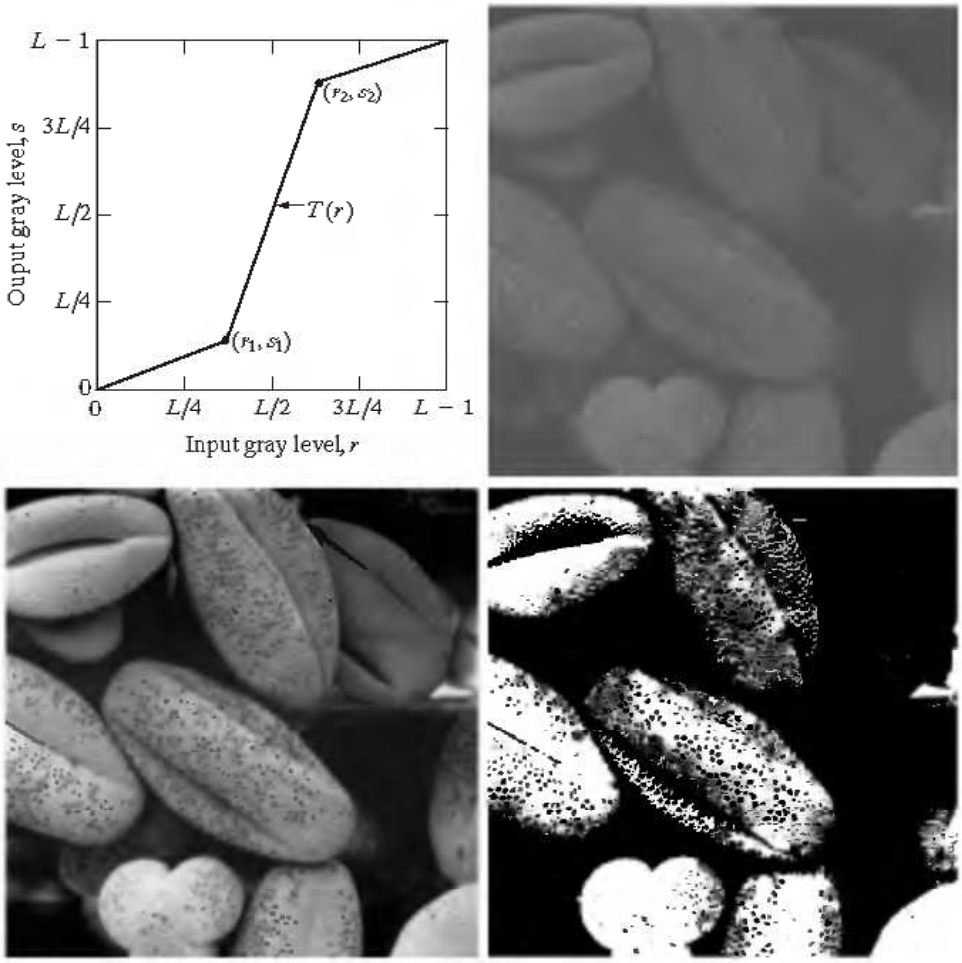
3.2.4 Piecewise-Linear Transformation Functions

A complementary approach to the methods discussed in the previous three sections is to use piecewise linear functions. The principal advantage of piecewise linear functions over the types of functions we have discussed thus far is that the form of piecewise functions can be arbitrarily complex. In fact, as we will see shortly, a practical implementation of some important transformations can be formulated only as piecewise functions. The principal disadvantage of piecewise functions is that their specification requires considerably more user input.

Contrast stretching

One of the simplest piecewise linear functions is a contrast-stretching transformation. Low-contrast images can result from poor illumination, lack of dynamic range in the imaging sensor, or even wrong setting of a lens aperture during image acquisition. The idea behind contrast stretching is to increase the dynamic range of the gray levels in the image being processed.

Figure 3.10(a) shows a typical transformation used for contrast stretching. The locations of points (r_1, s_1) and (r_2, s_2) control the shape of the transformation



a b
c d

FIGURE 3.10 Contrast stretching. (a) Form of transformation function. (b) A low-contrast image. (c) Result of contrast stretching. (d) Result of thresholding. (Original image courtesy of Dr. Roger Heady, Research School of Biological Sciences, Australian National University, Canberra, Australia.)

function. If $r_1 = s_1$ and $r_2 = s_2$, the transformation is a linear function that produces no changes in gray levels. If $r_1 = r_2, s_1 = 0$ and $s_2 = L - 1$, the transformation becomes a *thresholding function* that creates a binary image, as illustrated in Fig. 3.2(b). Intermediate values of (r_1, s_1) and (r_2, s_2) produce various degrees of spread in the gray levels of the output image, thus affecting its contrast. In general, $r_1 \leq r_2$ and $s_1 \leq s_2$ is assumed so that the function is single valued and monotonically increasing. This condition preserves the order of gray levels, thus preventing the creation of intensity artifacts in the processed image.

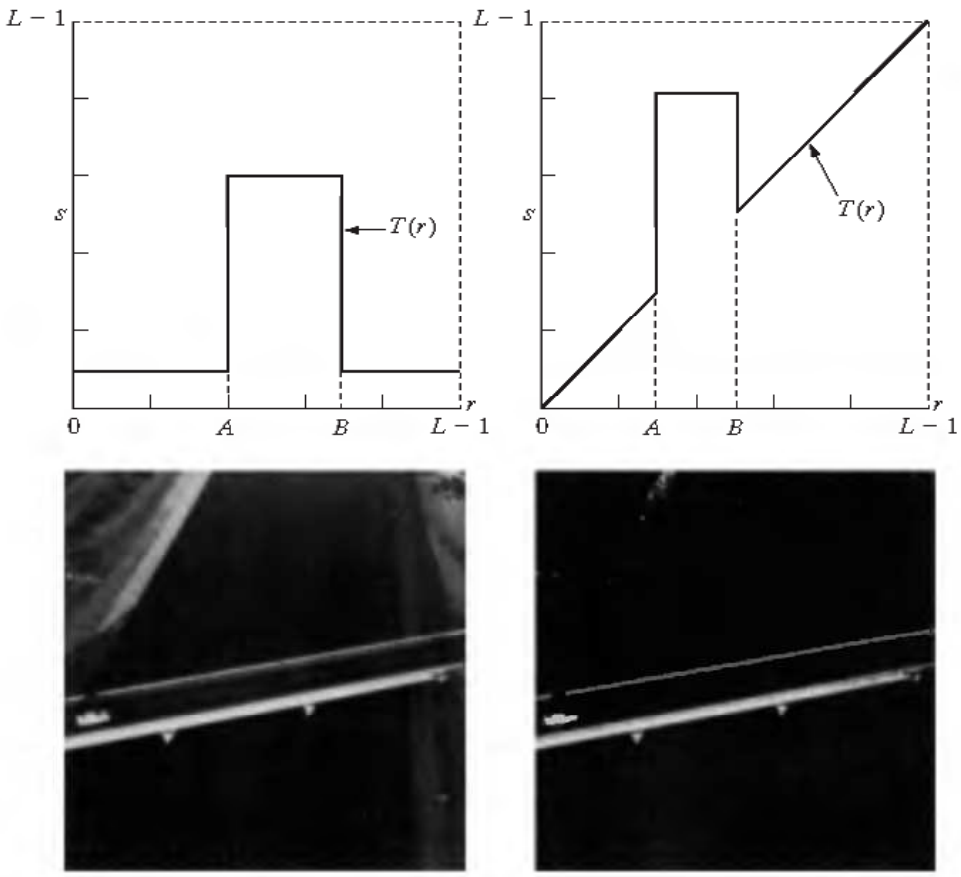
Figure 3.10(b) shows an 8-bit image with low contrast. Fig. 3.10(c) shows the result of contrast stretching, obtained by setting $(r_1, s_1) = (r_{\min}, 0)$ and $(r_2, s_2) = (r_{\max}, L - 1)$ where $r_{\min}$ and $r_{\max}$ denote the minimum and maximum gray levels in the image, respectively. Thus, the transformation function stretched the levels linearly from their original range to the full range $[0, L - 1]$. Finally, Fig. 3.10(d) shows the result of using the thresholding function defined previously, with $r_1 = r_2 = m$, the mean gray level in the image. The original image on which these results are based is a scanning electron microscope image of pollen, magnified approximately 700 times.

Gray-level slicing

Highlighting a specific range of gray levels in an image often is desired. Applications include enhancing features such as masses of water in satellite imagery and enhancing flaws in X-ray images. There are several ways of doing level slicing, but most of them are variations of two basic themes. One approach is to display a high value for all gray levels in the range of interest and a low value for all other gray levels. This transformation, shown in Fig. 3.11(a), produces a binary image. The second approach, based on the transformation shown in Fig. 3.11(b), brightens the desired range of gray levels but preserves the background and gray-level tonalities in the image. Figure 3.11(c) shows a gray-scale image, and Fig. 3.11(d) shows the result of using the transformation in Fig. 3.11(a). Variations of the two transformations shown in Fig. 3.11 are easy to formulate.

Bit-plane slicing

Instead of highlighting gray-level ranges, highlighting the contribution made to total image appearance by specific bits might be desired. Suppose that each pixel in an image is represented by 8 bits. Imagine that the image is composed of eight 1-bit planes, ranging from bit-plane 0 for the least significant bit to bit-plane 7 for the most significant bit. In terms of 8-bit bytes, plane 0 contains all the lowest order bits in the bytes comprising the pixels in the image and plane 7 contains all the high-order bits. Figure 3.12 illustrates these ideas, and Fig. 3.14 shows the various bit planes for the image shown in Fig. 3.13. Note that the higher-order bits (especially the top four) contain the majority of the visually significant data. The other bit planes contribute to more subtle details in the image. Separating a digital image into its bit planes is useful for analyzing the relative importance played by each bit of the image, a process that aids in determining the adequacy of the number of bits used to quantize each pixel. Also, this type of decomposition is useful for image compression, as discussed in Chapter 8.



a b
c d

FIGURE 3.11 (a) This transformation highlights range $[A, B]$ of gray levels and reduces all others to a constant level. (b) This transformation highlights range $[A, B]$ but preserves all other levels. (c) An image. (d) Result of using the transformation in (a).

In terms of bit-plane extraction for an 8-bit image, it is not difficult to show that the (binary) image for bit-plane 7 can be obtained by processing the input image with a thresholding gray-level transformation function that (1) maps all levels in the image between 0 and 127 to one level (for example, 0); and (2) maps all levels between 129 and 255 to another (for example, 255). The binary image for bit-plane 7 in Fig. 3.14 was obtained in just this manner. It is left as an exercise (Problem 3.3) to obtain the gray-level transformation functions that would yield the other bit planes.

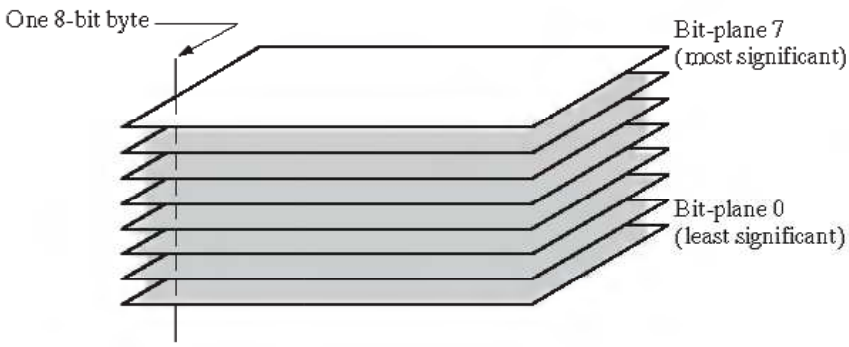


FIGURE 3.12 Bit-plane representation of an 8-bit image.

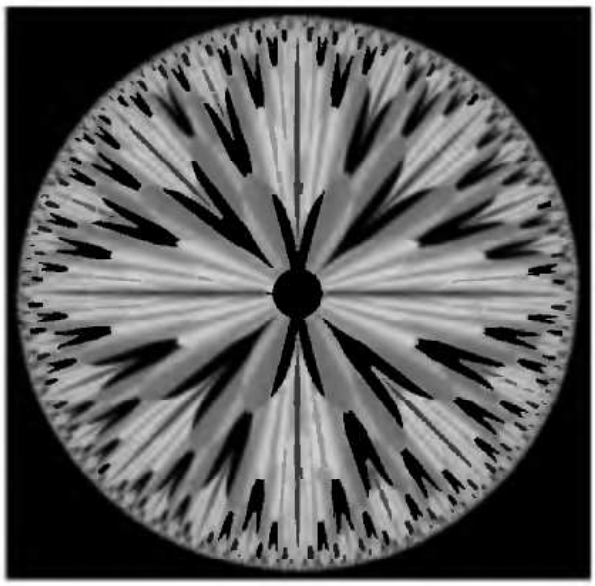


FIGURE 3.13 An 8-bit fractal image. (A fractal is an image generated from mathematical expressions). (Courtesy of Ms. Melissa D. Binde, Swarthmore College, Swarthmore, PA.)

3.3 Histogram Processing

See inside front cover
Consult the book web site for a review of basic probability theory.

The histogram of a digital image with gray levels in the range $[0, L-1]$ is a discrete function $h(r_k) = n_k$, where r_k is the kth gray level and n_k is the number of pixels in the image having gray level r_k. It is common practice to normalize a histogram by dividing each of its values by the total number of pixels in the image, denoted by n. Thus, a normalized histogram is given by $p(r_k) = n_k/n$, for $k = 0, 1, \ldots, L-1$. Loosely speaking, $p(r_k)$ gives an estimate of the probability of occurrence of gray level r_k. Note that the sum of all components of a normalized histogram is equal to 1.

Histograms are the basis for numerous spatial domain processing techniques. Histogram manipulation can be used effectively for image enhancement, as shown in this section. In addition to providing useful image statistics, we shall see in subsequent chapters that the information inherent in histograms also is quite useful in other image processing applications, such as image compression and segmentation. Histograms are simple to calculate in software and also lend themselves to economic hardware implementations, thus making them a popular tool for real-time image processing.

As an introduction to the role of histogram processing in image enhancement, consider Fig. 3.15, which is the pollen image of Fig. 3.10 shown in four basic gray-level characteristics: dark, light, low contrast, and high contrast. The right side of the figure shows the histograms corresponding to these images. The horizontal axis of each histogram plot corresponds to gray level values, r_k. The vertical axis corresponds to values of $h(r_k) = n_k$ or $p(r_k) = n_k/n$ if the values are normalized. Thus, as indicated previously, these histogram plots are simply plots of $h(r_k) = n_k$ versus r_k or $p(r_k) = n_k/n$ versus r_k.

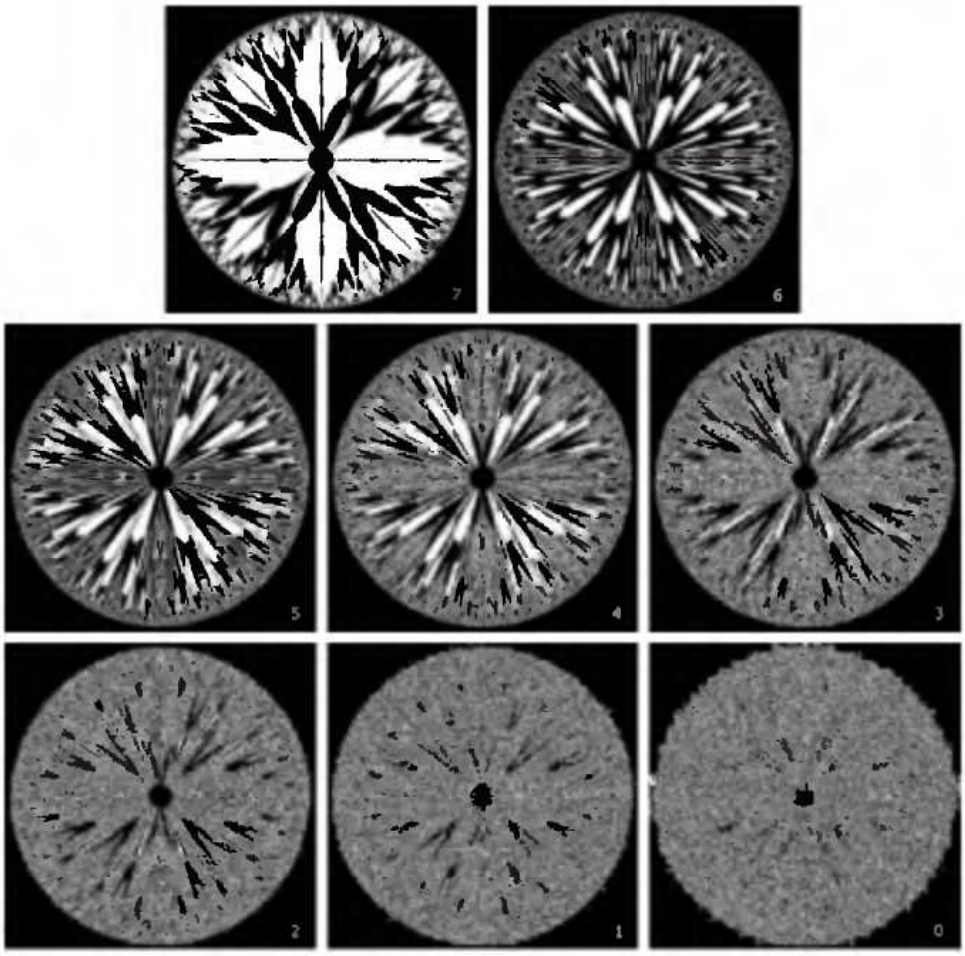
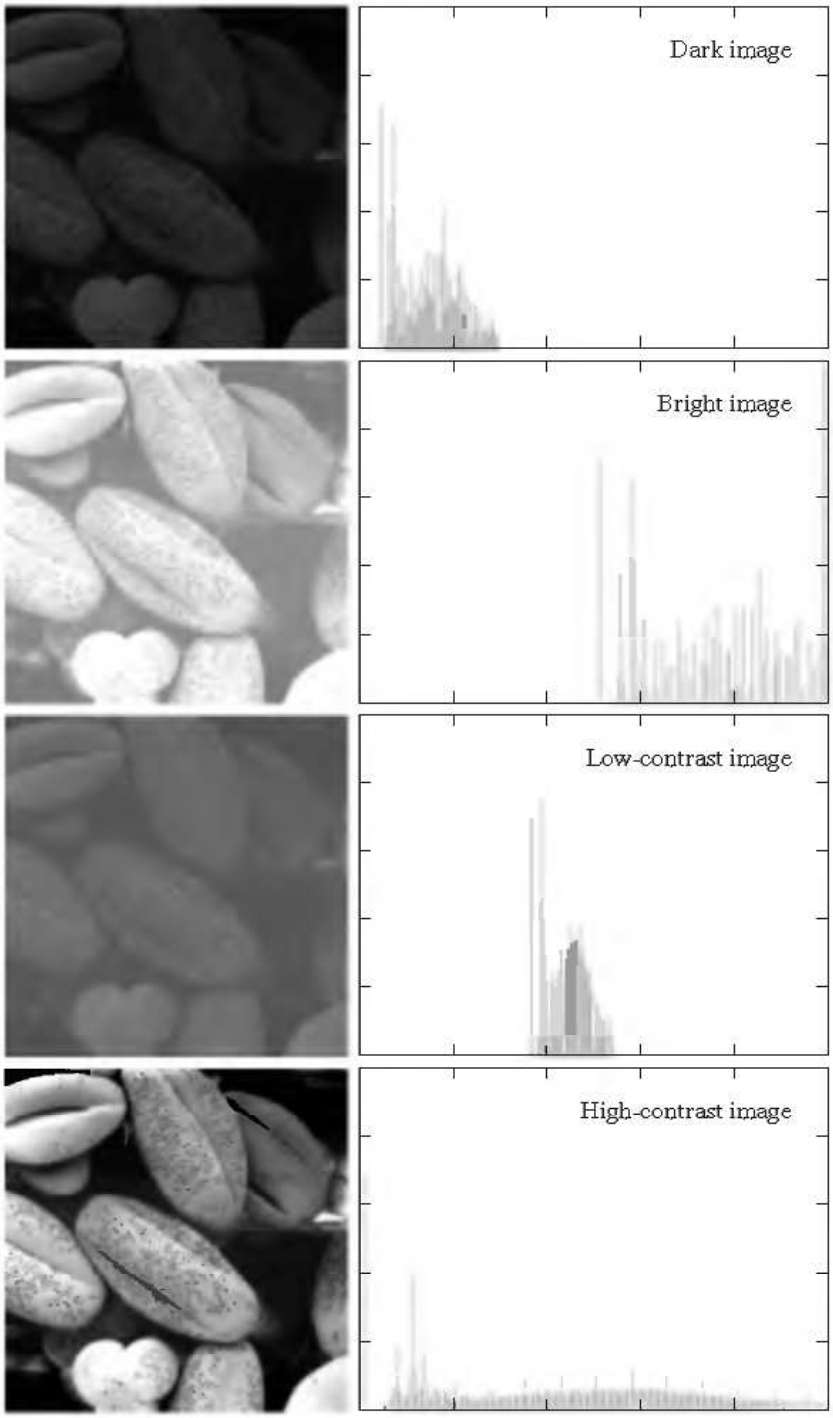


FIGURE 3.14 The eight bit planes of the image in Fig. 3.13. The number at the bottom, right of each image identifies the bit plane.

We note in the dark image that the components of the histogram are concentrated on the low (dark) side of the gray scale. Similarly, the components of the histogram of the bright image are biased toward the high side of the gray scale. An image with low contrast has a histogram that will be narrow and will be centered toward the middle of the gray scale. For a monochrome image this implies a dull, washed-out gray look. Finally, we see that the components of the histogram in the high-contrast image cover a broad range of the gray scale and, further, that the distribution of pixels is not too far from uniform, with very few vertical lines being much higher than the others. Intuitively, it is reasonable to conclude that an image whose pixels tend to occupy the entire range of possible gray levels and, in addition, tend to be distributed uniformly, will have an appearance of high contrast and will exhibit a large variety of gray tones. The net effect will be an image that shows a great deal of gray-level detail and has high dynamic range. It will be shown shortly that it is possible to develop a transformation function that can automatically achieve this effect, based only on information available in the histogram of the input image.



a b

FIGURE 3.15 Four basic image types: dark, light, low contrast, high contrast, and their corresponding histograms. (Original image courtesy of Dr. Roger Heady, Research School of Biological Sciences, Australian National University, Canberra, Australia.)

3.3.1 Histogram Equalization

Consider for a moment continuous functions, and let the variable r represent the gray levels of the image to be enhanced. In the initial part of our discussion we assume that r has been normalized to the interval $[0, 1]$, with $r = 0$ representing black and $r = 1$ representing white. Later, we consider a discrete formulation and allow pixel values to be in the interval $[0, L - 1]$.

For any r satisfying the aforementioned conditions, we focus attention on transformations of the form

$$s = T(r) \qquad 0 \leq r \leq 1 \tag{3.3-1}$$

that produce a level s for every pixel value r in the original image. For reasons that will become obvious shortly, we assume that the transformation function $T(r)$ satisfies the following conditions:

(a) $T(r)$ is single-valued and monotonically increasing in the interval $0 \leq r \leq 1$; and
(b) $0 \leq T(r) \leq 1$ for $0 \leq r \leq 1$.

The requirement in (a) that $T(r)$ be single valued is needed to guarantee that the inverse transformation will exist, and the monotonicity condition preserves the increasing order from black to white in the output image. A transformation function that is not monotonically increasing could result in at least a section of the intensity range being inverted, thus producing some inverted gray levels in the output image. While this may be a desirable effect in some cases, that is not what we are after in the present discussion. Finally, condition (b) guarantees that the output gray levels will be in the same range as the input levels. Figure 3.16 gives an example of a transformation function that satisfies these two conditions. The inverse transformation from s back to r is denoted

$$r = T^{-1}(s) \qquad 0 \leq s \leq 1. \tag{3.3-2}$$

It can be shown by example (Problem 3.8) that even if $T(r)$ satisfies conditions (a) and (b), it is possible that the corresponding inverse $T^{-1}(s)$ may fail to be single valued.

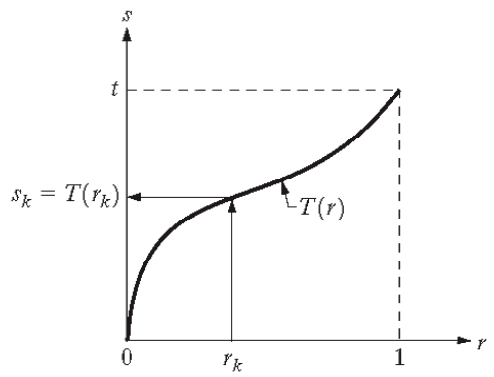


FIGURE 3.16 A gray-level transformation function that is both single valued and monotonically increasing.

The gray levels in an image may be viewed as random variables in the interval $[0, 1]$. One of the most fundamental descriptors of a random variable is its probability density function (PDF). Let $p_r(r)$ and $p_s(s)$ denote the probability density functions of random variables r and s, respectively, where the subscripts on p are used to denote that p_r and p_s are different functions. A basic result from an elementary probability theory is that, if $p_r(r)$ and $T(r)$ are known and $T^{-1}(s)$ satisfies condition (a), then the probability density function $p_s(s)$ of the transformed variable s can be obtained using a rather simple formula:

$$p_s(s) = p_r(r)\left|\frac{dr}{ds}\right|. \tag{3.3-3}$$

Thus, the probability density function of the transformed variable, s, is determined by the gray-level PDF of the input image and by the chosen transformation function.

A transformation function of particular importance in image processing has the form

$$s = T(r) = \int_0^r p_r(w)\, dw \tag{3.3-4}$$

where w is a dummy variable of integration. The right side of Eq. (3.3-4) is recognized as the cumulative distribution function (CDF) of random variable r. Since probability density functions are always positive, and recalling that the integral of a function is the area under the function, it follows that this transformation function is single valued and monotonically increasing, and, therefore, satisfies condition (a). Similarly, the integral of a probability density function for variables in the range $[0, 1]$ also is in the range $[0, 1]$, so condition (b) is satisfied as well.

Given transformation function $T(r)$, we find $p_s(s)$ by applying Eq. (3.3-3). We know from basic calculus (Leibniz's rule) that the derivative of a definite integral with respect to its upper limit is simply the integrand evaluated at that limit. In other words,

$$\begin{aligned}\frac{ds}{dr} &= \frac{dT(r)}{dr} \\ &= \frac{d}{dr}\left[\int_0^r p_r(w)\, dw\right] \\ &= p_r(r).\end{aligned} \tag{3.3-5}$$

Substituting this result for dr/ds into Eq. (3.3-3), and keeping in mind that all probability values are positive, yields

$$\begin{aligned}p_s(s) &= p_r(r)\left|\frac{dr}{ds}\right| \\ &= p_r(r)\left|\frac{1}{p_r(r)}\right| \\ &= 1 \qquad 0 \le s \le 1.\end{aligned} \tag{3.3-6}$$

Because $p_s(s)$ is a probability density function, it follows that it must be zero outside the interval $[0, 1]$ in this case because its integral over all values of s must equal 1. We recognize the form of $p_s(s)$ given in Eq. (3.3-6) as a *uniform* probability density function. Simply stated, we have demonstrated that performing the transformation function given in Eq. (3.3-4) yields a random variable s characterized by a uniform probability density function. It is important to note from Eq. (3.3-4) that $T(r)$ depends on $p_r(r)$, but, as indicated by Eq. (3.3-6), the resulting $p_s(s)$ *always* is uniform, *independent* of the form of $p_r(r)$.

For discrete values we deal with probabilities and summations instead of probability density functions and integrals. The probability of occurrence of gray level r_k in an image is approximated by

$$p_r(r_k) = \frac{n_k}{n} \qquad k = 0, 1, 2, \ldots, L - 1 \tag{3.3-7}$$

where, as noted at the beginning of this section, n is the total number of pixels in the image, n_k is the number of pixels that have gray level r_k, and L is the total number of possible gray levels in the image. The discrete version of the transformation function given in Eq. (3.3-4) is

$$\begin{aligned} s_k = T(r_k) &= \sum_{j=0}^{k} p_r(r_j) \\ &= \sum_{j=0}^{k} \frac{n_j}{n} \qquad k = 0, 1, 2, \ldots, L - 1. \end{aligned} \tag{3.3-8}$$

Thus, a processed (output) image is obtained by mapping each pixel with level r_k in the input image into a corresponding pixel with level s_k in the output image via Eq. (3.3-8). As indicated earlier, a plot of $p_r(r_k)$ versus r_k is called a *histogram*. The transformation (mapping) given in Eq. (3.3-8) is called *histogram equalization* or *histogram linearization*. It is not difficult to show (Problem 3.9) that the transformation in Eq. (3.3-8) satisfies conditions (a) and (b) stated previously in this section.

Unlike its continuos counterpart, it cannot be proved in general that this discrete transformation will produce the discrete equivalent of a uniform probability density function, which would be a uniform histogram. However, as will be seen shortly, use of Eq. (3.3-8) does have the general tendency of spreading the histogram of the input image so that the levels of the histogram-equalized image will span a fuller range of the gray scale.

We discussed earlier in this section the many advantages of having gray-level values that cover the entire gray scale. In addition to producing gray levels that have this tendency, the method just derived has the additional advantage that it is fully "automatic." In other words, given an image, the process of histogram equalization consists simply of implementing Eq. (3.3-8), which is based on information that can be extracted directly from the given image, without the need for further parameter specifications. We note also the simplicity of the computations that would be required to implement the technique.

The inverse transformation from s back to r is denoted by

$$r_k = T^{-1}(s_k) \qquad k = 0, 1, 2, \ldots, L - 1 \tag{3.3-9}$$

It can be shown (Problem 3.9) that the inverse transformation in Eq. (3.3-9) satisfies conditions (a) and (b) stated previously in this section only if none of the levels, $r_k, k = 0, 1, 2, \ldots, L - 1$, are missing from the input image. Although the inverse transformation is not used in histogram equalization, it plays a central role in the histogram-matching scheme developed in the next section. We also discuss in that section details of how to implement histogram processing techniques.

EXAMPLE 3.3: Histogram equalization.

■ Figure 3.17(a) shows the four images from Fig. 3.15, and Fig. 3.17(b) shows the result of performing histogram equalization on each of these images. The first three results (top to bottom) show significant improvement. As expected, histogram equalization did not produce a significant visual difference in the fourth image because the histogram of this image already spans the full spectrum of the gray scale. The transformation functions used to generate the images in Fig. 3.17(b) are shown in Fig. 3.18. These functions were generated from the histograms of the original images [see Fig. 3.15(b)] using Eq. (3.3-8). Note that transformation (4) has a basic linear shape, again indicating that the gray levels in the fourth input image are nearly uniformly distributed. As was just noted, we would expect histogram equalization in this case to have negligible effect on the appearance of the image.

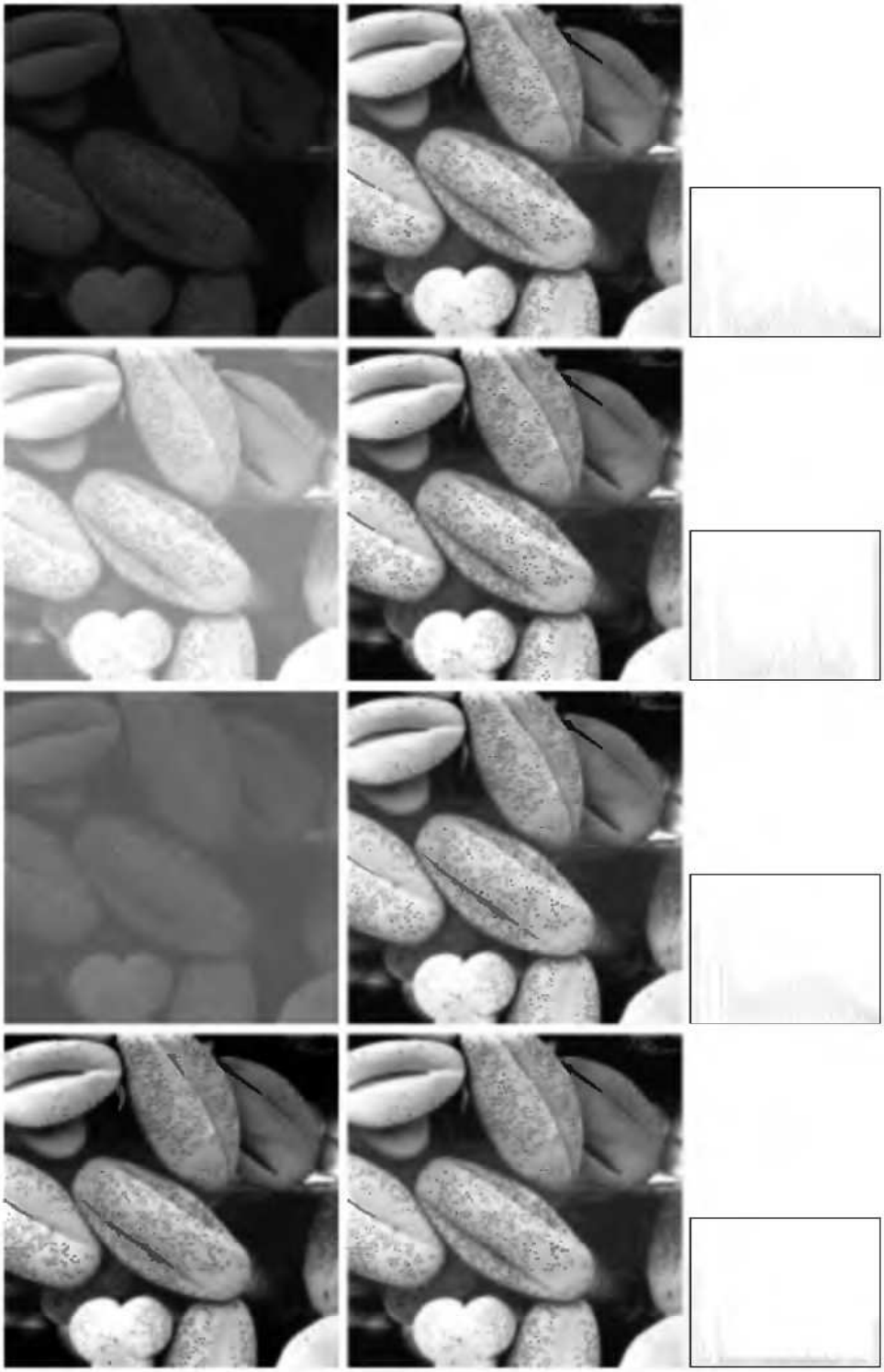
The histograms of the equalized images are shown in Fig. 3.17(c). It is of interest to note that, while all these histograms are different, the histogram-equalized images themselves are visually very similar. This is not unexpected because the difference between the images in the left column is simply one of contrast, not of content. In other words, since the images have the same content, the increase in contrast resulting from histogram equalization was enough to render any gray-level differences in the resulting images visually indistinguishable. Given the significant contrast differences of the images in the left column, this example illustrates the power of histogram equalization as an adaptive enhancement tool. ■

3.3.2 Histogram Matching (Specification)

As indicated in the preceding discussion, histogram equalization automatically determines a transformation function that seeks to produce an output image that has a uniform histogram. When automatic enhancement is desired, this is a good approach because the results from this technique are predictable and the method is simple to implement. We show in this section that there are applications in which attempting to base enhancement on a uniform histogram is not the best approach. In particular, it is useful sometimes to be able to specify the shape of the histogram that we wish the processed image to have. The method used to generate a processed image that has a specified histogram is called *histogram matching* or *histogram specification*.

Development of the method

Let us return for a moment to continuous gray levels r and z (considered continuous random variables), and let $p_r(r)$ and $p_z(z)$ denote their corresponding continuos probability density functions. In this notation, r and z denote



a b c

FIGURE 3.17 (a) Images from Fig. 3.15. (b) Results of histogram equalization. (c) Corresponding histograms.

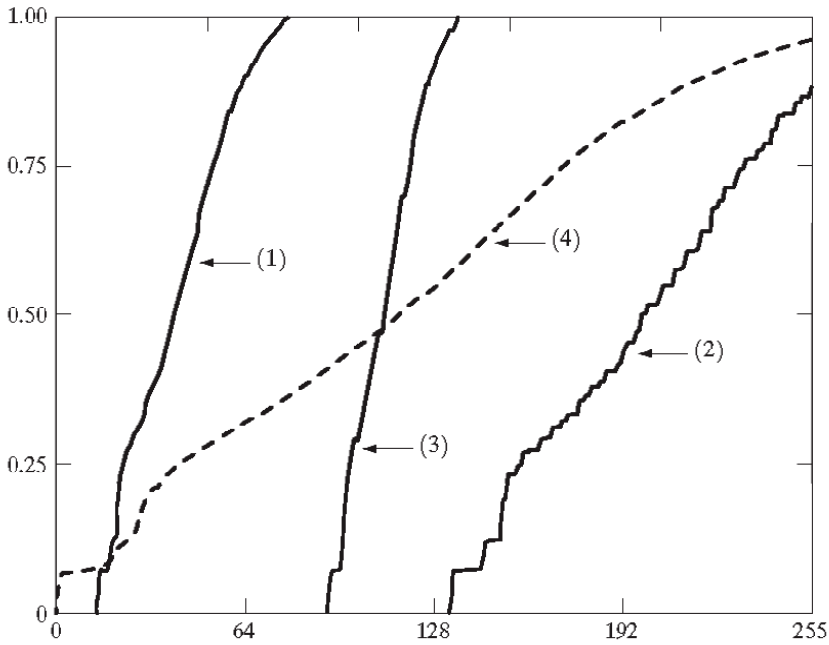


FIGURE 3.18 Transformation functions (1) through (4) were obtained from the histograms of the images in Fig.3.17(a), using Eq. (3.3-8).

the gray levels of the input and output (processed) images, respectively. We can estimate $p_r(r)$ from the given input image, while $p_z(z)$ is the *specified* probability density function that we wish the output image to have.

Let s be a random variable with the property

$$s = T(r) = \int_0^r p_r(w)\, dw \tag{3.3-10}$$

where w is a dummy variable of integration. We recognize this expression as the continuos version of histogram equalization given in Eq. (3.3-4). Suppose next that we define a random variable z with the property

$$G(z) = \int_0^z p_z(t)\, dt = s \tag{3.3-11}$$

where t is a dummy variable of integration. It then follows from these two equations that $G(z) = T(r)$ and, therefore, that z must satisfy the condition

$$z = G^{-1}(s) = G^{-1}[T(r)]. \tag{3.3-12}$$

The transformation $T(r)$ can be obtained from Eq. (3.3-10) once $p_r(r)$ has been estimated from the input image. Similarly, the transformation function $G(z)$ can be obtained using Eq. (3.3-11) because $p_z(z)$ is given.

Assuming that G^{-1} exists and that it satisfies conditions (a) and (b) in the previous section, Eqs. (3.3-10) through (3.3-12) show that an image with a specified probability density function can be obtained from an input image by using the following procedure: (1) Obtain the transformation function $T(r)$ using Eq. (3.3-10). (2) Use Eq. (3.3-11) to obtain the transformation function $G(z)$. (3) Obtain the inverse transformation function G^{-1}. (4) Obtain the output image

by applying Eq. (3.3-12) to all the pixels in the input image. The result of this procedure will be an image whose gray levels, z, have the specified probability density function $p_z(z)$.

Although the procedure just described is straightforward in principle, it is seldom possible in practice to obtain analytical expressions for $T(r)$ and for G^{-1}. Fortunately, this problem is simplified considerably in the case of discrete values. The price we pay is the same as in histogram equalization, where only an approximation to the desired histogram is achievable. In spite of this, however, some very useful results can be obtained even with crude approximations.

The discrete formulation of Eq. (3.3-10) is given by Eq. (3.3-8), which we repeat here for convenience:

$$\begin{aligned} s_k &= T(r_k) = \sum_{j=0}^{k} p_r(r_j) \\ &= \sum_{j=0}^{k} \frac{n_j}{n} \qquad k = 0, 1, 2, \ldots, L-1 \end{aligned} \tag{3.3-13}$$

where n is the total number of pixels in the image, n_j is the number of pixels with gray level r_j, and L is the number of discrete gray levels. Similarly, the discrete formulation of Eq. (3.3-11) is obtained from the given histogram $p_z(z_i), i = 0, 1, 2, \ldots, L-1$, and has the form

$$v_k = G(z_k) = \sum_{i=0}^{k} p_z(z_i) = s_k \qquad k = 0, 1, 2, \ldots, L-1. \tag{3.3-14}$$

As in the continuos case, we are seeking values of z that satisfy this equation. The variable v_k was added here for clarity in the discussion that follows. Finally, the discrete version of Eq. (3.3-12) is given by

$$z_k = G^{-1}[T(r_k)] \qquad k = 0, 1, 2, \ldots, L-1 \tag{3.3-15}$$

or, from Eq. (3.3-13),

$$z_k = G^{-1}(s_k) \qquad k = 0, 1, 2, \ldots, L-1. \tag{3.3-16}$$

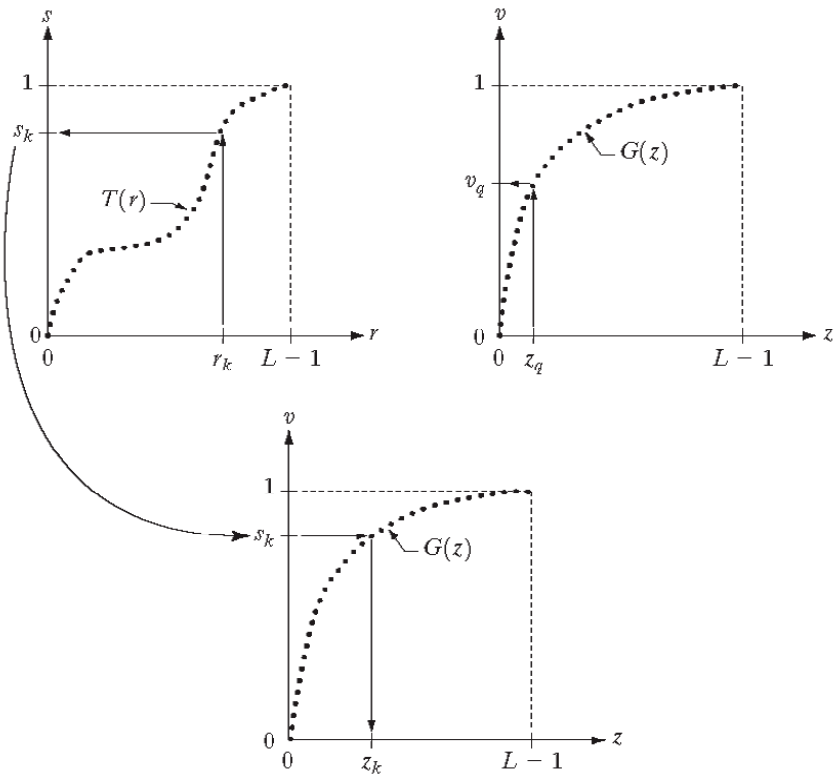
Equations (3.3-13) through (3.3-16) are the foundation for implementing histogram matching for digital images. Equation (3.3-13) is a mapping from the levels in the original image into corresponding levels s_k based on the histogram of the original image, which we compute from the pixels in the image. Equation (3.3-14) computes a transformation function G from the given histogram $p_z(z)$. Finally, Eq. (3.3-15) or its equivalent, Eq. (3.3-16), gives us (an approximation of) the desired levels of the image with that histogram. The first two equations can be implemented easily because all the quantities are known. Implementation of Eq. (3.3-16) is straightforward, but requires additional explanation.

Implementation

We start by noting the following: (1) Each set of gray levels $\{r_j\}$, $\{s_j\}$, and $\{z_j\}$, $j = 0, 1, 2, \ldots, L-1$, is a one-dimensional array of dimension $L \times 1$. (2) All mappings from r to s and from s to z are simple table lookups between a given

pixel value and these arrays. (3) Each of the elements of these arrays, for example, s_k, contains two important pieces of information: The subscript k denotes the location of the element in the array, and s denotes the value at that location. (4) We need to be concerned only with integer pixel values. For example, in the case of an 8-bit image, $L = 256$ and the elements of each of the arrays just mentioned are integers between 0 and 255. This implies that we now work with gray level values in the interval $[0, L-1]$ instead of the normalized interval $[0, 1]$ that we used before to simplify the development of histogram processing techniques.

In order to see how histogram matching actually can be implemented, consider Fig. 3.19(a), ignoring for a moment the connection shown between this figure and Fig. 3.19(c). Figure 3.19(a) shows a hypothetical discrete transformation function $s = T(r)$ obtained from a given image. The first gray level in the image, r_1, maps to s_1; the second gray level, r_2, maps to s_2; the kth level r_k maps to s_k; and so on (the important point here is the *ordered* correspondence between these values). Each value s_j in the array is precomputed using Eq. (3.3-13), so the process of mapping simply uses the actual value of a pixel as an index in an array to determine the corresponding value of s. This process is particularly easy because we are dealing with integers. For example, the s mapping for an 8-bit pixel with value 127 would be found in the 128th position in array $\{s_j\}$ (recall that we start at 0) out of the possible 256 positions. If we stopped here and mapped the value of each pixel of an input image by the



a b
c

FIGURE 3.19 (a) Graphical interpretation of mapping from r_k to s_k via $T(r)$. (b) Mapping of z_q to its corresponding value v_q via $G(z)$. (c) Inverse mapping from s_k to its corresponding value of z_k.

method just described, the output would be a histogram-equalized image, according to Eq. (3.3-8).

In order to implement histogram matching we have to go one step further. Figure 3.19(b) is a hypothetical transformation function G obtained from a given histogram $p_z(z)$ by using Eq. (3.3-14). For any z_q, this transformation function yields a corresponding value v_q. This mapping is shown by the arrows in Fig. 3.19(b). Conversely, given any value v_q, we would find the corresponding value z_q from G^{-1}. In terms of the figure, all this means graphically is that we would reverse the direction of the arrows to map v_q into its corresponding z_q. However, we know from the definition in Eq. (3.3-14) that $v = s$ for corresponding subscripts, so we can use exactly this process to find the z_k corresponding to any value s_k that we computed previously from the equation $s_k = T(r_k)$. This idea is shown in Fig. 3.19(c).

Since we really do not have the z's (recall that finding these values is precisely the objective of histogram matching), we must resort to some sort of iterative scheme to find z from s. The fact that we are dealing with integers makes this a particularly simple process. Basically, because $v_k = s_k$, we have from Eq. (3.3-14) that the z's for which we are looking must satisfy the equation $G(z_k) = s_k$, or $(G(z_k) - s_k) = 0$. Thus, all we have to do to find the value of z_k corresponding to s_k is to iterate on values of z such that this equation is satisfied for $k = 0, 1, 2, \ldots, L - 1$. This is the same thing as Eq. (3.3-16), except that we do not have to find the inverse of G because we are going to iterate on z. Since we are dealing with integers, the closest we can get to satisfying the equation $(G(z_k) - s_k) = 0$ is to let $z_k = \hat{z}$ for each value of k, where $\hat{z}$ is the *smallest* integer in the interval $[0, L - 1]$ such that

$$(G(\hat{z}) - s_k) \geq 0 \qquad k = 0, 1, 2, \ldots, L - 1. \tag{3.3-17}$$

Given a value s_k, all this means conceptually in terms of Fig. 3.19(c) is that we would start with $\hat{z} = 0$ and increase it in integer steps until Eq. (3.3-17) is satisfied, at which point we let $z_k = \hat{z}$. Repeating this process for all values of k would yield all the required mappings from s to z, which constitutes the implementation of Eq. (3.3-16). In practice, we would not have to start with $\hat{z} = 0$ each time because the values of s_k are known to increase monotonically. Thus, for $k = k + 1$, we would start with $\hat{z} = z_k$ and increment in integer values from there.

The procedure we have just developed for histogram matching may be summarized as follows:

1. Obtain the histogram of the given image.
2. Use Eq. (3.3-13) to precompute a mapped level s_k for each level r_k.
3. Obtain the transformation function G from the given $p_z(z)$ using Eq. (3.3-14).
4. Precompute z_k for each value of s_k using the iterative scheme defined in connection with Eq. (3.3-17).
5. For each pixel in the original image, if the value of that pixel is r_k, map this value to its corresponding level s_k; then map level s_k into the final level z_k. Use the precomputed values from Steps (2) and (4) for these mappings.

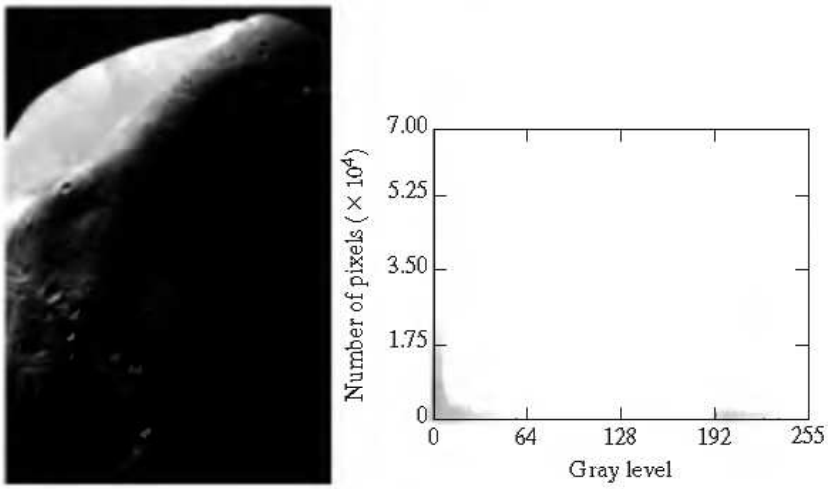
Note that Step (5) implements two mappings for each pixel in the image being processed. The first mapping is nothing more than histogram equalization. If the histogram-equalized image is not required, it obviously would be beneficial to combine both transformations into one in order to save an intermediate step.

Finally, we note that, even in the discrete case, we need to be concerned about G^{-1} satisfying conditions (a) and (b) of the previous section. It is not difficult to show (Problem 3.9) that the only way to guarantee that G^{-1} be single valued and monotonic is to require that G be strictly monotonic (i.e., always increasing), which means simply that none of the values of the specified histogram $p_z(z_i)$ in Eq. (3.3-14) can be zero.

EXAMPLE 3.4: Comparison between histogram equalization and histogram matching.

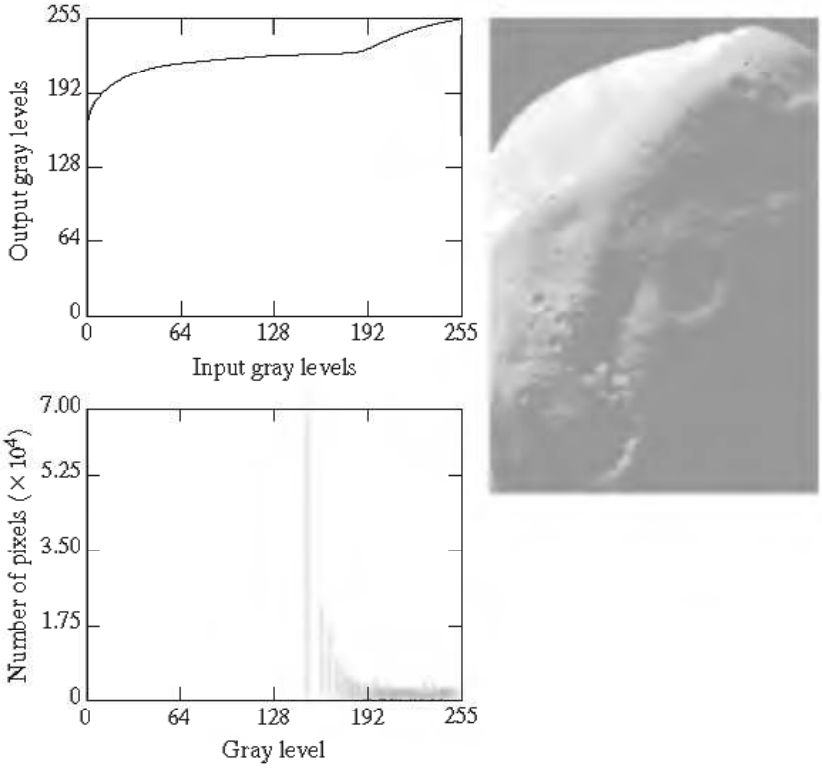
■ Figure 3.20(a) shows an image of the Mars moon, Phobos, taken by NASA's *Mars Global Surveyor*. Figure 3.20(b) shows the histogram of Fig. 3.20(a). The image is dominated by large, dark areas, resulting in a histogram characterized by a large concentration of pixels in the dark end of the gray scale. At first glance, one might conclude that histogram equalization would be a good approach to enhance this image, so that details in the dark areas become more visible. It is demonstrated in the following discussion that this is not so.

Figure 3.21(a) shows the histogram equalization transformation [Eq. (3.3-8) or (3.3-13)] obtained from the histogram shown in Fig. 3.20(b). The most relevant characteristic of this transformation function is how fast it rises from gray level 0 to a level near 190. This is caused by the large concentration of pixels in the input histogram having levels very near 0. When this transformation is applied to the levels of the input image to obtain a histogram-equalized result, the net effect is to map a very narrow interval of dark pixels into the upper end of the gray scale of the output image. Because numerous pixels in the input image have levels precisely in this interval, we would expect the result to be an



a b

FIGURE 3.20 (a) Image of the Mars moon Photos taken by NASA's *Mars Global Surveyor*. (b) Histogram. (Original image courtesy of NASA.)



a b
c

FIGURE 3.21 (a) Transformation function for histogram equalization. (b) Histogram-equalized image (note the washed-out appearance). (c) Histogram of (b).

image with a light, washed-out appearance. As shown in Fig. 3.21(b), this is indeed the case. The histogram of this image is shown in Fig. 3.21(c). Note how all the gray levels are biased toward the upper one-half of the gray scale.

Since the problem with the transformation function in Fig. 3.21(a) was caused by a large concentration of pixels in the original image with levels near 0, a reasonable approach is to modify the histogram of that image so that it does not have this property. Figure 3.22(a) shows a *manually specified* function that preserves the general shape of the original histogram, but has a smoother transition of levels in the dark region of the gray scale. Sampling this function into 256 equally spaced discrete values produced the desired specified histogram. The transformation function $G(z)$ obtained from this histogram using Eq. (3.3-14) is labeled transformation (1) in Fig. 3.22(b). Similarly, the inverse transformation $G^{-1}(s)$ from Eq. (3.3-16) [obtained using the iterative technique discussed in connection with Eq. (3.3-17)] is labeled transformation (2) in Fig. 3.22(b). The enhanced image in Fig. 3.22(c) was obtained by applying transformation (2) to the pixels of the histogram-equalized image in Fig. 3.21(b). The improvement of the histogram-specified image over the result obtained by histogram equalization is evident by comparing these two images. It is of interest to note that a rather modest change in the original histogram was all that was required to obtain a significant improvement in enhancement. The histogram of Fig. 3.22(c) is shown in Fig. 3.22(d). The most distinguishing feature of this histogram is how its low end has shifted right toward the lighter region of the gray scale, as desired. ■

a c
b
d

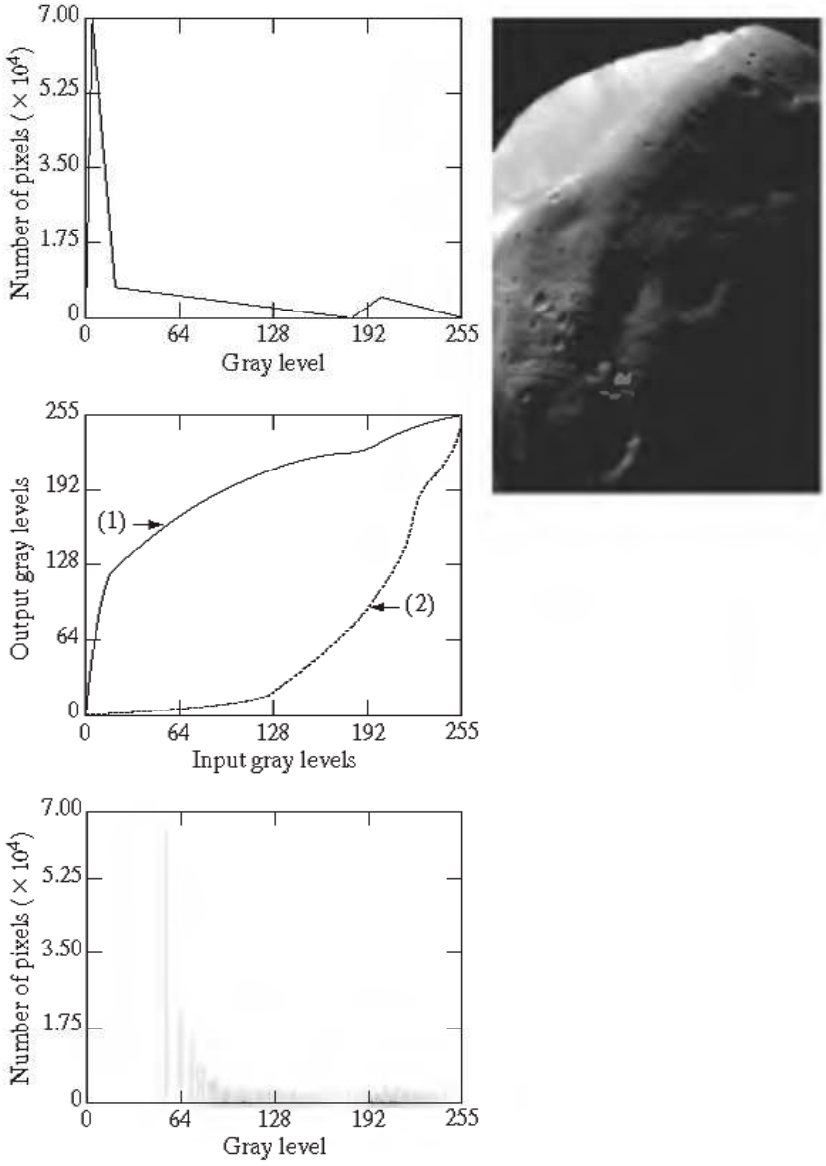


FIGURE 3.22 (a) Specified histogram. (b) Curve (1) is from Eq. (3.3-14), using the histogram in (a); curve (2) was obtained using the iterative procedure in Eq. (3.3-17). (c) Enhanced image using mappings from curve (2). (d) Histogram of (c).

Although it probably is obvious by now, we emphasize before leaving this section that histogram specification is, for the most part, a trial-and-error process. One can use guidelines learned from the problem at hand, just as we did in the preceding example. At times, there may be cases in which it is possible to formulate what an "average" histogram should look like and use that as the specified histogram. In cases such as these, histogram specification becomes a straightforward process. In general, however, there are no rules for specifying histograms, and one must resort to analysis on a case-by-case basis for any given enhancement task.

3.3.3 Local Enhancement

The histogram processing methods discussed in the previous two sections are *global*, in the sense that pixels are modified by a transformation function based on the gray-level content of an entire image. Although this global approach is suitable for overall enhancement, there are cases in which it is necessary to enhance details over small areas in an image. The number of pixels in these areas may have negligible influence on the computation of a global transformation whose shape does not necessarily guarantee the desired local enhancement. The solution is to devise transformation functions based on the gray-level distribution—or other properties—in the neighborhood of every pixel in the image. Although processing methods based on neighborhoods are the topic of Section 3.5, we discuss local histogram processing here for the sake of clarity and continuity. The reader will have no difficulty in following the discussion.

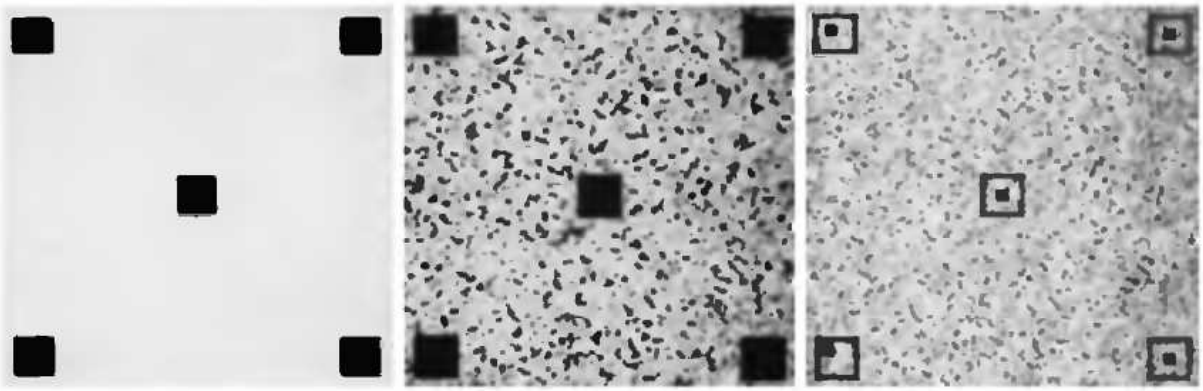
The histogram processing techniques previously described are easily adaptable to local enhancement. The procedure is to define a square or rectangular neighborhood and move the center of this area from pixel to pixel. At each location, the histogram of the points in the neighborhood is computed and either a histogram equalization or histogram specification transformation function is obtained. This function is finally used to map the gray level of the pixel centered in the neighborhood. The center of the neighborhood region is then moved to an adjacent pixel location and the procedure is repeated. Since only one new row or column of the neighborhood changes during a pixel-to-pixel translation of the region, updating the histogram obtained in the previous location with the new data introduced at each motion step is possible (Problem 3.11). This approach has obvious advantages over repeatedly computing the histogram over all pixels in the neighborhood region each time the region is moved one pixel location. Another approach used some times to reduce computation is to utilize nonoverlapping regions, but this method usually produces an undesirable checkerboard effect.

EXAMPLE 3.5: Enhancement using local histograms.

■ Figure 3.23(a) shows an image that has been slightly blurred to reduce its noise content (see Section 3.6.1 regarding blurring). Figure 3.23(b) shows the result of global histogram equalization. As is often the case when this technique is applied to smooth, noisy areas, Fig. 3.23(b) shows considerable enhancement of the noise, with a slight increase in contrast. Note that no new structural details were brought out by this method. However, local histogram equalization using a 7×7 neighborhood revealed the presence of small squares inside the larger dark squares. The small squares were too close in gray level to the larger ones, and their sizes were too small to influence global histogram equalization significantly. Note also the finer noise texture in Fig. 3.23(c), a result of local processing using relatively small neighborhoods. ■

3.3.4 Use of Histogram Statistics for Image Enhancement

Instead of using the image histogram directly for enhancement, we can use instead some statistical parameters obtainable directly from the histogram. Let r denote a discrete random variable representing discrete gray-levels in the range



a b c

FIGURE 3.23 (a) Original image. (b) Result of global histogram equalization. (c) Result of local histogram equalization using a 7×7 neighborhood about each pixel.

$[0, L-1]$, and let $p(r_i)$ denote the normalized histogram component corresponding to the ith value of r. As indicated previously in this section, we may view $p(r_i)$ as an estimate of the probability of occurrence of gray level r_i. The nth moment of r about its mean is defined as

$$\mu_n(r) = \sum_{i=0}^{L-1} (r_i - m)^n p(r_i) \tag{3.3-18}$$

where m is the mean value of r (its average gray level):

$$m = \sum_{i=0}^{L-1} r_i p(r_i). \tag{3.3-19}$$

It follows from Eqs. (3.3-18) and (3.3-19) that $\mu_0 = 1$ and $\mu_1 = 0$. The second moment is given by

$$\mu_2(r) = \sum_{i=0}^{L-1} (r_i - m)^2 p(r_i). \tag{3.3-20}$$

We recognize this expression as the variance of r, which is denoted conventionally by $\sigma^2(r)$. The standard deviation is defined simply as the square root of the variance. We will revisit moments in Chapter 11 in connection with image description. In terms of enhancement, however, we are interested primarily in the mean, which is a measure of average gray level in an image, and the variance (or standard deviation), which is a measure of average contrast.

We consider two uses of the mean and variance for enhancement purposes. The *global* mean and variance are measured over an entire image and are useful primarily for gross adjustments of overall intensity and contrast. A much more powerful use of these two measures is in local enhancement, where the *local* mean and variance are used as the basis for making changes that depend on image characteristics in a predefined region about each pixel in the image.

Let (x, y) be the coordinates of a pixel in an image, and let S_{xy} denote a neighborhood (subimage) of specified size, centered at (x, y). From Eq. (3.3-19) the mean value $m_{S_{xy}}$ of the pixels in S_{xy} can be computed using the expression

$$m_{S_{xy}} = \sum_{(s,t)\in S_{xy}} r_{s,t} p(r_{s,t}) \tag{3.3-21}$$

where $r_{s,t}$ is the gray level at coordinates (s, t) in the neighborhood, and $p(r_{s,t})$ is the neighborhood normalized histogram component corresponding to that value of gray level. Similarly, from Eq. (3.3-20), the gray-level variance of the pixels in region S_{xy} is given by

$$\sigma^2_{S_{xy}} = \sum_{(s,t)\in S_{xy}} [r_{s,t} - m_{S_{xy}}]^2 p(r_{s,t}). \tag{3.3-22}$$

The local mean is a measure of average gray level in neighborhood S_{xy}, and the variance (or standard deviation) is a measure of contrast in that neighborhood.

An important aspect of image processing using the local mean and variance is the flexibility they afford in developing simple, yet powerful enhancement techniques based on statistical measures that have a close, predictable correspondence with image appearance. We illustrate these characteristics by means of an example.

EXAMPLE 3.6: Enhancement based on local statistics.

■ Figure 3.24 shows an SEM (scanning electron microscope) image of a tungsten filament wrapped around a support. The filament in the center of the image and its support are quite clear and easy to study. There is another filament structure on the right side of the image, but it is much darker and its size and other features are not as easily discernable. Local enhancement by contrast manipulation is an ideal approach to try on problems such as this, where part of the image is acceptable, but other parts may contain hidden features of interest.

In this particular case, the problem is to enhance dark areas while leaving the light area as unchanged as possible since it does note require enhancement. We can use the concepts presented in this section to formulate an enhancement method that can tell the difference between dark and light and, at the same time, is capable of enhancing only the dark areas. A measure of whether an area is relatively light or dark at a point (x, y) is to compare the local average gray level $m_{S_{xy}}$ to the average image gray level, called the global mean and denoted M_G. This latter quantity is obtained by letting S encompass the entire image. Thus, we have the first element of our enhancement scheme: We will consider the pixel at a point (x, y) as a candidate for processing if $m_{S_{xy}} \leq k_0 M_G$, where k_0 is a positive constant with value less than 1.0. Since we are interested in enhancing areas that have low contrast, we also need a measure to determine whether the contrast of an area makes it a candidate for enhancement. Thus, we will consider the pixel at a point (x, y) as a candidate for enhancement if $\sigma_{S_{xy}} \leq k_2 D_G$, where D_G is the global standard deviation and k_2 is a positive constant. The value of this constant will be greater than 1.0 if we are interested in enhancing light areas and less than 1.0 for dark areas. Finally, we need to restrict

the lowest values of contrast we are willing to accept, otherwise the procedure would attempt to enhance even constant areas, whose standard deviation is zero. Thus, we also set a lower limit on the local standard deviation by requiring that $k_1 D_G \leq \sigma_{S_{xy}}$, with $k < k_2$. A pixel at (x, y) that meets all the conditions for local enhancement is processed simply by multiplying it by a specified constant, E, to increase (or decrease) the value of its gray level relative to the rest of the image. The values of pixels that do not meet the enhancement conditions are left unchanged.

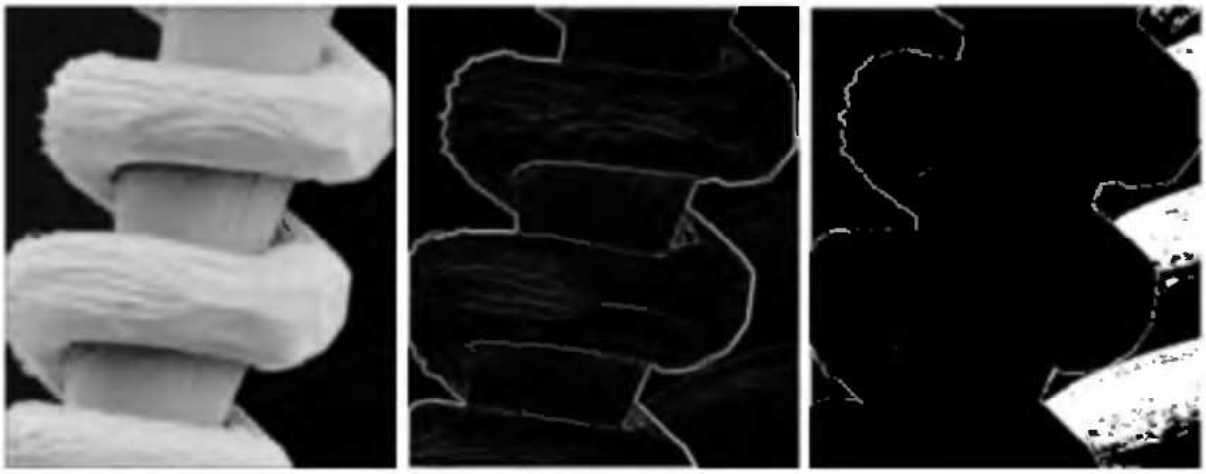
A summary of the enhancement method is as follows. Let $f(x, y)$ represent the value of an image pixel at any image coordinates (x, y), and let $g(x, y)$ represent the corresponding enhanced pixel at those coordinates. Then

$$g(x, y) = \begin{cases} E \cdot f(x, y) & \text{if } m_{S_{xy}} \leq k_0 M_G \quad \text{AND} \quad k_1 D_G \leq \sigma_{S_{xy}} \leq k_2 D_G \\ f(x, y) & \text{otherwise} \end{cases}$$

where, as indicated previously, E, k_0, k_1, and k_2 are specified parameters; M_G is the global mean of the input image; and D_G is its global standard deviation.

Normally, making a successful selection of parameters requires a bit of experimentation to gain familiarity with a given image or class of images. In this case, the following values were selected: $E = 4.0$, $k_0 = 0.4$, $k_1 = 0.02$, and $k_2 = 0.4$. The relatively low value of 4.0 for E was chosen so that, when it was multiplied by the levels in the areas being enhanced (which are dark), the result would still tend toward the dark end of the scale, and thus preserve the general visual balance of the image. The value of k_0 was chosen as somewhat less than half the global mean since it is obvious by looking at the image that the areas that require enhancement definitely are dark enough to be below half the global mean. A similar analysis led to the choice of values for k_1 and k_2. Choosing these constants is not a difficult task in general, but their choice

FIGURE 3.24 SEM image of a tungsten filament and support, magnified approximately 130×. (Original image courtesy of Mr. Michael Shaffer, Department of Geological Sciences, University of Oregon, Eugene).



a b c

FIGURE 3.25 (a) Image formed from all local means obtained from Fig. 3.24 using Eq. (3.3-21). (b) Image formed from all local standard deviations obtained from Fig. 3.24 using Eq. (3.3-22). (c) Image formed from all multiplication constants used to produce the enhanced image shown in Fig. 3.26.

definitely must be guided by a logical analysis of the enhancement problem at hand. Finally, the choice of size for the local area should be as small as possible in order to preserve detail and keep the computational burden as low as possible. We chose a small (3×3) local region.

Figure 3.25(a) shows the values of $m_{S_{xy}}$ for all values of (x, y). Since the value of $m_{S_{xy}}$ for each (x, y) is the average of the neighboring pixels in a 3×3 area centered at (x, y), we expect the result to be similar to the original image, but

FIGURE 3.26 Enhanced SEM image. Compare with Fig. 3.24. Note in particular the enhanced area on the right, bottom side of the image.

slightly blurred. This indeed is the case in Fig. 3.25(a). Figure 3.25(b) shows in image formed using all the values of $\sigma_{S_{xy}}$. Similarly, we can construct an image out the values that multiply $f(x, y)$ at each coordinate pair (x, y) to form $g(x, y)$. Since the values are either 1 or E, the image is binary, as shown in Fig. 3.25(c). The dark areas correspond to 1 and the light areas to E. Thus, any light point in Fig. 3.25(c) signifies a coordinate pair (x, y) at which the enhancement procedure multiplied $f(x, y)$ by E to produce an enhanced pixel. The dark points represent coordinates at which the procedure did not to modify the pixel values.

The enhanced image obtained with the method just described is shown in Fig. 3.26. In comparing this image with the original in Fig. 3.24, we note the obvious detail that has been brought out on the right side of the enhanced image. It is worthwhile to point out that the unenhanced portions of the image (the light areas) were left intact for the most part. We do note the appearance of some small bright dots in the shadow areas where the coil meets the support stem, and around some of the borders between the filament and the background. These are undesirable artifacts created by the enhancement technique. In other words, the points appearing as light dots met the criteria for enhancement and their values were amplified by factor E. Introduction of artifacts is a definite drawback of a method such as the one just described because of the nonlinear way in which they process an image. The key point here, however, is that the image was enhanced in a most satisfactory way as far as bringing out the desired detail. ■

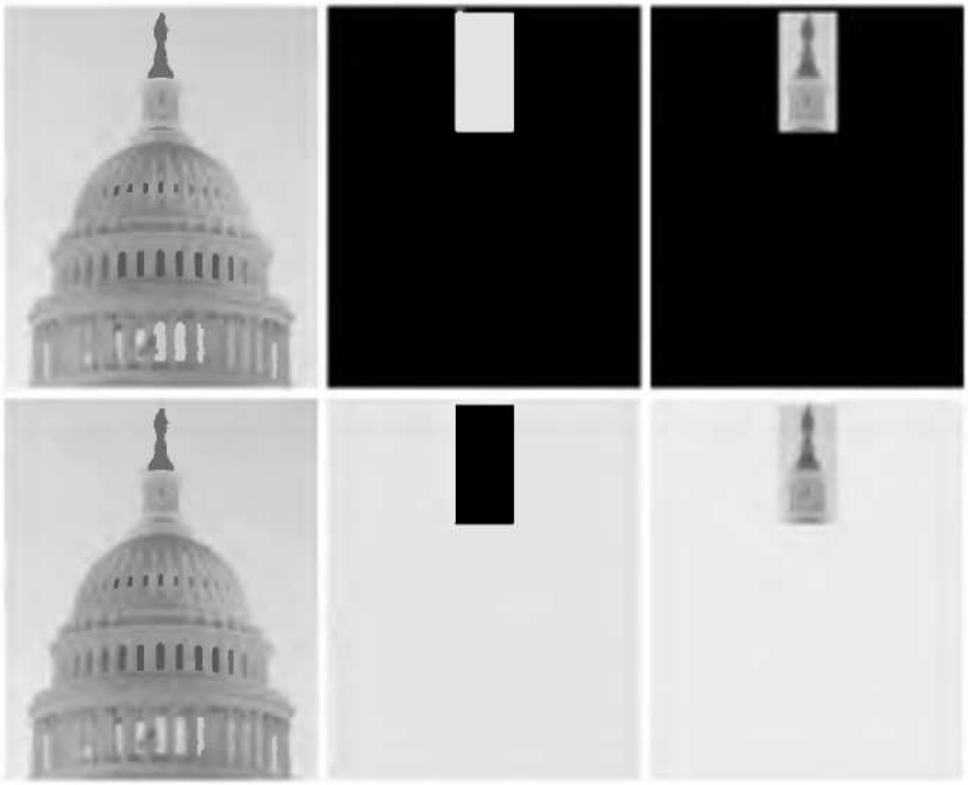
It is not difficult to imagine the numerous ways in which the example just given could be adapted or extended to other situations in which local enhancement is applicable.

3.4 Enhancement Using Arithmetic/Logic Operations

Arithmetic/logic operations involving images are performed on a pixel-by-pixel basis between two or more images (this excludes the logic operation NOT, which is performed on a single image). As an example, subtraction of two images results in a new image whose pixel at coordinates (x, y) is the difference between the pixels in that same location in the two images being subtracted. Depending on the hardware and/or software being used, the actual mechanics of implementing arithmetic/logic operations can be done sequentially, one pixel at a time, or in parallel, where all operations are performed simultaneously.

Logic operations similarly operate on a pixel-by-pixel basis[†]. We need only be concerned with the ability to implement the AND, OR, and NOT logic operators because these three operators are *functionally complete*. In other words, any other logic operator can be implemented by using only these three basic functions. When dealing with logic operations on gray-scale images, pixel values are processed as strings of binary numbers. For example, performing the NOT operation on a black, 8-bit pixel (a string of eight 0's) produces a white pixel

[†] Recall that, for two binary variables a and b: aANDb yields 1 only when both a and b are 1; otherwise the result is 0. Similarly, aORb is 0 when both variables are 0; otherwise the result is 1. Finally, if a is 1, NOT (a) is 0, and vice versa.



a b c
d e f

FIGURE 3.27 (a) Original image. (b) AND image mask. (c) Result of the AND operation on images (a) and (b). (d) Original image. (e) OR image mask. (f) Result of operation OR on images (d) and (e).

(a string of eight 1's). Intermediate values are processed the same way, changing all 1's to 0's and vice versa. Thus, the NOT logic operator performs the same function as the negative transformation of Eq. (3.2-1). The AND and OR operations are used for *masking*; that is, for selecting subimages in an image, as illustrated in Fig. 3.27. In the AND and OR image masks, light represents a binary 1 and dark represents a binary 0. Masking sometimes is referred to as *region of interest* (ROI) processing. In terms of enhancement, masking is used primarily to isolate an area for processing. This is done to highlight that area and differentiate it from the rest of the image. Logic operations also are used frequently in conjunction with morphological operations, as discussed in Chapter 9.

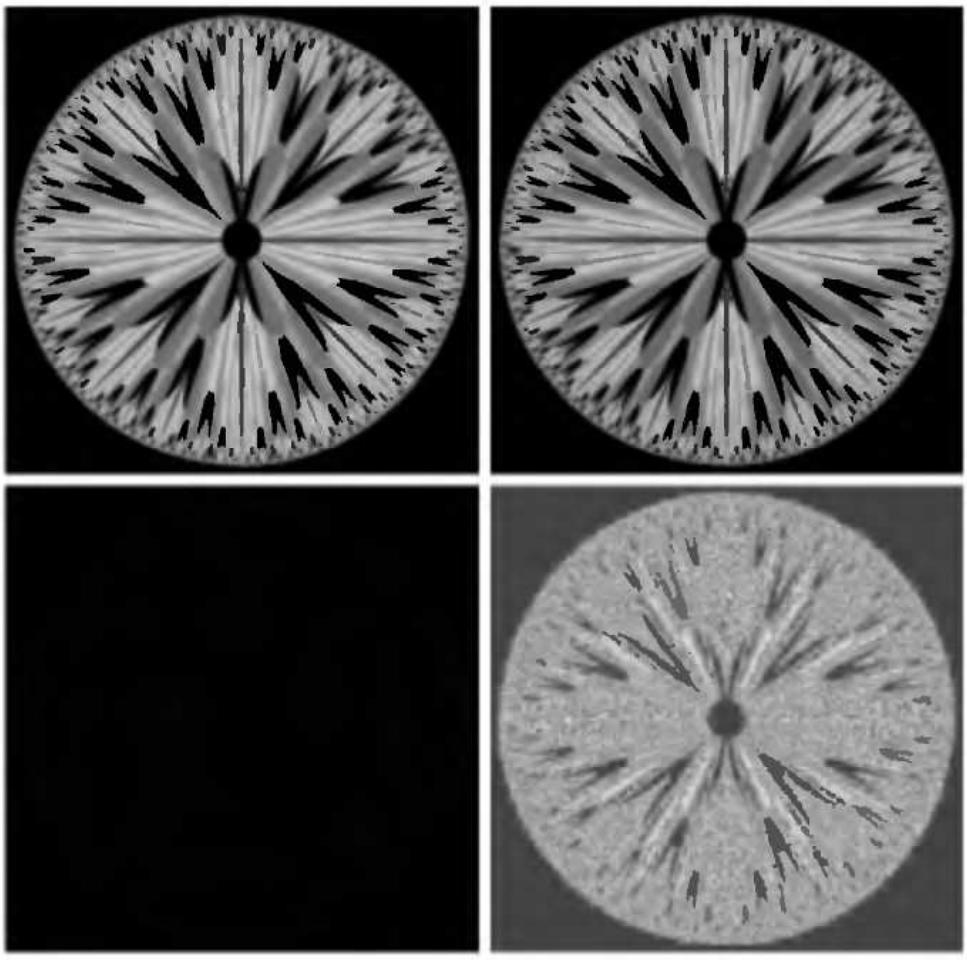
Of the four arithmetic operations, subtraction and addition (in that order) are the most useful for image enhancement. We consider division of two images simply as multiplication of one image by the reciprocal of the other. Aside from the obvious operation of multiplying an image by a constant to increase its average gray level, image multiplication finds use in enhancement primarily as a masking operation that is more general than the logical masks discussed in the previous paragraph. In other words, multiplication of one image by another can be used to implement gray-level, rather than binary, masks. We give an example in Section 3.8 of how such a masking operation can be a useful tool. In the remainder of this section, we develop and illustrate methods based on subtraction and addition for image enhancement. Other uses of image multiplication are discussed in Chapter 5, in the context of image restoration.

3.4.1 Image Subtraction

The difference between two images $f(x, y)$ and $h(x, y)$, expressed as

$$g(x, y) = f(x, y) - h(x, y), \tag{3.4-1}$$

is obtained by computing the difference between all pairs of corresponding pixels from f and h. The key usefulness of subtraction is the enhancement of *differences* between images. We illustrate this concept by returning briefly to the discussion in Section 3.2.4, where we showed that the higher-order bit planes of an image carry a significant amount of visually relevant detail, while the lower planes contribute more to fine (often imperceptible) detail. Figure 3.28(a) shows the fractal image used earlier to illustrate the concept of bit planes. Figure 3.28(b) shows the result of discarding (setting to zero) the four least significant bit planes of the original image. The images are nearly identical visually, with the exception of a very slight drop in overall contrast due to less variability of the gray-level values in the image of Fig. 3.28(b). The pixel-by-pixel difference between these two images is shown in Fig. 3.28(c). The differences in pixel values are so small that the difference image appears nearly black when displayed on an 8-bit



a b
c d

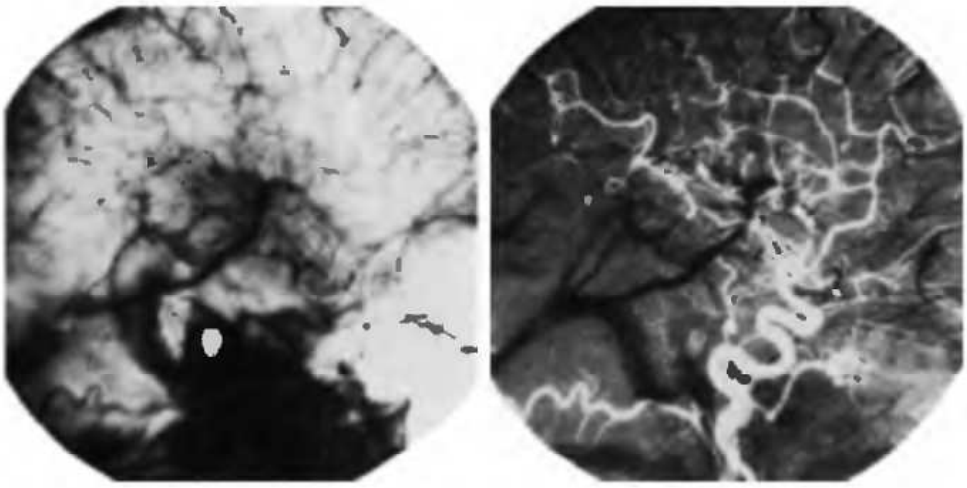
FIGURE 3.28 (a) Original fractal image. (b) Result of setting the four lower-order bit planes to zero. (c) Difference between (a) and (b). (d) Histogram-equalized difference image. (Original image courtesy of Ms. Melissa D. Binde, Swarthmore College, Swarthmore, PA).

display. In order to bring out more detail, we can perform a contrast stretching transformation, such as those discussed in Sections 3.2 or 3.3. We chose histogram equalization, but an appropriate power-law transformation would have done the job also. The result is shown in Fig. 3.28(d). This is a very useful image for evaluating the effect of setting to zero the lower-order planes.

EXAMPLE 3.7: Use of image subtraction in mask mode radiography.

■ One of the most commercially successful and beneficial uses of image subtraction is in the area of medical imaging called *mask mode radiography*. In this case $h(x, y)$, the *mask*, is an X-ray image of a region of a patient's body captured by an intensified TV camera (instead of traditional X-ray film) located opposite an X-ray source. The procedure consists of injecting a contrast medium into the patient's bloodstream, taking a series of images of the same anatomical region as $h(x, y)$, and subtracting this mask from the series of incoming images after injection of the contrast medium. The net effect of subtracting the mask from each sample in the incoming stream of TV images is that the areas that are different between $f(x, y)$ and $h(x, y)$ appear in the output image as enhanced detail. Because images can be captured at TV rates, this procedure in essence gives a movie showing how the contrast medium propagates through the various arteries in the area being observed.

Figure 3.29(a) shows an X-ray image of the top of a patient's head prior to injection of an iodine medium into the bloodstream. The camera yielding this image was positioned above the patient's head, looking down. As a reference point, the bright spot in the lower one-third of the image is the core of the spinal column. Figure 3.29(b) shows the difference between the mask (Fig. 3.29a) and an image taken some time after the medium was introduced into the bloodstream. The bright arterial paths carrying the medium are unmistakably enhanced in Fig. 3.29(b). These arteries appear quite bright because they are not subtracted out (that is, they are not part of the mask image). The overall background is much darker than that in Fig. 3.29(a) because differences between areas of little change yield low values, which in turn appear as dark shades of gray in the difference image. Note, for instance, that the spinal cord, which is bright in Fig. 3.29(a), appears quite dark in Fig. 3.29(b) as a result of subtraction. ■



a b

FIGURE 3.29 Enhancement by image subtraction. (a) Mask image. (b) An image (taken after injection of a contrast medium into the bloodstream) with mask subtracted out.

A few comments on implementation are an order before we leave this section. In practice, most images are displayed using 8 bits (even 24-bit color images consists of three separate 8-bit channels). Thus, we expect image values not to be outside the range from 0 to 255. The values in a difference image can range from a minimum of −255 to a maximum of 255, so some sort of scaling is required to display the results. There are two principal ways to scale a difference image. One method is to add 255 to every pixel and then divide by 2. It is not guaranteed that the values will cover the entire 8-bit range from 0 to 255, but all pixel values definitely will be within this range. This method is fast and simple to implement, but it has the limitations that the full range of the display may not be utilized and, potentially more serious, the truncation inherent in the division by 2 will generally cause loss in accuracy.

If more accuracy and full coverage of the 8-bit range are desired, then we can resort to another approach. First, the value of the minimum difference is obtained and its negative added to all the pixels in the difference image (this will create a modified difference image whose minimum values is 0). Then, all the pixels in the image are scaled to the interval $[0, 255]$ by multiplying each pixel by the quantity 255/Max, where Max is the maximum pixel value in the modified difference image. It is evident that this approach is considerably more complex and difficult to implement.

Before leaving this section we note also that change detection via image subtraction finds another major application in the area of segmentation, which is the topic of Chapter 10. Basically, segmentation techniques attempt to subdivide an image into regions based on a specified criterion. Image subtraction for segmentation is used when the criterion is "changes." For instance, in tracking (segmenting) moving vehicles in a sequence of images, subtraction is used to remove all stationary components in an image. What is left should be the moving elements in the image, plus noise.

3.4.2 Image Averaging

Consider a noisy image $g(x, y)$ formed by the addition of noise $\eta(x, y)$ to an original image $f(x, y)$; that is,

$$g(x, y) = f(x, y) + \eta(x, y) \tag{3.4-2}$$

where the assumption is that at every pair of coordinates (x, y) the noise is uncorrelated† and has zero average value. The objective of the following procedure is to reduce the noise content by adding a set of noisy images, $\{g_i(x, y)\}$.

If the noise satisfies the constraints just stated, it can be shown (Problem 3.15) that if an image $\bar{g}(x, y)$ is formed by averaging K different noisy images,

$$\bar{g}(x, y) = \frac{1}{K} \sum_{i=1}^{K} g_i(x, y) \tag{3.4-3}$$

†Recall that the variance of a random variable x with mean m is defined as $E[(x - m)^2]$, where $E\{\cdot\}$ is the expected value of the argument. The covariance of two random variables x_i and x_j is defined as $E[(x_i - m_i)(x_j - m_j)]$. If the variables are *uncorrelated*, their covariance is 0.

then it follows that

$$E\{\bar{g}(x, y)\} = f(x, y) \tag{3.4-4}$$

and

$$\sigma^2_{\bar{g}(x,y)} = \frac{1}{K}\sigma^2_{\eta(x,y)} \tag{3.4-5}$$

where $E\{\bar{g}(x, y)\}$ is the expected value of $\bar{g}$, and $\sigma^2_{\bar{g}(x,y)}$ and $\sigma^2_{\eta(x,y)}$ are the variances of $\bar{g}$ and η, all at coordinates (x, y). The standard deviation at any point in the average image is

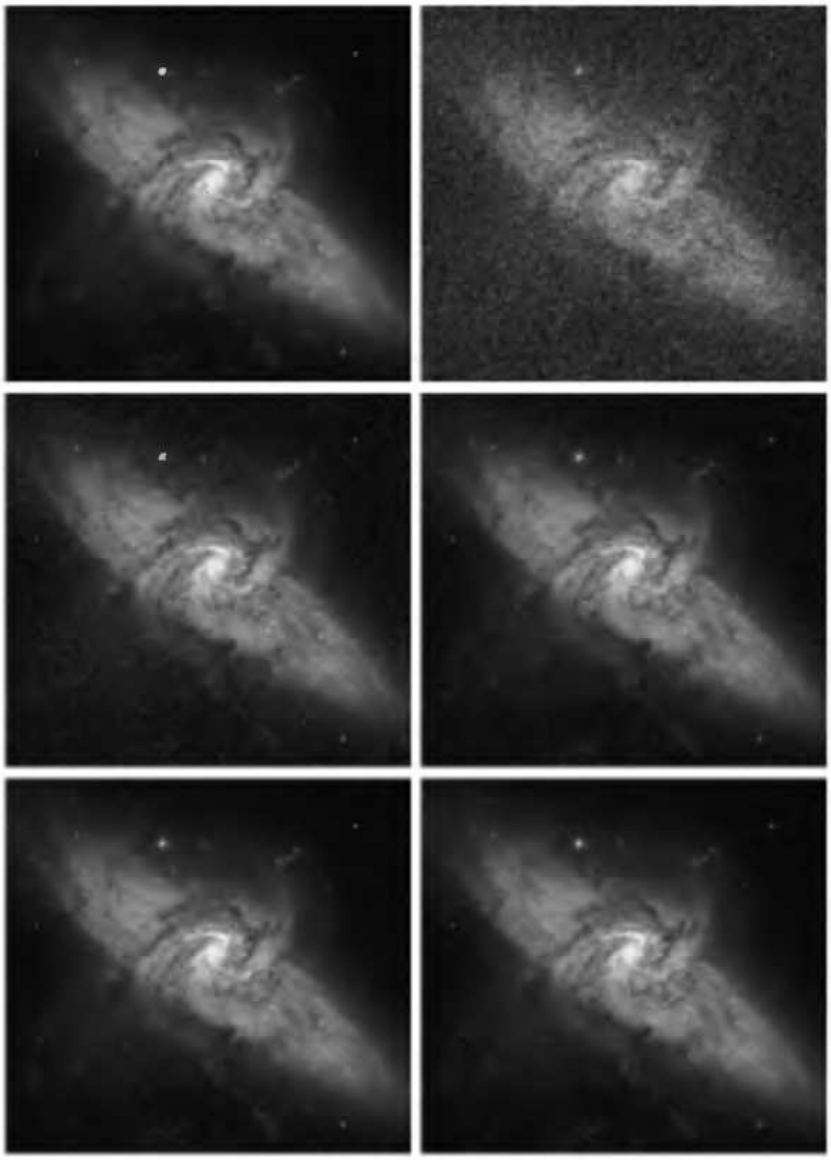
$$\sigma_{\bar{g}(x,y)} = \frac{1}{\sqrt{K}}\sigma_{\eta(x,y)}. \tag{3.4-6}$$

As K increases, Eqs. (3.4-5) and (3.4-6) indicate that the variability (noise) of the pixel values at each location (x, y) decreases. Because $E\{\bar{g}(x, y)\} = f(x, y)$, this means that $\bar{g}(x, y)$ approaches $f(x, y)$ as the number of noisy images used in the averaging process increases. In practice, the images $g_i(x, y)$ must be registered (aligned) in order to avoid the introduction of blurring and other artifacts in the output image.

EXAMPLE 3.8: Noise reduction by image averaging.

■ An important application of image averaging is in the field of astronomy, where imaging with very low light levels is routine, causing sensor noise frequently to render single images virtually useless for analysis. Figure 3.30(a) shows an image of a galaxy pair called NGC 3314, taken by NASA's Hubble Space Telescope with a wide field planetary camera. NGC 3314 lies about 140 million light-years from Earth, in the direction of the southern-hemisphere constellation Hydra. The bright stars forming a pinwheel shape near the center of the front galaxy have formed recently from interstellar gas and dust. Figure 3.30(b) shows the same image, but corrupted by uncorrelated Gaussian noise with zero mean and a standard deviation of 64 gray levels. This image is useless for all practical purposes. Figures 3.30(c) through (f) show the results of averaging 8, 16, 64, and 128 images, respectively. We see that the result obtained with $K = 128$ is reasonably close to the original in visual appearance.

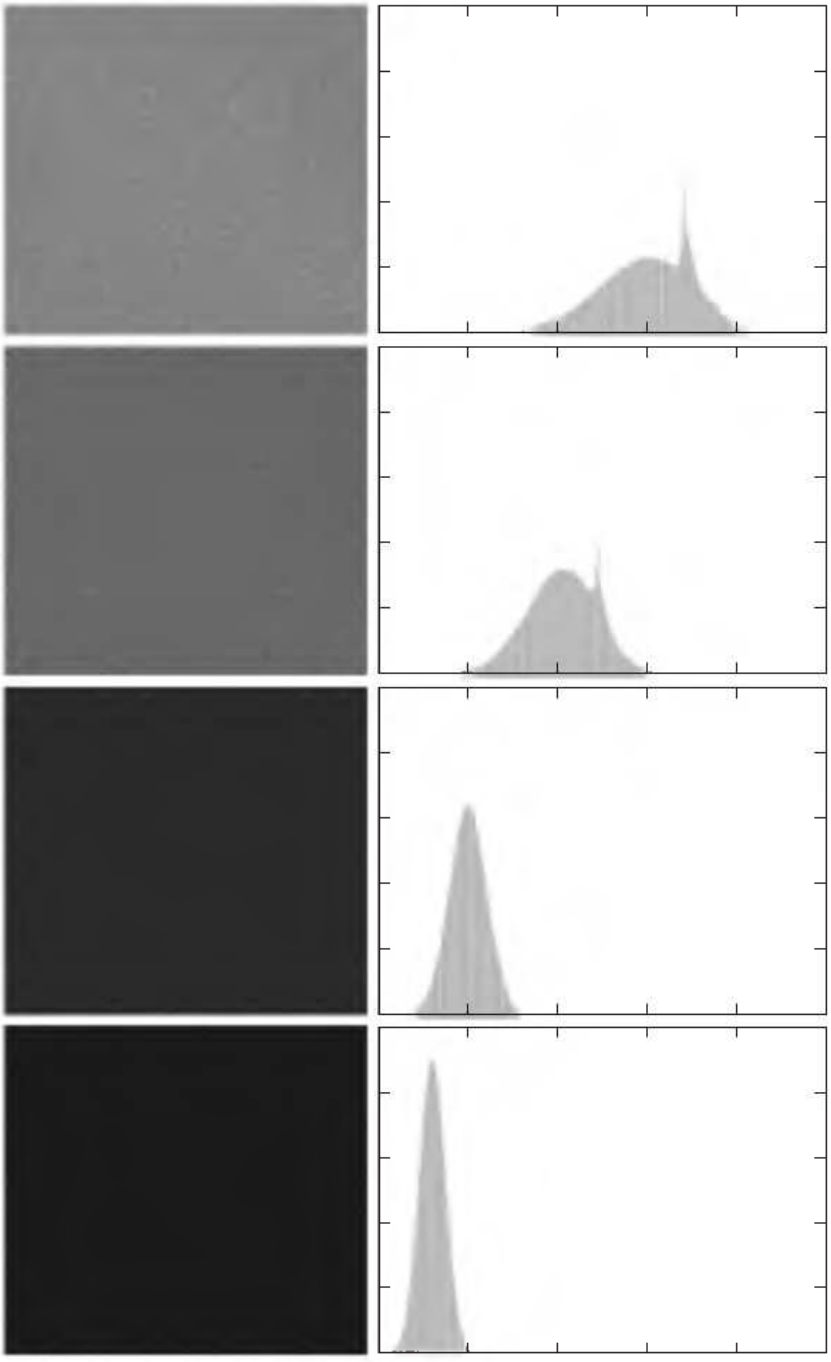
We can get a better appreciation from Fig. 3.31 for how reduction in the visual appearance of noise takes place as a function of increasing K. This figure shows the difference images between the original [Fig. 3.30(a)] and each of the averaged images in Figs. 3.30(c) through (f). The histograms corresponding to the difference images are also shown in the figure. As usual, the vertical scale in the histograms represents number of pixels and is in the range $[0, 2.6 \times 10^4]$. The horizontal scale represents gray level and is in the range $[0, 255]$. Notice in the histograms that the mean and standard deviation of the difference images decrease as K increases. This is as expected because, according to Eqs. (3.4-3) and (3.4-4), the average image should approach the original as K increases. We can also see the effect of a decreasing mean in the difference images on the left column of Fig. 3.31, which become darker as the K increases.



a b
c d
e f

FIGURE 3.30 (a) Image of Galaxy Pair NGC 3314. (b) Image corrupted by additive Gaussian noise with zero mean and a standard deviation of 64 gray levels. (c)–(f) Results of averaging $K = 8, 16, 64$, and 128 noisy images. (Original image courtesy of NASA.)

Addition is the discrete formulation of continuous integration. In astronomical observations, a process equivalent to the method just described is to use the integrating capabilities of CCD or similar sensors for noise reduction by observing the same scene over long periods of time. The net effect, however, is analogous to the procedure just discussed. Cooling the sensor further reduces its noise level. ■



a b

FIGURE 3.31 (a) From top to bottom: Difference images between Fig. 3.30(a) and the four images in Figs. 3.30(c) through (f), respectively. (b) Corresponding histograms.

As in the case of image subtraction, adding two or more 8-bit images requires special care when it comes to displaying the result on an 8-bit display. The values in the sum of K, 8-bit images can range from 0 to $255 \times K$. Scaling back to 8 bits in this case consists simply of dividing the result by K. Naturally, some accuracy will be lost in the process, but this is unavoidable if the display has to be limited to 8 bits.

It is possible in some implementations of image averaging to have negative values when noise is added to an image. In fact, in the example just given, this was precisely the case because Gaussian random variables with zero mean and nonzero variance have negative as well as positive values. The images in the example were scaled using the second scaling method discussed at the end of the previous section. That is, the minimum value in a given average image was obtained and its negative was added to the image. Then all the pixels in the modified image were scaled to the range $[0, 255]$ by multiplying each pixel in the modified image by the quantity 255/Max, where Max was the maximum pixel value in that image.

3.5 Basics of Spatial Filtering

As mentioned in Section 3.1, some neighborhood operations work with the values of the image pixels in the neighborhood *and* the corresponding values of a subimage that has the same dimensions as the neighborhood. The subimage is called a *filter, mask, kernel, template,* or *window*, with the first three terms being the most prevalent terminology. The values in a filter subimage are referred to as *coefficients*, rather than pixels.

The concept of filtering has its roots in the use of the Fourier transform for signal processing in the so-called *frequency domain*. This topic is discussed in more detail in Chapter 4. In the present chapter, we are interested in filtering operations that are performed directly on the pixels of an image. We use the term *spatial filtering* to differentiate this type of process from the more traditional frequency domain filtering.

The mechanics of spatial filtering are illustrated in Fig. 3.32. The process consists simply of moving the filter mask from point to point in an image. At each point (x, y), the *response* of the filter at that point is calculated using a predefined relationship. For *linear* spatial filtering (see Section 2.6 regarding linearity), the response is given by a sum of products of the filter coefficients and the corresponding image pixels in the area spanned by the filter mask. For the 3×3 mask shown in Fig. 3.32, the result (or response), R, of linear filtering with the filter mask at a point (x, y) in the image is

$$\begin{aligned} R = {} & w(-1,-1)f(x-1, y-1) + w(-1,0)f(x-1, y) + \cdots \\ & + w(0,0)f(x, y) + \cdots + w(1,0)f(x+1, y) + w(1,1)f(x+1, y+1), \end{aligned}$$

which we see is the sum of products of the mask coefficients with the corresponding pixels directly under the mask. Note in particular that the coefficient $w(0,0)$ coincides with image value $f(x, y)$, indicating that the mask is centered at (x, y) when the computation of the sum of products takes place. For a mask of size $m \times n$, we assume that $m = 2a + 1$ and $n = 2b + 1$, where a and b are nonnegative integers. All this says is that our focus in the following discussion will be on masks of *odd* sizes, with the smallest meaningful size being 3×3 (we exclude from our discussion the trivial case of a 1×1 mask).

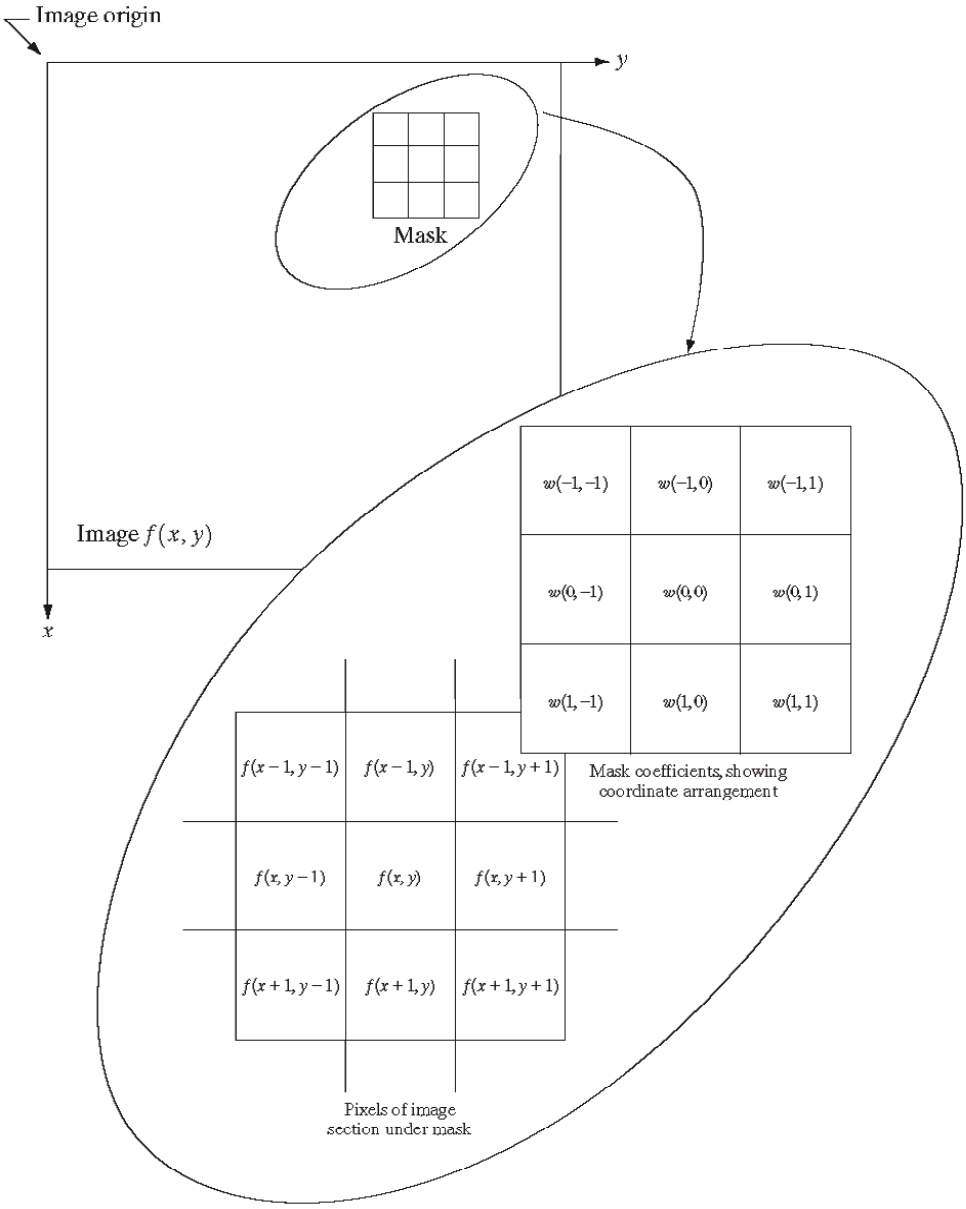


FIGURE 3.32 The mechanics of spatial filtering. The magnified drawing shows a 3×3 mask and the image section directly under it; the image section is shown displaced out from under the mask for ease of readability.

In general, linear filtering of an image f of size $M \times N$ with a filter mask of size $m \times n$ is given by the expression:

$$g(x, y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s, t) f(x + s, y + t) \tag{3.5-1}$$

where, from the previous paragraph, $a = (m - 1)/2$ and $b = (n - 1)/2$. To generate a complete filtered image this equation must be applied for $x = 0, 1, 2, \ldots, M - 1$ and $y = 0, 1, 2, \ldots, N - 1$. In this way, we are assured that the

mask processes all pixels in the image. It is easily verified when $m = n = 3$ that this expression reduces to the example given in the previous paragraph.

As discussed in Chapter 4, the process of linear filtering given in Eq. (3.5-1) is similar to a frequency domain concept called *convolution*. For this reason, linear spatial filtering often is referred to as "convolving a mask with an image." Similarly, filter masks are sometimes called *convolution masks*. The term *convolution kernel* also is in common use.

When interest lies on the response, R, of an $m \times n$ mask at any point (x, y), and not on the mechanics of implementing mask convolution, it is common practice to simplify the notation by using the following expression:

$$\begin{aligned} R &= w_1 z_1 + w_2 z_2 + \ldots + w_{mn} z_{mn} \\ &= \sum_{i=1}^{mn} w_i z_i \end{aligned} \tag{3.5-2}$$

where the w's are mask coefficients, the z's are the values of the image gray levels corresponding to those coefficients, and mn is the total number of coefficients in the mask. For the 3×3 general mask shown in Fig. 3.33 the response at any point (x, y) in the image is given by

$$\begin{aligned} R &= w_1 z_1 + w_2 z_2 + \ldots w_9 z_9 \\ &= \sum_{i=1}^{9} w_i z_i. \end{aligned} \tag{3.5-3}$$

We make special mention of this simple formula because it is seen frequently in the published literature on image processing.

Nonlinear spatial filters also operate on neighborhoods, and the mechanics of sliding a mask past an image are the same as was just outlined. In general, however, the filtering operation is based conditionally on the values of the pixels in the neighborhood under consideration, and they do not explicitly use coefficients in the sum-of-products manner described in Eqs. (3.5-1) and (3.5-2). As shown in Section 3.6.2, for example, noise reduction can be achieved effectively with a nonlinear filter whose basic function is to compute the median gray-level value in the neighborhood in which the filter is located. Computation of the median is a nonlinear operation, as is computation of the variance, which we used in Section 3.3.4.

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

FIGURE 3.33 Another representation of a general 3×3 spatial filter mask.

An important consideration in implementing neighborhood operations for spatial filtering is the issue of what happens when the center of the filter approaches the border of the image. Consider for simplicity a square mask of size $n \times n$. At least one edge of such a mask will coincide with the border of the image when the center of the mask is at a distance of $(n - 1)/2$ pixels away from the border of the image. If the center of the mask moves any closer to the border, one or more rows or columns of the mask will be located outside the image plane. There are several ways to handle this situation. The simplest is to limit the excursions of the center of the mask to be at a distance no less than $(n - 1)/2$ pixels from the border. The resulting filtered image will be smaller than the original, but all the pixels in the filtered imaged will have been processed with the full mask. If the result is required to be the same size as the original, then the approach typically employed is to filter all pixels only with the section of the mask that is fully contained in the image. With this approach, there will be bands of pixels near the border that will have been processed with a partial filter mask. Other approaches include "padding" the image by adding rows and columns of 0's (or other constant gray level), or padding by replicating rows or columns. The padding is then stripped off at the end of the process. This keeps the size of the filtered image the same as the original, but the values of the padding will have an effect near the edges that becomes more prevalent as the size of the mask increases. The only way to obtain a perfectly filtered result is to accept a somewhat smaller filtered image by limiting the excursions of the center of the filter mask to a distance no less than $(n - 1)/2$ pixels from the border of the original image.

3.6 Smoothing Spatial Filters

Smoothing filters are used for blurring and for noise reduction. Blurring is used in preprocessing steps, such as removal of small details from an image prior to (large) object extraction, and bridging of small gaps in lines or curves. Noise reduction can be accomplished by blurring with a linear filter and also by nonlinear filtering.

3.6.1 Smoothing Linear Filters

The output (response) of a smoothing, linear spatial filter is simply the average of the pixels contained in the neighborhood of the filter mask. These filters sometimes are called *averaging filters*. For reasons explained in Chapter 4, they also are referred to a *lowpass filters*.

The idea behind smoothing filters is straightforward. By replacing the value of every pixel in an image by the average of the gray levels in the neighborhood defined by the filter mask, this process results in an image with reduced "sharp" transitions in gray levels. Because random noise typically consists of sharp transitions in gray levels, the most obvious application of smoothing is noise reduction. However, edges (which almost always are desirable features of an image) also are characterized by sharp transitions in gray levels, so averaging filters have the undesirable side effect that they blur edges. Another application of this type of process includes the smoothing of false contours that result

b

FIGURE 3.34 Two 3×3 smoothing (averaging) filter masks. The constant multiplier in front of each mask is equal to the sum of the values of its coefficients, as is required to compute an average.

$$\frac{1}{9} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad \frac{1}{16} \times \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

from using an insufficient number of gray levels, as discussed in Section 2.4.3. A major use of averaging filters is in the reduction of "irrelevant" detail in an image. By "irrelevant" we mean pixel regions that are small with respect to the size of the filter mask. This latter application is illustrated later in this section.

Figure 3.34 shows two 3×3 smoothing filters. Use of the first filter yields the standard average of the pixels under the mask. This can best be seen by substituting the coefficients of the mask into Eq. (3.5-3):

$$R = \frac{1}{9} \sum_{i=1}^{9} z_i,$$

which is the average of the gray levels of the pixels in the 3×3 neighborhood defined by the mask. Note that, instead of being 1/9, the coefficients of the filter are all 1's. The idea here is that it is computationally more efficient to have coefficients valued 1. At the end of the filtering process the entire image is divided by 9. An $m \times n$ mask would have a normalizing constant equal to $1/mn$. A spatial averaging filter in which all coefficients are equal is sometimes called a *box filter*.

The second mask shown in Fig. 3.34 is a little more interesting. This mask yields a so-called *weighted average*, terminology used to indicate that pixels are multiplied by different coefficients, thus giving more importance (weight) to some pixels at the expense of others. In the mask shown in Fig. 3.34(b) the pixel at the center of the mask is multiplied by a higher value than any other, thus giving this pixel more importance in the calculation of the average. The other pixels are inversely weighted as a function of their distance from the center of the mask. The diagonal terms are further away from the center than the orthogonal neighbors (by a factor of $\sqrt{2}$) and, thus, are weighed less than these immediate neighbors of the center pixel. The basic strategy behind weighing the center point the highest and then reducing the value of the coefficients as a function of increasing distance from the origin is simply an attempt to reduce blurring in the smoothing process. We could have picked other weights to accomplish the same general objective. However, the sum of all the coefficients in the mask of Fig. 3.34(b) is equal to 16, an attractive feature for computer implementation because it has an integer power of 2. In practice, it is difficult in general to see differences between images smoothed by using either of the masks in Fig. 3.34, or similar arrangements, because the area these masks span at any one location in an image is so small.

With reference to Eq. (3.5-1), the general implementation for filtering an $M \times N$ image with a weighted averaging filter of size $m \times n$ (m and n odd) is given by the expression

$$g(x, y) = \frac{\sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x + s, y + t)}{\sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t)} \tag{3.6-1}$$

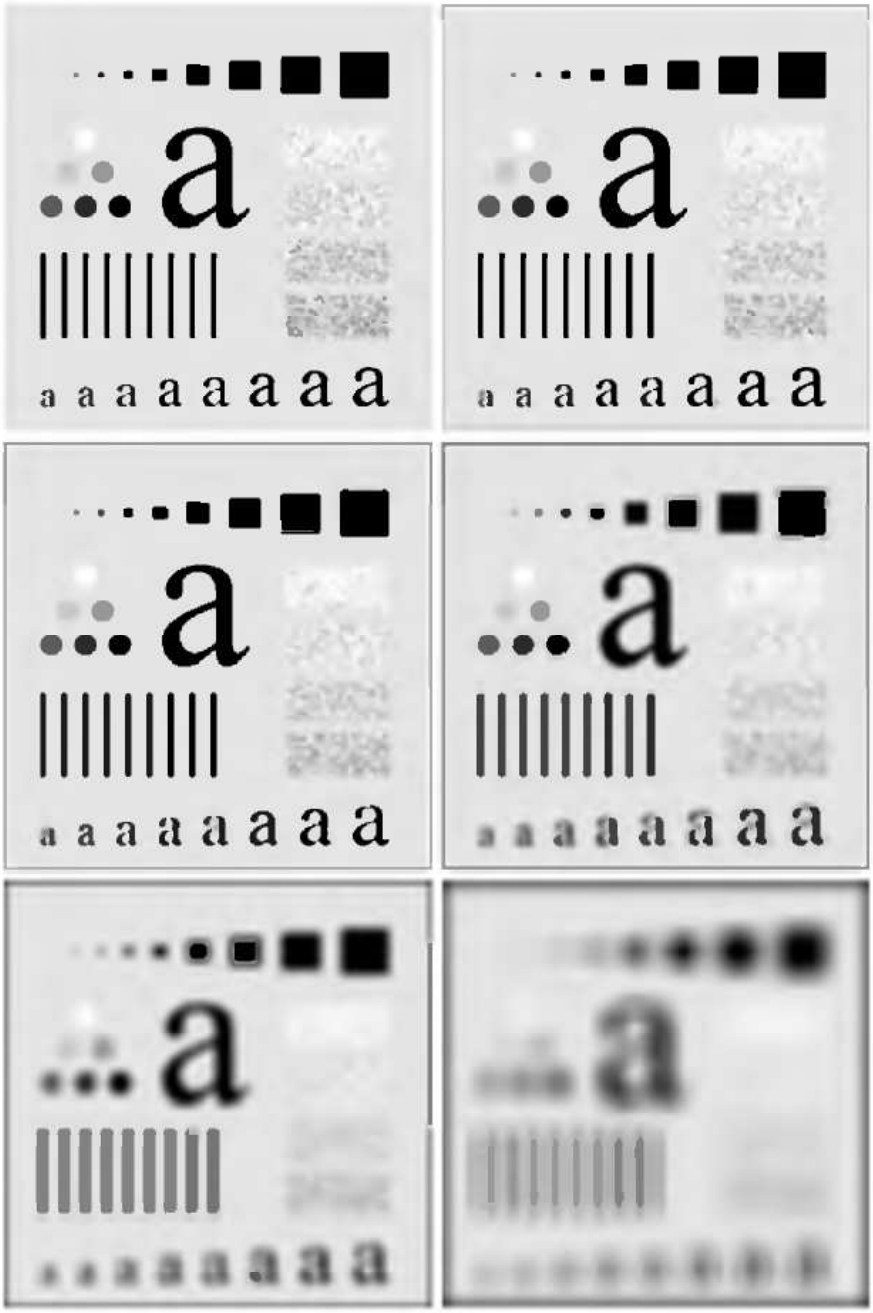
The parameters in this equation are as defined in Eq. (3.5-1). As before, it is understood that the complete filtered image is obtained by applying Eq. (3.6-1) for $x = 0, 1, 2, \ldots, M - 1$ and $y = 0, 1, 2, \ldots, N - 1$. The denominator in Eq. (3.6-1) is simply the sum of the mask coefficients and, therefore, it is a constant that needs to be computed only once. Typically, this scale factor is applied to all the pixels of the output image after the filtering process is completed.

EXAMPLE 3.9: Image smoothing with masks of various sizes.

■ The effects of smoothing as a function of filter size are illustrated in Fig. 3.35, which shows an original image and the corresponding smoothed results obtained using square averaging filters of sizes $n = 3, 5, 9, 15$, and 35 pixels, respectively. The principal features of these results are as follows: For $n = 3$, we note a general slight blurring throughout the entire image but, as expected, details that are of approximately the same size as the filter mask are affected considerably more. For example, the 3×3 and 5×5 squares, the small letter "a," and the fine grain noise show significant blurring when compared to the rest of the image. A positive result is that the noise is less pronounced. Note that the jagged borders of the characters and gray circles have been pleasingly smoothed.

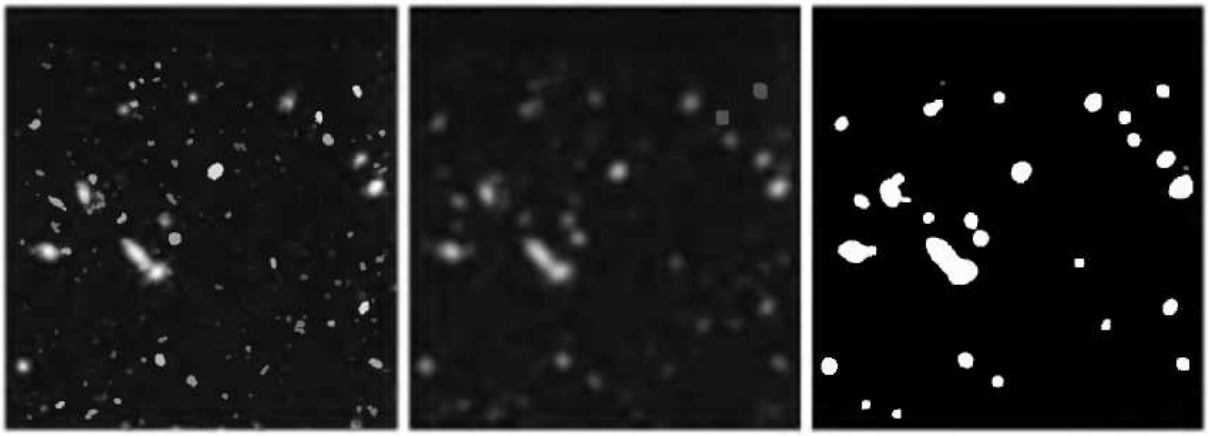
The result for $n = 5$ is somewhat similar, with a slight further increase in blurring. For $n = 9$ we see considerably more blurring, and the 20% black circle is not nearly as distinct from the background as in the previous three images, illustrating the blending effect that blurring has on objects whose gray level content is close to that of its neighboring pixels. Note the significant further smoothing of the noisy rectangles. The results for $n = 15$ and 35 are extreme with respect to the sizes of the objects in the image. This type of excessive blurring is generally used to eliminate small objects from an image. For instance, the three small squares, two of the circles, and most of the noisy rectangle areas have been blended into the background of the image in Fig. 3.35(f). Note also in this figure the pronounced black border. This is a result of padding the border of the original image with 0's (black) and then trimming off the padded area. Some of the black was blended into all filtered images, but became truly objectionable for the images smoothed with the larger filters. ■

As mentioned earlier, an important application of spatial averaging is to blur an image for the purpose getting a gross representation of objects of interest, such that the intensity of smaller objects blends with the background and larger objects become "bloblike" and easy to detect. The size of the mask establishes the relative size of the objects that will be blended with the background. As an illustration, consider Fig. 3.36(a), which is an image from the Hubble telescope in orbit around the Earth. Figure 3.36(b) shows the result of applying a



a b
c d
e f

FIGURE 3.35 (a) Original image, of size 500 × 500 pixels. (b)–(f) Results of smoothing with square averaging filter masks of sizes $n = 3, 5, 9, 15$, and 35, respectively. The black squares at the top are of sizes 3, 5, 9, 15, 25, 35, 45, and 55 pixels, respectively; their borders are 25 pixels apart. The letters at the bottom range in size from 10 to 24 points, in increments of 2 points; the large letter at the top is 60 points. The vertical bars are 5 pixels wide and 100 pixels high; their separation is 20 pixels. The diameter of the circles is 25 pixels, and their borders are 15 pixels apart; their gray levels range from 0% to 100% black in increments of 20%. The background of the image is 10% black. The noisy rectangles are of size 50 × 120 pixels.



a b c

FIGURE 3.36 (a) Image from the Hubble Space Telescope. (b) Image processed by a 15 × 15 averaging mask. (c) Result of thresholding (b). (Original image courtesy of NASA.)

15 × 15 averaging mask to this image. We see that a number of objects have either blended with the background or their intensity has diminished considerably. It is typical to follow an operation like this with thresholding to eliminate objects based on their intensity. The result of using the thresholding function of Fig. 3.2(b) with a threshold value equal to 25% of the highest intensity in the blurred image is shown in Fig. 3.36(c). Comparing this result with the original image, we see that it is a reasonable representation of what we would consider to be the largest, brightest objects in that image.

3.6.2 Order-Statistics Filters

Order-statistics filters are nonlinear spatial filters whose response is based on ordering (ranking) the pixels contained in the image area encompassed by the filter, and then replacing the value of the center pixel with the value determined by the ranking result. The best-known example in this category is the *median filter*, which, as its name implies, replaces the value of a pixel by the median of the gray levels in the neighborhood of that pixel (the original value of the pixel is included in the computation of the median). Median filters are quite popular because, for certain types of random noise, they provide excellent noise-reduction capabilities, with considerably less blurring than linear smoothing filters of similar size. Median filters are particularly effective in the presence of *impulse noise*, also called *salt-and-pepper noise* because of its appearance as white and black dots superimposed on an image.

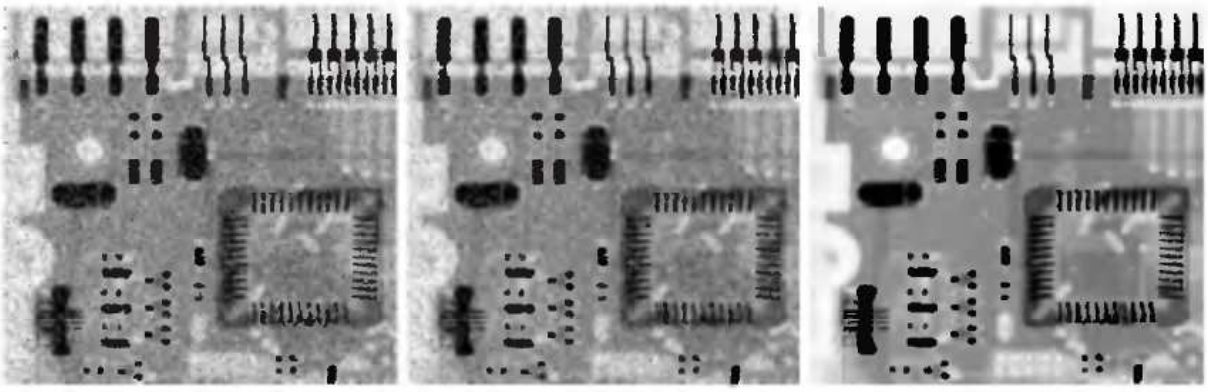
The median, ξ, of a set of values is such that half the values in the set are less than or equal to ξ, and half are greater than or equal to ξ. In order to perform median filtering at a point in an image, we first sort the values of the pixel in question and its neighbors, determine their median, and assign this value to that pixel. For example, in a 3 × 3 neighborhood the median is the 5th largest value, in a 5 × 5 neighborhood the 13th largest value, and so on. When several values

in a neighborhood are the same, all equal values are grouped. For example, suppose that a 3×3 neighborhood has values (10, 20, 20, 20, 15, 20, 20, 25, 100). These values are sorted as (10, 15, 20, 20, 20, 20, 20, 25, 100), which results in a median of 20. Thus, the principal function of median filters is to force points with distinct gray levels to be more like their neighbors. In fact, isolated clusters of pixels that are light or dark with respect to their neighbors, and whose area is less than $n^2/2$ (one-half the filter area), are eliminated by an $n \times n$ median filter. In this case "eliminated" means forced to the median intensity of the neighbors. Larger clusters are affected considerably less.

Although the median filter is by far the most useful order-statistics filter in image processing, it is by no means the only one. The median represents the 50th percentile of a ranked set of numbers, but the reader will recall from basic statistics that ranking lends itself to many other possibilities. For example, using the 100th percentile results in the so-called *max filter*, which is useful in finding the brightest points in an image. The response of a 3×3 max filter is given by $R = \max\{z_k | k = 1, 2, \ldots, 9\}$. The 0th percentile filter is the *min filter*, used for the opposite purpose. Median, max, and mean filters are considered in more detail in Chapter 5.

EXAMPLE 3.10: Use of median filtering for noise reduction.

■ Figure 3.37(a) shows an X-ray image of a circuit board heavily corrupted by salt-and-pepper noise. To illustrate the point about the superiority of median filtering over average filtering in situations such as this, we show in Fig. 3.37(b) the result of processing the noisy image with a 3×3 neighborhood averaging mask, and in Fig. 3.37(c) the result of using a 3×3 median filter. The image processed with the averaging filter has less visible noise, but the price paid is significant blurring. The superiority in all respects of median over average filtering in this case is quite evident. In general, median filtering is much better suited than averaging for the removal of additive salt-and-pepper noise. ■



a b c

FIGURE 3.37 (a) X-ray image of circuit board corrupted by salt-and-pepper noise. (b) Noise reduction with a 3×3 averaging mask. (c) Noise reduction with a 3×3 median filter. (Original image courtesy of Mr. Joseph E. Pascente, Lixi, Inc.)

3.7 Sharpening Spatial Filters

The principal objective of sharpening is to highlight fine detail in an image or to enhance detail that has been blurred, either in error or as a natural effect of a particular method of image acquisition. Uses of image sharpening vary and include applications ranging from electronic printing and medical imaging to industrial inspection and autonomous guidance in military systems.

In the last section, we saw that image blurring could be accomplished in the spatial domain by pixel averaging in a neighborhood. Since averaging is analogous to integration, it is logical to conclude that sharpening could be accomplished by spatial differentiation. This, in fact, is the case, and the discussion in this section deals with various ways of defining and implementing operators for sharpening by digital differentiation. Fundamentally, the strength of the response of a derivative operator is proportional to the degree of discontinuity of the image at the point at which the operator is applied. Thus, image differentiation enhances edges and other discontinuities (such as noise) and deemphasizes areas with slowly varying gray-level values.

3.7.1 Foundation

In the two sections that follow, we consider in some detail sharpening filters that are based on first- and second-order derivatives, respectively. Before proceeding with that discussion, however, we stop to look at some of the fundamental properties of these derivatives in a digital context. To simplify the explanation, we focus attention on one-dimensional derivatives. In particular, we are interested in the behavior of these derivatives in areas of constant gray level (flat segments), at the onset and end of discontinuities (step and ramp discontinuities), and along gray-level ramps. These types of discontinuities can be used to model noise points, lines, and edges in an image. The behavior of derivatives during transitions into and out of these image features also is of interest.

The derivatives of a digital function are defined in terms of differences. There are various ways to define these differences. However, we require that any definition we use for a first derivative (1) must be zero in flat segments (areas of constant gray-level values); (2) must be nonzero at the onset of a gray-level step or ramp; and (3) must be nonzero along ramps. Similarly, any definition of a second derivative (1) must be zero in flat areas; (2) must be nonzero at the onset and end of a gray-level step or ramp; and (3) must be zero along ramps of constant slope. Since we are dealing with digital quantities whose values are finite, the maximum possible gray-level change also is finite, and the shortest distance over which that change can occur is between adjacent pixels.

A basic definition of the first-order derivative of a one-dimensional function $f(x)$ is the difference

$$\frac{\partial f}{\partial x} = f(x + 1) - f(x).$$

We used a partial derivative here in order to keep the notation the same as when we consider an image function of two variables, $f(x, y)$, at which time we

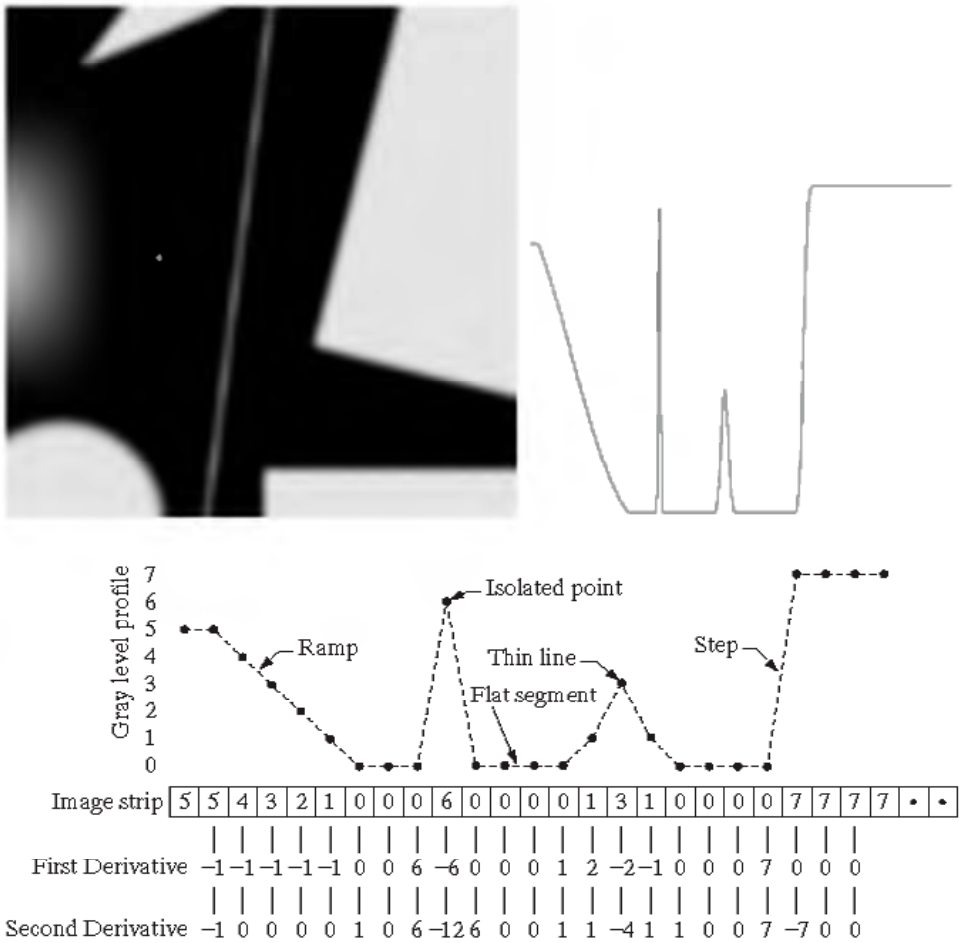
will be dealing with partial derivatives along the two spatial axes. Use of a partial derivative in the present discussion does not affect in any way the nature of what we are trying to accomplish.

Similarly, we define a second-order derivative as the difference

$$\frac{\partial^2 f}{\partial x^2} = f(x+1) + f(x-1) - 2f(x).$$

It is easily verified that these two definitions satisfy the conditions stated previously regarding derivatives of the first and second order. To see this, and also to highlight the fundamental similarities and differences between first- and second-order derivatives in the context of image processing, consider the example shown in Fig. 3.38.

Figure 3.38(a) shows a simple image that contains various solid objects, a line, and a single noise point. Figure 3.38(b) shows a horizontal gray-level profile (scan line) of the image along the center and including the noise point. This profile is the one-dimensional function we will use for illustrations regarding this figure. Figure 3.38(c) shows a simplification of the profile, with just enough num-



a b
c

FIGURE 3.38 (a) A simple image. (b) 1-D horizontal gray-level profile along the center of the image and including the isolated noise point. (c) Simplified profile (the points are joined by dashed lines to simplify interpretation).

bers to make it possible for us to analyze how the first- and second-order derivatives behave as they encounter a noise point, a line, and then the edge of an object. In our simplified diagram the transition in the ramp spans four pixels, the noise point is a single pixel, the line is three pixels thick, and the transition into the gray-level step takes place between adjacent pixels. The number of gray levels was simplified to only eight levels.

Let us consider the properties of the first and second derivatives as we traverse the profile from left to right. First, we note that the first-order derivative is nonzero along the entire ramp, while the second-order derivative is nonzero only at the onset and end of the ramp. Because edges in an image resemble this type of transition, we conclude that first-order derivatives produce "thick" edges and second-order derivatives, much finer ones. Next we encounter the isolated noise point. Here, the response at and around the point is much stronger for the second- than for the first-order derivative. Of course, this is not unexpected. A second-order derivative is much more aggressive than a first-order derivative in enhancing sharp changes. Thus, we can expect a second-order derivative to enhance fine detail (including noise) much more than a first-order derivative. The thin line is a fine detail, and we see essentially the same difference between the two derivatives. If the maximum gray level of the line had been the same as the isolated point, the response of the second derivative would have been stronger for the latter. Finally, in this case, the response of the two derivatives is the same at the gray-level step (in most cases when the transition into a step is not from zero, the second derivative will be weaker). We also note that the second derivative has a transition from positive back to negative. In an image, this shows as a thin double line. This "double-edge" effect is an issue that will be important in Chapter 10, where we use derivatives for edge detection. It is of interest also to note that if the gray level of the thin line had been the same as the step, the response of the second derivative would have been stronger for the line than for the step.

In summary, comparing the response between first- and second-order derivatives, we arrive at the following conclusions. (1) First-order derivatives generally produce thicker edges in an image. (2) Second-order derivatives have a stronger response to fine detail, such as thin lines and isolated points. (3) First-order derivatives generally have a stronger response to a gray-level step. (4) Second-order derivatives produce a double response at step changes in gray level. We also note of second-order derivatives that, for similar changes in gray-level values in an image, their response is stronger to a line than to a step, and to a point than to a line.

In most applications, the second derivative is better suited than the first derivative for image enhancement because of the ability of the former to enhance fine detail. For this, and for reasons of simpler implementation and extensions, we will focus attention initially on uses of the second derivative for enhancement. First-order derivatives are discussed in Section 3.7.3. Although the principle of use of first derivatives in image processing is for edge extraction, they do have important uses in image enhancement. In fact, we show in Section 3.8 that they can be used in conjunction with the second derivative to obtain some impressive enhancement results.

3.7.2 Use of Second Derivatives for Enhancement–The Laplacian

In this section we consider in some detail the use of two-dimensional, second-order derivatives for image enhancement. The approach basically consists of defining a discrete formulation of the second-order derivative and then constructing a filter mask based on that formulation. We are interested in *isotropic* filters, whose response is independent of the direction of the discontinuities in the image to which the filter is applied. In other words, isotropic filters are *rotation invariant*, in the sense that rotating the image and then applying the filter gives the same result as applying the filter to the image first and then rotating the result.

Development of the method

It can be shown (Rosenfeld and Kak [1982]) that the simplest isotropic derivative operator is the *Laplacian*, which, for a function (image) $f(x, y)$ of two variables, is defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}. \tag{3.7-1}$$

Because derivatives of any order are linear operations, the Laplacian is a linear operator.

In order to be useful for digital image processing, this equation needs to be expressed in discrete form. There are several ways to define a digital Laplacian using neighborhoods. Whatever the definition, however, it has to satisfy the properties of a second derivative outlined in Section 3.7.1. The definition of the digital second derivative given in that section is one of the most used. Taking into account that we now have two variables, we use the following notation for the partial second-order derivative in the x-direction:

$$\frac{\partial^2 f}{\partial^2 x^2} = f(x + 1, y) + f(x - 1, y) - 2f(x, y) \tag{3.7-2}$$

and, similarly in the y-direction, as

$$\frac{\partial^2 f}{\partial^2 y^2} = f(x, y + 1) + f(x, y - 1) - 2f(x, y) \tag{3.7-3}$$

The digital implementation of the two-dimensional Laplacian in Eq. (3.7-1) is obtained by summing these two components:

$$\begin{aligned}\nabla^2 f = {} & \left[f(x + 1, y) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1)\right] \\ & - 4f(x, y).\end{aligned} \tag{3.7-4}$$

This equation can be implemented using the mask shown in Fig. 3.39(a), which gives an isotropic result for rotations in increments of 90°. The mechanics of implementation are given in Eq. (3.5-1) and are illustrated in Section 3.6.1 for the linear smoothing filters. We simply are using different coefficients here.

The diagonal directions can be incorporated in the definition of the digital Laplacian by adding two more terms to Eq. (3.7-4), one for each of the two diagonal directions. The form of each new term is the same as either Eq. (3.7-2)

0	1	0
1	−4	1
0	1	0

1	1	1
1	−8	1
1	1	1

0	−1	0
−1	4	−1
0	−1	0

−1	−1	−1
−1	8	−1
−1	−1	−1

a b
c d

FIGURE 3.39 (a) Filter mask used to implement the digital Laplacian, as defined in Eq. (3.7-4). (b) Mask used to implement an extension of this equation that includes the diagonal neighbors. (c) and (d) Two other implementations of the Laplacian.

or (3.7-3), but the coordinates are along the diagonals. Since each diagonal term also contains a $-2f(x, y)$ term, the total subtracted from the difference terms now would be $-8f(x, y)$. The mask used to implement this new definition is shown in Fig. 3.39(b). This mask yields isotropic results for increments of 45°. The other two masks shown in Fig. 3.39 also are used frequently in practice. They are based on a definition of the Laplacian that is the negative of the one we used here. As such, they yield equivalent results, but the difference in sign must be kept in mind when combining (by addition or subtraction) a Laplacian-filtered image with another image.

Because the Laplacian is a derivative operator, its use highlights gray-level discontinuities in an image and deemphasizes regions with slowly varying gray levels. This will tend to produce images that have grayish edge lines and other discontinuities, all superimposed on a dark, featureless background. Background features can be "recovered" while still preserving the sharpening effect of the Laplacian operation simply by adding the original and Laplacian images. As noted in the previous paragraph, it is important to keep in mind which definition of the Laplacian is used. If the definition used has a negative center coefficient, then we *subtract*, rather than add, the Laplacian image to obtain a sharpened result. Thus, the basic way in which we use the Laplacian for image enhancement is as follows:

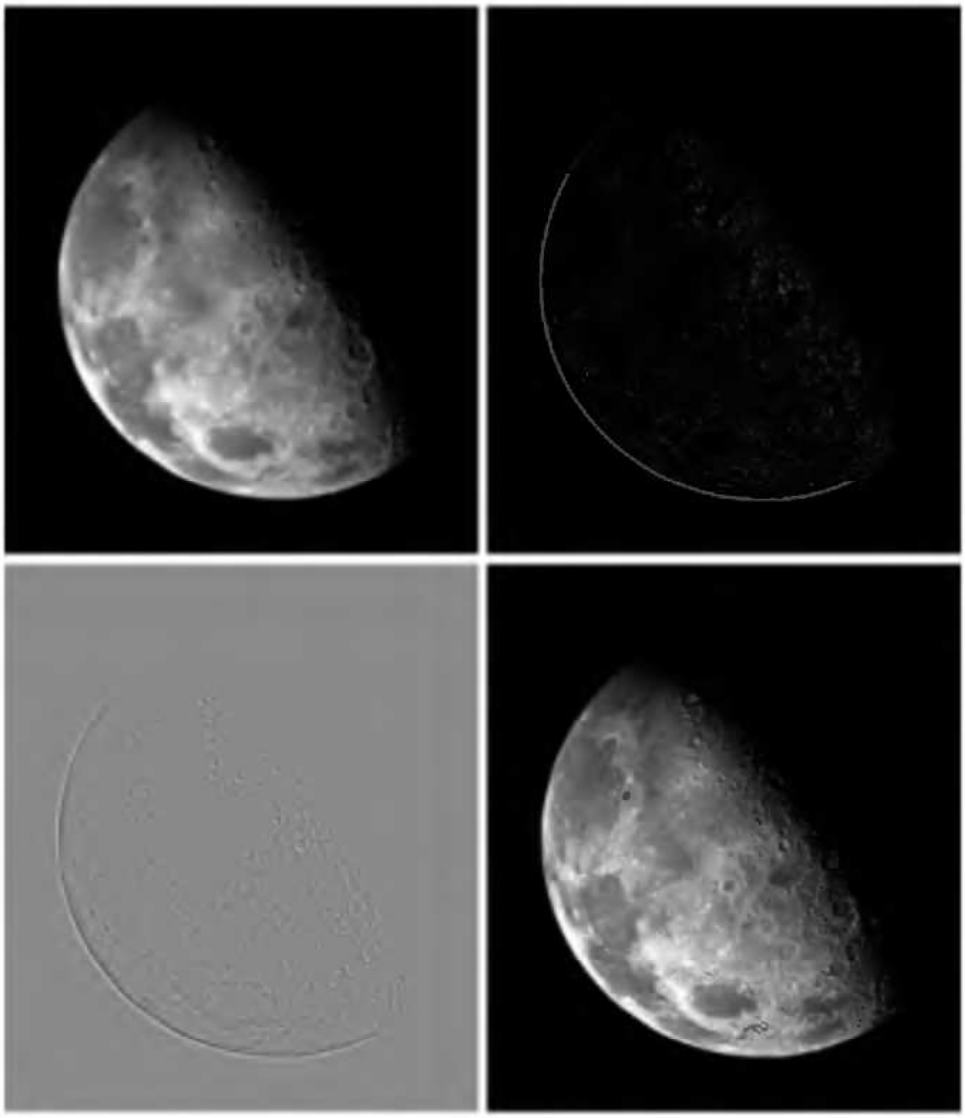
$$g(x, y) = \begin{cases} f(x, y) - \nabla^2 f(x, y) & \text{if the center coefficient of the Laplacian mask is negative} \\ f(x, y) + \nabla^2 f(x, y) & \text{if the center coefficient of the Laplacian mask is positive.} \end{cases} \tag{3.7-5}$$

Use of this equation is illustrated next.

EXAMPLE 3.11: Imaging sharpening with the Laplacian.

■ Figure 3.40(a) shows an image of the North Pole of the moon. Figure 3.40(b) shows the result of filtering this image with the Laplacian mask in Fig. 3.39(b). Since the Laplacian image contains both positive and negative values, a typical way to scale it is to use the approach discussed at the end of Section 3.4.1. Sometimes one encounters the absolute value being used for this purpose, but this really is not correct because it produces double lines of nearly equal magnitude, which can be confusing.

The image shown in Fig. 3.40(c) was scaled in the manner just described for display purposes. Note that the dominant features of the image are edges and sharp gray-level discontinuities of various gray-level values. The background, previously near black, is now gray due to the scaling. This grayish appearance is typical of Laplacian images that have been scaled properly. Finally, Fig. 3.40(d)



a b
c d

FIGURE 3.40 (a) Image of the North Pole of the moon. (b) Laplacian-filtered image. (c) Laplacian image scaled for display purposes. (d) Image enhanced by using Eq. (3.7-5). (Original image courtesy of NASA.)

shows the result obtained using Eq. (3.7-5). The detail in this image is unmistakably clearer and sharper than in the original image. Adding the image to the Laplacian restored the overall gray level variations in the image, with the Laplacian increasing the contrast at the locations of gray-level discontinuities. The net result is an image in which small details were enhanced and the background tonality was perfectly preserved. Results like these have made Laplacian-based enhancement a fundamental tool used frequently for sharpening digital images. ■

Simplifications

In the previous example, we implemented Eq. (3.7-5) by first computing the Laplacian-filtered image and then subtracting it from the original image. This was done for instructional purposes to illustrate each step in the procedure. In practice, Eq. (3.7-5) is usually implemented with one pass of a single mask. The coefficients of the single mask are easily obtained by substituting Eq. (3.7-4) for $\nabla^2 f(x, y)$ in the first line of Eq. (3.7-5):

$$\begin{aligned} g(x, y) &= f(x, y) - \big[f(x+1, y) + f(x-1, y) \\ &\quad + f(x, y+1) + f(x, y-1)\big] + 4f(x, y) \\ &= 5f(x, y) - \big[f(x+1, y) + f(x-1, y) \\ &\quad + f(x, y+1) + f(x, y-1)\big]. \end{aligned} \tag{3.7-6}$$

This equation can be implemented using the mask shown in Fig. 3.41(a). The mask shown in Fig. 3.41(b) would be used if the diagonal neighbors also were included in the calculation of the Laplacian. Identical masks would have resulted if we had substituted the negative of Eq. (3.7-4) into the second line of Eq. (3.7-5).

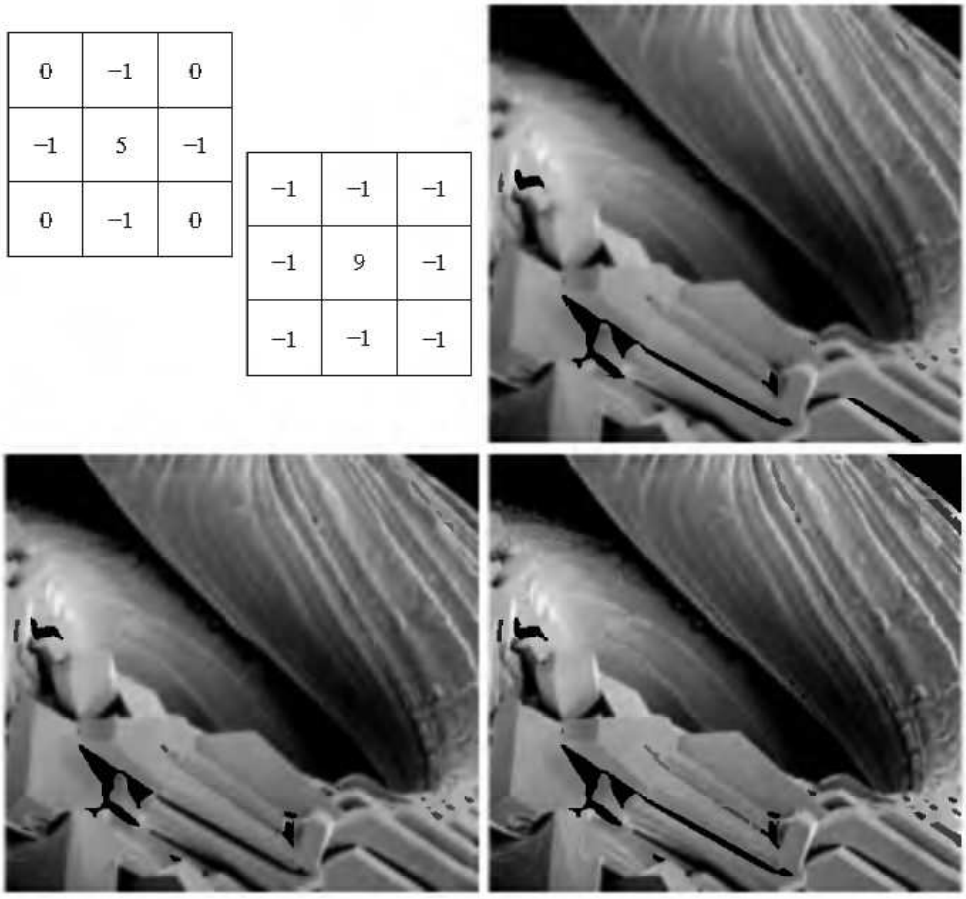
EXAMPLE 3.12: Image enhancement using a composite Laplacian mask.

■ The results obtainable with the mask containing the diagonal terms usually are a little sharper than those obtained with the more basic mask of Fig. 3.41(a). This property is illustrated by the Laplacian-filtered images shown in Figs. 3.41(d) and (e), which were obtained by using the masks in Figs. 3.41(a) and (b), respectively. By comparing the filtered images with the original image shown in Fig. 3.41(c), we note that both masks produced effective enhancement, but the result using the mask in Fig. 3.41(b) is visibly sharper. Figure 3.41(c) is a scanning electron microscope (SEM) image of a tungsten filament following thermal failure; the magnification is approximately 250×.) ■

Because the Laplacian is a linear operator, we could have arrived at the same composite masks in Figs. 3.41(a) and (b) by noting that Eq. (3.7-5) is the difference between (sum of) two linear processes. That is, $f(x, y)$ be may viewed as itself processed with a mask that has a unit coefficient in the center and zeros elsewhere. The second term in the equation is the same image processed with one of the Laplacian masks of Fig. 3.39. Due to linearity, the result obtained in Eq. (3.7-5) with the unit-center mask and one of those Laplacian masks would be the same as the result obtained with a single mask formed by subtracting (adding) the Laplacian mask from (to) the unity-center mask.

0	−1	0
−1	5	−1
0	−1	0

−1	−1	−1
−1	9	−1
−1	−1	−1



a b c
d e

FIGURE 3.41 (a) Composite Laplacian mask. (b) A second composite mask. (c) Scanning electron microscope image. (d) and (e) Results of filtering with the masks in (a) and (b), respectively. Note how much sharper (e) is than (d). (Original image courtesy of Mr. Michael Shaffer, Department of Geological Sciences, University of Oregon, Eugene.)

Unsharp masking and high-boost filtering

A process used for many years in the publishing industry to sharpen images consists of subtracting a blurred version of an image from the image itself. This process, called *unsharp masking*, is expressed as

$$f_s(x, y) = f(x, y) - \bar{f}(x, y) \tag{3.7-7}$$

where $f_s(x, y)$ denotes the sharpened image obtained by unsharp masking, and $\bar{f}(x, y)$ is a blurred version of $f(x, y)$. The origin of unsharp masking is in darkroom photography, where it consists of clamping together a blurred negative to a corresponding positive film and then developing this combination to produce a sharper image.

A slight further generalization of unsharp masking is called *high-boost filtering*. A high-boost filtered image, f_{hb}, is defined at any point (x, y) as

$$f_{hb}(x, y) = Af(x, y) - \bar{f}(x, y) \tag{3.7-8}$$

0	−1	0
−1	$A + 4$	−1
0	−1	0

−1	−1	−1
−1	$A + 8$	−1
−1	−1	−1

a b

FIGURE 3.42 The high-boost filtering technique can be implemented with either one of these masks, with $A \geq 1$.

where $A \geq 1$ and, as before, $\bar{f}$ is a blurred version of f. This equation may be written as

$$f_{\text{hb}}(x, y) = (A - 1)f(x, y) + f(x, y) - \bar{f}(x, y). \tag{3.7-9}$$

By using Eq. (3.7-7), we obtain

$$f_{\text{hb}}(x, y) = (A - 1)f(x, y) + f_s(x, y) \tag{3.7-10}$$

as the expression for computing a high-boost-filtered image.

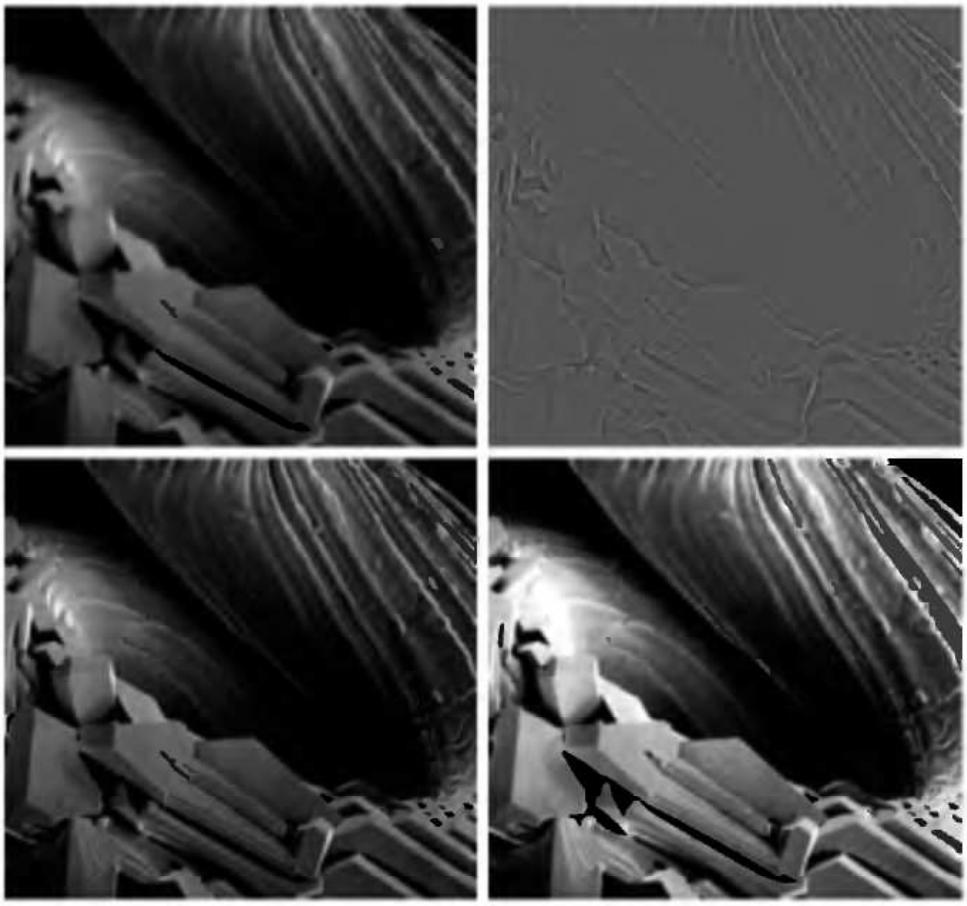
Equation (3.7-10) is applicable in general and does not state explicitly how the sharp image is obtained. If we elect to use the Laplacian, then we know that $f_s(x, y)$ can be obtained using Eq. (3.7-5). In this case, Eq. (3.7-10) becomes

$$f_{\text{hb}} = \begin{cases} Af(x, y) - \nabla^2 f(x, y) & \text{if the center coefficient of the Laplacian mask is negative} \\ Af(x, y) + \nabla^2 f(x, y) & \text{if the center coefficient of the Laplacian mask is positive.} \end{cases} \tag{3.7-11}$$

High-boost filtering can be implemented with one pass using either of the two masks shown in Fig. 3.42. Note that, when $A = 1$, high-boost filtering becomes "standard" Laplacian sharpening. As the value of A increases past 1, the contribution of the sharpening process becomes less and less important. Eventually, if A is large enough, the high-boost image will be approximately equal to the original image multiplied by a constant.

EXAMPLE 3.13: Image enhancement with a high-boost filter.

■ One of the principal applications of boost filtering is when the input image is darker than desired. By varying the boost coefficient, it generally is possible to obtain an overall increase in average gray level of the image, thus helping to brighten the final result. Figure 3.43 shows such an application. Part (a) of this figure is a darker version of the image in Fig. 3.41(c). Figure 3.43(b) shows the Laplacian computed using the mask in Fig. 3.42(b), with $A = 0$. Figure 3.43(c) was obtained using the mask in Fig. 3.42(b) with $A = 1$. As expected, the image has been sharpened, but it is still as dark as the original. Finally, Fig. 3.43(d) shows the result of using $A = 1.7$. This is a much more acceptable result, in which the average gray level has increased, thus making the image lighter and more natural. ■



a b
c d

FIGURE 3.43 (a) Same as Fig. 3.41(c), but darker. (a) Laplacian of (a) computed with the mask in Fig. 3.42(b) using $A = 0$. (c) Laplacian enhanced image using the mask in Fig. 3.42(b) with $A = 1$. (d) Same as (c), but using $A = 1.7$.

3.7.3 Use of First Derivatives for Enhancement—The Gradient

First derivatives in image processing are implemented using the magnitude of the gradient. For a function $f(x, y)$, the gradient of f at coordinates (x, y) is defined as the two-dimensional column *vector*

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix}. \tag{3.7-12}$$

The magnitude of this vector is given by

$$\begin{aligned} \nabla f &= \text{mag}(\nabla \mathbf{f}) \\ &= \left[G_x^2 + G_y^2\right]^{1/2} \\ &= \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right]^{1/2}. \end{aligned} \tag{3.7-13}$$

The components of the gradient vector itself are linear operators, but the magnitude of this vector obviously is not because of the squaring and square root

operations. On the other hand, the partial derivatives in Eq. (3.7-12) are not rotation invariant (isotropic), but the magnitude of the gradient vector is. Although it is not strictly correct, the magnitude of the gradient vector often is referred to as the *gradient*. In keeping with tradition, we will use this term in the following discussions, explicitly referring to the vector or its magnitude only in cases where confusion is likely.

The computational burden of implementing Eq. (3.7-13) over an entire image is not trivial, and it is common practice to approximate the magnitude of the gradient by using absolute values instead of squares and square roots:

$$\nabla f \approx |G_x| + |G_y|. \tag{3.7-14}$$

This equation is simpler to compute and it still preserves relative changes in gray levels, but the isotropic feature property is lost in general. However, as in the case of the Laplacian, the isotropic properties of the digital gradient defined in the following paragraph are preserved only for a limited number of rotational increments that depend on the masks used to approximate the derivatives. As it turns out, the most popular masks used to approximate the gradient give the same result only for vertical and horizontal edges and thus the isotropic properties of the gradient are preserved only for multiples of 90°. These results are independent of whether Eq. (3.7-13) or (3.7-14) is used, so nothing of significance is lost in using the simpler of the two equations.

As in the case of the Laplacian, we now define digital approximations to the preceding equations, and from there formulate the appropriate filter masks. In order to simplify the discussion that follows, we will use the notation in Fig. 3.44(a) to denote image points in a 3×3 region. For example, the center point, z_5, denotes $f(x, y)$, z_1 denotes $f(x - 1, y - 1)$, and so on. As indicated in Section 3.7.1, the simplest approximations to a first-order derivative that satisfy the conditions stated in that section are $G_x = (z_8 - z_5)$ and $G_y = (z_6 - z_5)$. Two other definitions proposed by Roberts [1965] in the early development of digital image processing use cross differences:

$$G_x = (z_9 - z_5) \qquad \text{and} \qquad G_y = (z_8 - z_6). \tag{3.7-15}$$

If we elect to use Eq. (3.7-13), then we compute the gradient as

$$\nabla f = [(z_9 - z_5)^2 + (z_8 - z_6)^2]^{1/2} \tag{3.7-16}$$

If we use absolute values, then substituting the quantities in Eq. (3.7-15) into Eq. (3.7-14) gives us the following approximation to the gradient:

$$\nabla f \approx |z_9 - z_5| + |z_8 - z_6|. \tag{3.7-17}$$

This equation can be implemented with the two masks shown in Figs. 3.44(b) and (c). These masks are referred to as the *Roberts cross-gradient operators*.

Masks of even size are awkward to implement. The smallest filter mask in which we are interested is of size 3×3. An approximation using absolute values, still at point z_5, but using a 3×3 mask, is

$$\begin{aligned}\nabla f \approx\ & |(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)| \\ & + |(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)|.\end{aligned} \tag{3.7-18}$$

a
b c
d e

FIGURE 3.44 A 3 × 3 region of an image (the z's are gray-level values) and masks used to compute the gradient at point labeled z_5. All masks coefficients sum to zero, as expected of a derivative operator.

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

−1	0
0	1

0	−1
1	0

−1	−2	−1
0	0	0
1	2	1

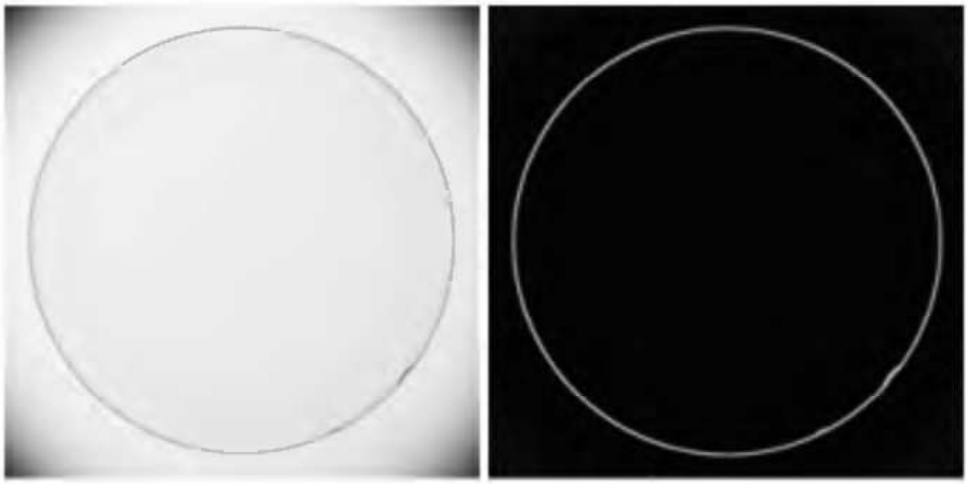
−1	0	1
−2	0	2
−1	0	1

The difference between the third and first rows of the 3 × 3 image region approximates the derivative in the x-direction, and the difference between the third and first columns approximates the derivative in the y-direction. The masks shown in Figs. 3.44(d) and (e), called the *Sobel operators*, can be used to implement Eq. (3.7-18) via the mechanics given in Eq. (3.5-1). The idea behind using a weight value of 2 is to achieve some smoothing by giving more importance to the center point (we discuss this in more detail in Chapter 10). Note that the coefficients in all the masks shown in Fig. 3.44 sum to 0, indicating that they would give a response of 0 in an area of constant gray level, as expected of a derivative operator.

EXAMPLE 3.14: Use of the gradient for edge enhancement.

■ The gradient is used frequently in industrial inspection, either to aid humans in the detection of defects or, what is more common, as a preprocessing step in automated inspection. We will have more to say about this in Chapters 10 and 11. However, it will be instructive at this point to consider a simple example to show how the gradient can be used to enhance defects and eliminate slowly changing background features. In this particular example, the enhancement is used as a preprocessing step for automated inspection, rather than for human analysis.

Figure 3.45(a) shows an optical image of a contact lens, illuminated by a lighting arrangement designed to highlight imperfections, such as the two edge



a b

FIGURE 3.45 Optical image of contact lens (note defects on the boundary at 4 and 5 o'clock). (b) Sobel gradient. (Original image courtesy of Mr. Pete Sites, Perceptics Corporation.)

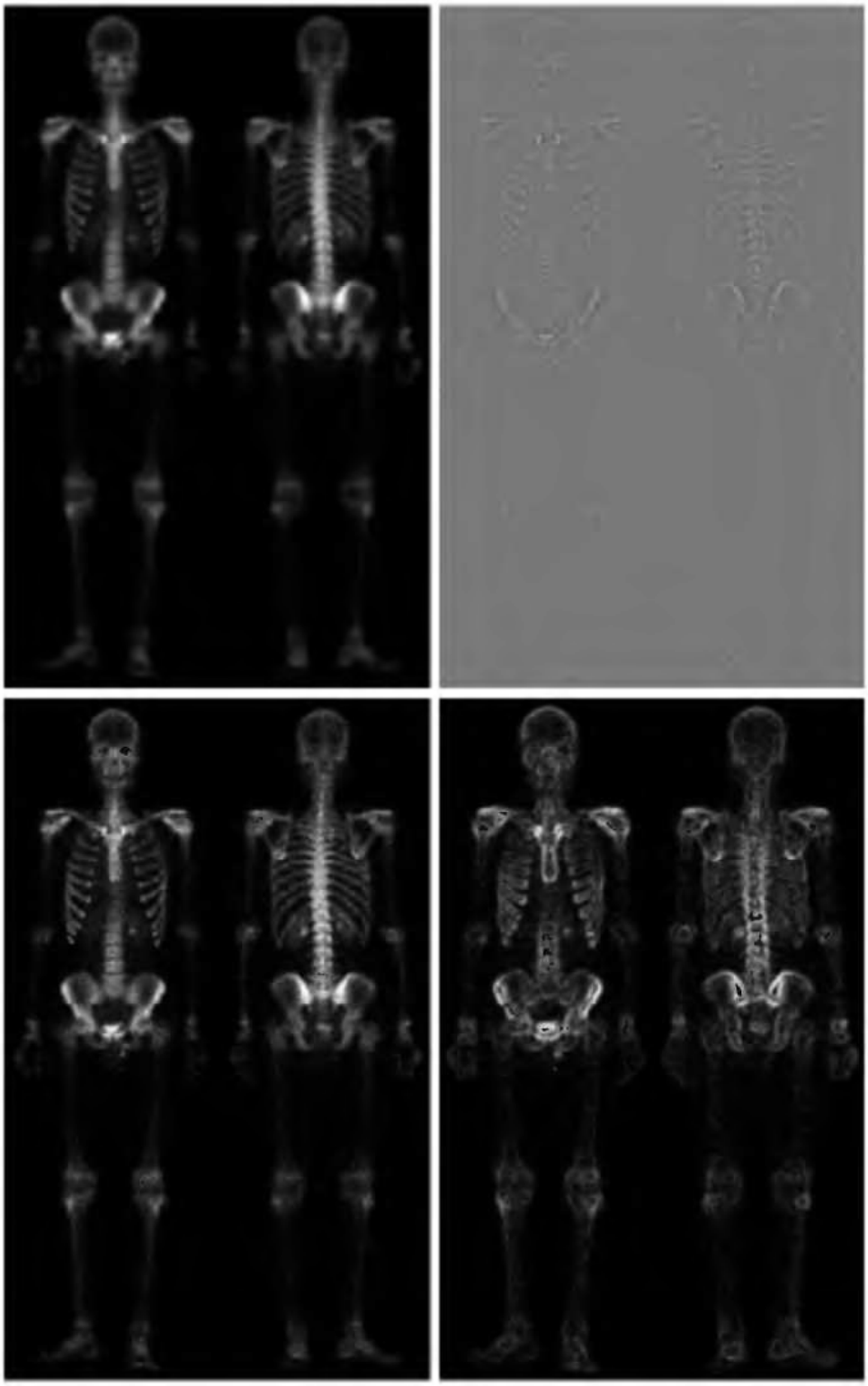
defects in the lens boundary seen at 4 and 5 o'clock. Figure 3.45(b) shows the gradient obtained using Eq. (3.7-14) with the two Sobel masks in Figs. 3.44(d) and (e). The edge defects also are quite visible in this image, but with the added advantage that constant or slowly varying shades of gray have been eliminated, thus simplifying considerably the computational task required for automated inspection. Note also that the gradient process highlighted small specs that are not readily visible in the gray-scale image (specs like these can be foreign matter, air pockets in a supporting solution, or miniscule imperfections in the lens). The ability to enhance small discontinuities in an otherwise flat gray field is another important feature of the gradient. ■

3.8 Combining Spatial Enhancement Methods

With a few exceptions, like combining blurring with thresholding in Section 3.6.1, we have focused attention thus far on individual enhancement approaches. Frequently, a given enhancement task will require application of several complementary enhancement techniques in order to achieve an acceptable result. In this section we illustrate by means of an example how to combine several of the approaches developed in this chapter to address a difficult enhancement task.

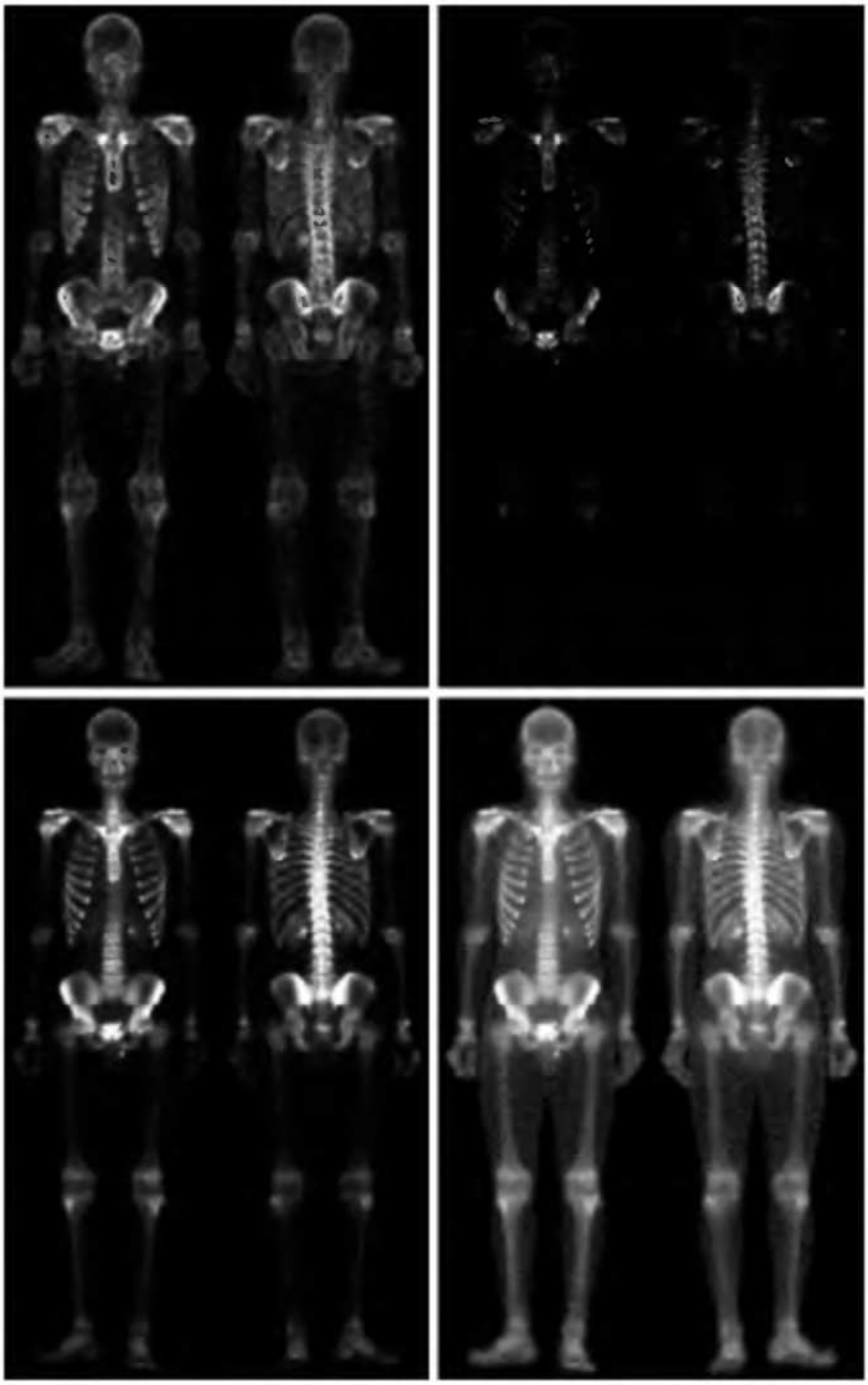
The image shown in Fig. 3.46(a) is a nuclear whole body bone scan, used to detect diseases such as bone infection and tumors. Our objective is to enhance this image by sharpening it and by bringing out more of the skeletal detail. The narrow dynamic range of the gray levels and high noise content make this image difficult to enhance. The strategy we will follow is to utilize the Laplacian to highlight fine detail, and the gradient to enhance prominent edges. For reasons that will be explained shortly, a smoothed version of the gradient image will be used to mask the Laplacian image (see Section 3.4 regarding masking). Finally, we will attempt to increase the dynamic range of the gray levels by using a gray-level transformation.

Figure 3.46 (b) shows the Laplacian of the original image, obtained using the mask in Fig. 3.39(d). This image was scaled (for display only) using the same technique as in Fig. 3.40. We can obtain a sharpened image at this point



a b
c d

FIGURE 3.46 (a) Image of whole body bone scan. (b) Laplacian of (a). (c) Sharpened image obtained by adding (a) and (b). (d) Sobel of (a).



e f
g h

FIGURE 3.46 *(Continued)* (e) Sobel image smoothed with a 5×5 averaging filter. (f) Mask image formed by the product of (c) and (e). (g) Sharpened image obtained by the sum of (a) and (f). (h) Final result obtained by applying a power-law transformation to (g). Compare (g) and (h) with (a). (Original image courtesy of G.E. Medical Systems.)

simply by adding Figs. 3.46(a) and (b), which are an implementation of the second line in Eq. (3.7-5) (we used a mask with a positive center coefficient). Just by looking at the noise level in (b), we would expect a rather noisy sharpened image if we added Figs. 3.46(a) and (b), a fact that is confirmed by the result shown in Fig. 3.46(c). One way that comes immediately to mind to reduce the noise is to use a median filter. However, median filtering is a nonlinear process capable of removing image features. This is unacceptable in medical image processing.

An alternate approach is to use a mask formed from a smoothed version of the gradient of the original image. The motivation behind this is straightforward and is based on the properties of first- and second-order derivatives explained in Section 3.7.1. The Laplacian, being a second-order derivative operator, has the definite advantage that it is superior in enhancing fine detail. However, this causes it to produce noisier results than the gradient. This noise is most objectionable in smooth areas, where it tends to be more visible. The gradient has a stronger response in areas of significant gray-level transitions (gray-level ramps and steps) than does the Laplacian. The response of the gradient to noise and fine detail is lower than the Laplacian's and can be lowered further by smoothing the gradient with an averaging filter. The idea, then, is to smooth the gradient and multiply it by the Laplacian image. In this context, we may view the smoothed gradient as a mask image. The product will preserve details in the strong areas while reducing noise in the relatively flat areas. This process can be viewed roughly as combining the best features of the Laplacian and the gradient. The result is added to the original to obtain a final sharpened image, and could even be used in boost filtering.

Figure 3.46(d) shows the Sobel gradient of the original image, computed using Eq. (3.7-14). Components G_x and G_y were obtained using the masks in Figs. 3.44(d) and (e), respectively. As expected from our discussion in Section 3.7.1, edges are much more dominant in this image than in the Laplacian image. The smoothed gradient image shown in Fig. 3.46(e) was obtained by using an averaging filter of size 5×5. The two gradient images were scaled for display in the same manner as the two Laplacian images. Because the smallest possible value of a gradient image is 0, the background is black in the scaled gradient images, rather than gray as in the scaled Laplacian. The fact that Figs. 3.46(d) and (e) are much brighter than Fig. 3.46(b) is again evidence that the gradient of an image with significant edge content has values that are higher in general than in a Laplacian image.

The product of the Laplacian and smoothed-gradient image is shown in Fig. 3.46(f). Note the dominance of the strong edges and the relative lack of visible noise, which is the key objective behind masking the Laplacian with a smoothed gradient image. Adding the product image to the original resulted in the sharpened image shown in Fig. 3.46(g). The significant increase in sharpness of detail in this image over the original is evident in most parts of the image, including the ribs, spinal chord, pelvis, and skull. This type of improvement would not have been possible by using the Laplacian or gradient alone.

The sharpening procedure just discussed does not affect in an appreciable way the dynamic range of the gray levels in an image. Thus, the final step in our

enhancement task is to increase the dynamic range of the sharpened image. As we discussed in some detail in Sections 3.2 and 3.3, there are a number of gray-level transformation functions that can accomplish this objective. We do know from the results in Section 3.3.2 that histogram equalization is not likely to work well on images that have dark gray-level distributions like our images have here. Histogram specification could be a solution, but the dark characteristics of the images with which we are dealing lend themselves much better to a power-law transformation. Since we wish to spread the gray levels, the value of γ in Eq. (3.2-3) has to be less than 1. After a few trials with this equation we arrived at the result shown in Fig. 3.46(h), obtained with $\gamma = 0.5$ and $c = 1$. Comparing this image with Fig. 3.46(g), we see that significant new detail is visible in Fig. 3.46(h). The areas around the wrists, hands, ankles, and feet are good examples of this. The skeletal bone structure also is much more pronounced, including the arm and leg bones. Note also the faint definition of the outline of the body, and of body tissue. Bringing out detail of this nature by expanding the dynamic range of the gray levels also enhanced noise, but Fig. 3.46(h) represents a significant visual improvement over the original image.

The approach just discussed is representative of the types of processes that can be linked in order to achieve results that are not possible with a single technique. The way in which the results are used depends on the application. The final user of the type of images shown in this section is likely to be a radiologist. For a number of reasons that are beyond the scope of our discussion, physicians are unlikely to rely on enhanced results to arrive at a diagnosis. However, enhanced images are quite useful in highlighting details that can serve as clues for further analysis in the original image or sequence of images. In other areas, the enhanced result may indeed be the final product. Examples are found in the printing industry, in image-based product inspection, in forensics, in microscopy, in surveillance, and in a host of other areas where the principal objective of enhancement is to obtain an image with a higher content of visual detail.

Summary

The material presented in this chapter is representative of spatial domain techniques commonly used in practice for image enhancement. This area of image processing is a dynamic field, and new techniques and applications are reported routinely in professional literature and in new product announcements. For this reason, the topics included in this chapter were selected for their value as fundamental material that would serve as a foundation for understanding the state of the art in enhancement techniques, as well as for further study in this field. In addition to enhancement, this chapter served the purpose of introducing a number of concepts, such as filtering with spatial masks, that will be used in numerous occasions throughout the remainder of the book. In the following chapter, we deal with enhancement from a complementary viewpoint in the frequency domain. Between these two chapters, the reader will have developed a solid foundation for the terminology and some of the most fundamental tools used in image processing. The fact that these tools were introduced in the context of image enhancement is likely to aid in the understanding of how they operate on digital images.

References and Further Reading

The material in Section 3.1 is from Gonzalez [1986]. Additional reading for the material in Section 3.2 may be found in Schowengerdt [1983], Poyton [1996], and Russ [1999]. See also the paper by Tsujii et al. [1998] regarding the optimization of image displays. Early references on histogram processing are Hummel [1974], Gonzalez and Fittes [1977], and Woods and Gonzalez [1981]. Stark [2000] gives some interesting generalizations of histogram equalization for adaptive contrast enhancement. Other approaches for contrast enhancement are exemplified by Centeno and Haertel [1997] and Cheng and Xu [2000]. For enhancement based on an ideal image model, see Highnam and Brady [1997]. For extensions of the local histogram equalization method, see Caselles et al. [1999], and Zhu et al. [1999]. See Narendra and Fitch [1981] on the use and implementation of local statistics for image enhancement. Kim et al. [1997] present an interesting approach combining the gradient with local statistics for image enhancement.

Image subtraction (Section 3.4.1) is a generic image processing tool widely used for change detection. As noted in that section, one of the principal applications of digital image subtraction is in mask mode radiography, where patient motion is a problem because motion smears the results. The problem of motion during image subtraction has received significant attention over the years, as exemplified in the survey article by Meijering et al. [1999]. The method of noise reduction by image averaging (Section 3.4.2) was first proposed by Kohler and Howell [1963]. See Peebles [1993] regarding the expected value of the mean and variance of a sum of random variables.

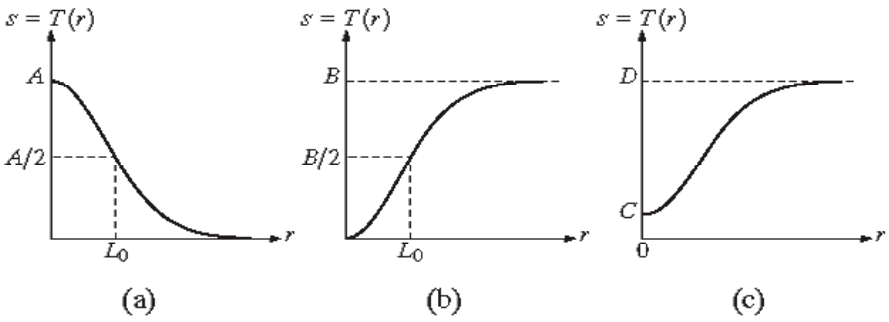
For additional reading on linear spatial filters and their implementation, see Umbaugh [1998], Jain [1989], and Rosenfeld and Kak [1982]. Rank-order filters are discussed in these references as well. Wilburn [1998] discusses generalizations of rank-order filters. The book by Pitas and Venetsanopoulos [1990] also deals with median and other nonlinear spatial filters. A special issue of *IEEE Transactions in Image Processing* [1996] is dedicated to the topic of nonlinear image processing. The material on high-boost filtering is from Schowengerdt [1983]. We will encounter again many of the spatial filters introduced in this chapter in discussions dealing with image restoration (Chapter 5) and edge detection (Chapter 10).

Problems

See inside front cover

Detailed solutions to the problems marked with a star can be found in the book web site. The site also contains suggested projects based on the material in this chapter.

3.1 Exponentials of the form $e^{-\alpha r^2}$, with α a positive constant, are useful for constructing smooth gray-level transformation functions. Start with this basic function and construct transformation functions having the general shapes shown in the following figures. The constants shown are *input* parameters, and your proposed transformations must include them in their specification. (For simplicity in your answers, L_0 is not a required parameter in the third curve.)



3.2 ★ **(a)** Give a continuous function for implementing the contrast stretching transformation shown in Fig. 3.2(a). In addition to m, your function must include a parameter, E, for controlling the slope of the function as it transitions from low to high gray-level values. Your function should be normalized so that its minimum and maximum values are 0 and 1, respectively.

(b) Sketch a family of transformations as a function of parameter E, for a fixed value $m = L/2$, where L is the number of gray levels in the image.

(c) What is the smallest value of s that will make your function *effectively* perform as the function in Fig. 3.2(b)? In other words, your function does not have to be identical to Fig. 3.2(b). It just has to yield the same result of producing a binary image. Assume that you are working with 8-bit images, and let $m = 128$. Also, let C be the smallest positive number representable in the computer you are using.

3.3 Propose a set of gray-level-slicing transformations capable of producing all the individual bit planes of an 8-bit monochrome image. (For example, a transformation function with the property $T(r) = 0$ for r in the range $[0, 127]$, and $T(r) = 255$ for r in the range $[128, 255]$ produces an image of the 7th bit plane in an 8-bit image.)

3.4 ★ **(a)** What effect would setting to zero the lower-order bit planes have on the histogram of an image in general?

(b) What would be the effect on the histogram if we set to zero the higher-order bit planes instead?

★ **3.5** Explain why the discrete histogram equalization technique does not, in general, yield a flat histogram.

3.6 Suppose that a digital image is subjected to histogram equalization. Show that a second pass of histogram equalization will produce exactly the same result as the first pass.

3.7 In some applications it is useful to model the histogram of input images as Gaussian probability density functions of the form

$$p_r(r) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(r-m)^2}{2\sigma^2}}$$

where m and σ are the mean and standard deviation of the Gaussian PDF. The approach is to let m and σ be measures of average gray level and contrast of a given image. What is the transformation function you would use for histogram equalization?

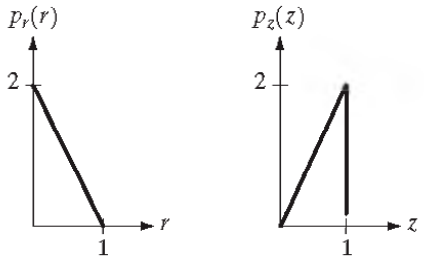
★ **3.8** Assuming continuous values, show by example that it is possible to have a case in which the transformation function given in Eq. (3.3-4) satisfies Conditions (a) and (b) in Section 3.3.1, but its inverse may fail to be single valued.

3.9 **(a)** Show that the discrete transformation function given in Eq. (3.3-8) for histogram equalization satisfies conditions (a) and (b) in Section 3.3.1.

(b) Show by example that this does not hold in general for the inverse discrete transformation function given in Eq. (3.3-9).

★ **(c)** Show that the inverse discrete transformation in Eq. (3.3-9) satisfies Conditions (a) and (b) in Section 3.3.1 if none of the gray levels $r_k, k = 0, 1, \ldots, L - 1$, are missing.

3.10 An image has the gray level PDF $p_r(r)$ shown in the following diagram. It is desired to transform the gray levels of this image so that they will have the specified $p_z(z)$ shown. Assume continuous quantities and find the transformation (in terms of r and z) that will accomplish this.



★3.11 Propose a method for updating the local histogram for use in the local enhancement technique discussed in Section 3.3.3.

3.12 Two images, $f(x, y)$ and $g(x, y)$, have histograms h_f and h_g. Give the conditions under which you can determine the histograms of

(a) $f(x, y) + g(x, y)$
(b) $f(x, y) - g(x, y)$
(c) $f(x, y) \times g(x, y)$
(d) $f(x, y) \div g(x, y)$

in terms of h_f and h_g. Explain how to obtain the histogram in each case.

3.13 Consider two 8-bit images whose gray levels span the full range from 0 to 255.

(a) Discuss the limiting effect of repeatedly subtracting image (b) from image (a).

(b) Would reversing the order of the images yield a different result?

★3.14 Image subtraction is used often in industrial applications for detecting missing components in product assembly. The approach is to store a "golden" image that corresponds to a correct assembly; this image is then subtracted from incoming images of the same product. Ideally, the differences would be zero if the new products are assembled correctly. Difference images for products with missing components would be nonzero in the area where they differ from the golden image. What conditions do you think have to be met in practice for this method to work?

3.15 Prove the validity of Eqs. (3.4-4) and (3.4-5).

3.16 In an industrial application, X-ray imaging is to be used to inspect the inside of certain composite castings. The objective is to look for voids in the castings, which typically appear as small blobs in the image. However, due to properties in of the casting material and X-ray energy used, high noise content often makes inspection difficult, so the decision is made to use image averaging to reduce the noise and thus improve visible contrast. In computing the average, it is important to keep the number of images as small as possible to reduce the time the parts have to remain stationary during imaging. After numerous experiments, it is concluded that decreasing the noise variance by a factor of 10 is sufficient. If the imaging device can produce 30 frames/s, how long would the castings have to remain stationary during imaging to achieve the desired decrease in variance? Assume that the noise is uncorrelated and has zero mean.

3.17 The implementation of linear spatial filters requires moving the center of a mask throughout an image and, at each location, computing the sum of products of the mask coefficients with the corresponding pixels at that location (see Section 3.5). In the case of lowpass filtering, all coefficients are 1, allowing use of a so-called *box-filter* or *moving-average* algorithm, which consists of updating only the part of the computation that changes from one location to the next.

★ **(a)** Formulate such an algorithm for an $n \times n$ filter, showing the nature of the computations involved and the scanning sequence used for moving the mask around the image.

(b) The ratio of the number of computations performed by a brute-force implementation to the number of computations performed by the box-filter algorithm is called the *computational advantage*. Obtain the computational advantage in this case and plot it as a function of n for $n > 1$. The $1/n^2$ scaling factor is common to both approaches, so you need not consider it in obtaining the computational advantage. Assume that the image has an outer border of zeros that is thick enough to allow you to ignore border effects in your analysis.

3.18 Discuss the limiting effect of repeatedly applying a 3×3 lowpass spatial filter to a digital image. You may ignore border effects.

3.19 ★ **(a)** It was stated in Section 3.6.2 that isolated clusters of dark or light (with respect to the background) pixels whose area is less than one-half the area of a median filter are eliminated (forced to the median value of the neighbors) by the filter. Assume a filter of size $n \times n$, with n odd, and explain why this is so.

(b) Consider an image having various sets of pixel clusters. Assume that all points in a cluster are lighter or darker than the background (but not both simultaneously in the same cluster), and that the area of each cluster is less than or equal to $n^2/2$. In terms of n, under what condition would one or more of these clusters cease to be isolated in the sense described in part (a)?

★ **3.20** **(a)** Develop a procedure for computing the median of an $n \times n$ neighborhood.

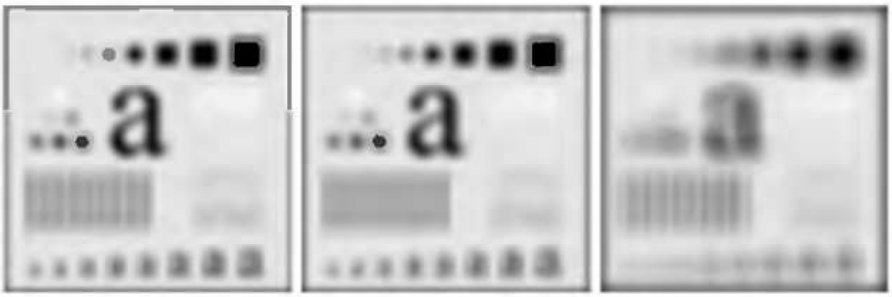
(b) Propose a technique for updating the median as the center of the neighborhood is moved from pixel to pixel.

3.21 **(a)** In a character recognition application, text pages are reduced to binary form using a thresholding transformation function of the form shown in Fig. 3.2(b). This is followed by a procedure that thins the characters until they become strings of binary 1's on a background of 0's. Due to noise, the binarization and thinning processes result in broken strings of characters with gaps ranging from 1 to 3 pixels. One way to "repair" the gaps is to run an averaging mask over the binary image to blur it, and thus create bridges of nonzero pixels between gaps. Give the (odd) size of the smallest averaging mask capable of performing this task.

(b) After bridging the gaps, it is desired to threshold the image in order to convert it back to binary form. For your answer in (a), what is the minimum value of the threshold required to accomplish this, without causing the segments to break up again?

★ **3.22** The three images shown were blurred using square averaging masks of sizes $n = 23, 25$, and 45, respectively. The vertical bars on the left lower part of (a) and (c) are blurred, but a clear separation exists between them. However, the bars

have merged in image (b), in spite of the fact that the mask that produced this image is significantly smaller than the mask that produced image (c). Explain this.



(a) (b) (c)

3.23 Consider an application such as the one shown in Fig. 3.36, in which it is desired to eliminate objects smaller than those enclosed in a square of size $q \times q$ pixels. Suppose that we want to reduce the average gray level of those objects to one-tenth of their original average gray level. In this way, those objects will be closer to the gray level of the background and they can then be eliminated by thresholding. Give the (odd) size of the smallest averaging mask that will accomplish the desired reduction in average gray level in only one pass of the mask over the image.

3.24 In a given application an averaging mask is applied to input images to reduce noise, and then a Laplacian mask is applied to enhance small details. Would the result be the same if the order of these operations were reversed?

★3.25 Show that the Laplacian operation defined in Eq. (3.7-1) is isotropic (invariant to rotation). You will need the following equations relating coordinates after axis rotation by an angle θ:

$$x = x' \cos\theta - y' \sin\theta$$
$$y = x' \sin\theta + y' \cos\theta$$

where (x, y) are the unrotated and (x', y') are the rotated coordinates.

3.26 Give a 3×3 mask for performing unsharp masking in a single pass through an image.

★3.27 Show that subtracting the Laplacian from an image is proportional to unsharp masking. Use the definition for the Laplacian given in Eq. (3.7-4).

3.28 **(a)** Show that the magnitude of the gradient given in Eq. (3.7-13) is an isotropic operation. (See Problem 3.25.)

(b) Show that the isotropic property is lost in general if the gradient is computed using Eq. (3.7-14).

3.29 A CCD TV camera is used to perform a long-term study by observing the same area 24 hours a day, for 30 days. Digital images are captured and transmitted to a central location every 5 minutes. The illumination of the scene changes from natural daylight to artificial lighting. At no time is the scene without illumination, so it is always possible to obtain an image. Because the range of illumination is such that it is always in the linear operating range of the camera, it is decided not to employ any compensating mechanisms on the camera itself. Rather, it is decided to use digital techniques to postprocess, and thus normalize, the images to the equivalent of constant illumination. Propose a method to do this. You are at liberty to use any method you wish, but state clearly all the assumptions you made in arriving at your design.

4 Image Enhancement in the Frequency Domain

Enhance: To make greater ⟨as in value, desirability, or attractiveness⟩.

Frequency: The number of times that a periodic function repeats the same sequence of values during a unit variation of the independent variable.

Webster's New Collegiate Dictionary

Preview

Although significant effort was devoted in the previous chapter to spatial techniques for image enhancement, a thorough understanding of this area is impossible without having at least a working knowledge of how the Fourier transform and the frequency domain can be used for image processing. A solid understanding of these topics can be developed without having to become a signal processing expert. The key lies in focusing on the fundamentals and their relevance to digital image processing. The notation, usually a source of trouble for the beginner, is clarified significantly in this chapter by emphasizing the connection between image characteristics and the mathematical tools used to represent them. This chapter is concerned primarily with helping the reader develop a basic understanding of the Fourier transform and the frequency domain, and how they apply to image enhancement. Later, in Chapters 5, 8, 10, and 11, we discuss other applications of the Fourier transform.

We begin the discussion with a brief outline of the origins of the Fourier transform and its impact on countless branches of mathematics, science, and engineering. Next, we give an introduction to the Fourier transform and the frequency domain, and establish the notation and reasons why these tools are so useful for image enhancement. This is followed by sections that parallel the spatial smoothing and sharpening filtering techniques discussed in Chapter 3, except that all filtering is done in the frequency domain. After discussing other uses of the Fourier transform for image enhancement, we conclude the chapter with

a discussion of issues related to implementing the Fourier transform in the context of image processing.

4.1 Background

The French mathematician Jean Baptiste Joseph Fourier was born in 1768 in the town of Auxerre, about midway between Paris and Dijon. The contribution for which he is most remembered was outlined in a memoir in 1807 and published in 1822 in his book, *La Théorie Analitique de la Chaleur* (The Analytic Theory of Heat). This book was translated into English 55 years later by Freeman (see Freeman [1878]). Basically, Fourier's contribution in this particular field states that any function that periodically repeats itself can be expressed as the sum of sines and/or cosines of different frequencies, each multiplied by a different coefficient (we now call this sum a *Fourier series*). It does not matter how complicated the function is; as long as it is periodic and meets some mild mathematical conditions, it can be represented by such a sum. This is now taken for granted, but at the time it first appeared it was a revolutionary concept to which it took mathematicians all over the world over a century to "adjust." At that time, regularity in functions was a mainstay of mathematical thinking. With this type of cultural mindset, the concept that complicated functions could be represented as a sum of simple sines and cosines was not at all intuitive (Fig. 4.1), so it is not surprising that Fourier's ideas in this regard were met with skepticism.

Even functions that are not periodic (but whose area under the curve is finite) can be expressed as the integral of sines and/or cosines multiplied by a weighing function. The formulation in this case is the *Fourier transform*, and its utility is even greater than the Fourier series in most practical problems. Both representations share the important characteristic that a function, expressed in either a Fourier series or transform, can be reconstructed (recovered) completely via an inverse process, with no loss of information. This is one of the most important characteristics of these representations because they allow us to work in the "Fourier domain" and then return to the original domain of the function without losing any information.

Ultimately, it was the utility of the Fourier series and transform in solving practical problems that made them widely used and studied as fundamental tools. The application of Fourier initial ideas was in the field of heat diffusion, where they allowed the formulation of differential equations representing heat flow in such a way that solutions could be obtained for the first time. During the past century, and especially in the past 50 years, entire industries and academic disciplines have flourished as a result of Fourier's ideas. The advent of digital computation and the "discovery" of a fast Fourier transform (FFT) algorithm in the late 1950s (more about this later) revolutionized the field of signal processing. These two core technologies allowed for the first time practical processing and meaningful interpretation of a host of signals of exceptional human and industrial importance, from medical monitors and scanners to modern electronic communications.

We will be dealing only with functions (images) of finite duration, so the Fourier transform is the tool in which we are interested. The material in the fol-

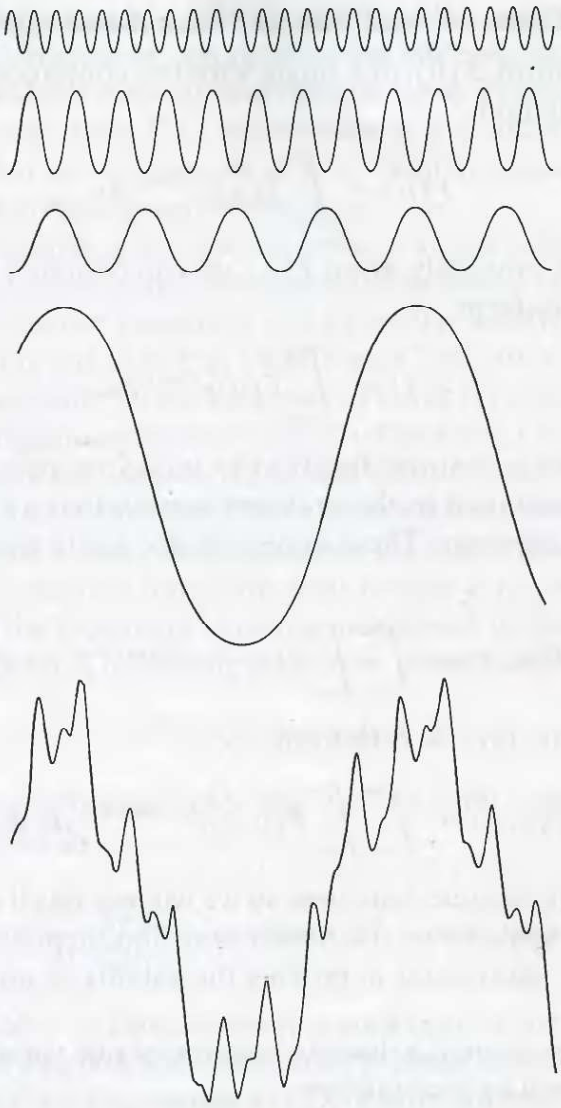


FIGURE 4.1 The function at the bottom is the sum of the four functions above it. Fourier's idea in 1807 that periodic functions could be represented as a weighted sum of sines and cosines was met with skepticism.

lowing section introduces the Fourier transform and the frequency domain. It is shown that Fourier techniques provide a meaningful and practical way to study and implement a host of image enhancement approaches. In some cases, these approaches are similar to the ones we developed in Chapter 3. In others, they are complementary.

4.2 Introduction to the Fourier Transform and the Frequency Domain

This section introduces the Fourier transform in one and two dimensions. The focus is mostly on a discrete formulation of the continuous transform and some of its properties.

4.2.1 The One-Dimensional Fourier Transform and its Inverse

The Fourier transform, $F(u)$, of a single variable, continuous function, $f(x)$, is defined by the equation

$$F(u) = \int_{-\infty}^{\infty} f(x)e^{-j2\pi ux}\,dx \tag{4.2-1}$$

where $j = \sqrt{-1}$. Conversely, given $F(u)$, we can obtain $f(x)$ by means of the *inverse* Fourier transform

$$f(x) = \int_{-\infty}^{\infty} F(u)e^{j2\pi ux}\,du. \tag{4.2-2}$$

These two equations comprise the *Fourier transform pair*. They indicate the important fact mentioned in the previous section that a function can be recovered from its transform. These equations are easily extended to two variables, u and v:

$$F(u, v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)e^{-j2\pi(ux+vy)}\,dx\,dy \tag{4.2-3}$$

and, similarly for the inverse transform,

$$f(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u, v)e^{j2\pi(ux+vy)}\,du\,dv. \tag{4.2-4}$$

Our interest is in discrete functions, so we will not dwell on these equations here. However, in some cases, the reader may find them easier to manipulate than their discrete equivalents in proving the validity of properties of the 2-D Fourier transform.

The Fourier transform of a discrete function of one variable, $f(x)$, $x = 0, 1, 2, \ldots, M-1$, is given by the equation

$$F(u) = \frac{1}{M}\sum_{x=0}^{M-1} f(x)e^{-j2\pi ux/M} \qquad \text{for } u = 0, 1, 2, \ldots, M-1. \tag{4.2-5}$$

This *discrete Fourier transform* (DFT) is the foundation for most of the work in this chapter. Similarly, given $F(u)$, we can obtain the original function back using the inverse DFT:

$$f(x) = \sum_{u=0}^{M-1} F(u)e^{j2\pi ux/M} \qquad \text{for } x = 0, 1, 2, \ldots, M-1. \tag{4.2-6}$$

The $1/M$ multiplier in front of the Fourier transform sometimes is placed in front of the inverse instead. Other times (not as often) both equations are multiplied by $1/\sqrt{M}$. The location of the multiplier does not matter. If two multipliers are used, the only requirement is that their product be equal to $1/M$. Considering their importance, these equations really are very simple.

In order to compute $F(u)$ in Eq. (4.2-5) we start by substituting $u = 0$ in the exponential term and then summing for *all* values of x. We then substitute $u = 1$ in the exponential and repeat the summation over all values of x. We repeat

this process for all M values of u in order to obtain the complete Fourier transform. It takes approximately M^2 summations and multiplications to compute the discrete Fourier transform (reduction of this number is an important topic of discussion in Section 4.6). Like $f(x)$, the transform is a discrete quantity, and it has the same number of components as $f(x)$. Similar comments apply to the computation of the inverse Fourier transform.

An important property of the discrete transform pair is that, unlike the continuos case, we need not be concerned about the existence of the DFT or its inverse. The discrete Fourier transform and its inverse always exist. This can be shown by substituting either of Eqs. (4.2-5) or (4.2-6) into the other and making use of the orthogonality of the exponential terms (Problem 4.1). We will get an identity, indicating the existence of the two functions. Of course, there is always the question of what happens if $f(x)$ has infinite values, but we deal strictly with finite quantities in this book. These comments are directly applicable to two-dimensional (and higher) functions. Thus, for digital image processing, existence of either the discrete transform or its inverse is not an issue.

The concept of the frequency domain, mentioned numerous times in this chapter and in Chapter 3, follows directly from Euler's formula:

$$e^{j\theta} = \cos\theta + j\sin\theta. \tag{4.2-7}$$

Substituting this expression into Eq. (4.2-5), and using the fact that $\cos(-\theta) = \cos\theta$, gives us

$$F(u) = \frac{1}{M}\sum_{x=0}^{M-1} f(x)[\cos 2\pi ux/M - j\sin 2\pi ux/M] \tag{4.2-8}$$

for $u = 0, 1, 2, \ldots, M-1$. Thus, we see that *each* term of the Fourier transform [that is, the value of $F(u)$ for each value of u] is composed of the sum of *all* values of the function $f(x)$. The values of $f(x)$, in turn, are multiplied by sines and cosines of various frequencies. The domain (values of u) over which the values of $F(u)$ range is appropriately called the *frequency domain*, because u determines the frequency of the components of the transform. (The x's also affect the frequencies, but they are summed out and they all make the same contributions for each value of u.) Each of the M terms of $F(u)$ is called a *frequency component* of the transform. Use of the terms *frequency domain* and *frequency components* is really no different from the terms *time domain* and *time components*, which we would use to express the domain and values of $f(x)$ if x were a time variable.

A useful analogy is to compare the Fourier transform to a glass prism. The prism is a physical device that separates light into various color components, each depending on its wavelength (or frequency) content. The Fourier transform may be viewed as a "mathematical prism" that separates a function into various components, also based on frequency content. When we consider light, we talk about its spectral or frequency content. Similarly, the Fourier transform lets us characterize a function by its frequency content. This is a powerful concept that lies at the heart of linear filtering.

In general, we see from Eqs. (4.2-5) or (4.2-8) that the components of the Fourier transform are complex quantities. As in the analysis of complex numbers, we find it convenient sometimes to express $F(u)$ in polar coordinates:

$$F(u) = |F(u)|e^{-j\phi(u)} \tag{4.2-9}$$

where

$$|F(u)| = \left[R^2(u) + I^2(u)\right]^{1/2} \tag{4.2-10}$$

is called the *magnitude* or *spectrum* of the Fourier transform, and

$$\phi(u) = \tan^{-1}\left[\frac{I(u)}{R(u)}\right] \tag{4.2-11}$$

is called the *phase angle* or *phase spectrum* of the transform. In Eqs. (4.2-10) and (4.2-11), $R(u)$ and $I(u)$ are the real and imaginary parts of $F(u)$, respectively. In terms of image enhancement we are concerned primarily with properties of the spectrum. Another quantity that is used later in this chapter is the *power spectrum*, defined as the square of the Fourier spectrum:

$$\begin{aligned} P(u) &= |F(u)|^2 \\ &= R^2(u) + I^2(u). \end{aligned} \tag{4.2-12}$$

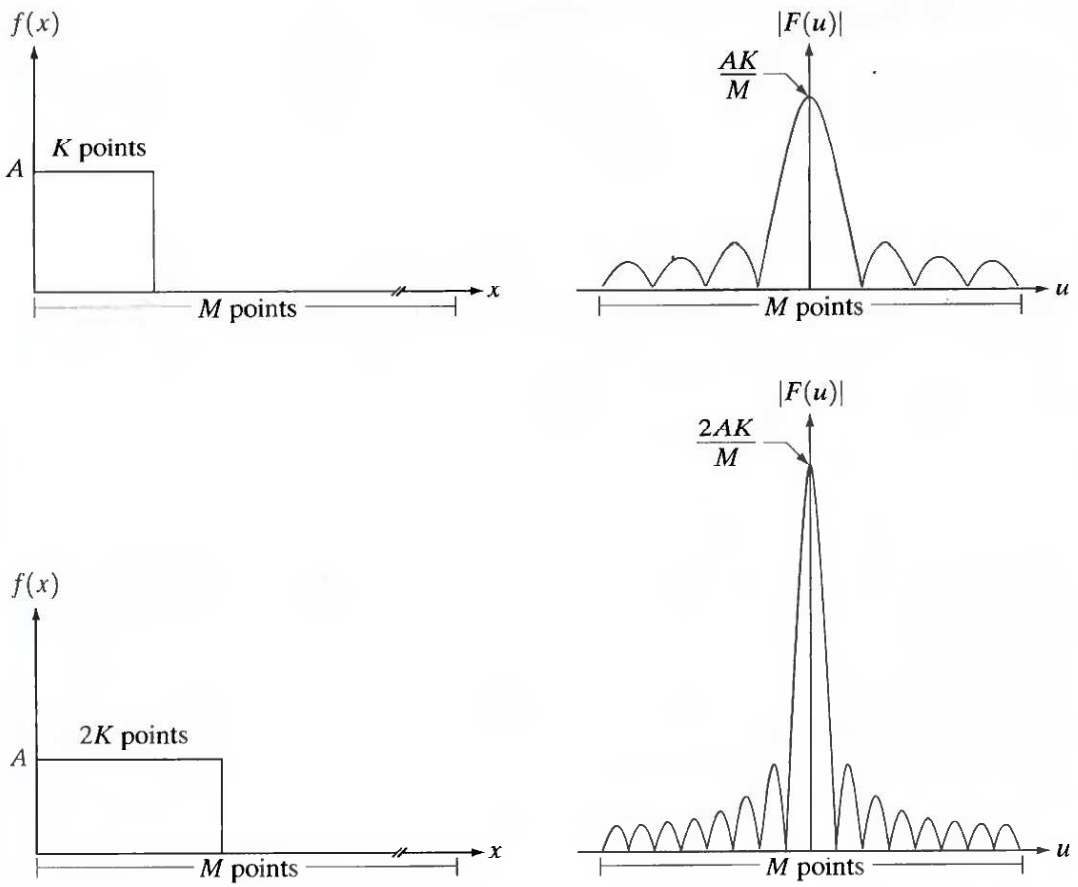
The term *spectral density* also is used to refer to the power spectrum.

EXAMPLE 4.1: Fourier spectra of two simple 1-D functions.

■ Before proceeding, it will be helpful to consider a simple one-dimensional example of the DFT. Figure 4.2(a) shows a function and Fig. 4.2(b) shows its Fourier spectrum. Both $f(x)$ and $F(u)$ are discrete quantities, but the points in the plots are linked to make them easier to follow visually. In this example, $M = 1024$, $A = 1$, and K is only 8 points. Also note that the spectrum is centered at $u = 0$. As shown in the following section, this is accomplished by multiplying $f(x)$ by $(-1)^x$ before taking the transform. The next two figures depict basically the same thing, but with $K = 16$ points. The important features to note are that (1) the height of the spectrum doubled as the area under the curve in the x-domain doubled, and (2) the number of zeros in the spectrum in the same interval doubled as the length of the function doubled. This "reciprocal" nature of the Fourier transform pair is most useful in interpreting results of image processing in the frequency domain. ■

In the discrete transform of Eq. (4.2-5), the function $f(x)$ for $x = 0, 1, 2, \ldots, M - 1$, represents M samples from its continuous counterpart. It is important to keep in mind that these samples are *not* necessarily always taken at integer values of x in the interval $[0, M - 1]$. They are taken at equally spaced, but otherwise arbitrary, points. This is usually represented by letting x_0 denote the first (arbitrarily located) point in the sequence. The first value of the sampled function is then $f(x_0)$. The next sample has taken a fixed interval Δx units away to give $f(x_0 + \Delta x)$. The kth sample gives us $f(x_0 + k\Delta x)$, and the final sample is $f(x_0 + [M - 1]\Delta x)$. Thus, in the discrete case, when we write $f(k)$, it is understood that we are utilizing shorthand notation that really means $f(x_0 + k\Delta x)$. In terms of the notation we have used thus far, $f(x)$ is then understood to mean

$$f(x) \triangleq f(x_0 + x\Delta x) \tag{4.2-13}$$



a b
c d

FIGURE 4.2 (a) A discrete function of M points, and (b) its Fourier spectrum. (c) A discrete function with twice the number of nonzero points, and (d) its Fourier spectrum.

when dealing with discrete variables. The variable u has a similar interpretation, but the sequence always starts at true zero frequency. Thus, the sequence for the values of u is $0, \Delta u, 2\Delta u, \ldots, [M-1]\Delta u$. Then, $F(u)$ is understood to mean

$$F(u) \triangleq F(u\Delta u) \tag{4.2-14}$$

for $u = 0, 1, 2, \ldots, M-1$. This type of shorthand notation simplifies equations considerably and is much easier to follow.

Given the inverse relationship between a function and its transform illustrated in Fig. 4.2, it is not surprising that Δx and Δu are inversely related by the expression

$$\Delta u = \frac{1}{M\Delta x}. \tag{4.2-15}$$

This relationship is useful when measurements are an issue in the images being processed. For instance, in an application of electron microscopy the image samples may be spaced 1 micron apart, and certain characteristics in the frequency domain (like periodicity components) may have implications regarding the structure of the physical sample. For the most part in subsequent discussions in this book we use the variables x and u without making reference to specific sampling or other measurement considerations.

4.2.2 The Two-Dimensional DFT and Its Inverse

Extension of the one-dimensional discrete Fourier transform and its inverse to two dimensions is straightforward. The discrete Fourier transform of a function (image) $f(x, y)$ of size $M \times N$ is given by the equation

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M+vy/N)}. \tag{4.2-16}$$

As in the 1-D case, this expression must be computed for values of $u = 0, 1, 2, \ldots, M - 1$, and also for $v = 0, 1, 2, \ldots, N - 1$. Similarly, given $F(u, v)$, we obtain $f(x, y)$ via the *inverse* Fourier transform, given by the expression

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux/M+vy/N)} \tag{4.2-17}$$

for $x = 0, 1, 2, \ldots, M - 1$ and $y = 0, 1, 2, \ldots, N - 1$. Equations (4.2-16) and (4.2-17) comprise the *two-dimensional, discrete Fourier transform (DFT) pair*. The variables u and v are the *transform* or *frequency variables*, and x and y are the *spatial* or *image variables*. As in the one-dimensional case, the location of the $1/MN$ constant is not important. Sometimes it is located in front of the inverse transform. Other times it is found split into two equal terms of $1/\sqrt{MN}$ multiplying the transform and its inverse.

We define the Fourier spectrum, phase angle, and power spectrum as in the previous section:

$$|F(u, v)| = \left[R^2(u, v) + I^2(u, v)\right]^{1/2} \tag{4.2-18}$$

$$\phi(u, v) = \tan^{-1}\left[\frac{I(u, v)}{R(u, v)}\right] \tag{4.2-19}$$

and

$$\begin{aligned} P(u, v) &= |F(u, v)|^2 \\ &= R^2(u, v) + I^2(u, v) \end{aligned} \tag{4.2-20}$$

where $R(u, v)$ and $I(u, v)$ are the real and imaginary parts of $F(u, v)$, respectively.

It is common practice to multiply the input image function by $(-1)^{x+y}$ prior to computing the Fourier transform. Due to the properties of exponentials, it is not difficult to show (see Section 4.6) that

$$\Im[f(x, y)(-1)^{x+y}] = F(u - M/2, v - N/2) \tag{4.2-21}$$

where $\Im[\cdot]$ denotes the Fourier transform of the argument. This equation states that the origin of the Fourier transform of $f(x, y)(-1)^{x+y}$ [that is, $F(0, 0)$] is located at $u = M/2$ and $v = N/2$. In other words, multiplying $f(x, y)$ by $(-1)^{x+y}$ shifts the origin of $F(u, v)$ to frequency coordinates $(M/2, N/2)$, which is the center of the $M \times N$ area occupied by the 2-D DFT. We refer to this area of the frequency domain as the *frequency rectangle*. It extends from $u = 0$ to $u = M - 1$, and from $v = 0$ to $v = N - 1$ (keep in mind that u and v are integers). In order to guarantee that these shifted coordinates are integers, we require that M and N be even numbers. When implementing the Fourier transform in a computer, the limits of

summations are from $u = 1$ to M and $v = 1$ to N. The actual center of the transform will then be at $u = (M/2) + 1$ and $v = (N/2) + 1$.

The value of the transform at $(u, v) = (0, 0)$ is, from Eq. (4.2-16),

$$F(0,0) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y), \tag{4.2-22}$$

which we see is the average of $f(x, y)$. In other words, if $f(x, y)$ is an image, the value of the Fourier transform at the origin is equal to the average gray level of the image. Because both frequencies are zero at the origin, $F(0, 0)$ sometimes is called the *dc* component of the spectrum. This terminology is from electrical engineering, where "dc" signifies direct current (i.e., current of zero frequency).

If $f(x, y)$ is real, its Fourier transform is conjugate symmetric; that is,

$$F(u, v) = F^*(-u, -v) \tag{4.2-23}$$

where "*" indicates the standard conjugate operation on a complex number. From this, it follows that

$$|F(u, v)| = |F(-u, -v)|, \tag{4.2-24}$$

which says that the spectrum of the Fourier transform is symmetric. Conjugate symmetry and the centering property discussed previously truly simplify the specification of circularly symmetric filters in the frequency domain, as shown in the following section.

Finally, as in the 1-D case, we have the following relationships between samples in the spatial and frequency domains:

$$\Delta u = \frac{1}{M\Delta x} \tag{4.2-25}$$

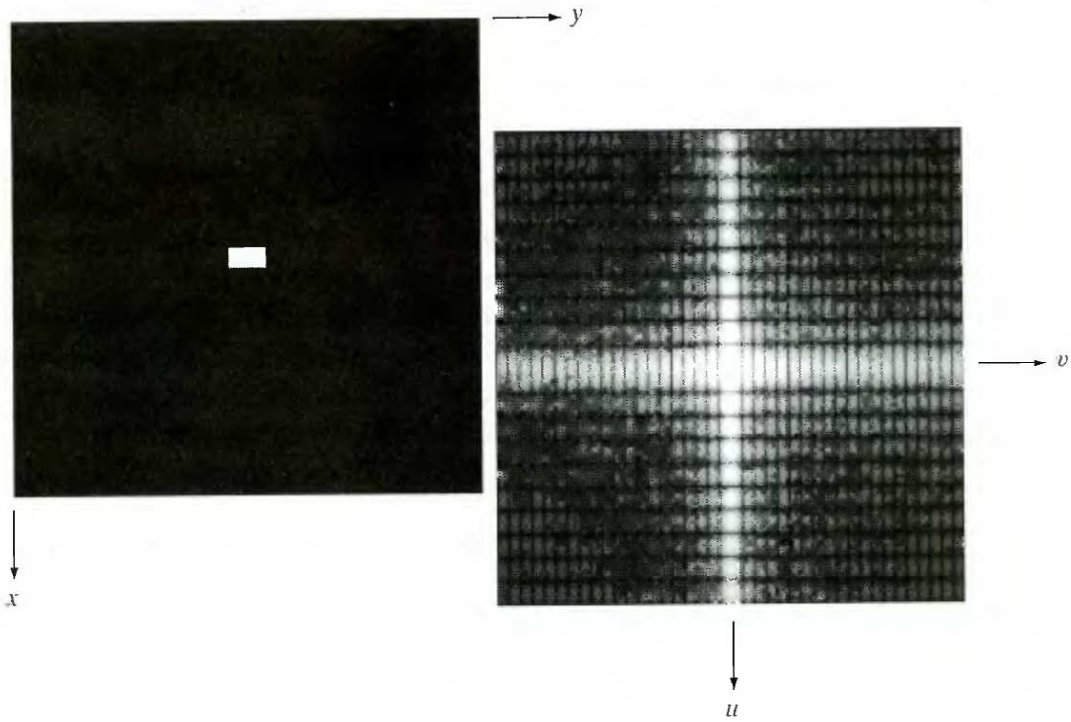
and

$$\Delta v = \frac{1}{N\Delta y}. \tag{4.2-26}$$

The significance of these variables is identical in meaning to the explanation given in Section 4.2.1 for 1-D variables.

EXAMPLE 4.2: Centered spectrum of a simple 2-D function.

■ Figure 4.3(a) shows a white rectangle of size 20×40 pixels superimposed on a black background of size 512×512 pixels. This image was multiplied by $(-1)^{x+y}$ prior to computing the Fourier transform in order to center the spectrum, which is shown in Fig. 4.3(b). (Note the location, labels, and origin of the axes in both figures. We follow this convention throughout all discussions of images and their corresponding Fourier spectra.) In Fig. 4.3(b), the separation of spectrum zeros in the u-direction is exactly twice the separation of zeros in the v direction. This corresponds inversely to the 1-to-2 size ratio of the rectangle in the image. The spectrum was processed prior to displaying by using the log transformation in Eq. (3.2-2) to enhance gray-level detail. A value of $c = 0.5$ was used in the transformation in order to decrease overall intensity. Most Fourier spectra shown in this chapter are similarly processed by a log transformation. ■



a b

FIGURE 4.3 (a) Image of a 20×40 white rectangle on a black background of size 512×512 pixels. (b) Centered Fourier spectrum shown after application of the log transformation given in Eq. (3.2-2). Compare with Fig. 4.2.

4.2.3 Filtering in the Frequency Domain

As noted in the past two sections, the frequency domain is nothing more than the space defined by values of the Fourier transform and its frequency variables (u, v). In this section, we attach "meaning" to the frequency domain, as it relates to image processing.

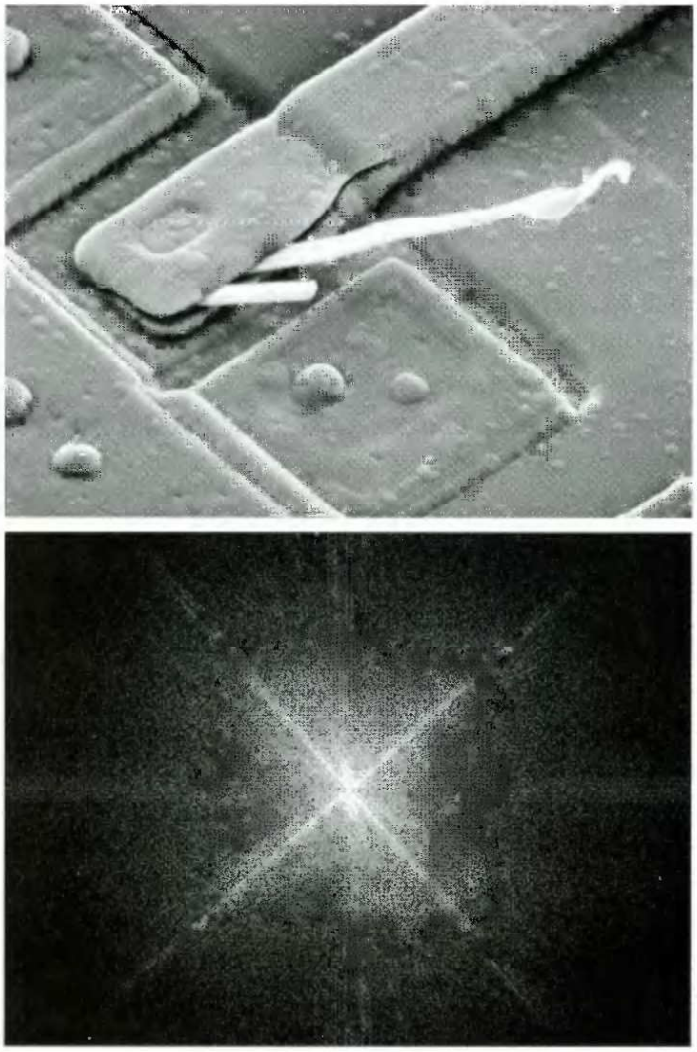
Some basic properties of the frequency domain

We start by observing in Eq. (4.2-16) that *each* term of $F(u, v)$ contains *all* values of $f(x, y)$, modified by the values of the exponential terms. Thus, with the exception of trivial cases, it usually is impossible to make direct associations between specific components of an image and its transform. However, some general statements can be made about the relationship between the frequency components of the Fourier transform and spatial characteristics of an image. For instance, since frequency is directly related to rate of change, it is not difficult intuitively to associate frequencies in the Fourier transform with patterns of intensity variations in an image. We showed in the previous section that the slowest varying frequency component $(u = v = 0)$ corresponds to the average gray level of an image. As we move away from the origin of the transform, the low frequencies correspond to the slowly varying components of an image. In an image of a room, for example, these might correspond to smooth gray-level variations on the walls and floor. As we move further away from the origin, the higher frequencies begin to correspond to faster and faster gray level changes in the image. These are the edges of objects and other components of an image characterized by abrupt changes in gray level, such as noise.

EXAMPLE 4.3: An image and its Fourier spectrum, showing some important features.

■ An illustration will help fix these ideas. The image shown in Fig. 4.4(a) is a scanning electron miscroscope image of an integrated circuit, magnified approximately 2500 times. Aside from the interesting construction of the device itself, we note two principal features: strong edges that run approximately at $\pm 45°$, and the two white oxide protrusions resulting from thermally induced failure. The Fourier spectrum in Fig. 4.4(b) shows prominent components along the $\pm 45°$ directions that correspond to the edges just mentioned. Looking carefully along the vertical axis, we see a vertical component that is off-axis slightly to the left. This component was caused by the edges of the oxide protrusions. Note how the off-axis angle of the frequency component corresponds to the inclination off horizontal of the long white element, and note also the zeros in the vertical frequency component, corresponding to the narrow vertical span of the oxide protrusions.

This example is typical of the types of associations that can be made in general between the frequency and spatial domains. As we show throughout this chapter, even these types of gross associations, along with the relationships mentioned previously between frequency content and rate of change of gray levels in an image, can lead to some very useful enhancement results. ■



a
b

FIGURE 4.4 (a) SEM image of a damaged integrated circuit. (b) Fourier spectrum of (a). (Original image courtesy of Dr. J. M. Hudak, Brockhouse Institute for Materials Research, McMaster University, Hamilton, Ontario, Canada.)

Basics of filtering in the frequency domain

Filtering in the frequency domain is straightforward. It consists of the following steps:

1. Multiply the input image by $(-1)^{x+y}$ to center the transform, as indicated in Eq. (4.2-21).
2. Compute $F(u, v)$, the DFT of the image from (1).
3. Multiply $F(u, v)$ by a *filter* function $H(u, v)$.
4. Compute the inverse DFT of the result in (3).
5. Obtain the real part of the result in (4).
6. Multiply the result in (5) by $(-1)^{x+y}$.

The reason that $H(u, v)$ is called a *filter* (the term *filter transfer function* also is used commonly) is because it suppresses certain frequencies in the transform while leaving others unchanged. The analogy from everyday life is a screen filter that passes certain objects and suppresses others, based strictly on their size.

In equation form, let $f(x, y)$ represent the input image in Step 1 and $F(u, v)$ its Fourier transform. Then the Fourier transform of the output image is given by

$$G(u, v) = H(u, v)F(u, v). \tag{4.2-27}$$

The multiplication of H and F involves two-dimensional functions and is defined on an element-by-element basis. That is, the first element of H multiplies the first element of F, the second element of H multiplies the second element of F, and so on. In general, the components of F are complex quantities, but the filters with which we deal in this book typically are real. In this case, each component of H multiplies *both* the real and imaginary parts of the corresponding component in F. Such filters are called *zero-phase-shift* filters. As their name implies, these filters do not change the phase of the transform, a fact that can be seen in Eq. (4.2-19) by noting that the multiplier of the real and imaginary parts would cancel out because they have the same value.

The filtered image is obtained simply by taking the inverse Fourier transform of $G(u, v)$:

$$\text{Filtered Image} = \Im^{-1}[G(u, v)]. \tag{4.2-28}$$

The final image is obtained by taking the real part of this result and multiplying it by $(-1)^{x+y}$ to cancel the multiplication of the input image by this quantity. The inverse Fourier transform is, in general, complex. However, when the input image and the filter function are real, the imaginary components of the inverse transform should all be zero. In practice, the inverse DFT generally has parasitic imaginary components due to computational round-off errors. These components are ignored.

The filtering procedure just outlined is summarized in Fig. 4.5 in a slightly more general form that includes pre- and postprocessing stages. In addition to the $(-1)^{x+y}$ process, examples of other processes might include cropping of the input image to its closest even dimensions (required for proper transform centering), gray-level scaling, conversion to floating point on input, and conversion to an 8-bit integer format on the output. Multiple filtering stages and other pre- and postprocessing functions are possible. There are numerous variations

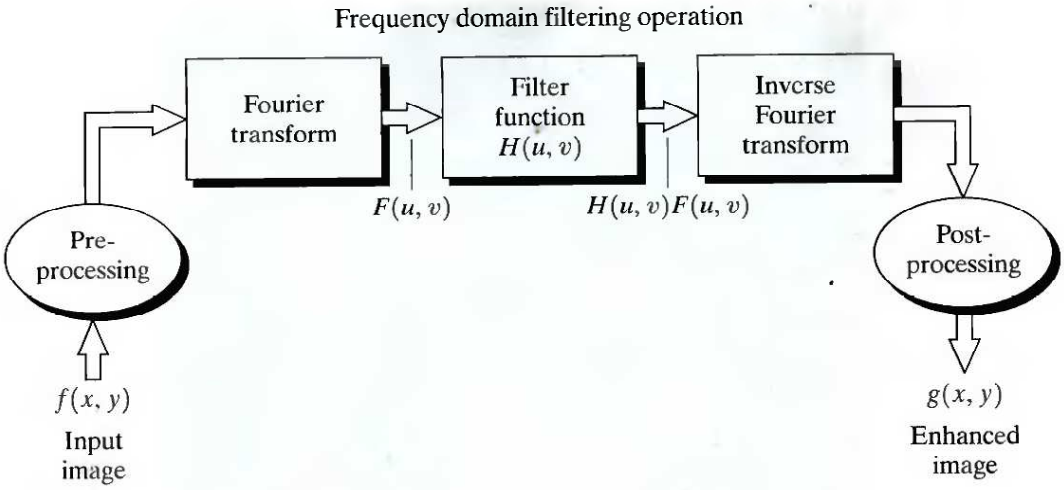


FIGURE 4.5 Basic steps for filtering in the frequency domain.

of this basic theme. The important point to keep in mind is that the filtering process is based on modifying the transform of an image in some way via a filter function, and then taking the inverse of the result to obtain the processed output image.

Some basic filters and their properties

At this point we have established the foundation for filtering in the frequency domain. The next logical step is to look at some specific filters and see how they affect images. The earlier discussion of Eq. (4.2-22) gives us a perfect lead into an introductory example of filtering. Suppose that we wish to force the average value of an image to zero. According to Eq. (4.2-22), the average value of an image is given by $F(0, 0)$. If we set this term to zero in the frequency domain and take the inverse transform, then the average value of the resulting image will be zero. Assuming that the transform has been centered as discussed in Eq. (4.2-21), we can do this operation by multiplying all values of $F(u, v)$ by the filter function:

$$H(u, v) = \begin{cases} 0 & \text{if} \quad (u, v) = (M/2, N/2) \\ 1 & \text{otherwise.} \end{cases} \tag{4.2-29}$$

All this filter would do is set $F(0, 0)$ to zero and leave all other frequency components of the Fourier transform untouched, as desired. The processed image (with zero average value) can then be obtained by taking the inverse Fourier transform of $H(u, v)F(u, v)$, as indicated in Eq. (4.2-28). As stated earlier, both the real and imaginary parts of $F(u, v)$ are multiplied by the filter function $H(u, v)$.

The filter just discussed is called a *notch filter* because it is a constant function with a hole (notch) at the origin. The result of processing the image in Fig. 4.4(a) with this filter is shown in Fig. 4.6. Note the drop in overall average gray level resulting from forcing the average value to zero; note also the "byproduct" result of making prominent edges stand out. (In reality the average of the *displayed* image cannot be zero because the image has to have

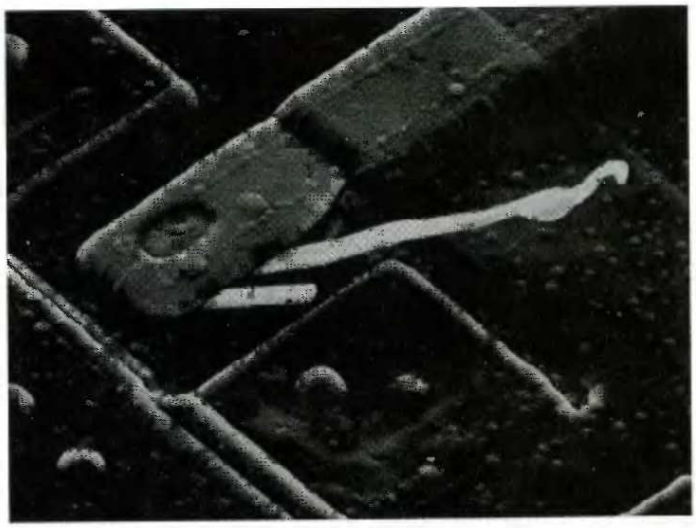
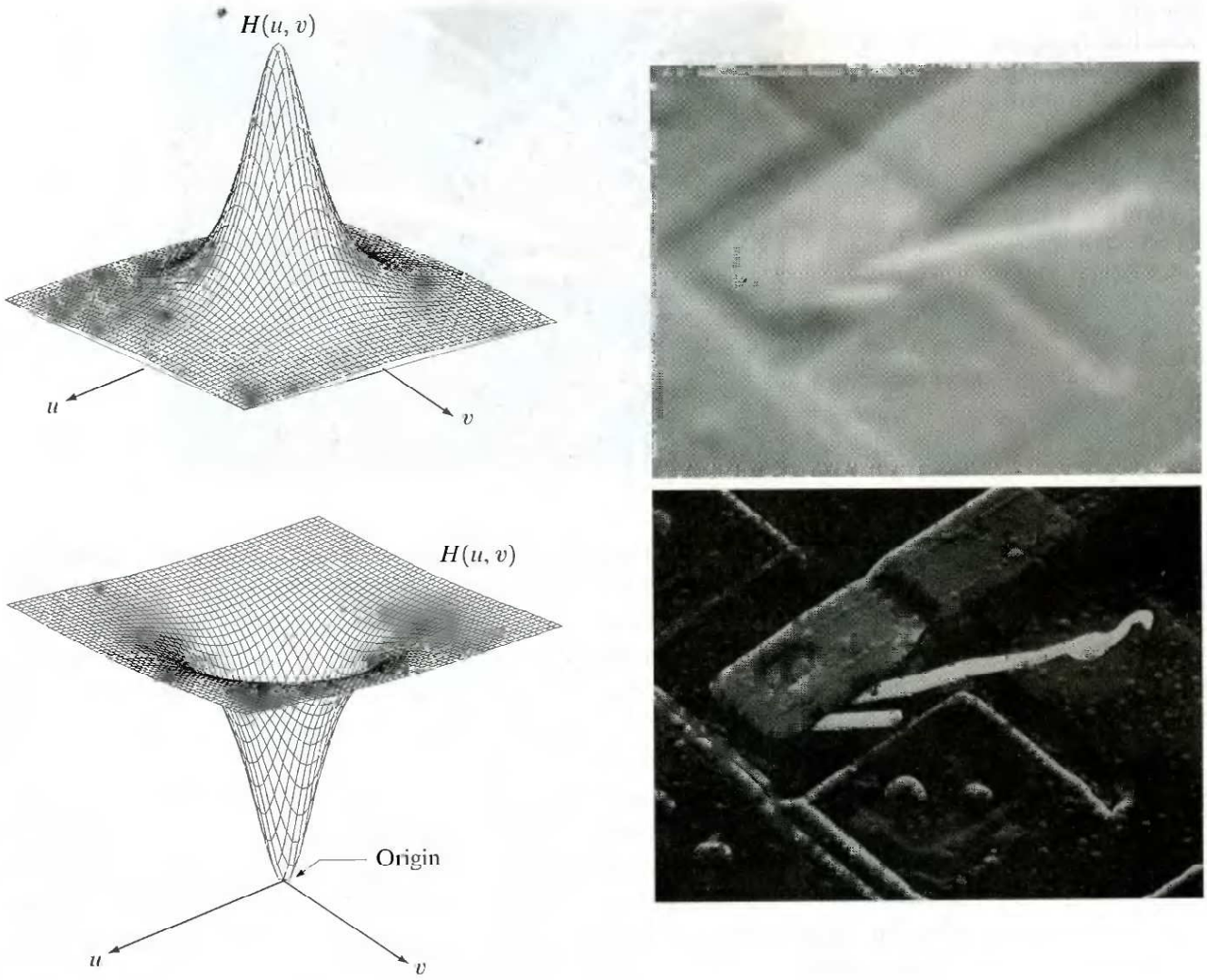


FIGURE 4.6 Result of filtering the image in Fig. 4.4(a) with a notch filter that set to 0 the $F(0, 0)$ term in the Fourier transform.

negative values for its average gray level to be zero and displays cannot handle negative quantities. Figure 4.6 was displayed in the "standard" way, which is to display the most negative value as 0, or black, with all other values scaled up from that.) As shown in Section 5.4.3, notch filters are exceptionally useful tools when it is possible to identify spatial image effects caused by specific, localized frequency domain components.

Low frequencies in the Fourier transform are responsible for the general gray-level appearance of an image over smooth areas, while high frequencies are responsible for detail, such as edges and noise. These ideas are discussed in more detail in the sections that follow, but it will be instructive to complement our illustration of the notch filter with an example of filters in these other two categories. A filter that attenuates high frequencies while "passing" low frequencies is called a *lowpass filter*. A filter that has the opposite characteristic is appropriately called a *highpass filter*. We would expect a lowpass-filtered image to have less sharp detail than the original because the high frequencies have been attenuated. Similarly, a highpass-filtered image would have less gray level variations in smooth areas and emphasized transitional (e.g., edge) gray-level detail. Such an image will appear sharper.

Figure 4.7 illustrates the effects of lowpass and highpass filtering the image in Fig. 4.4(a). The left part of the figure shows the filters and the right part shows the results of filtering using the procedure summarized in Fig. 4.5. The filters, $H(u, v)$, shown are both circularly symmetric. After shifting their origin to the center of the frequency rectangle occupied by $F(u, v)$, they were multiplied by the centered transform, as outlined in our discussion of Eqs. (4.2-27), (4.2-28), and Fig. 4.5. Taking the real part of each result and multiplying it by $(-1)^{x+y}$ yielded the images on the right. As expected, the image in Fig. 4.7(b) is blurred, and the image in Fig. 4.7(d) is sharp, with little smooth gray-level detail because the $F(0, 0)$ term has been set to zero. This is typical of highpassed results, and a procedure often followed is to add a constant to the filter so that it will not completely eliminate $F(0, 0)$. The result of using this procedure is shown in Fig. 4.8. The improvement over Fig. 4.7(d) is evident.



a b
c d

FIGURE 4.7 (a) A two-dimensional lowpass filter function. (b) Result of lowpass filtering the image in Fig. 4.4(a). (c) A two-dimensional highpass filter function. (d) Result of highpass filtering the image in Fig. 4.4(a).

4.2.4 Correspondence between Filtering in the Spatial and Frequency Domains

In the previous chapter we arrived at forms for various spatial filters using intuition and/or a mathematical formulation, such as the Laplacian. In this section we establish a direct link between some of those spatial filters and their frequency domain counterparts.

The most fundamental relationship between the spatial and frequency domains is established by a well-known result called the *convolution theorem*. The reader is already familiar with the basic concepts and mechanics of convolution in the spatial domain, which were introduced and illustrated in Section 3.5.

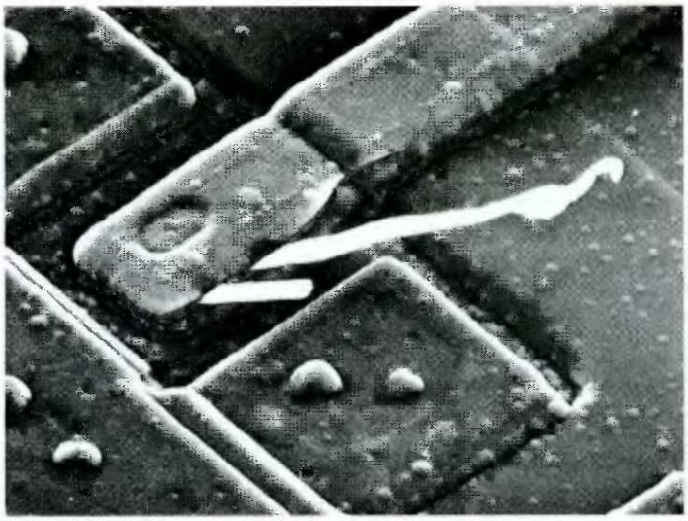


FIGURE 4.8 Result of highpass filtering the image in Fig. 4.4(a) with the filter in Fig. 4.7(c), modified by adding a constant of one-half the filter height to the filter function. Compare with Fig. 4.4(a).

The process by which we move a mask from pixel to pixel in an image, and compute a predefined quantity at each pixel, is the foundation of the convolution process. Formally, the discrete convolution of two functions $f(x, y)$ and $h(x, y)$ of size $M \times N$ is denoted by $f(x, y) * h(x, y)$ and is defined by the expression

$$f(x, y) * h(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n)h(x - m, y - n). \quad (4.2\text{-}30)$$

With the exception of the leading constant, the minus signs, and the limits of the summation, this expression is similar in *form* to Eq. (3.5-1). The minus signs, in particular, simply mean that function h is mirrored about the origin. This is inherent in the definition of convolution. Equation (4.2-30) is really nothing more than an implementation for (1) flipping one function about the origin; (2) shifting that function with respect to the the other by changing the values of (x, y); and (3) computing a sum of products over all values of m and n, for *each* displacement (x, y). The displacements (x, y) are integer increments that stop when the functions no longer overlap.

Letting $F(u, v)$ and $H(u, v)$ denote the Fourier transforms of $f(x, y)$ and $h(x, y)$, respectively, one-half of the convolution theorem simply states that $f(x, y) * h(x, y)$ and $F(u, v)H(u, v)$ constitute a Fourier transform pair. This result is formally stated as

$$f(x, y) * h(x, y) \Leftrightarrow F(u, v)H(u, v). \quad (4.2\text{-}31)$$

The double arrow is used to indicate that the expression on the left (spatial convolution) can be obtained by taking the *inverse* Fourier transform of the expression on the right [the product $F(u, v)H(u, v)$ in the frequency domain]. Conversely, the expression on the right can be obtained by taking the *forward* Fourier transform of the expression on the left. An analogous result is that convolution in the frequency domain reduces to multiplication in the spatial domain, and vice versa; that is,

$$f(x, y)h(x, y) \Leftrightarrow F(u, v) * H(u, v). \quad (4.2\text{-}32)$$

These two results comprise the convolution theorem. It is important to keep in mind that there is nothing complicated about what has just been stated. We already

know what the convolution operation is all about. The other part of the process, multiplication, is simply the element-by-element product of the two functions.

We need one more concept before completing the tie between the spatial and frequency domains. An *impulse function* of *strength* A, located at coordinates (x_0, y_0), is denoted by $A\delta(x - x_0, y - y_0)$ and is *defined* by the expression

$$\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} s(x, y) A\delta(x - x_0, y - y_0) = As(x_0, y_0). \tag{4.2-33}$$

In words, this equation states that the summation of a function $s(x, y)$ multiplied by an impulse is simply the value of the function at the location of the impulse, multiplied by the strength of the impulse. It is understood that the limits of the summation are the same as the limits spanned by the function. We point out that $A\delta(x - x_0, y - y_0)$ also is an image of size $M \times N$. It is composed of all zeros, except at coordinates (x_0, y_0), where the value of the image is A.

By letting either f or h in Eq. (4.2-30) be an impulse function, and using the definition in Eq. (4.2-33), we would conclude after a little manipulation that convolution of a function with an impulse "copies" the value of that function at the location of the impulse. This characteristic is called the *sifting property* of the impulse function. Of particular importance at the moment is the case of a unit impulse located at the origin, which is denoted as $\delta(x, y)$. In this case,

$$\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} s(x, y)\delta(x, y) = s(0, 0). \tag{4.2-34}$$

Armed with these simple tools, we are now in a position to establish a most interesting and useful tie between filtering in the spatial and frequency domains. From Eq. (4.2-16), we can compute the Fourier transform of a unit impulse at the origin,

$$\begin{aligned} F(u, v) &= \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \delta(x, y) e^{-j2\pi(ux/M + vy/N)} \\ &= \frac{1}{MN} \end{aligned} \tag{4.2-35}$$

where the second step follows from Eq. (4.2-34). Thus, we see that the Fourier transform of an impulse at the origin of the spatial domain is a *real* constant (this means that the phase angle is zero). If the impulse were located elsewhere, the transform would have complex components. The magnitude would be the same, with the translation of the impulse being reflected in a nonzero phase angle in the transform.

Now suppose that we let $f(x, y) = \delta(x, y)$ and carry out the convolution defined in Eq. (4.2-30). Using Eq. (4.2-34) again gives us

$$\begin{aligned} f(x, y) * h(x, y) &= \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \delta(m, n) h(x - m, y - n) \\ &= \frac{1}{MN} h(x, y) \end{aligned} \tag{4.2-36}$$

where the last step follows from Eq. (4.2-34) by noting that the variables in the summation are m and n. By combining the results of Eqs. (4.2-35) and (4.2-36) with Eq. (4.2-31), we obtain

$$\begin{aligned} f(x, y) * h(x, y) &\Leftrightarrow F(u, v)H(u, v) \\ \delta(x, y) * h(x, y) &\Leftrightarrow \Im[\delta(x, y)]H(u, v) \\ h(x, y) &\Leftrightarrow H(u, v). \end{aligned} \tag{4.2-37}$$

Using only the properties of the impulse function and the convolution theorem, we have established that filters in the spatial and frequency domains constitute a Fourier transform pair. Thus, given a filter in the frequency domain, we can obtain the corresponding filter in the spatial domain by taking the inverse Fourier transform of the former. The reverse also is true.

Note that *all* functions in the preceding development are of the same size, $M \times N$. Therefore, in practice, specifying a filter in the frequency domain and then taking the inverse transform to compute an equivalent spatial domain filter of the same size does not really help matters from a computational point of view. As discussed in Section 4.6, if both filters are of the same size, it generally is more efficient computationally to do the filtering in the frequency domain. But we use much smaller filters in the spatial domain. This is precisely the connection in which we are interested. Filtering often is more intuitive in the frequency domain. However, whenever possible, it makes more sense to filter in the spatial domain using small filter masks. Equation (4.2-37) tells us that we can specify filters in the frequency domain, take their inverse transform, and then use the resulting filter in the spatial domain as a guide for constructing smaller spatial filter masks (more formal approaches are discussed in Section 4.6.7). This is illustrated next. Keep in mind during the following discussion that the Fourier transform and its inverse are linear processes (Problem 4.2), so the discussion is by definition limited to linear filtering.

Filters based on Gaussian functions are of particular importance because their shapes are easily specified and because both the forward and inverse Fourier transforms of a Gaussian function are real Gaussian functions. We will limit the discussion here to one variable to simplify the notation. Two-dimensional functions are discussed later in this chapter.

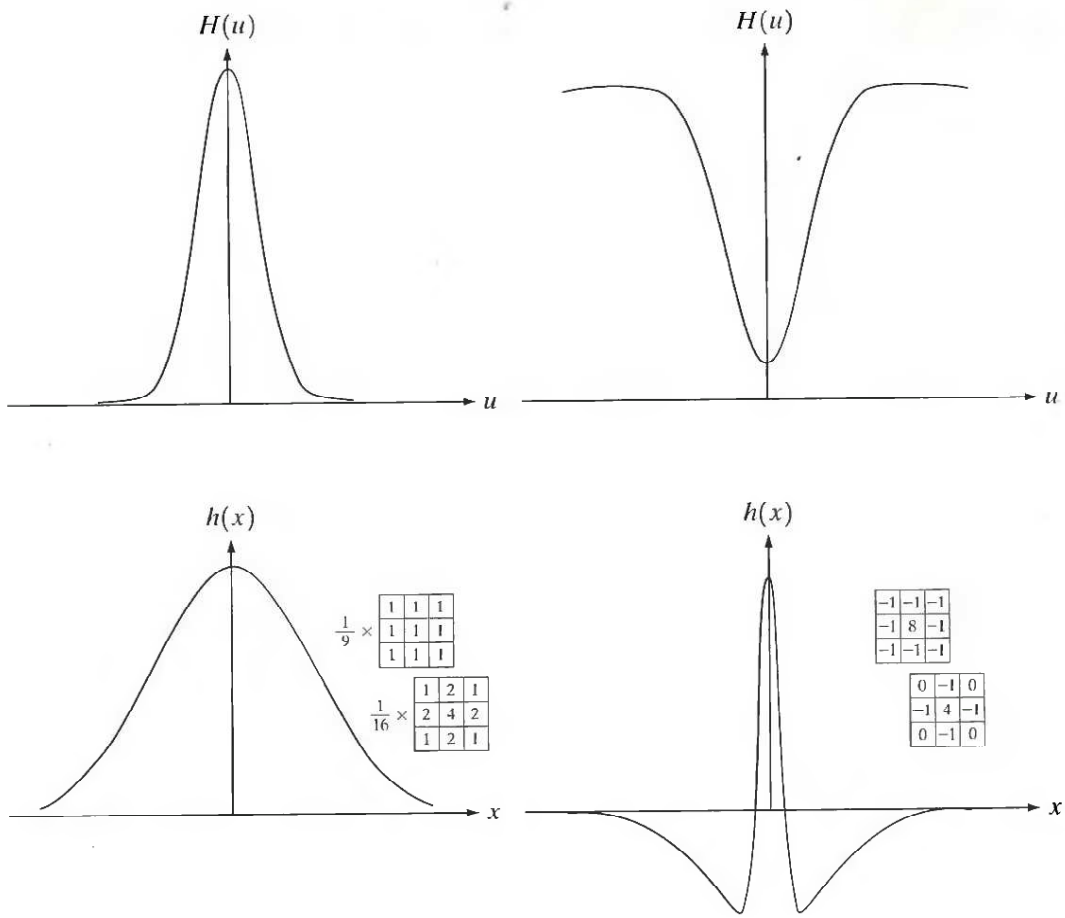
Let $H(u)$ denote a frequency domain, Gaussian filter function given by the equation

$$H(u) = Ae^{-u^2/2\sigma^2} \tag{4.2-38}$$

where σ is the standard deviation of the Gaussian curve. It can be shown (Problem 4.4) that the corresponding filter in the spatial domain is

$$h(x) = \sqrt{2\pi}\sigma A e^{-2\pi^2\sigma^2 x^2}. \tag{4.2-39}$$

These two equations represent an important result for two reasons: (1) They constitute a Fourier transform pair, both components of which are Gaussian *and* real. This facilitates analysis considerably because we do not have to be concerned with complex numbers. In addition, Gaussian curves are intuitive and easy to manipulate. (2) These functions behave reciprocally with respect to one another. In other words, when $H(u)$ has a broad profile (large value of σ), $h(x)$ has a narrow profile, and vice versa. In fact, when σ approaches infinity, $H(u)$ tends toward a constant function and $h(x)$ tends toward an impulse. This is exactly the type of reciprocal behavior we saw in Section 4.2, in connection with



a b
c d

FIGURE 4.9 (a) Gaussian frequency domain lowpass filter. (b) Gaussian frequency domain highpass filter. (c) Corresponding lowpass spatial filter. (d) Corresponding highpass spatial filter. The masks shown are used in Chapter 3 for lowpass and highpass filtering.

Figs. 4.2 and 4.3. These two properties help considerably in developing a solid understanding of the properties of filtering in both the spatial and frequency domains because they lend themselves to familiar analytical interpretations.

A plot of a Gaussian filter in the frequency domain is shown in Fig. 4.9(a). The reader will recognize this shape of $H(u)$ as a lowpass filter. The corresponding lowpass filter in the spatial domain is shown in Fig. 4.9(c). Our interest is in the general shape of $h(x)$, which we generally want to use as a guide to specify the coefficients of a smaller filter in the spatial domain. A glaring similarity between the two filters is that all the values are positive in both domains. Thus, we arrive at the conclusion that we can implement lowpass filtering in the spatial domain by using a mask with all positive coefficients, just as we did in Section 3.6.1. Two of the masks from that section are shown in Fig. 4.9(c) for reference. Another important characteristic is the reciprocal relationship discussed in the previous paragraph. The narrower the frequency domain filter, the more it will attenuate the low frequencies, resulting in increased blurring. In the spatial domain this means a wider filter, which in turn implies a larger mask, as illustrated in Example 3.9.

More complex filters can be constructed from the basic Gaussian function of Eq. (4.2-38). For instance, we can construct a highpass filter as a difference of Gaussians, as follows:

$$H(u) = Ae^{-u^2/2\sigma_1^2} - Be^{-u^2/2\sigma_2^2} \tag{4.2-40}$$

with $A \geq B$ and $\sigma_1 > \sigma_2$. The corresponding filter in the spatial domain is

$$h(x) = \sqrt{2\pi}\sigma_1 A e^{-2\pi^2\sigma_1^2 x^2} - \sqrt{2\pi}\sigma_2 B e^{-2\pi^2\sigma_2^2 x^2}. \tag{4.2-41}$$

Plots of these two functions are shown in Figs. 4.9(b) and (d), respectively. We note again the reciprocity in width, but the most important feature here is that the spatial filter has both negative and positive values. In fact, it is interesting to note that once the values turn negative, they never turn positive again. Two of the masks we used in Chapter 3 for highpass filtering are shown in Fig. 4.9(d). The similarity in form between the spatial curve and the filters is unmistakable.

In Chapter 3, we specified the shapes of lowpass and highpass filters based strictly on spatial domain considerations. It is important to note that we could have arrived at the basic shapes of all the small spatial filter masks shown in Fig. 4.9 by following the alternate path provided by the frequency domain analysis we have just completed. Although we have gone through significant effort to get here, the reader is assured that it is impossible to truly understand filtering in the frequency domain without the foundation we have just established.

A question that often arises at this point in the development of frequency domain techniques is the issue of computational complexity. Why do in the frequency domain what could be done (at least partially) in the spatial domain using small spatial masks? The basic answer is twofold. First, as we have seen, the frequency domain carries with it a significant degree of intuitiveness regarding how to specify filters. The second part of the answer depends on the size of the spatial masks and is usually answered with respect to comparable implementations.

A benchmark used frequently for this purpose is implementation of convolution in the spatial and frequency domains. Spatial convolution is given in Eq. (4.2-30), and we know from the convolution theorem that we can obtain the same result via the frequency domain by taking the inverse transform of the product of the transforms of the two functions. Suppose that we implemented both approaches in software on the same machine [using the fast Fourier transform (FFT) algorithm discussed in Section 4.6.6 for frequency domain computations]. We would find that the frequency domain implementation runs faster for surprisingly small values of M and N. For instance, a comparison by Brigham [1988] showed that, for the 1-D case, the FFT approach is faster if the number of points is greater than 32. Although this number is somewhat dependent on other factors, such as the machine and algorithms used, it certainly is well below the values that we encounter in image processing.

The frequency domain may be viewed as a "laboratory" in which we take advantage of the correspondence between frequency content and image appearance. As is demonstrated numerous times later in this chapter, some enhancement tasks that would be exceptionally difficult or impossible to formulate directly in the spatial domain become almost trivial in the frequency domain. Once we have selected a specific filter via experimentation in the frequency domain, the actual implementation of the method usually is done in the spatial domain. One approach is to specify small spatial masks that attempt to capture the "essence" of the full filter function in the spatial domain, as was explained

in Fig. 4.9. A more formal approach is to design a 2-D digital filter by using approximations based on mathematical or statistical criteria. We touch on this point again in Section 4.6.7.

4.3 Smoothing Frequency-Domain Filters

As indicated in Section 4.2.3, edges and other sharp transitions (such as noise) in the gray levels of an image contribute significantly to the high-frequency content of its Fourier transform. Hence smoothing (blurring) is achieved in the frequency domain by attenuating a specified range of high-frequency components in the transform of a given image.

Our basic "model" for filtering in the frequency domain is given by Eq. (4.2-27), which we repeat here for convenience:

$$G(u, v) = H(u, v)F(u, v) \tag{4.3-1}$$

where $F(u, v)$ is the Fourier transform of the image to be smoothed. The objective is to select a filter transfer function $H(u, v)$ that yields $G(u, v)$ by attenuating the high-frequency components of $F(u, v)$. All filtering done in this section is based on the procedure outlined in Section 4.2.3, including the use of zero-phase-shift filters.

We consider three types of lowpass filters: ideal, Butterworth, and Gaussian filters. These three filters cover the range from very sharp (ideal) to very smooth (Gaussian) filter functions. The Butterworth filter has a parameter, called the filter *order*. For high values of this parameter the Butterworth filter approaches the form of the ideal filter. For lower-order values, the Butterworth filter has a smooth form similar to the Gaussian filter. Thus, the Butterworth filter may be viewed as a transition between two "extremes."

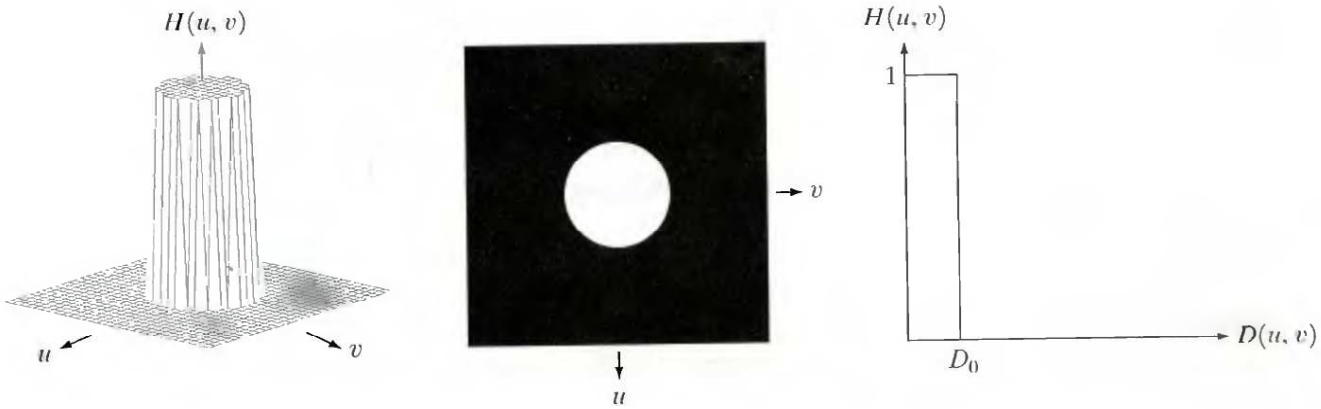
4.3.1 Ideal Lowpass Filters

The simplest lowpass filter we can envision is a filter that "cuts off" all high-frequency components of the Fourier transform that are at a distance greater than a specified distance D_0 from the origin of the (centered) transform. Such a filter is called a two-dimensional (2-D) *ideal lowpass filter* (ILPF) and has the transfer function

$$H(u, v) = \begin{cases} 1 & \text{if } D(u, v) \le D_0 \\ 0 & \text{if } D(u, v) > D_0 \end{cases} \tag{4.3-2}$$

where D_0 is a specified nonnegative quantity, and $D(u, v)$ is the distance from point (u, v) to the center of the frequency rectangle. If the image in question is of size $M \times N$, we know that its transform also is of this size, so the center of the frequency rectangle is at $(u, v) = (M/2, N/2)$ due to the fact that the transform has been centered, as discussed in connection with Eq. (4.2-21). In this case, the distance from any point (u, v) to the center (origin) of the Fourier transform is given by

$$D(u, v) = \left[(u - M/2)^2 + (v - N/2)^2\right]^{1/2}. \tag{4.3-3}$$



a b c

FIGURE 4.10 (a) Perspective plot of an ideal lowpass filter transfer function. (b) Filter displayed as an image. (c) Filter radial cross section.

Figure 4.10(a) shows a 3-D perspective plot of $H(u, v)$ as a function of u and v, and Fig. 4.10(b) shows $H(u, v)$ displayed as an image. The name *ideal filter* indicates that *all* frequencies inside a circle of radius D_0 are passed with no attenuation, whereas *all* frequencies outside this circle are completely attenuated. The lowpass filters considered in this chapter are radially symmetric about the origin. This means that a cross section extending as a function of distance from the origin along a radial line is sufficient to specify the filter, as Fig. 4.10(c) shows. The complete filter transfer function can be visualized by rotating the cross section 360° about the origin.

For an ideal lowpass filter cross section, the point of transition between $H(u, v) = 1$ and $H(u, v) = 0$ is called the *cutoff frequency*. In the case of Fig. 4.10, for example, the cutoff frequency is D_0. The sharp cutoff frequencies of an ideal lowpass filter cannot be realized with electronic components, although they can certainly be implemented in a computer. The effects of using these "nonphysical" filters on a digital image are discussed later in this section.

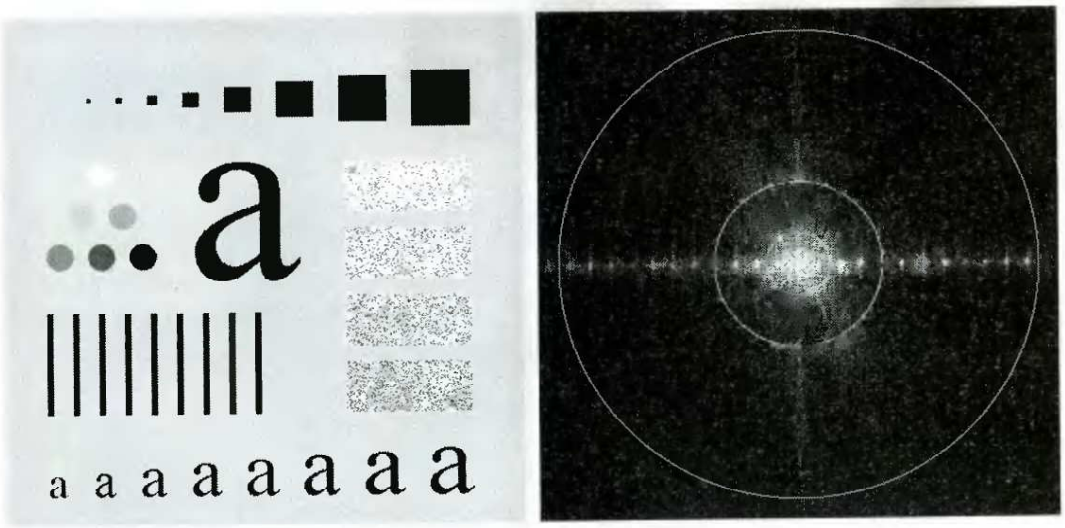
The lowpass filters introduced in this section are compared by studying their behavior as a function of the same cutoff frequencies. One way to establish a set of standard cutoff frequency loci is to compute circles that enclose specified amounts of total image power P_T. This quantity is obtained by summing the components of the power spectrum at each point (u, v), for $u = 0, 1, 2, \ldots, M - 1$ and $v = 0, 1, 2, \ldots, N - 1$; that is,

$$P_T = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} P(u, v) \tag{4.3-4}$$

where $P(u, v)$ is given in Eq. (4.2-20). If the transform has been centered, a circle of radius r with origin at the center of the frequency rectangle encloses α percent of the power, where

$$\alpha = 100\left[\sum_u \sum_v P(u, v)/P_T\right] \tag{4.3-5}$$

and the summation is taken over the values of (u, v) that lie inside the circle or on its boundary.



a b

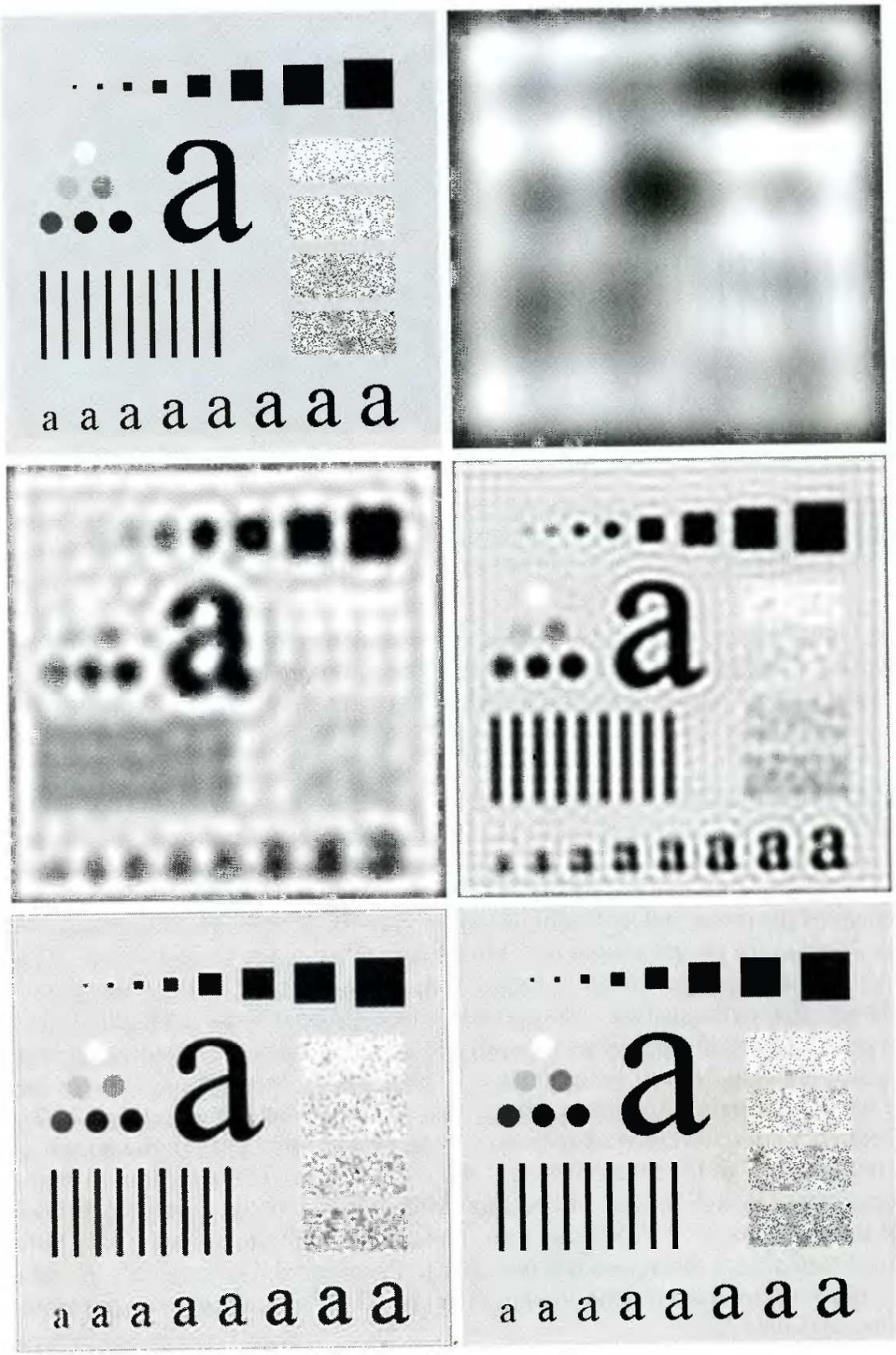
FIGURE 4.11 (a) An image of size 500 × 500 pixels and (b) its Fourier spectrum. The superimposed circles have radii values of 5, 15, 30, 80, and 230, which enclose 92.0, 94.6, 96.4, 98.0, and 99.5% of the image power, respectively.

EXAMPLE 4.4: Image power as a function of distance from the origin of the DFT.

Figure 4.11(a) shows the test pattern we used in Fig. 3.35 to illustrate spatial blurring. The Fourier spectrum of this image is shown in Fig. 4.11(b). The circles superimposed on the spectrum have radii of 5, 15, 30, 80, and 230 pixels (the circle of radius 5 is not easily visible). These circles enclose α percent of the image power, for $\alpha = 92.0, 94.6, 96.4, 98$, and 99.5%, respectively. The spectrum falls off rapidly, with 92% of the total power being enclosed by a relatively small circle of radius 5.

Figure 4.12 shows the results of applying ideal lowpass filters with cutoff frequencies at the radii shown in Fig. 4.11(b). Figure 4.12(b) is useless for all practical purposes, unless the objective of blurring in this case is to eliminate all detail in the image, except the "blobs" representing the largest objects. The severe blurring in this image is a clear indication that most of the sharp detail information in the picture is contained in the 8% power removed by the filter. As the filter radius increases, less and less power is removed, resulting in less severe blurring. Note that the images in Figs. 4.12(c) through (e) are characterized by "ringing," which becomes finer in texture as the amount of high-frequency content removed decreases. Ringing is evident even in the image in which only 2% of the total power was removed. This ringing behavior is a characteristic of ideal filters, as will be explained shortly. Finally, close observation of the result for $\alpha = 99.5$ shows very slight blurring in the noisy squares but, for the most part, this image is quite close to the original. This indicates that little edge information is contained in the upper 0.5% of the spectrum power in this particular case.

It is clear from this example that ideal lowpass filtering is not very practical. However, because ideal filters can be implemented in a computer, it is useful to study their behavior as part of our development of filtering concepts. Also, as shown in the discussion that follows, some interesting insight is gained by attempting to explain the ringing property of ILPFs in the spatial domain.



a b
c d
e f

FIGURE 4.12 (a) Original image. (b)–(f) Results of ideal lowpass filtering with cutoff frequencies set at radii values of 5, 15, 30, 80, and 230, as shown in Fig. 4.11(b). The power removed by these filters was 8, 5.4, 3.6, 2, and 0.5% of the total, respectively.

The blurring and ringing properties of the ILPF can be explained by reference to the convolution theorem discussed in Section 4.2.4. The Fourier transforms of the original image $f(x, y)$ and the blurred image $g(x, y)$ are related in the frequency domain by the equation

$$G(u, v) = H(u, v)F(u, v)$$

where, as before, $H(u, v)$ is the filter function and F and G are the Fourier transforms of the two images just mentioned. The convolution theorem tells us that the corresponding process in the spatial domain is

$$g(x, y) = h(x, y) * f(x, y)$$

where $h(x, y)$ is the inverse Fourier transform of the filter transfer function $H(u, v)$.

The key to understanding blurring as a convolution process in the spatial domain lies in the nature of $h(x, y)$. For instance, the ILPF of radius 5 that caused so much blurring in the preceding example is shown in Fig. 4.13(a). This is the function $H(u, v)$ in the *frequency* domain. The spatial filter function $h(x, y)$ was obtained in the standard way: (1) $H(u, v)$ was multiplied by $(-1)^{u+v}$ for centering; (2) this was followed by the inverse DFT; and (3) the real part of the inverse DFT was multiplied by $(-1)^{x+y}$. Figure 4.13(b) shows the result of this process.

We see that the filter $h(x, y)$ has two major distinctive characteristics: a dominant component at the origin, and concentric, circular components about the center component. The center component is primarily responsible for blurring. The concentric components are responsible primarily for the ringing characteristic of ideal filters. Both the radius of the center component and the number of circles per unit distance from the origin are inversely proportional to the value of the cutoff frequency of the ideal filter. The insert at the top is a gray-level profile of a horizontal scan line through the center of the spatial filter. The axis shown indicates zero amplitude, so we see that the spatial filter has negative values. This normally is not a serious problem because the larger center component dominates the convolution result. However, the filtered image can have negative values, so scaling normally is required.

Suppose next that $f(x, y)$ is a simple image composed of five bright pixels on a black background, as Fig. 4.13(c) shows. These bright points may be viewed as approximations to impulses, whose strength depends on the intensity of the points. Then the convolution of $h(x, y)$ and $f(x, y)$ is simply a process of "copying" $h(x, y)$ at the location of each impulse, as noted in Section 4.2.4. The result of this operation, shown in Fig. 4.13(d), explains how the original points are blurred as a consequence of convolving $f(x, y)$ with the blurring filter function $h(x, y)$. Note also that ringing was introduced during the same process. In fact, the ringing is so severe in this case that distortion is caused by their interference with one another. These concepts are extended conceptually to more complex images by considering each pixel as an impulse whose strength is proportional to the gray level of the pixel. The insert at the bottom of Fig. 4.13 shows the gray-level profile of a diagonal scan line through the center of the filtered image.

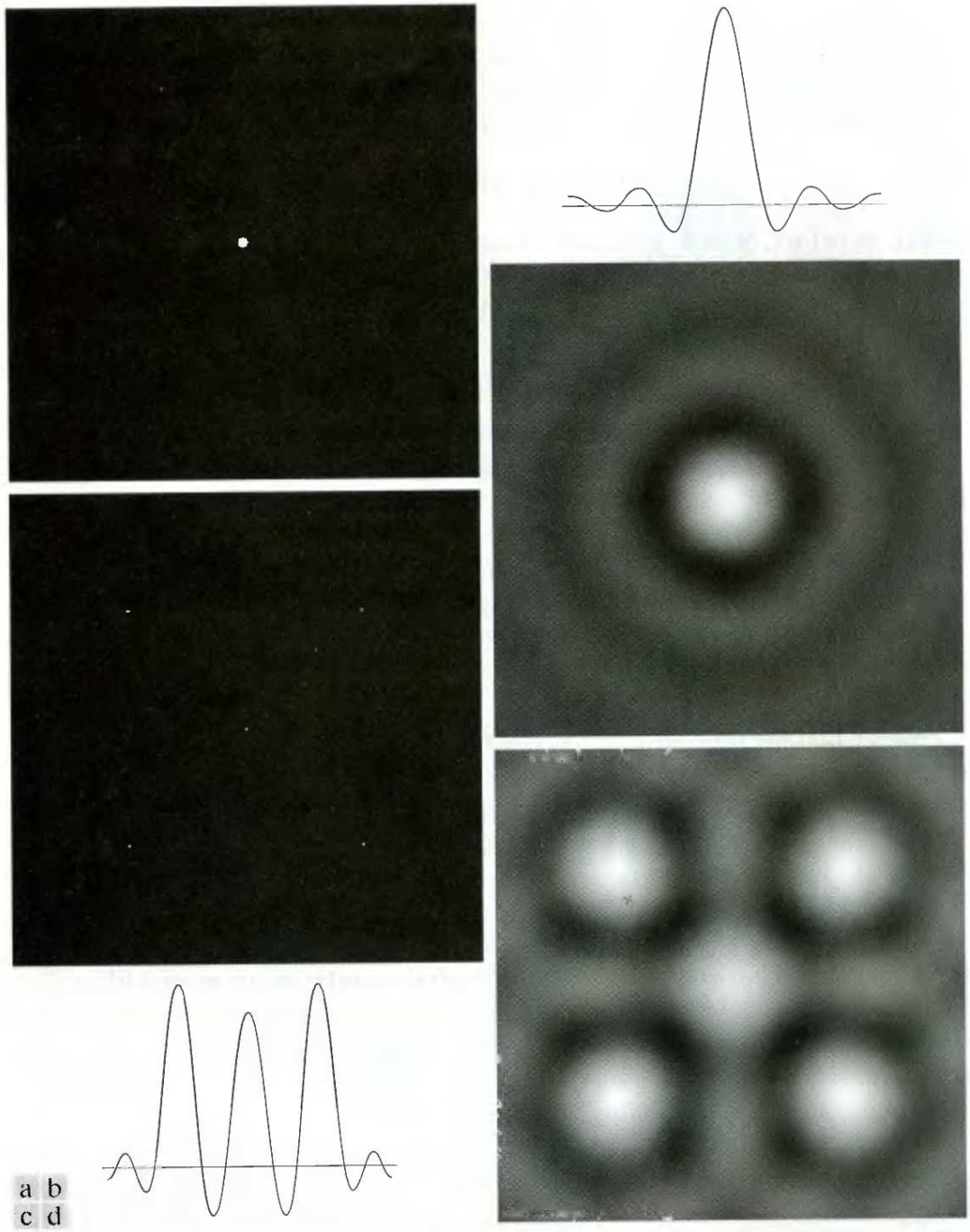
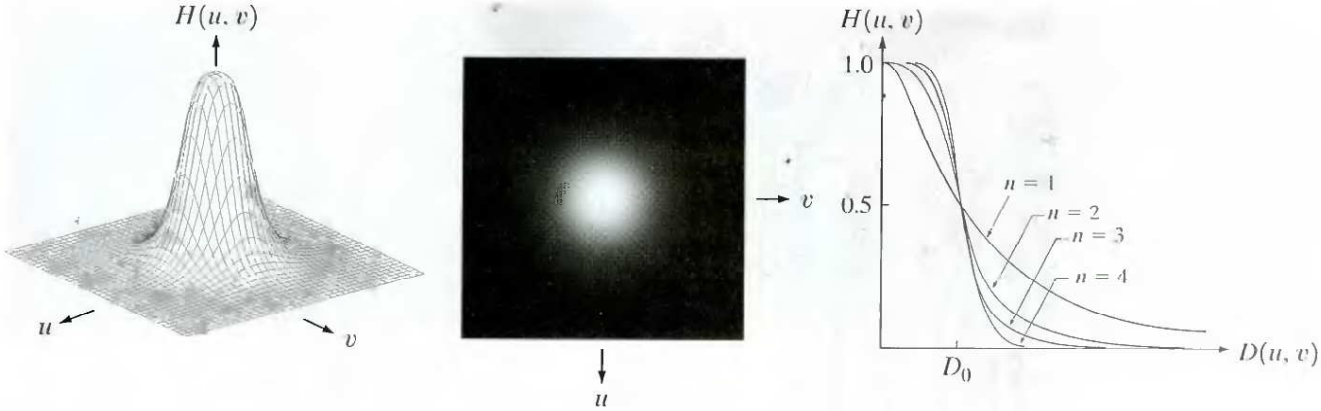


FIGURE 4.13 (a) A frequency-domain ILPF of radius 5. (b) Corresponding spatial filter (note the ringing). (c) Five impulses in the spatial domain, simulating the values of five pixels. (d) Convolution of (b) and (c) in the spatial domain.

The reciprocal nature between $H(u, v)$ and $h(x, y)$, along with the convolution process just discussed, explains mathematically why the blurring and ringing are more severe the narrower the filter in the frequency domain. This type of reciprocal behavior should be routine to the reader by now. In the next two sections we show that it is possible to achieve blurring with little or no ringing, which is our main objective.



a b c

FIGURE 4.14 (a) Perspective plot of a Butterworth lowpass filter transfer function. (b) Filter displayed as an image. (c) Filter radial cross sections of orders 1 through 4.

4.3.2 Butterworth Lowpass Filters

The transfer function of a Butterworth lowpass filter (BLPF) of order n, and with cutoff frequency at a distance D_0 from the origin, is defined as

$$H(u, v) = \frac{1}{1 + \left[D(u, v)/D_0\right]^{2n}} \tag{4.3-6}$$

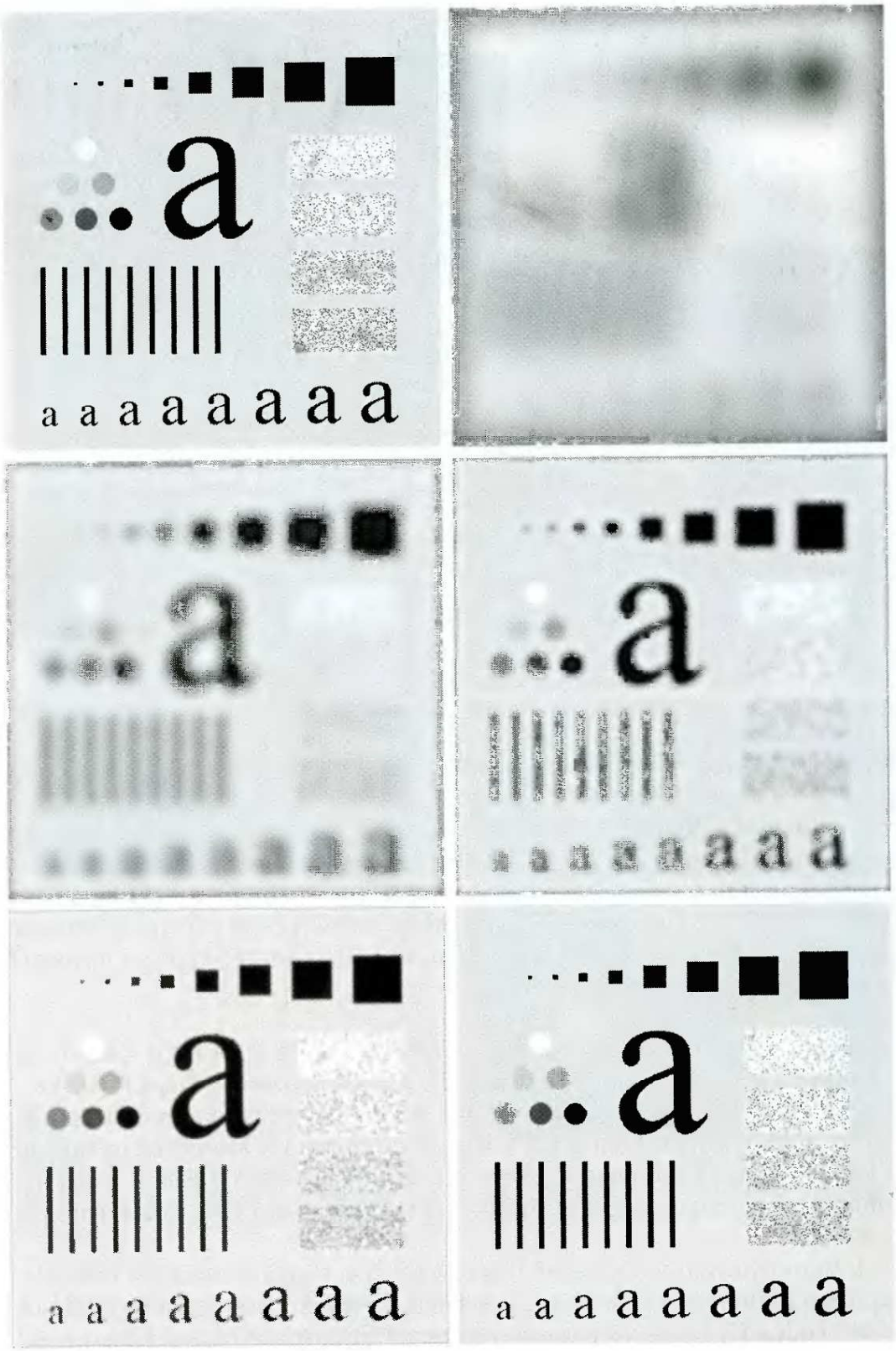
where $D(u, v)$ is given by Eq. (4.3-3). A perspective plot, image display, and radial cross sections of the BLPF function are shown in Fig. 4.14.

Unlike the ILPF, the BLPF transfer function does not have a sharp discontinuity that establishes a clear cutoff between passed and filtered frequencies. For filters with smooth transfer functions, defining a cutoff frequency locus at points for which $H(u, v)$ is down to a certain fraction of its maximum value is customary. In the case of Eq. (4.3-6), $H(u, v) = 0.5$ (down 50% from its maximum value of 1) when $D(u, v) = D_0$.

EXAMPLE 4.5: Butterworth lowpass filtering.

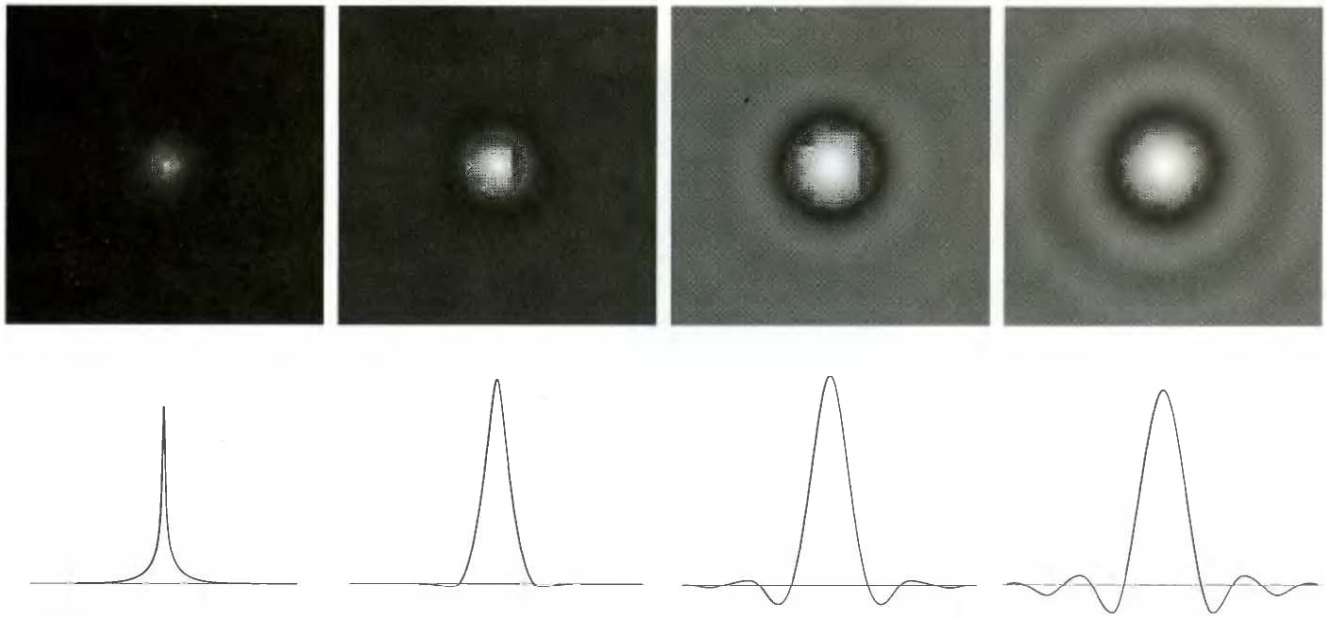
Figure 4.15 shows the results of applying the BLPF of Eq. (4.3-6) to Fig. 4.15(a), with $n = 2$ and D_0 equal to the five radii shown in Fig. 4.11(b). Unlike the results shown in Fig. 4.12 for the ILPF, we note here a smooth transition in blurring as a function of increasing cutoff frequency. Moreover, no ringing is visible in any of the images processed with this particular BLPF, a fact attributed to the filter's smooth transition between low and high frequencies.

A Butterworth filter of order 1 has no ringing. Ringing generally is imperceptible in filters of order 2, but can become a significant factor in filters of higher order. Figure 4.16 shows an interesting comparison between the *spatial* representation of BLPFs of various orders (with cutoff frequency of 5 pixels). Shown also is the gray-level profile along a horizontal scan line through the center of each filter. These filters were obtained and displayed by using the same procedure we used to generate Fig. 4.13(b). In order to facilitate comparisons,



a b
c d
e f

FIGURE 4.15 (a) Original image. (b)–(f) Results of filtering with BLPFs of order 2, with cutoff frequencies at radii of 5, 15, 30, 80, and 230, as shown in Fig. 4.11(b). Compare with Fig. 4.12.



a b c d

FIGURE 4.16 (a)–(d) Spatial representation of BLPFs of order 1, 2, 5, and 20, and corresponding gray-level profiles through the center of the filters (all filters have a cutoff frequency of 5). Note that ringing increases as a function of filter order.

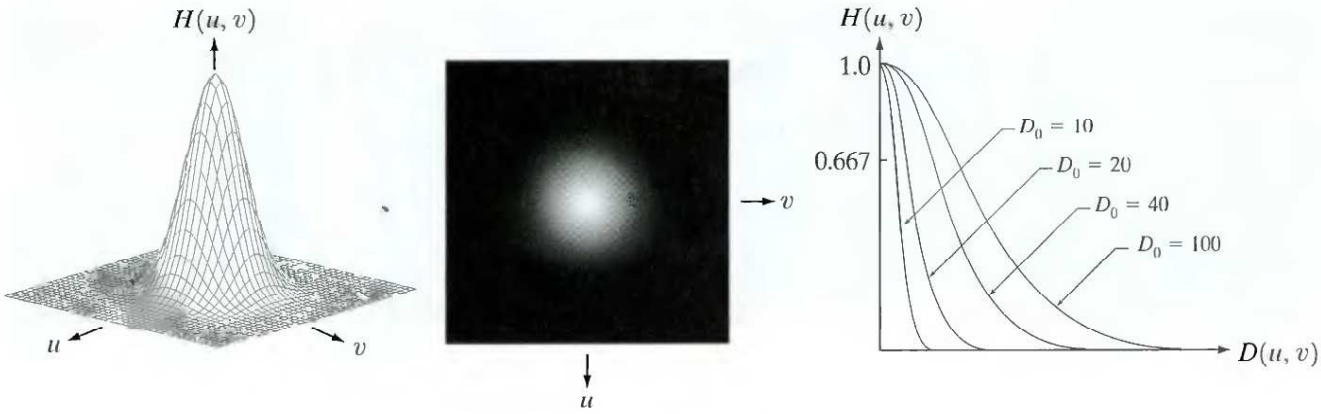
additional enhancing with a gamma transformation [see Eq. (3.2-3)] was applied to the images of Fig. 4.16 to accentuate even more the components further away from the origin. The BLPF of order 1 [Fig. 4.16(a)] has neither ringing nor negative values. The filter of order 2 does show mild ringing and small negative values, but they certainly are less pronounced than in the ILPF. As the remaining images show, ringing in the BLPF becomes significant for higher-order filters. A Butterworth filter of order 20 already exhibits the characteristics of the ILPF, which can be seen by comparing Figs. 4.16(d) and 4.13(b). In the limit, both filters are identical. In general, BLPFs of order 2 are a good compromise between effective lowpass filtering and acceptable ringing characteristics.

4.3.3 Gaussian Lowpass Filters

Gaussian lowpass filters (GLPFs) of one dimension were introduced in Section 4.2.4 as an aid in exploring some important relationships between the spatial and frequency domains. The form of these filters in two dimensions is given by

$$H(u, v) = e^{-D^2(u,v)/2\sigma^2} \tag{4.3-7}$$

where, as in Eq. (4.3-3), $D(u, v)$ is the distance from the origin of the Fourier transform, which we assume has been shifted to the center of the frequency rectangle using the procedure outlined in Section 4.2.3. We did not use a constant in front of the filter as in Section 4.2.4 to be consistent with all the other filters discussed in the present section, which have a value of 1 at the origin. As before,



a b c

FIGURE 4.17 (a) Perspective plot of a GLPF transfer function. (b) Filter displayed as an image. (c) Filter radial cross sections for various values of D_0.

σ is a measure of the spread of the Gaussian curve. By letting $\sigma = D_0$, we can express the filter in a more familiar form in terms of the notation in this section:

$$H(u, v) = e^{-D^2(u,v)/2D_0^2} \tag{4.3-8}$$

where D_0 is the cutoff frequency. When $D(u, v) = D_0$, the filter is down to 0.607 of its maximum value.

As discussed in Section 4.2.4, the inverse Fourier transform of the Gaussian lowpass filter also is Gaussian. We already saw in that section the advantages of this property as an analysis tool. In terms of our current interest, this also means that a *spatial* Gaussian filter, obtained by computing the inverse Fourier transform of Eq. (4.3-7) or (4.3-8), will have no ringing. A perspective plot, image display, and radial cross sections of a GLPF function are shown in Fig. 4.17.

EXAMPLE 4.6: Gaussian lowpass filtering.

■ Figure 4.18 shows the results of applying the GLPF of Eq. (4.3-8) to Fig. 4.18(a), with D_0 equal to the five radii shown in Fig. 4.11(b). As in the case of the BLPF (Fig. 4.15), we note a smooth transition in blurring as a function of increasing cutoff frequency. The GLPF did not achieve as much smoothing as the BLPF of order 2 for the same value of cutoff frequency, as can be seen, for example, by comparing Figs. 4.15(c) and 4.18(c). This is expected, because the profile of the GLPF is not as "tight" as the profile of the BLPF of order 2. However, the results are quite comparable in general, and we are assured of no ringing in the case of the GLPF. This is an important characteristic in practice, especially in situations where any type of artifact (e.g., in medical imaging) is not acceptable. In cases where tight control of the transition between low and high frequencies about the cutoff frequency are needed, then the BLPF presents a more suitable choice. The price of this additional control over the filter profile is the possibility of ringing. ■

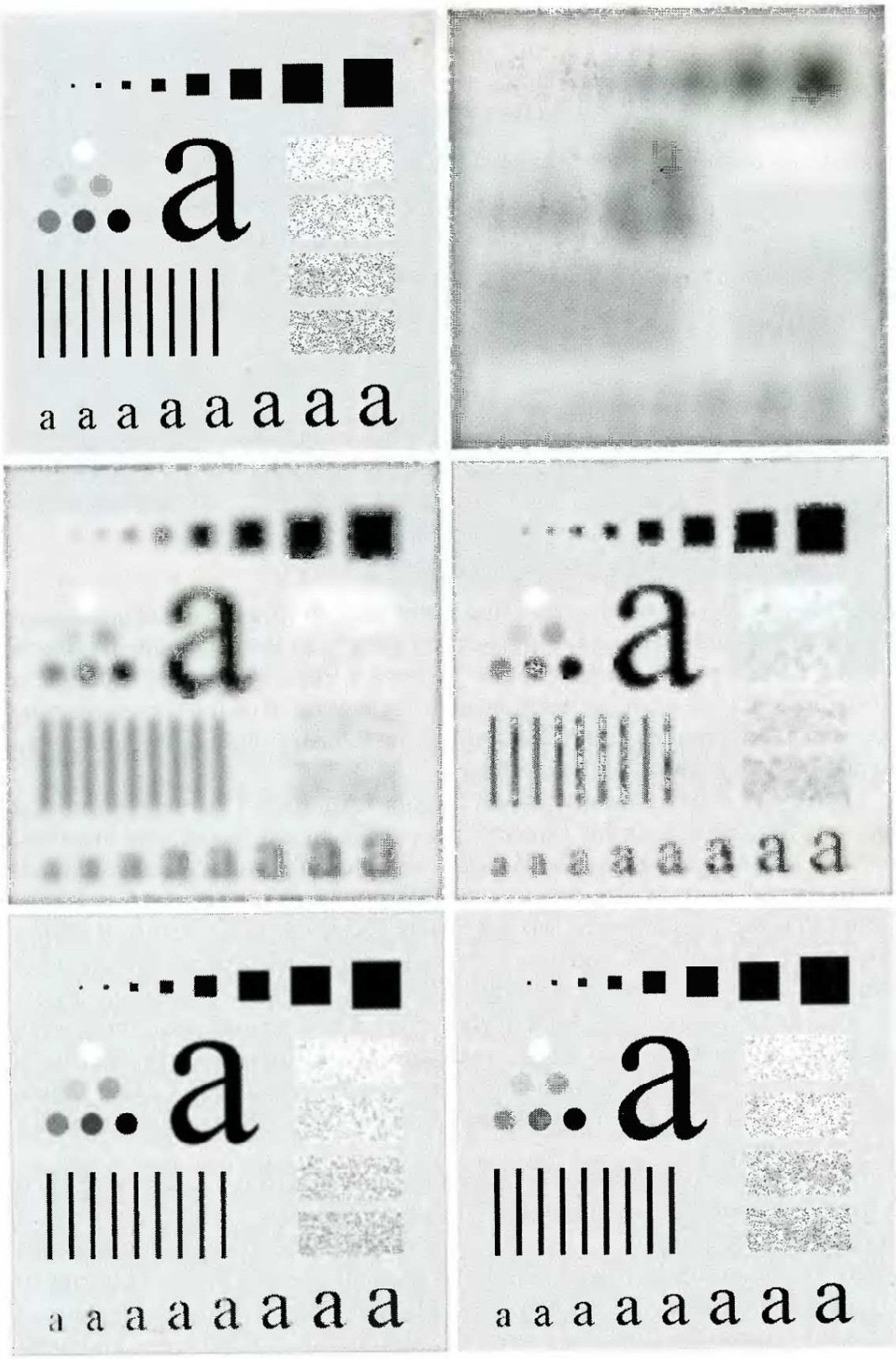
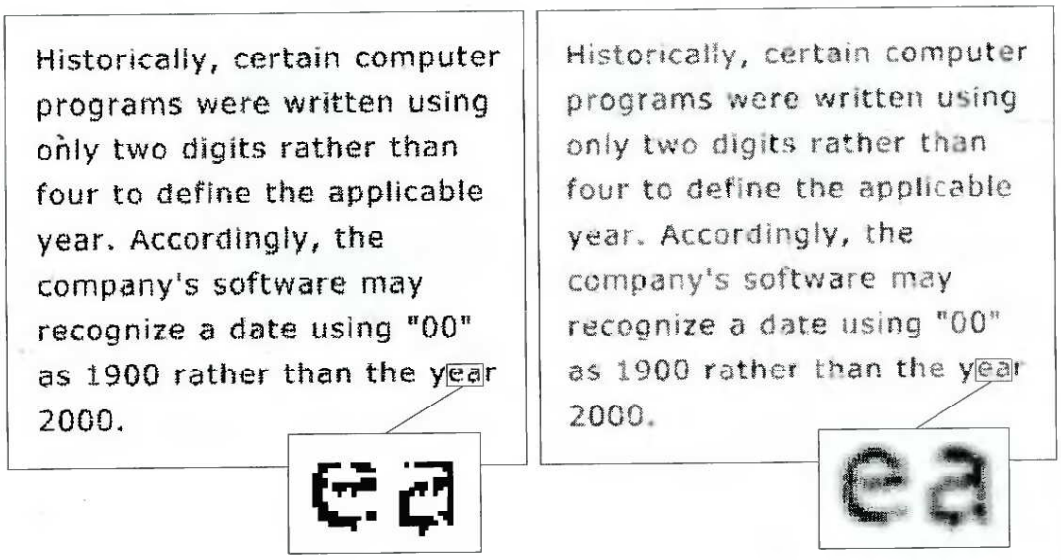


FIGURE 4.18 (a) Original image. (b)–(f) Results of filtering with Gaussian lowpass filters with cutoff frequencies set at radii values of 5, 15, 30, 80, and 230, as shown in Fig. 4.11(b). Compare with Figs. 4.12 and 4.15.

a b
c d
e f



a b

FIGURE 4.19 (a) Sample text of poor resolution (note broken characters in magnified view). (b) Result of filtering with a GLPF (broken character segments were joined).

4.3.4 Additional Examples of Lowpass Filtering

The lowpass filtering results given thus far have been with images of good quality in order to illustrate and compare filter effects. In the following discussion we show a few practical applications of lowpass filtering. The first example is from the field of machine perception, with application to character recognition; the second is from the printing and publishing industry; and the third is related to processing satellite and aerial images.

Figure 4.19 shows a sample of text of poor resolution. One encounters text like this, for example, in fax transmissions, duplicated material, and historical records. As poor text goes, this particular sample is free of additional difficulties like smudges, creases, and torn sections. The magnified section in Fig. 4.19(a) shows that the characters in this particular document have distorted shapes due to lack of resolution, and many of the characters are broken. Although humans fill these gaps visually without difficulty, a machine recognition system has real difficulties reading broken characters. The approach used most often to handle this problem is to bridge small gaps in the input image by blurring it. Figure 4.19(b) shows how well characters can be "repaired" by this simple process using a Gaussian lowpass filter with $D_0 = 80$. The images are of size 444×508 pixels.

Lowpass filtering is a staple in the printing and publishing industry, where it is used for numerous preprocessing functions, including unsharp masking, as discussed in Section 3.7.2. "Cosmetic" processing is another use of lowpass filtering prior to printing. Figure 4.20 shows an application of lowpass filtering to produce a smoother, softer-looking result from a sharp original. For human faces, the typical objective is to reduce the sharpness of fine skin lines and small blemishes. The magnified sections in Figs. 4.20(b) and (c) clearly show a significant reduction in fine skin lines around the eyes in this particular case. In fact, the smoothed images look quite soft and pleasing.

a b c

FIGURE 4.20 (a) Original image (1028 × 732 pixels). (b) Result of filtering with a GLPF with $D_0 = 100$. (c) Result of filtering with a GLPF with $D_0 = 80$. Note reduction in skin fine lines in the magnified sections of (b) and (c).

Figure 4.21 shows two applications of lowpass filtering on the same image, but with totally different objectives. Figure 4.21(a) is a 588 × 600 very high resolution radiometer (VHRR) image showing part of the Gulf of Mexico (dark) and Florida (light), taken from a NOAA satellite (note horizontal sensor scan lines). The boundaries between water bodies were caused by loop currents. This image is illustrative of remotely sensed images in which sensors (for a number of reasons beyond the present discussion) have the tendency to produce pronounced scan lines along the direction in which the scene is being scanned. Lowpass filtering is a crude but simple way to reduce the effect of these lines, as Fig. 4.21(b) shows (we consider more effective approaches in Chapter 5). This image was obtained using a Gaussian lowpass filter with $D_0 = 30$. The resulting reduction in the effect of the scan lines can simplify the detection of features like the interface boundaries between ocean currents. Figure 4.21(c) shows the result of considerably more aggressive Gaussian lowpass filtering $(D_0 = 10)$. Here the objective is to blur out as much detail as possible while leaving large features recognizable. For instance, this type of filtering would be part of a preprocessing stage for an image analysis system that was

a b c

FIGURE 4.21 (a) Image showing prominent scan lines. (b) Result of using a GLPF with $D_0 = 30$. (c) Result of using a GLPF with $D_0 = 10$. (Original image courtesy of NOAA.)

looking for features in an image bank. An example of such features could be lakes of a given size [such as Lake Okeechobee in the lower eastern region of Florida, shown as a nearly round dark region in Fig. 4.21(c)]. Lowpass filtering helps simplify the analysis by averaging out features smaller than the ones of interest.

4.4 Sharpening Frequency Domain Filters

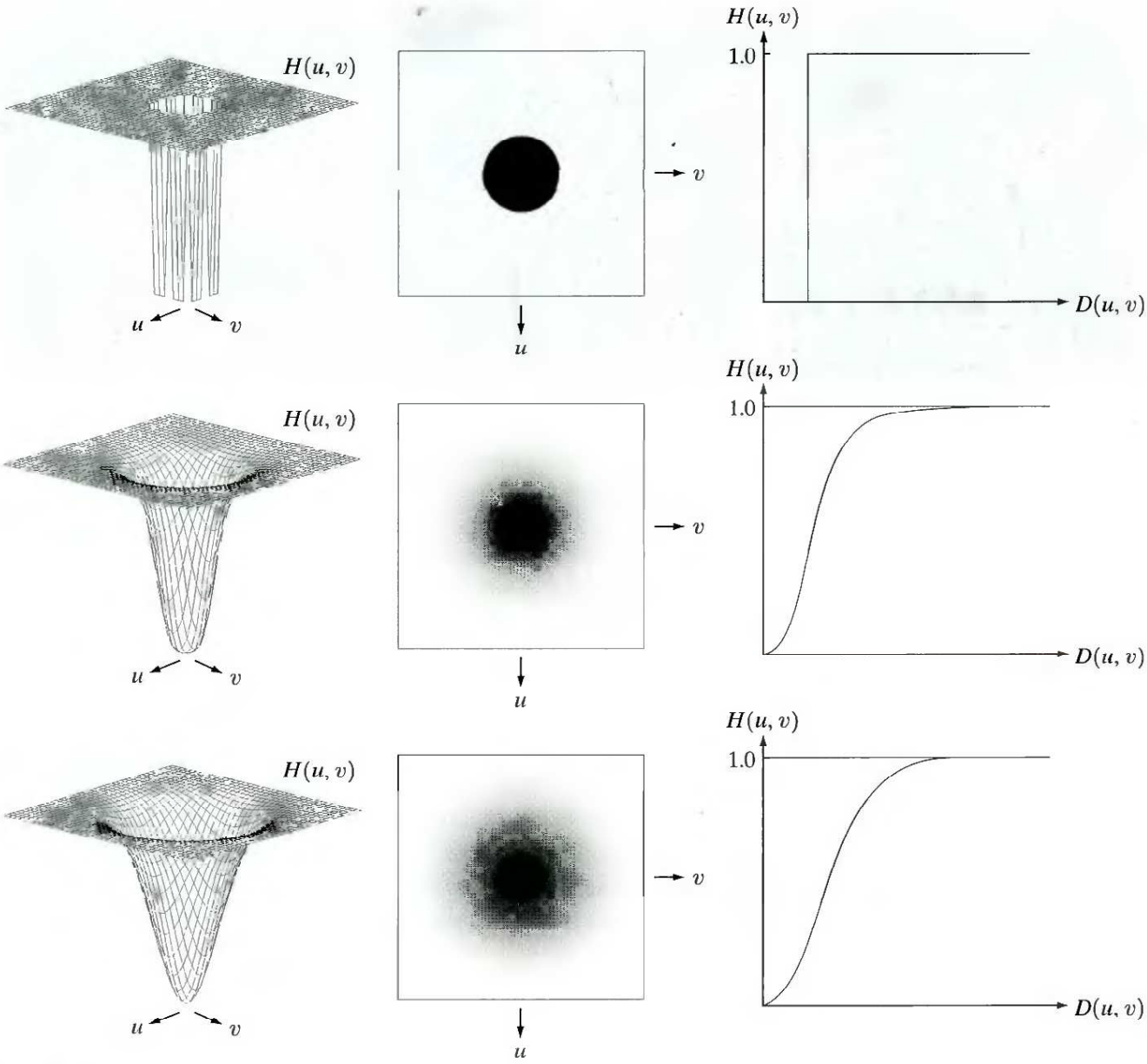
In the previous section we showed that an image can be blurred by attenuating the high-frequency components of its Fourier transform. Because edges and other abrupt changes in gray levels are associated with high-frequency components, image sharpening can be achieved in the frequency domain by a *highpass filtering* process, which attenuates the low-frequency components without disturbing high-frequency information in the Fourier transform. As in Section 4.3, we consider only zero-phase-shift filters that are radially symmetric. All filtering in this section is based on the procedure outlined in Section 4.2.3.

Because the intended function of the filters in this section is to perform precisely the reverse operation of the ideal lowpass filters discussed in the previous section, the transfer function of the highpass filters discussed in this section can be obtained using the relation

$$H_{hp}(u, v) = 1 - H_{lp}(u, v) \tag{4.4-1}$$

where $H_{lp}(u, v)$ is the transfer function of the corresponding lowpass filter. That is, when the lowpass filter attenuates frequencies, the highpass filter passes them, and vice versa.

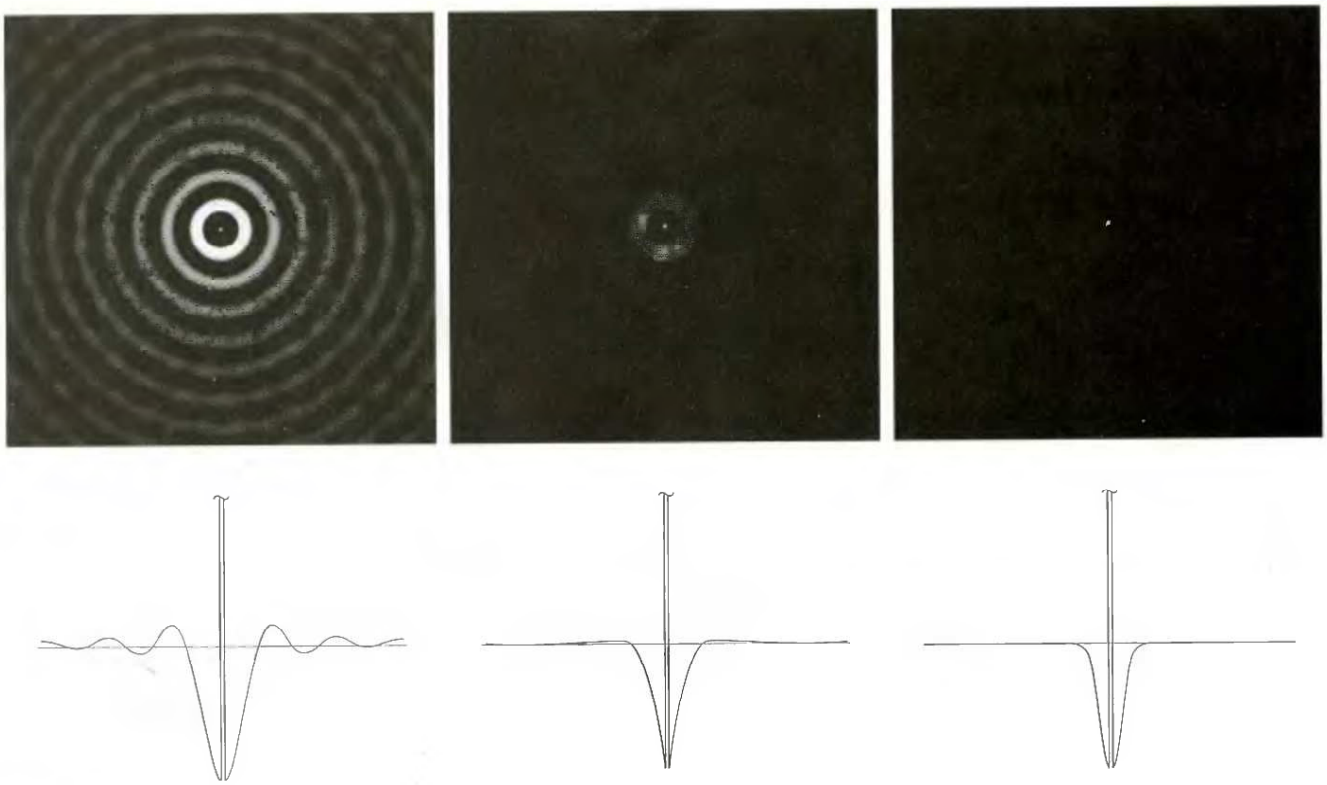
In this section we consider ideal, Butterworth, and Gaussian highpass filters. As in the previous section, we illustrate the characteristics of these filters in both the frequency and spatial domains. Figure 4.22 shows typical 3-D plots, image representations, and cross sections for these filters. As before, we see that



a b c
d e f
g h i

FIGURE 4.22 Top row: Perspective plot, image representation, and cross section of a typical ideal highpass filter. Middle and bottom rows: The same sequence for typical Butterworth and Gaussian highpass filters.

the Butterworth filter represents a transition between the sharpness of the ideal filter and the total smoothness of the Gaussian filter. Figure 4.23 illustrates what these filters look like in the spatial domain. Recall that a spatial representation of a frequency domain filter is obtained by (1) multiplying $H(u, v)$ by $(-1)^{u+v}$ for centering; (2) computing the inverse DFT; and (3) multiplying the real part of the inverse DFT by $(-1)^{x+y}$. Important characteristics of these figures are discussed in the sections that follow.



a b c

FIGURE 4.23 Spatial representations of typical (a) ideal, (b) Butterworth, and (c) Gaussian frequency domain highpass filters, and corresponding gray-level profiles.

4.4.1 Ideal Highpass Filters

A 2-D ideal highpass filter (IHPF) is defined as

$$H(u, v) = \begin{cases} 0 & \text{if } D(u, v) \leq D_0 \\ 1 & \text{if } D(u, v) > D_0 \end{cases} \tag{4.4-2}$$

where D_0 is the cutoff distance measured from the origin of the frequency rectangle, and $D(u, v)$ is given in Eq. (4.3-3). This expression follows directly from Eqs. (4.4-1) and (4.4-2). As intended, this filter is the opposite of the ideal lowpass filter in the sense that it sets to zero all frequencies inside a circle of radius D_0 while passing, without attenuation, all frequencies outside the circle. As in the case of the ideal lowpass filter, the IHPF is not physically realizable with electronic components. However, since it can be implemented in a computer, we consider it for completeness. The discussion will be brief.

Because of the way they are related [Eq. (4.4-1)], we can expect IHPFs to have the same ringing properties as ILPFs [see Fig. 4.23(a)]. This is demonstrated clearly in Fig. 4.24, which consists of various IHPF results using the original image in Fig. 4.11(a) with D_0 set to 15, 30, and 80 pixels, respectively. The ringing in Fig. 4.24(a) is so severe that it produced distorted, thickened object boundaries (e.g., look at the large letter "a"). Edges of the top three circles do not show well because they are not as strong as the other edges in the image (the gray level of