

**FIGURE 10.37** (a) Noisy image. (b) Intensity ramp in the range  $[0.2, 0.6]$ . (c) Product of (a) and (b). (d)–(f) Corresponding histograms.

where separation of the modes without additional processing (see Sections 10.3.4 and 10.3.5) is no longer possible. Similar results would be obtained if the illumination was perfectly uniform, but the reflectance of the image was not, due, for example, to natural reflectivity variations in the surface of objects and/or background.

The key point in the preceding paragraph is that illumination and reflectance play a central role in the success of image segmentation using thresholding or other segmentation techniques. Therefore, controlling these factors when it is possible to do so should be the first step considered in the solution of a segmentation problem. There are three basic approaches to the problem when control over these factors is not possible. One is to correct the shading pattern directly. For example, nonuniform (but fixed) illumination can be corrected by multiplying the image by the inverse of the pattern, which can be obtained by imaging a flat surface of constant intensity. The second approach is to attempt to correct the global shading pattern via processing using, for example, the top-hat transformation introduced in Section 9.6.3. The third approach is to “work around” nonuniformities using variable thresholding, as discussed in Section 10.3.7.

### 10.3.2 Basic Global Thresholding

As noted in the previous section, when the intensity distributions of objects and background pixels are sufficiently distinct, it is possible to use a single (global) threshold applicable over the entire image. In most applications, there

In theory, the histogram of a ramp image is uniform. In practice, achieving perfect uniformity depends on the size of the image and number of intensity bits. For example, a  $256 \times 256$ , 256-level ramp image has a uniform histogram, but a  $256 \times 257$  ramp image with the same number of intensities does not.



is usually enough variability between images that, even if global thresholding is a suitable approach, an algorithm capable of estimating automatically the threshold value for each image is required. The following iterative algorithm can be used for this purpose:

1. Select an initial estimate for the global threshold,  $T$ .
2. Segment the image using  $T$  in Eq. (10.3-1). This will produce two groups of pixels:  $G_1$  consisting of all pixels with intensity values  $> T$ , and  $G_2$  consisting of pixels with values  $\leq T$ .
3. Compute the average (mean) intensity values  $m_1$  and  $m_2$  for the pixels in  $G_1$  and  $G_2$ , respectively.
4. Compute a new threshold value:

$$T = \frac{1}{2}(m_1 + m_2)$$

5. Repeat Steps 2 through 4 until the difference between values of  $T$  in successive iterations is smaller than a predefined parameter  $\Delta T$ .

This simple algorithm works well in situations where there is a reasonably clear valley between the modes of the histogram related to objects and background. Parameter  $\Delta T$  is used to control the number of iterations in situations where speed is an important issue. In general, the larger  $\Delta T$  is, the fewer iterations the algorithm will perform. The initial threshold must be chosen greater than the minimum and less than maximum intensity level in the image (Problem 10.28). The average intensity of the image is a good initial choice for  $T$ .

**EXAMPLE 10.15:**  
Global  
thresholding.

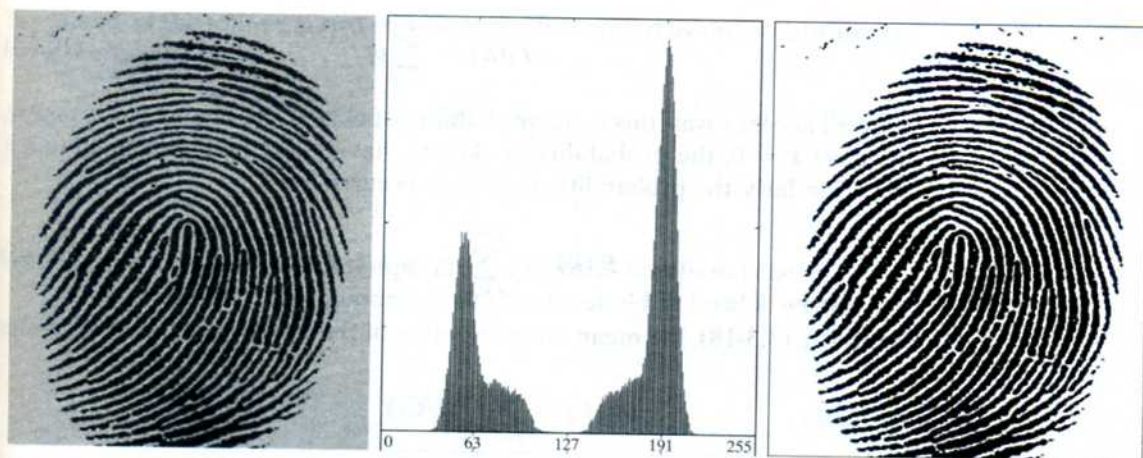
■ Figure 10.38 shows an example of segmentation based on a threshold estimated using the preceding algorithm. Figure 10.38(a) is the original image, and Fig. 10.38(b) is the image histogram, showing a distinct valley. Application of the preceding iterative algorithm resulted in the threshold  $T = 125.4$  after three iterations, starting with  $T = m$  (the average image intensity), and using  $\Delta T = 0$ . Figure 10.38(c) shows the result obtained using  $T = 125$  to segment the original image. As expected from the clear separation of modes in the histogram, the segmentation between object and background was quite effective. ■

The preceding algorithm was stated in terms of successively thresholding the input image and calculating the means at each step because it is more intuitive to introduce it in this manner. However, it is possible to develop a more efficient procedure by expressing all computations in the terms of the image histogram, which has to be computed only once (Problem 10.26).

### 10.3.3 Optimum Global Thresholding Using Otsu's Method

Thresholding may be viewed as a statistical-decision theory problem whose objective is to minimize the average error incurred in assigning pixels to two or more groups (also called *classes*). This problem is known to have an elegant closed-form solution known as the *Bayes decision rule* (see Section 12.2.2). The solution is based on only two parameters: the probability density function (PDF) of the intensity levels of each class and the probability that each class occurs in a given application. Unfortunately, estimating PDFs is not a trivial





**FIGURE 10.38** (a) Noisy fingerprint. (b) Histogram. (c) Segmented result using a global threshold (the border was added for clarity). (Original courtesy of the National Institute of Standards and Technology.)

matter, so the problem usually is simplified by making workable assumptions about the form of the PDFs, such as assuming that they are Gaussian functions. Even with simplifications, the process of implementing solutions using these assumptions can be complex and not always well-suited for practical applications.

The approach discussed in this section, called *Otsu's method* (Otsu [1979]), is an attractive alternative. The method is optimum in the sense that it maximizes the *between-class variance*, a well-known measure used in statistical discriminant analysis. The basic idea is that well-thresholded classes should be distinct with respect to the intensity values of their pixels and, conversely, that a threshold giving the best separation between classes in terms of their intensity values would be the best (optimum) threshold. In addition to its optimality, Otsu's method has the important property that it is based entirely on computations performed on the histogram of an image, an easily obtainable 1-D array.

Let  $\{0, 1, 2, \dots, L - 1\}$  denote the  $L$  distinct intensity levels in a digital image of size  $M \times N$  pixels, and let  $n_i$  denote the number of pixels with intensity  $i$ . The total number,  $MN$ , of pixels in the image is  $MN = n_0 + n_1 + n_2 + \dots + n_{L-1}$ . The normalized histogram (see Section 3.3) has components  $p_i = n_i/MN$ , from which it follows that

$$\sum_{i=0}^{L-1} p_i = 1, \quad p_i \geq 0 \quad (10.3-3)$$

Now, suppose that we select a threshold  $T(k) = k$ ,  $0 < k < L - 1$ , and use it to threshold the input image into two classes,  $C_1$  and  $C_2$ , where  $C_1$  consists of all the pixels in the image with intensity values in the range  $[0, k]$  and  $C_2$  consists of the pixels with values in the range  $[k + 1, L - 1]$ . Using this threshold, the probability,  $P_1(k)$ , that a pixel is assigned to (i.e., thresholded into) class  $C_1$  is given by the cumulative sum

$$P_1(k) = \sum_{i=0}^k p_i \quad (10.3-4)$$

Viewed another way, this is the probability of class  $C_1$  occurring. For example, if we set  $k = 0$ , the probability of class  $C_1$  having any pixels assigned to it is zero. Similarly, the probability of class  $C_2$  occurring is

$$P_2(k) = \sum_{i=k+1}^{L-1} p_i = 1 - P_1(k) \quad (10.3-5)$$

From Eq. (3.3-18), the mean intensity value of the pixels assigned to class  $C_1$  is

$$\begin{aligned} m_1(k) &= \sum_{i=0}^k iP(i/C_1) \\ &= \sum_{i=0}^k iP(C_1/i)P(i)/P(C_1) \\ &= \frac{1}{P_1(k)} \sum_{i=0}^k ip_i \end{aligned} \quad (10.3-6)$$

where  $P_1(k)$  is given in Eq. (10.3-4). The term  $P(i/C_1)$  in the first line of Eq. (10.3-6) is the probability of value  $i$ , given that  $i$  comes from class  $C_1$ . The second line in the equation follows from Bayes' formula:

$$P(A/B) = P(B/A)P(A)/P(B)$$

The third line follows from the fact that  $P(C_1/i)$ , the probability of  $C_1$  given  $i$ , is 1 because we are dealing only with values of  $i$  from class  $C_1$ . Also,  $P(i)$  is the probability of the  $i$ th value, which is simply the  $i$ th component of the histogram,  $p_i$ . Finally,  $P(C_1)$  is the probability of class  $C_1$ , which we know from Eq. (10.3-4) is equal to  $P_1(k)$ .

Similarly, the mean intensity value of the pixels assigned to class  $C_2$  is

$$\begin{aligned} m_2(k) &= \sum_{i=k+1}^{L-1} iP(i/C_2) \\ &= \frac{1}{P_2(k)} \sum_{i=k+1}^{L-1} ip_i \end{aligned} \quad (10.3-7)$$

The cumulative mean (average intensity) up to level  $k$  is given by

$$m(k) = \sum_{i=0}^k ip_i \quad (10.3-8)$$

and the average intensity of the entire image (i.e., the *global mean*) is given by

$$m_G = \sum_{i=0}^{L-1} ip_i \quad (10.3-9)$$



The validity of the following two equations can be verified by direct substitution of the preceding results:

$$P_1 m_1 + P_2 m_2 = m_G \quad (10.3-10)$$

and

$$P_1 + P_2 = 1 \quad (10.3-11)$$

where we have omitted the  $k$ s temporarily in favor of notational clarity.

In order to evaluate the “goodness” of the threshold at level  $k$  we use the normalized, dimensionless metric

$$\eta = \frac{\sigma_B^2}{\sigma_G^2} \quad (10.3-12)$$

where  $\sigma_G^2$  is the *global variance* [i.e., the intensity variance of all the pixels in the image, as given in Eq. (3.3-19)],

$$\sigma_G^2 = \sum_{i=0}^{L-1} (i - m_G)^2 p_i \quad (10.3-13)$$

and  $\sigma_B^2$  is the *between-class variance*, defined as

$$\sigma_B^2 = P_1(m_1 - m_G)^2 + P_2(m_2 - m_G)^2 \quad (10.3-14)$$

This expression can be written also as

$$\begin{aligned} \sigma_B^2 &= P_1 P_2 (m_1 - m_2)^2 \\ &= \frac{(m_G P_1 - m)^2}{P_1(1 - P_1)} \end{aligned} \quad (10.3-15)$$

where  $m_G$  and  $m$  are as stated earlier. The first line of this equation follows from Eqs. (10.3-14), (10.3-10), and (10.3-11). The second line follows from Eqs. (10.3-5) through (10.3-9). This form is slightly more efficient computationally because the global mean,  $m_G$ , is computed only once, so only two parameters,  $m$  and  $P_1$ , need to be computed for any value of  $k$ .

We see from the first line in Eq. (10.3-15) that the farther the two means  $m_1$  and  $m_2$  are from each other the larger  $\sigma_B^2$  will be, indicating that the between-class variance is a measure of *separability* between classes. Because  $\sigma_G^2$  is a constant, it follows that  $\eta$  also is a measure of separability, and maximizing this metric is equivalent to maximizing  $\sigma_B^2$ . The objective, then, is to determine the threshold value,  $k$ , that maximizes the between-class variance, as stated at the beginning of this section. Note that Eq. (10.3-12) assumes implicitly that  $\sigma_G^2 > 0$ . This variance can be zero only when all the intensity levels in the image are the same, which implies the existence of only one class of pixels. This in turn means that  $\eta = 0$  for a constant image since the separability of a single class from itself is zero.

The second step in Eq. (10.3-15) makes sense only if  $P_1$  is greater than 0 and less than 1, which, in view of Eq. (10.3-11), implies that  $P_2$  must satisfy the same condition.

Reintroducing  $k$ , we have the final results:

$$\eta(k) = \frac{\sigma_B^2(k)}{\sigma_G^2} \quad (10.3-16)$$

and

$$\sigma_B^2(k) = \frac{[m_G P_1(k) - m(k)]^2}{P_1(k)[1 - P_1(k)]} \quad (10.3-17)$$

Then, the optimum threshold is the value,  $k^*$ , that maximizes  $\sigma_B^2(k)$ :

$$\sigma_B^2(k^*) = \max_{0 \leq k \leq L-1} \sigma_B^2(k) \quad (10.3-18)$$

In other words, to find  $k^*$  we simply evaluate Eq. (10.3-18) for all integer values of  $k$  (such that the condition  $0 < P_1(k) < 1$  holds) and select that value of  $k$  that yielded the maximum  $\sigma_B^2(k)$ . If the maximum exists for more than one value of  $k$ , it is customary to average the various values of  $k$  for which  $\sigma_B^2(k)$  is maximum. It can be shown (Problem 10.33) that a maximum always exists, subject to the condition that  $0 < P_1(k) < 1$ . Evaluating Eqs. (10.3-17) and (10.3-18) for all values of  $k$  is a relatively inexpensive computational procedure, because the maximum number of integer values that  $k$  can have is  $L$ .

Once  $k^*$  has been obtained, the input image  $f(x, y)$  is segmented as before:

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) > k^* \\ 0 & \text{if } f(x, y) \leq k^* \end{cases} \quad (10.3-19)$$

for  $x = 0, 1, 2, \dots, M - 1$  and  $y = 0, 1, 2, \dots, N - 1$ . Note that all the quantities needed to evaluate Eq. (10.3-17) are obtained using only the histogram of  $f(x, y)$ . In addition to the optimum threshold, other information regarding the segmented image can be extracted from the histogram. For example,  $P_1(k^*)$  and  $P_2(k^*)$ , the class probabilities evaluated at the optimum threshold, indicate the portions of the areas occupied by the classes (groups of pixels) in the thresholded image. Similarly, the means  $m_1(k^*)$  and  $m_2(k^*)$  are estimates of the average intensity of the classes in the original image.

The normalized metric  $\eta$ , evaluated at the optimum threshold value,  $\eta(k^*)$ , can be used to obtain a quantitative estimate of the separability of classes, which in turn gives an idea of the ease of thresholding a given image. This measure has values in the range

$$0 \leq \eta(k^*) \leq 1 \quad (10.3-20)$$

The lower bound is attainable only by images with a single, constant intensity level, as mentioned earlier. The upper bound is attainable only by 2-valued images with intensities equal to 0 and  $L - 1$  (Problem 10.34).

Although our interest is in the value of  $\eta$  at the optimum threshold,  $k^*$ , this inequality holds in general for any value of  $k$  in the range  $[0, L - 1]$ .



Otsu's algorithm may be summarized as follows:

1. Compute the normalized histogram of the input image. Denote the components of the histogram by  $p_i$ ,  $i = 0, 1, 2, \dots, L - 1$ .
2. Compute the cumulative sums,  $P_1(k)$ , for  $k = 0, 1, 2, \dots, L - 1$ , using Eq. (10.3-4).
3. Compute the cumulative means,  $m(k)$ , for  $k = 0, 1, 2, \dots, L - 1$ , using Eq. (10.3-8).
4. Compute the global intensity mean,  $m_G$ , using (10.3-9).
5. Compute the between-class variance,  $\sigma_B^2(k)$ , for  $k = 0, 1, 2, \dots, L - 1$ , using Eq. (10.3-17).
6. Obtain the Otsu threshold,  $k^*$ , as the value of  $k$  for which  $\sigma_B^2(k)$  is maximum. If the maximum is not unique, obtain  $k^*$  by averaging the values of  $k$  corresponding to the various maxima detected.
7. Obtain the separability measure,  $\eta^*$ , by evaluating Eq. (10.3-16) at  $k = k^*$ .

The following example illustrates the preceding concepts.

■ Figure 10.39(a) shows an optical microscope image of polymersome cells, and Fig. 10.39(b) shows its histogram. The objective of this example is to segment the molecules from the background. Figure 10.39(c) is the result of using the basic global thresholding algorithm developed in the previous section. Because the histogram has no distinct valleys and the intensity difference between the background and objects is small, the algorithm failed to achieve the desired segmentation. Figure 10.39(d) shows the result obtained using Otsu's method. This result obviously is superior to Fig. 10.39(c). The threshold value computed by the basic algorithm was 169, while the threshold computed by Otsu's method was 181, which is closer to the lighter areas in the image defining the cells. The separability measure  $\eta$  was 0.467.

As a point of interest, applying Otsu's method to the fingerprint image in Example 10.15 yielded a threshold of 125 and a separability measure of 0.944. The threshold is identical to the value (rounded to the nearest integer) obtained with the basic algorithm. This is not unexpected, given the nature of the histogram. In fact, the separability measure is high due primarily to the relatively large separation between modes and the deep valley between them. ■

**EXAMPLE 10.16:**  
Optimum global thresholding using Otsu's method.

Polymersomes are cells artificially engineered using polymers. Polymersomes are invisible to the human immune system and can be used, for example, to deliver medication to targeted regions of the body.

### 10.3.4 Using Image Smoothing to Improve Global Thresholding

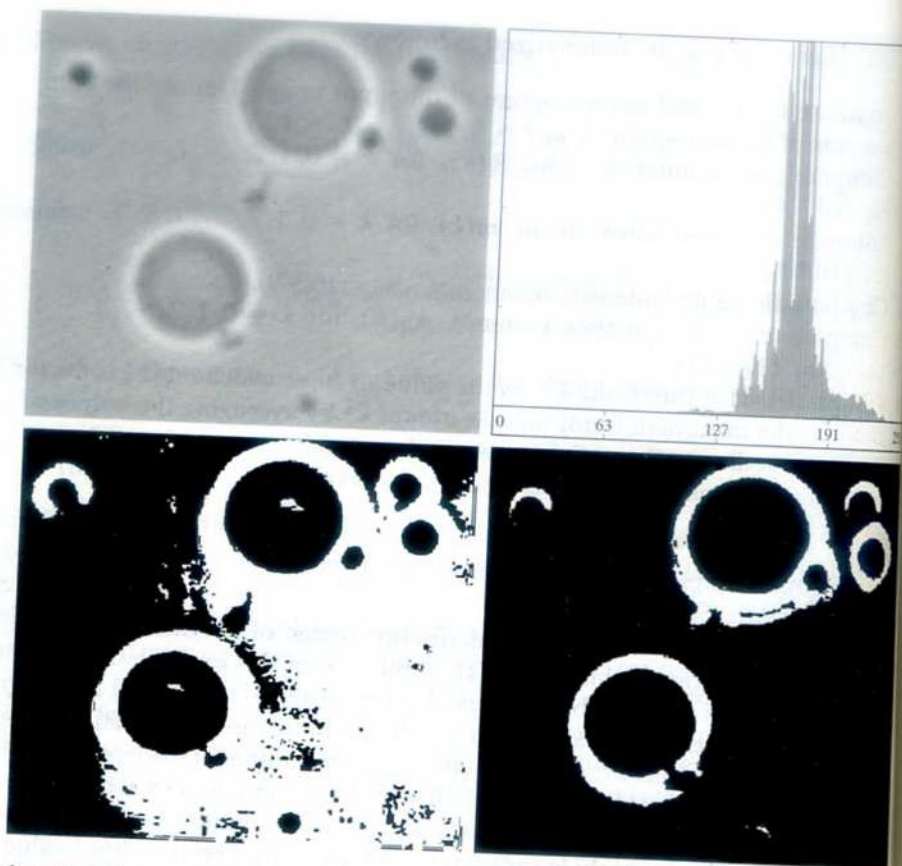
As noted in Fig. 10.36, noise can turn a simple thresholding problem into an unsolvable one. When noise cannot be reduced at the source, and thresholding is the segmentation method of choice, a technique that often enhances performance is to smooth the image prior to thresholding. We illustrate the approach with an example.

Figure 10.40(a) is the image from Fig. 10.36(c), Fig. 10.40(b) shows its histogram, and Fig. 10.40(c) is the image thresholded using Otsu's method. Every black point in the white region and every white point in the black region is a



a b  
c d**FIGURE 10.39**

(a) Original image.  
 (b) Histogram (high peaks were clipped to highlight details in the lower values).  
 (c) Segmentation result using the basic global algorithm from Section 10.3.2.  
 (d) Result obtained using Otsu's method. (Original image courtesy of Professor Daniel A. Hammer, the University of Pennsylvania.)



thresholding error, so the segmentation was highly unsuccessful. Figure 10.40(d) shows the result of smoothing the noisy image with an averaging mask of size  $5 \times 5$  (the image is of size  $651 \times 814$  pixels), and Fig. 10.40(e) is its histogram. The improvement in the shape of the histogram due to smoothing is evident, and we would expect thresholding of the smoothed image to be nearly perfect. As Fig. 10.40(f) shows, this indeed was the case. The slight distortion of the boundary between object and background in the segmented, smoothed image was caused by the blurring of the boundary. In fact, the more aggressively we smooth an image, the more boundary errors we should anticipate in the segmented result.

Next we consider the effect of reducing the size of the region in Fig. 10.40(a) with respect to the background. Figure 10.41(a) shows the result. The noise in this image is additive Gaussian noise with zero mean and a standard deviation of 10 intensity levels (as opposed to 50 in the previous example). As Fig. 10.41(b) shows, the histogram has no clear valley, so we would expect segmentation to fail, a fact that is confirmed by the result in Fig. 10.41(c). Figure 10.41(d) shows the image smoothed with an averaging mask of size  $5 \times 5$ , and Fig. 10.40(e) is the corresponding histogram. As expected, the net effect was to reduce the spread of the histogram, but the distribution still is unimodal. As Fig. 10.40(f) shows, segmentation failed again. The reason for the failure can be traced to the fact that the region is so small that its contribution to the histogram is insignificant compared to the intensity spread caused by noise. In