

10.3.1 基础知识

前一节采用首先寻找边缘线段，然后将这些线段连接为边界的方法来识别区域。本节讨论根据灰度值和/或灰度值的性质来将图像直接划分为多个区域的技术。

1. 灰度阈值处理基础

假设图 10.32(a)中的灰度直方图对应于图像 $f(x, y)$ ，该图像由暗色背景上的亮目标组成，其中目标像素和背景像素的灰度值组合成了两种

记住， $f(x, y)$ 表示 f 在坐标 (x, y) 处的灰度。

主导模式。从背景中提取目标的一种明显方法是，选择一个分隔这些模式的阈值 T 。然后，图像中 $f(x, y) > T$ 的任何点 (x, y) 称为一个目标点；否则该点称为背景点。换句话说，分割后的图像 $g(x, y)$ 为

$$g(x, y) = \begin{cases} 1, & f(x, y) > T \\ 0, & f(x, y) \leq T \end{cases} \quad (10.46)$$

当 T 是一个适合于整个图像的常数时，上式中给出的处理称为全局阈值处理。当 T 值在一幅图像上变化时，上式中给出的处理称为可变阈值处理。有时，我们使用局部阈值处理或区域阈值处理表示可变阈值处理，此时图像中任意一点 (x, y) 处的 T 值取决于 (x, y) 的邻域的性质（如邻域中像素的平均灰度）。若 T 取决于空间坐标 (x, y) 本身，则可变阈值处理通常称为动态阈值处理或自适应阈值处理。这些术语的应用并不普遍。

虽然我们按照约定为背景使用 0 灰度，为目标像素使用 1 灰度，但在式(10.46)中可以使用任何两个不同的值。

图 10.32(b)显示了一个更难的阈值处理问题，它包含一个具有 3 个主导模式的直方图，这 3 个主导模式分别对应于暗色背景和两类亮目标。这里，若 $f(x, y) \leq T_1$ ，则多阈值处理把点 (x, y) 分类为背景；若 $T_1 < f(x, y) \leq T_2$ ，则把点 (x, y) 分类为目标；若 $f(x, y) > T_2$ ，则把点分类为另一个目标。也就是说，分割后的图像为

$$g(x, y) = \begin{cases} a, & f(x, y) > T_2 \\ b, & T_1 < f(x, y) \leq T_2 \\ c, & f(x, y) \leq T_1 \end{cases} \quad (10.47)$$

式中， a 、 b 和 c 是任意三个不同的灰度值。本节稍后将讨论双阈值处理。要求两个以上阈值的分割问题很难求解（或通常不可能求解），而使用其他方法通常能得到较好的结果，如本节稍后将要讨论的可变阈值处理方法或 10.4 节中将要讨论的区域生长方法。

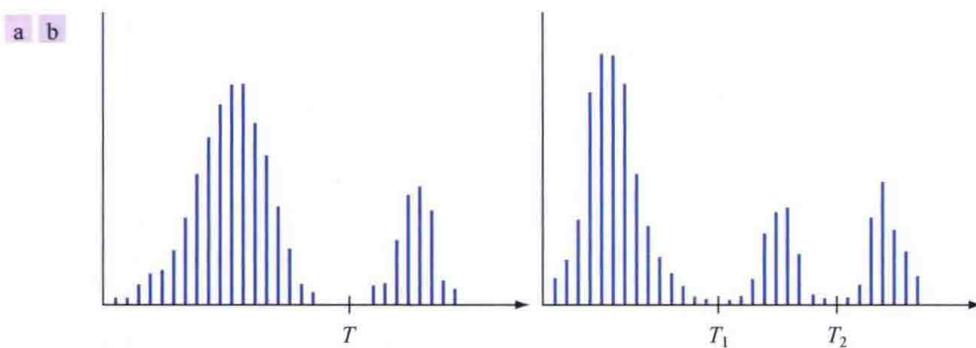


图 10.32 可被(a)单阈值和(b)双阈值分割的灰度直方图

根据前面的讨论，我们可以凭直觉推断灰度阈值处理成功与否，与分隔直方图模式的波谷的宽度和深度有关。而影响波谷特性的关键因素依次是：(1) 波峰间的间隔（波峰离得越远，分隔这些模式的机会越好）；(2) 图像中的噪声内容（噪声越大，模式越宽）；(3) 目标和背景的相对大小；(4) 光源的均匀性；(5) 图像反射性质的均匀性。

2. 图像阈值处理中噪声的作用

图 10.33(a)中的简单合成图像中没有噪声,因此其直方图由两个波峰模式组成,如图 10.33(d)所示。将该图像分割为两个区域很简单:我们只需在两个模式之间的任何位置选择一个阈值。图 10.33(b)显示了被均值为零、标准差为 10 个灰度级的高斯噪声污染的原图像。现在,模式较宽[见图 10.33(e)],但它们的间隔足以让它们之间的波谷的深度轻易地分隔两个模式。放在两个波峰之间的中间位置的一个阈值就能很好地分割该图像。图 10.33(c)显示了该图像被均值为零、标准差为 50 个灰度级的高斯噪声污染的结果。如图 10.33(f)中所示的直方图那样,现在的情况严重到无法区分两个模式。若不做其他处理(如本节稍后讨论的方法),那么就没有希望为分割这幅图像找到一个合适的阈值。

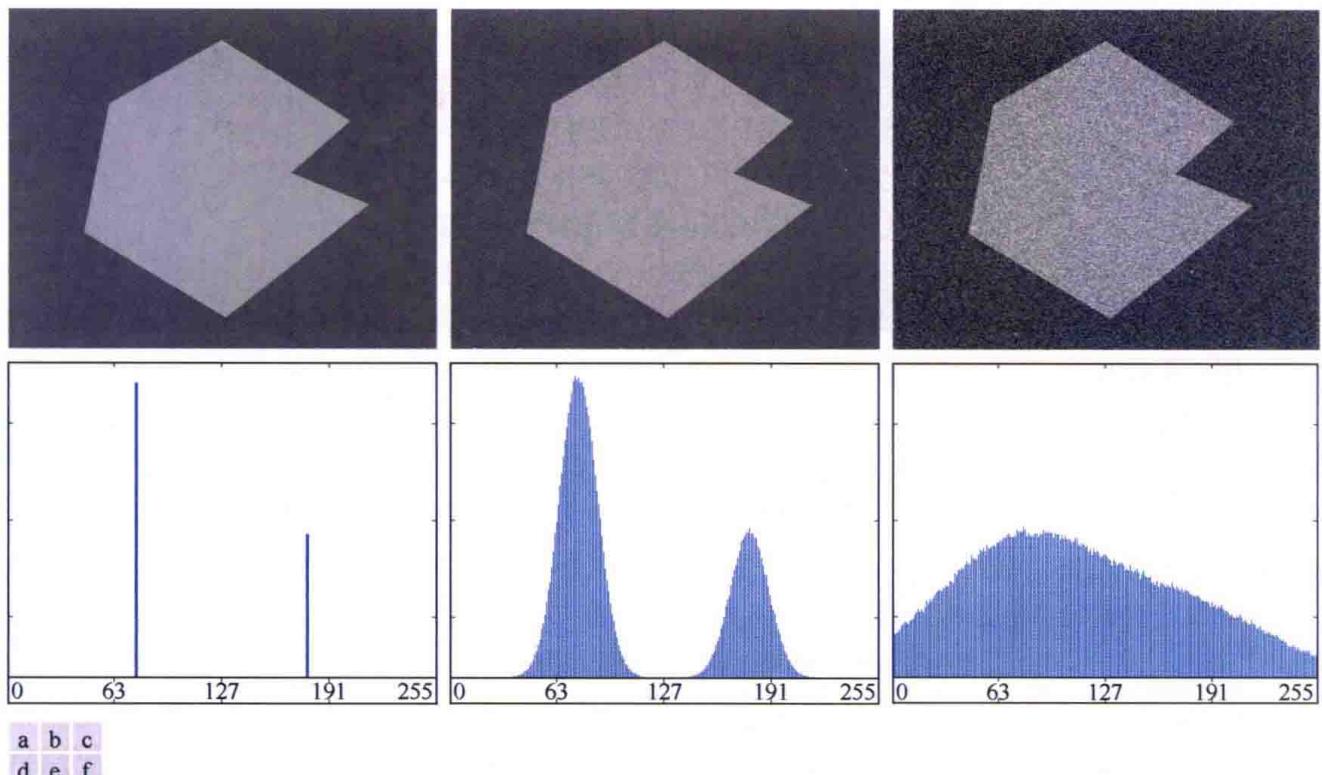


图 10.33 (a)无噪声 8 比特图像; (b)被均值为零、标准差为 10 个灰度级的加性高斯噪声污染的图像; (c)被均值为零、标准差为 50 个灰度级的加性高斯噪声污染的图像; (d)~(f)对应的直方图

3. 图像直方图处理中光照和反射的作用

图 10.34 说明了光照对图像直方图的影响。图 10.34(a)是来自图 10.33(b)的带噪声图像,图 10.34(d)显示了它的直方图。如前所述,使用单个阈值很容易分割这幅图像。参考 2.3 节讨论的图像形成模型,假设图 10.34(a)中的图像乘以一个不均匀的灰度函数,如图 10.37(b)中的灰度斜坡,其直方图如图 10.34(e)所示。图 10.34(c)显示了这两幅图的乘积,图 10.34(f)是得到的直方图。如果不进行额外的处理(将在本节稍后讨论),那么波峰之间的深谷会破坏到模式无法分隔的程度。如果光照非常均匀,但(目标和/或背景表面自然反射率的变化导致的)图像反射不均匀,那么也会得到类似的结果。

理论上,斜坡图像的直方图是均匀的。在实践中,均匀的程度取决于图像的大小和灰度级的数量。

重点是,在使用阈值处理或其他分割技术分割图像时,照明和反射起核心作用。因此,只要可能,就要首先在求解分割问题时考虑如何控制这些参数。无法控制这些参数时,解决这一问题有三种基本方法。第一种方法是直接校正这种阴影模式。例如,不均匀(但固定的)光照可以用相反的模式与图像相乘来校正,相反的模式可以通过对一个恒定灰度的平坦表面成像来得到。第二种方法是用诸

如 9.8 节中介绍的顶帽变换处理来校正全局阴影模式。第三种方法是用可变阈值处理来“绕过”不均匀性，见本节稍后的讨论。

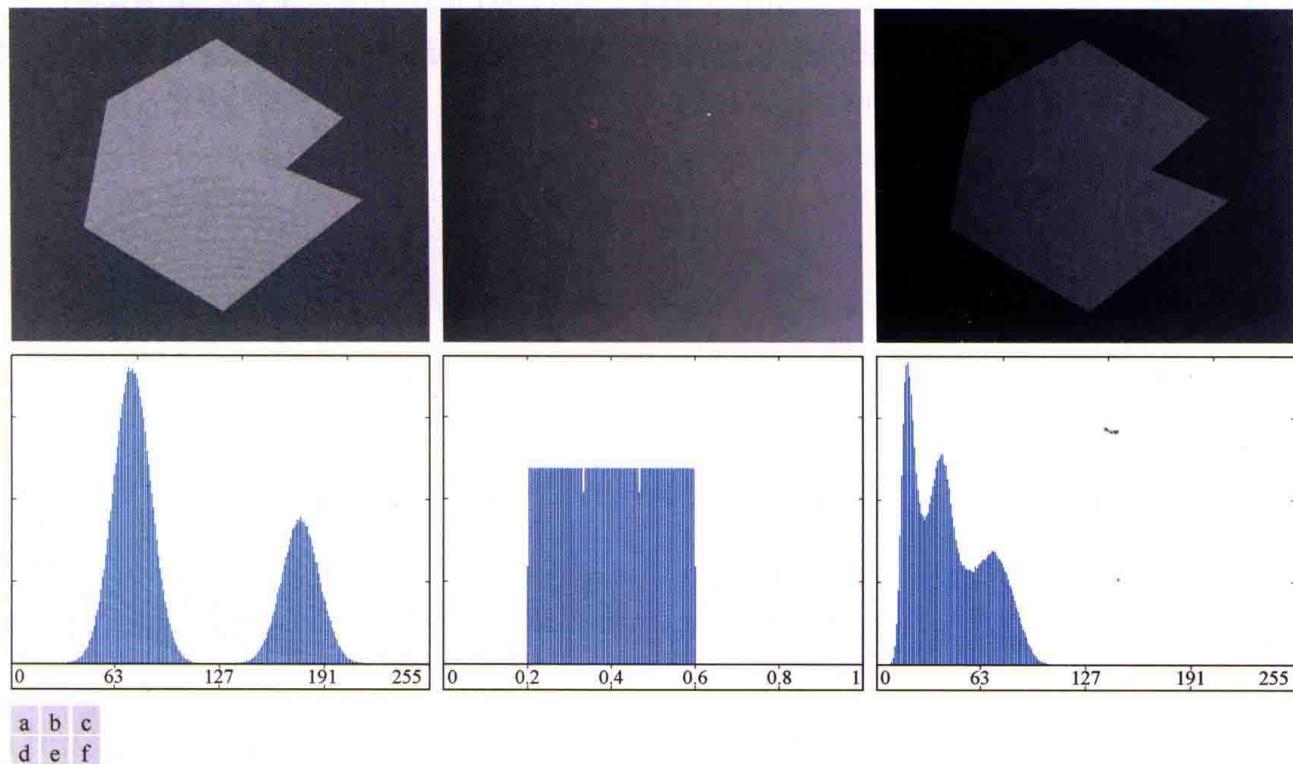


图 10.34 (a)带噪声图像; (b)值域为[0.2, 0.6]的灰度斜坡; (c)图(a)和图(b)的乘积; (d)~(f)对应的直方图

10.3.2 基本的全局阈值处理

当目标和背景像素的灰度分布非常不同时，可对整个图像使用单个（全局）阈值。在大多数应用中，图像之间通常存在足够的变化，即使全局阈值是一种合适的方法，也需要有能对每幅图像估计阈值的算法。下面的迭代算法适用于这一目的：

1. 为全局阈值 T 选择一个初始估计值。
2. 在式(10.46)中用 T 分割图像。这将产生两组像素：由灰度值大于 T 的所有像素组成的 G_1 ，由所有小于等于 T 的像素组成的 G_2 。
3. 对 G_1 和 G_2 中的像素分别计算平均灰度值（均值） m_1 和 m_2 。
4. 在 m_1 和 m_2 之间计算一个新的阈值：

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

5. 重复步骤 2 到步骤 4，直到连续迭代中的两个 T 值间的差小于某个预定义的值 ΔT 为止。

这里是从对输入图像进行连续阈值处理并在每一步中计算均值的角度来描述算法的，因为以这种方式引入它更直观。然而，通过将所有计算表达为图像直方图，可以开发一个等效（且更有效）的过程，这个过程只需要计算一次（见习题 10.29）。

当与目标和背景相关的直方图模式之间存在一个非常清晰的波谷时，上述算法很有效。参数 ΔT 用于在阈值变化不大时停止迭代。初始阈值必须大于图像中的最小灰度级、小于图像中的最大灰度级（选择图像的平均灰度作为 T 初始值最好）。满足这个条件时，无论模式是否可分，算法都会在有限数量的步骤内收敛（见习题 10.30）。

例 10.13 全局阈值处理。

图 10.35 显示了用前述迭代算法的一个分割例子。图 10.35(a)是原图像，图 10.35(b)是该图像的直方图，直方图中有一个明显的波谷。令 T 等于图像的平均灰度并令 $\Delta T = 0$ 开始，应用前述全局迭代算法经过 3 次迭代后，得到阈值 $T = 125.4$ 。图 10.35(c)是用阈值 $T = 125$ 分割原图像后的结果。如期望的那样，直方图中的模式已明显分离，目标和背景间的分割非常完美。

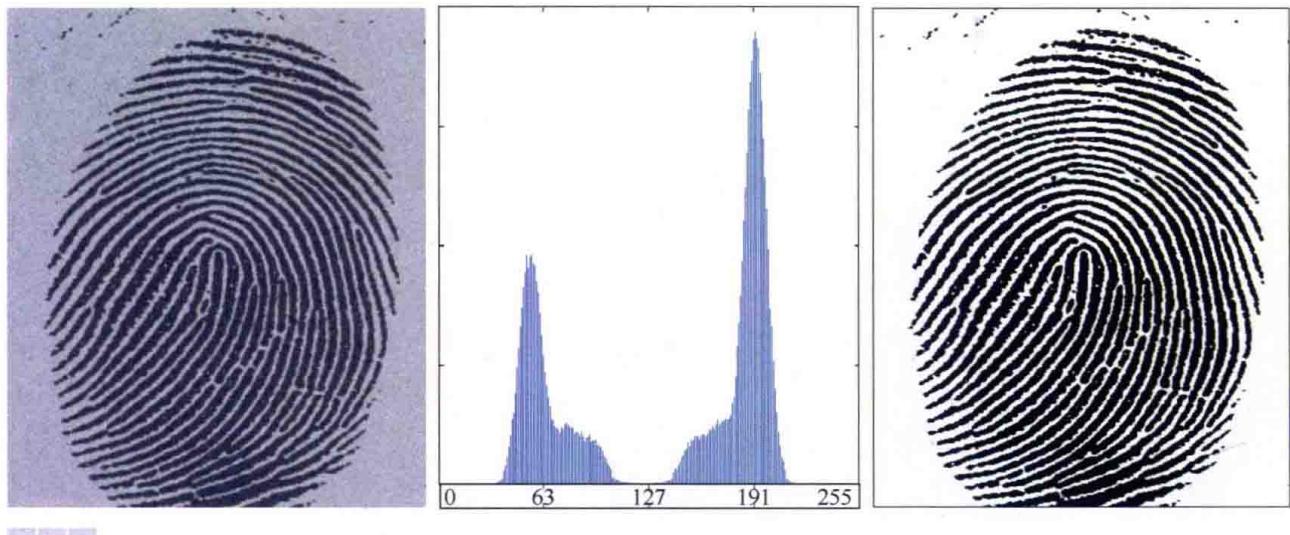


图 10.35 (a)带噪声指纹图像；(b)直方图；(c)使用全局阈值对图像分割后的结果（为清晰起见，添加了边界）(原图像由美国国家标准和技术研究所提供)

10.3.3 使用 Otsu 方法的最优全局阈值处理

阈值处理可视为一种统计决策理论问题，其目的是在把像素分配给两组或多组（也称分类）的过程中，使引入的平均误差最小。对于这个问题，已知有一个解析解，称为贝叶斯决策函数（见 12.4 节）。这个解析解仅基于两个参数：每类灰度级的概率密度函数（PDF）和已知应用中每类出现的概率。遗憾的是，估计 PDF 并不容易，因此通常采用一种假设的 PDF 形式来简化这一问题，如假设它们是高斯函数。即使采用了这一形式的简化，使用这些假设求解的过程也很复杂，并且对实时应用来说也并非总是合适的。

下边讨论的方法称为 Otsu 方法（Otsu[1979]），它是另一种具有吸引力的替代方案。这种方法在最大化类间方差方面是最优的，因此是在统计判别分析中使用的一种著名测度。这种方法的基本思想是，经过正确阈值处理后的类别相对于它们的像素的灰度值而言应是不同的，而根据灰度值给出类间最优分离的一个阈值会是最好（最优）的阈值。除最优化外，Otsu 方法还有一个重要的性质，即它完全基于对图像的直方图（一个容易得到的一维阵列，见 3.3 节）进行计算。

令 $\{0, 1, 2, \dots, L-1\}$ 表示一幅大小为 $M \times N$ 像素的数字图像中的 L 个不同的整数灰度级，并令 n_i 表示灰度级 i 的像素数。图像中的像素总数为 $MN = n_0 + n_1 + n_2 + \dots + n_{L-1}$ 。归一化后的直方图（见 3.3 节）有 $p_i = n_i / MN$ 个分量，由此有

$$\sum_{i=0}^{L-1} p_i = 1, \quad p_i \geq 0 \quad (10.48)$$

现在，假设我们选择一个阈值 $T(k) = k$, $0 < k < L-1$ ，并用它将输入图像阈值化为两类 c_1 和 c_2 ，其中 c_1 由图像中灰度值在区间 $[0, k]$ 内的所有像素组成， c_2 由灰度值在区间 $[k+1, L-1]$ 内的所有像素组成。使

用这个阈值，像素被分配给类 c_1 的概率 $P_1(k)$ 由如下的累积和给出：

$$P_1(k) = \sum_{i=0}^k p_i \quad (10.49)$$

换一个角度看，这是类 c_1 发生的概率。例如，如果令 $k = 0$ ，那么任何像素分配给类 c_1 的概率为零。

类似地，类 c_2 发生的概率为

$$P_2(k) = \sum_{i=k+1}^{L-1} p_i = 1 - P_1(k) \quad (10.50)$$

由式(3.25)可得类 c_1 中的像素的平均灰度值为

$$\begin{aligned} m_1(k) &= \sum_{i=0}^k i P(i/c_1) = \sum_{i=0}^k i P(c_1/i) P(i)/P(c_1) \\ &= \frac{1}{P_1(k)} \sum_{i=0}^k i p_i \end{aligned} \quad (10.51)$$

式中， $P_1(k)$ 由式(10.49)给出。式(10.51)中的 $P(i/c_1)$ 项是灰度值 i 的概率， i 来自类 c_1 。式中第一行的最右侧项来自贝叶斯公式：

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

第二行是根据如下事实得到的：已知 i 时， c_1 的概率 $P(c_1/i)$ 为 1，因为我们只处理来自类 c_1 的 i 的值。此外， $P(i)$ 是第 i 个值的概率，即直方图 p_i 的第 i 个分量。最后， $P(c_1)$ 是类 c_1 的概率，根据式(10.49)可知它等于 $P_1(k)$ 。

类似地，分配给类 c_2 的像素的平均灰度值为

$$m_2(k) = \sum_{i=k+1}^{L-1} i P(i/c_2) = \frac{1}{P_2(k)} \sum_{i=k+1}^{L-1} i p_i \quad (10.52)$$

高达 k 级的累积平均值（平均灰度）为

$$m(k) = \sum_{i=0}^k i p_i \quad (10.53)$$

整个图像的平均灰度（即全局均值）为

$$m_G = \sum_{i=0}^{L-1} i p_i \quad (10.54)$$

直接代入前面的结果中，可以验证如下两个公式成立：

$$P_1 m_1 + P_2 m_2 = m_G \quad (10.55)$$

和

$$P_1 + P_2 = 1 \quad (10.56)$$

为表达清楚起见，式中暂时省略了 k 。

为了评估 k 级的阈值的有效性，我们使用归一化的、无量纲的测度：

$$\eta = \frac{\sigma_B^2}{\sigma_G^2} \quad (10.57)$$

式中， σ_G^2 是全局方差 [即图像中所有像素的灰度方差，见式(3.26)]，

$$\sigma_G^2 = \sum_{i=0}^{L-1} (i - m_G)^2 p_i \quad (10.58)$$

σ_B^2 是类间方差，它定义为

$$\sigma_B^2 = P_1(m_1 - m_G)^2 + P_2(m_2 - m_G)^2 \quad (10.59)$$

这个表达式也可写为

$$\sigma_B^2 = P_1 P_2 (m_1 - m_2)^2 = \frac{(m_G P_1 - m)^2}{P_1 (1 - P_1)} \quad (10.60)$$

式中的第一个等式由式(10.55)、式(10.56)和式(10.59)得到，第二个等式由式(10.50)~式(10.54)到得。这种形式计算上更有效，因为全局均值 m_G 只计算一次，所以对 k 的任何值只需计算两个参数 m_1 和 P_1 。

式(10.60)中的第二步仅在 P_1 大于 0 并且小于 1 时才有意义，根据式(10.56)，这表明 P_2 必须满足相同的条件。

式(10.60)中的第一行指出，两个均值 m_1 和 m_2 彼此隔得越远， σ_B^2 越大，这说明类间方差是类间的可分离性测度。因为 σ_G^2 是一个常数，由此得出 η 也是一个可分离性测度，并且最大化这一测度等价于最大化 σ_B^2 。然后，如本节之初所述，目标是求最大化类间方差的阈值 k 。注意，式(10.57)隐式地假设 $\sigma_G^2 > 0$ 。仅当图像中的所有灰度级相同时，这一方差才为零，这意味着仅存在一类像素。同样，这也意味着对于常数图像有 $\eta = 0$ ，因为来自其自身单个类的可分离性为零。

再次引入 k ，得到最终结果：

$$\eta(k) = \frac{\sigma_B^2(k)}{\sigma_G^2} \quad (10.61)$$

和

$$\sigma_B^2(k) = \frac{[m_G P_1(k) - m(k)]^2}{P_1(k)[1 - P_1(k)]} \quad (10.62)$$

从而最优阈值是 k^* ，它最大化 $\sigma_B^2(k)$ ：

$$\sigma_B^2(k^*) = \max_{0 \leq k \leq L-1} \sigma_B^2(k) \quad (10.63)$$

要求 k^* ，只需对 k 的所有整数值 [满足条件 $0 < P_1(k) < 1$] 计算这个公式，并选取使得 $\sigma_B^2(k)$ 最大的 k 值。如果这个极大值对应于多个 k 值，那么习惯上取使得 $\sigma_B^2(k)$ 最大的各个 k 值的平均值。可以证明(见习题 10.36)，满足条件 $0 < P_1(k) < 1$ 时，总存在一个极大值。对所有 k 值计算式(10.62)和式(10.63)的开销相对较小，因为 k 能够具有的整数值的最大数量是 L ，对 8 比特图像而言它只是 256。

求得 k^* 后，就可像之前那样对输入图像 $f(x, y)$ 进行分割：

$$g(x, y) = \begin{cases} 1, & f(x, y) > k^* \\ 0, & f(x, y) \leq k^* \end{cases} \quad (10.64)$$

式中， $x = 0, 1, 2, \dots, M-1$ 和 $y = 0, 1, 2, \dots, N-1$ 。注意，只使用 $f(x, y)$ 的直方图就可得到计算式(10.62)所需的所有参量。除最优阈值外，与分割图像有关的其他信息可从直方图中提取。例如，在最优阈值位置计算的类的概率 $P_1(k^*)$ 和 $P_2(k^*)$ ，指出了阈值处理后的图像中由该类(像素组)占据的部分区域。类似地，均值 $m_1(k^*)$ 和 $m_2(k^*)$ 是原图像中类的平均灰度的估计。

一般来说，对于区间 $[0, L-1]$ 内的 k 值，式(10.61)中的测度的值域为

$$0 \leq \eta(k) \leq 1 \quad (10.65)$$

在最优阈值 k^* 处计算时，这个测度是类的可分离性的定量估计，于是为我们提供了使用 k^* 阈值化已知图像的精度概念。式(10.65)的下界只能由单个恒定灰度级的图像达到，上界只能由灰度为 0 和 $L-1$ 的二值图像达到(见习题 10.37)。

Otsu 算法小结如下：

1. 计算输入图像的归一化直方图。使用 $p_i, i = 0, 1, 2, \dots, L-1$ 表示该直方图的各个分量。
2. 用式(10.49)计算累积和 $P_i(k), k = 0, 1, 2, \dots, L-1$ 。
3. 用式(10.53)计算累积均值 $m(k), k = 0, 1, 2, \dots, L-1$ 。
4. 用式(10.54)计算全局灰度均值 m_G 。
5. 用式(10.62)计算类间方差 $\sigma_B^2(k), k = 0, 1, 2, \dots, L-1$ 。
6. 得到 Otsu 阈值 k^* ，即 $\sigma_B^2(k)$ 最大时的 k 值。如果极大值不唯一，那么取对应各个极大值的各个 k 值的平均值来得到 k^* 。
7. 用式(10.58)计算全局方差 σ_G^2 ，然后令 $k = k^*$ 来计算式(10.61)，得到可分离性测度 η^* 。

下面的例子说明了这个算法的应用。

例 10.14 使用 Otsu 方法的最优全局阈值处理。

图 10.36(a)是聚合细胞的光学显微图像，聚合细胞是用聚合物人工设计的细胞，它们对人体免疫系统而言是不可见的，例如，我们可用它们向身体的目标区域递送药物。图 10.36(b)显示了图像的直方图。本例的目的是从背景中分割出分子。图 10.36(c)是用前节给出的基本全局阈值处理算法得到的结果。因为直方图没有明显的波谷，并且背景和目标之间的灰度差很小，所以算法未实现期望的分割。图 10.36(d)是用 Otsu 方法得到的结果，这一结果明显要好于图 10.36(c)中的结果。使用基本算法算出的阈值是 169，而使用 Otsu 方法算出的阈值是 182，后者更接近图像中定义为细胞的较亮区域。可分离性测度 η^* 是 0.467。

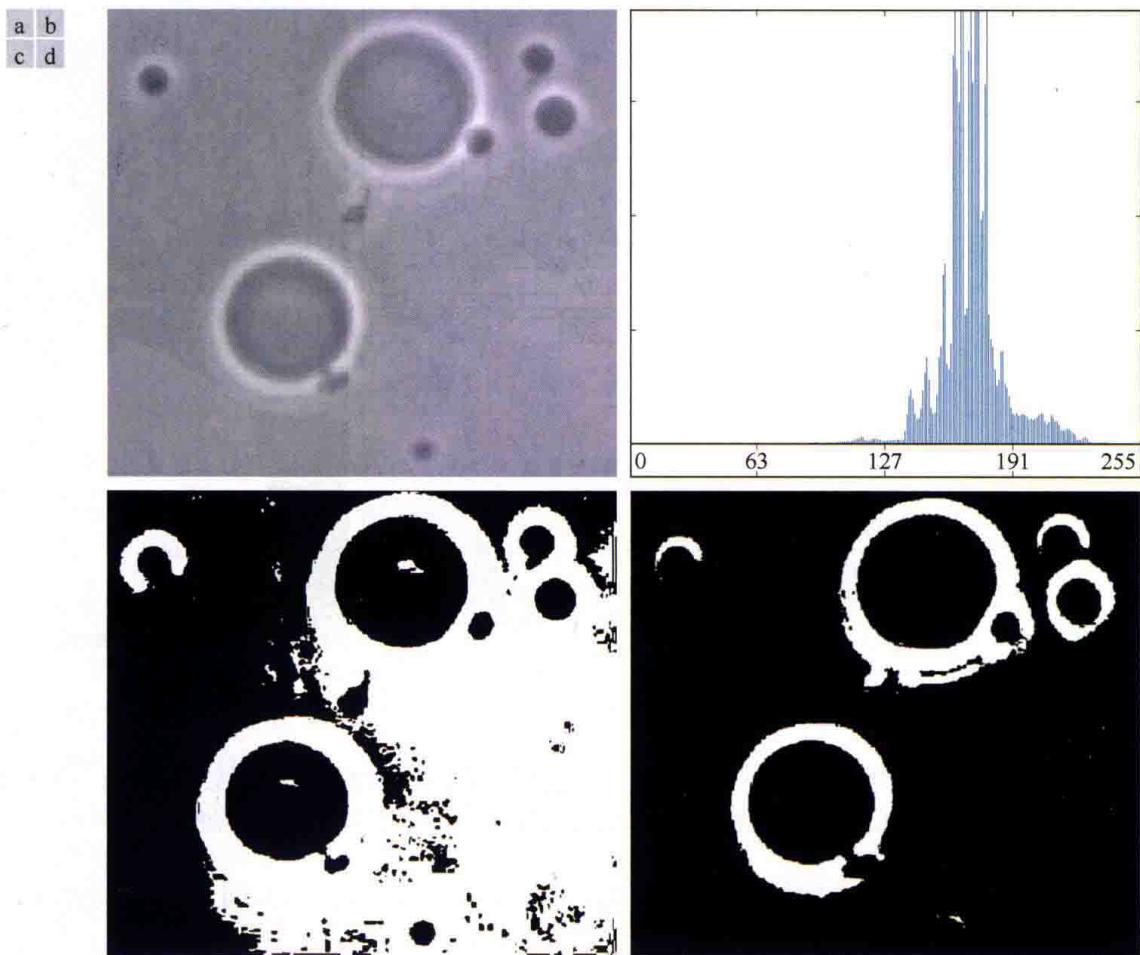


图 10.36 (a)原图像；(b)直方图（为突出低值中的细节，高峰值已被裁剪）；(c)用 10.3 节的基本全局算法得到的分割结果；(d)用 Otsu 方法得到的结果（原图像由宾夕法尼亚大学的 Daniel A. Hammer 教授提供）