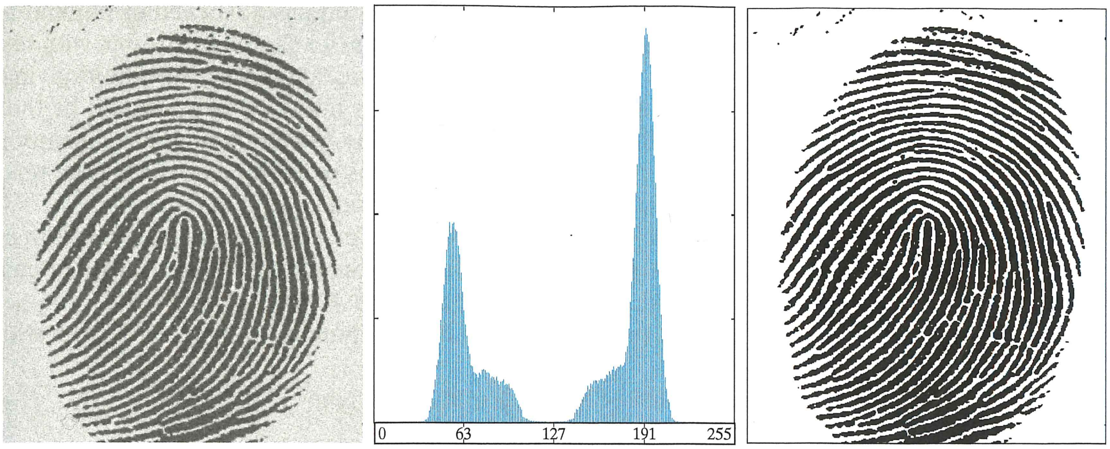

a b c

FIGURE 10.35 (a) Noisy fingerprint. (b) Histogram. (c) Segmented result using a global threshold (thin image border added for clarity). (Original image courtesy of the National Institute of Standards and Technology.).

initial choice for T). If this condition is met, the algorithm converges in a finite number of steps, whether or not the modes are separable (see Problem 10.30).

EXAMPLE 10.13: Global thresholding.

Figure 10.35 shows an example of segmentation using the preceding iterative algorithm. Figure 10.35(a) is the original image and Fig. 10.35(b) is the image histogram, showing a distinct valley. Application of the basic global algorithm resulted in the threshold $T = 125.4$ after three iterations, starting with T equal to the average intensity of the image, and using $\Delta T = 0$. Figure 10.35(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 perfect.

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 function* (see Section 12.4). 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 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 real-time applications.

The approach in the following discussion, 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 properly 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 (see Section 3.3).

Let $\{0, 1, 2, \ldots, L-1\}$ denote the set of L distinct integer 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 $MN = n_0 + n_1 + n_2 + \cdots + n_{L-1}$. The normalized histogram (see Section 3.3) has components $p_i = n_i/MN$, from which it follows that

$$\sum_{i=0}^{L-1} p_i = 1 \qquad p_i \geq 0 \tag{10-48}$$

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

$$P_1(k) = \sum_{i=0}^{k} p_i \tag{10-49}$$

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

$$P_2(k) = \sum_{i=k+1}^{L-1} p_i = 1 - P_1(k) \tag{10-50}$$

From Eq. (3-25), the *mean intensity* value of the pixels in c_1 is

$$\begin{aligned} m_1(k) &= \sum_{i=0}^{k} iP(i/c_1) = \sum_{i=0}^{k} iP(c_1/i)P(i)/P(c_1) \\ &= \frac{1}{P_1(k)} \sum_{i=0}^{k} i p_i \end{aligned} \tag{10-51}$$

where $P_1(k)$ is given by Eq. (10-49). The term $P(i/c_1)$ in Eq. (10-51) is the probability of intensity value i, given that i comes from class c_1. The rightmost term in the first line of the equation follows from Bayes' formula:

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

The second 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 ith value, which is the ith component of the histogram, p_i. Finally, $P(c_1)$ is the probability of class c_1 which, from Eq. (10-49), is equal to $P_1(k)$.

Similarly, the *mean intensity* value of the pixels assigned to class c_2 is

$$\begin{aligned} m_2(k) &= \sum_{i=k+1}^{L-1} iP(i/c_2) \\ &= \frac{1}{P_2(k)} \sum_{i=k+1}^{L-1} ip_i \end{aligned} \tag{10-52}$$

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

$$m(k) = \sum_{i=0}^{k} ip_i \tag{10-53}$$

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

$$m_G = \sum_{i=0}^{L-1} ip_i \tag{10-54}$$

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

$$P_1 m_1 + P_2 m_2 = m_G \tag{10-55}$$

and

$$P_1 + P_2 = 1 \tag{10-56}$$

where we have omitted the ks temporarily in favor of notational clarity.

In order to evaluate the effectiveness of the threshold at level k, we use the normalized, dimensionless measure

$$\eta = \frac{\sigma_B^2}{\sigma_G^2} \tag{10-57}$$

where σ_G^2 is the *global variance* [i.e., the intensity variance of all the pixels in the image, as given in Eq. (3-26)],

$$\sigma_G^2 = \sum_{i=0}^{L-1} (i - m_G)^2 p_i \tag{10-58}$$

and σ_B^2 is the *between-class variance*, defined as

$$\sigma_B^2 = P_1(m_1 - m_G)^2 + P_2(m_2 - m_G)^2 \tag{10-59}$$

This expression can also be written as

The second step in this equation makes sense only if P_1 is greater than 0 and less than 1, which, in view of Eq. (10-56), implies that P_2 must satisfy the same condition.

$$\begin{aligned} \sigma_B^2 &= P_1 P_2 (m_1 - m_2)^2 \\ &= \frac{(m_G P_1 - m)^2}{P_1(1 - P_1)} \end{aligned} \tag{10-60}$$

The first line of this equation follows from Eqs. (10-55), (10-56), and (10-59). The second line follows from Eqs. (10-50) through (10-54). This form is slightly more efficient computationally because the global mean, m_G, is computed only once, so only two parameters, m_1 and P_1, need to be computed for any value of k.

The first line in Eq. (10-60) indicates that the farther the two means m_1 and m_2 are from each other, the larger σ_B^2 will be, implying that the between-class variance is a measure of separability between classes. Because σ_G^2 is a constant, it follows that η also is a measure of separability, and maximizing this metric is equivalent to maximizing σ_B^2. The objective, then, is to determine the threshold value, k, that maximizes the between-class variance, as stated earlier. Note that Eq. (10-57) 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 because the separability of a single class from itself is zero.

Reintroducing k, we have the final results:

$$\eta(k) = \frac{\sigma_B^2(k)}{\sigma_G^2} \tag{10-61}$$

and

$$\sigma_B^2(k) = \frac{\left[m_G P_1(k) - m(k)\right]^2}{P_1(k)\left[1 - P_1(k)\right]} \tag{10-62}$$

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

$$\sigma_B^2\left(k^*\right) = \max_{0 \le k \le L-1} \sigma_B^2(k) \tag{10-63}$$

To find k^* we simply evaluate this equation for all *integer* values of k (subject to the condition $0 < P_1(k) < 1$) and select the 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 (see Problem 10.36) that a maximum always exists, subject to the condition $0 < P_1(k) < 1$. Evaluating Eqs. (10-62) and (10-63) for all values of k is a relatively inexpensive computational procedure, because the maximum number of integer values that k can have is L, which is only 256 for 8-bit images.

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

$$g(x,y) = \begin{cases} 1 & \text{if } f(x,y) > k^* \\ 0 & \text{if } f(x,y) \le k^* \end{cases} \tag{10-64}$$

for $x = 0,1,2,\ldots,M-1$ and $y = 0,1,2,\ldots,N-1$. Note that all the quantities needed to evaluate Eq. (10-62) 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.

In general, the measure in Eq.(10-61) has values in the range

$$0 \leq \eta(k) \leq 1 \tag{10-65}$$

for values of k in the range $[0, L-1]$. When evaluated at the optimum threshold k^*, this measure is a quantitative estimate of the separability of classes, which in turn gives us an idea of the accuracy of thresholding a given image with k^*. The lower bound in Eq. (10-65) is attainable only by images with a single, constant intensity level. The upper bound is attainable only by two-valued images with intensities equal to 0 and $L-1$ (see Problem 10.37).

Otsu's algorithm may be summarized as follows:

1. Compute the normalized histogram of the input image. Denote the components of the histogram by $p_i, i = 0, 1, 2, \ldots, L-1$.
2. Compute the cumulative sums, $P_1(k)$, for $k = 0, 1, 2, \ldots, L-1$, using Eq. (10-49).
3. Compute the cumulative means, $m(k)$, for $k = 0, 1, 2, \ldots, L-1$, using Eq. (10-53).
4. Compute the global mean, m_G, using Eq. (10-54).
5. Compute the between-class variance term, $\sigma_B^2(k)$, for $k = 0, 1, 2, \ldots, L-1$, using Eq. (10-62).
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. Compute the global variance, σ_G^2, using Eq. (10-58), and then obtain the separability measure, η^*, by evaluating Eq. (10-61) with $k = k^*$.

The following example illustrates the use of this algorithm.

EXAMPLE 10.14: Optimum global thresholding using Otsu's method.

Figure 10.36(a) shows an optical microscope image of polymersome cells. These are cells artificially engineered using polymers. They are invisible to the human immune system and can be used, for example, to deliver medication to targeted regions of the body. Figure 10.36(b) shows the image histogram. The objective of this example is to segment the molecules from the background. Figure 10.36(c) is the result of using the basic global thresholding algorithm discussed earlier. 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 10.36(d) shows the result obtained using Otsu's method. This result obviously is superior to Fig. 10.36(c). The threshold value computed by the basic algorithm was 169, while the threshold computed by Otsu's method was 182, which is closer to the lighter areas in the image defining the cells. The separability measure η^* was 0.467.

As a point of interest, applying Otsu's method to the fingerprint image in Example 10.13 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 because of the relatively large separation between modes and the deep valley between them.