

# Introdução ao Processamento de Imagem Digital (MO443/MC920)

Prof. Alexandre Xavier Falcão

Segundo semestre de 2003

## 1 Segmentação por limiarização

Analisando o histograma de uma imagem cinza  $\hat{I} = (D_I, I)$  percebemos que muitas vezes os objetos de interesse são representados por elevações separáveis por intervalos  $[l_i, h_i]$ ,  $i = 1, 2, \dots, k$ , de cinza disjuntos. Nessas situações, a segmentação por limiarização gera uma imagem  $\hat{L} = (D_I, L)$ , onde  $L(p) = i$  se  $l_i \leq I(p) \leq h_i$ .

A única característica usada na limiarização é o brilho dos pixels. Múltiplas características são usadas quando os intervalos de brilho não são disjuntos. A extensão desta abordagem para  $n > 1$  características—denominada método do paralelepípedo—requer a análise de um histograma  $n$ -dimensional, onde os objetos devem ser separáveis por hipercubos com  $n$  arestas e uma aresta de hipercubo é definida por um intervalo no eixo da característica correspondente.

O problema fica ainda mais complicado se a superfície de separação tiver forma mais complexa que a de um hipercubo. Técnicas de classificação estatística são aplicadas nessas situações, onde cada elevação no histograma é interpretada como uma aproximação da densidade de probabilidade do vetor de características no objeto correspondente. Outra alternativa seria aplicar uma transformada dual de watershed no espaço  $\mathfrak{R}^n$  para separar as elevações. Observe que tanto a classificação estatística quanto a transformada de watershed requerem que o aumento no número de características seja acompanhado de um aumento considerável no número de amostras que contribuem para o histograma  $n$ -dimensional. Porém, o número de amostras (pixels da imagem) é fixo. Por outro lado, essas amostras no espaço  $\mathfrak{R}^n$  de características formam aglomerados (*clusters*) e os objetos podem ser identificados pela separação desses aglomerados usando técnicas de *clustering*.

Para simplificar, vamos considerar nesta seção problemas envolvendo a segmentação da imagem em duas classes: objeto e fundo. Este é o caso, por exemplo, da segmentação de imagens de texto, imagens de células, e imagens de realce de bordas.

### 1.1 Limiarização ótima

Considere uma imagem cinza  $\hat{I} = (D_I, I)$  com  $0 \leq I(p) \leq L - 1$  para todo  $p \in D_I$ , duas classes de interesse,  $C_1$  e  $C_2$ , representando objeto e fundo, e as seguintes definições.

- $P_1$  é a probabilidade *a priori* de ocorrer  $C_1$ .
- $P_2$  é a probabilidade *a priori* de ocorrer  $C_2$ .
- $p_1(l)$  é a densidade de probabilidade do nível de cinza  $l$  em  $C_1$ .
- $p_2(l)$  é a densidade de probabilidade do nível de cinza  $l$  em  $C_2$ .
- $p(l)$  é a densidade de probabilidade do nível de cinza  $l$  na imagem.

Portanto,

$$p(l) = P_1 p_1(l) + P_2 p_2(l), \quad (1)$$

e pela regra de decisão de Bayes, um pixel  $p$  deve ser classificado como classe 1, se  $P_1 p_1(l) > P_2 p_2(l)$ , e como classe 2, se  $P_1 p_1(l) < P_2 p_2(l)$ , ficando a igualdade a critério da implementação. Esta regra minimiza a probabilidade de erro assumindo como limiar de decisão  $P_1 p_1(l) = P_2 p_2(l)$ .

Supondo que  $p_1(l)$  e  $p_2(l)$  são distribuições Gaussianas com médias  $\mu_1$  e  $\mu_2$ , e variâncias  $\sigma_1^2$  e  $\sigma_2^2$  conhecidas,

$$p_1(l) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(l - \mu_1)^2}{2\sigma_1^2}\right] \quad (2)$$

$$p_2(l) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(l - \mu_2)^2}{2\sigma_2^2}\right], \quad (3)$$

e que  $0 \leq \mu_1 < \mu_2 \leq L - 1$ , os intervalos de brilho que identificam cada classe podem ser encontrados da seguinte forma.

$$\begin{aligned} P_1 p_1(T) &= P_2 p_2(T) \\ \frac{P_1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(T - \mu_1)^2}{2\sigma_1^2}\right] &= \frac{P_2}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(T - \mu_2)^2}{2\sigma_2^2}\right] \\ \ln\left(\frac{P_1}{\sigma_1}\right) - \frac{(T - \mu_1)^2}{2\sigma_1^2} &= \ln\left(\frac{P_2}{\sigma_2}\right) - \frac{(T - \mu_2)^2}{2\sigma_2^2} \end{aligned}$$

Continuando o desenvolvimento chegamos à equação:

$$AT^2 + BT + C = 0, \quad (4)$$

onde

$$\begin{aligned} A &= (\sigma_1^2 - \sigma_2^2) \\ B &= 2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2) \\ C &= \left[ \mu_2^2\sigma_1^2 - \mu_1^2\sigma_2^2 + 2\sigma_1^2\sigma_2^2 \ln\left(\frac{P_1\sigma_2}{\sigma_1 P_2}\right) \right], \end{aligned}$$

cujas raízes são

$$\begin{aligned} T' &= \frac{-B - \sqrt{B^2 - 4AC}}{2A} \\ T'' &= \frac{-B + \sqrt{B^2 - 4AC}}{2A}. \end{aligned}$$

Temos, portanto, 5 possibilidades.

- Caso 1:  $B^2 - 4AC < 0$ . O problema não tem solução. Isto ocorre, por exemplo quando  $P_2 p_2(l) \ll P_1 p_1(l)$  para todo  $l \in [0, L - 1]$ .
- Caso 2:  $T' \in [0, L - 1]$  e  $T''$  não. Neste caso,  $C_1 : [0, T']$  e  $C_2 : [T' + 1, L - 1]$ .
- Caso 3:  $T'' \in [0, L - 1]$  e  $T'$  não. Neste caso,  $C_1 : [0, T'']$  e  $C_2 : [T'' + 1, L - 1]$ .
- Caso 4:  $T' < T''$  e  $T', T'' \in [0, L - 1]$ . Se  $T' < \mu_1$ , então  $C_1 : [T' + 1, T'']$  e  $C_2 : [0, T'] \cup [T'' + 1, L - 1]$ . Caso contrário,  $C_1 : [0, T'] \cup [T'' + 1, L - 1]$  e  $C_2 : [T' + 1, T'']$ .
- Caso 5:  $T = T' = T''$  e  $T', T'' \in [0, L - 1]$ .
  - Se  $B^2 - 4AC = 0$ , então  $T = \frac{-B}{2A} = \frac{\mu_2 \sigma_1^2 - \mu_1 \sigma_2^2}{\sigma_1^2 - \sigma_2^2}$ . Neste caso,  $C_1 : [0, T]$  e  $C_2 : [T + 1, L - 1]$ .
  - Se  $\sigma_1 = \sigma_2 = \sigma$ , então  $A = 0$  e  $BT + C = 0$  implica que  $T = \frac{-C}{B}$ . Isto é,

$$T = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2 \ln(P_1/P_2)}{\mu_2 - \mu_1}. \quad (5)$$

Observe que se  $P_1 = P_2$ ,  $T = \frac{\mu_1 + \mu_2}{2}$ .

## 1.2 Limiarização por modelo de aproximação

O único problema da abordagem anterior é que não conhecemos os valores de  $\mu_1, \mu_2, \sigma_1, \sigma_2, P_1$  e  $P_2$ . Sabemos apenas que  $P_1 + P_2 = 1$ . Uma alternativa é aproximar as distribuições Gaussianas usando otimização numérica. Considerando  $h(l)$  o histograma normalizado da imagem. Isto é, uma aproximação de  $p(l)$ . Temos que o erro quadrático médio é

$$E(P_1, \mu_1, \mu_2, \sigma_1, \sigma_2) = \frac{1}{L} \sum_{l=0}^{L-1} |p(l) - h(l)|^2. \quad (6)$$

Seja  $v = \begin{bmatrix} P_1 \\ \mu_1 \\ \mu_2 \\ \sigma_1 \\ \sigma_2 \end{bmatrix}$ . A derivada  $\frac{dE}{dv} = \frac{2}{L} \sum_{l=0}^{L-1} |p(l) - h(l)| \begin{bmatrix} dp(l)/dP_1 \\ dp(l)/d\mu_1 \\ dp(l)/d\mu_2 \\ dp(l)/d\sigma_1 \\ dp(l)/d\sigma_2 \end{bmatrix}$ . Conhecemos o histograma  $h(l)$ , e  $p(l)$  é dada pela Equação 1, onde  $P_2$  pode ser substituída por  $(1 - P_1)$ . As

derivadas da Equação 1 com relação aos parâmetros são

$$dp(l)/dP_1 = p_1(l) - p_2(l) \quad (7)$$

$$dp(l)/d\mu_1 = P_1 p_1(l) \frac{(l - \mu_1)}{\sigma_1^2} \quad (8)$$

$$dp(l)/d\mu_2 = (1 - P_1) p_2(l) \frac{(l - \mu_2)}{\sigma_2^2} \quad (9)$$

$$dp(l)/d\sigma_1 = P_1 p_1(l) \frac{1}{\sigma_1} \left[ \frac{(l - \mu_1)^2}{\sigma_1^2} - 1 \right] \quad (10)$$

$$dp(l)/d\sigma_2 = (1 - P_1) p_2(l) \frac{1}{\sigma_2} \left[ \frac{(l - \mu_2)^2}{\sigma_2^2} - 1 \right], \quad (11)$$

onde  $p_1(l)$  e  $p_2(l)$  são dadas pelas Equações 2 e 3. A idéia é partir de um valor inicial  $v^{(0)}$ , para a iteração  $t = 0$ , e aplicar iterativamente a equação

$$v^{(t+1)} = v^{(t)} - \alpha \frac{dE}{dv}, \quad (12)$$

onde  $0 < \alpha < 1$  é uma constante, até minimizarmos  $E(P_1, \mu_1, \mu_2, \sigma_1, \sigma_2)$ . Por exemplo, podemos usar o método de limiarização de Otsu para encontrar  $v^{(0)}$ .

Otsu utiliza uma função critério

$$J(T) = \frac{P_1(T)P_2(T)[\mu_1(T) - \mu_2(T)]^2}{P_1(T)\sigma_1^2(T) + P_2(T)\sigma_2^2(T)}, \quad (13)$$

onde

$$\begin{aligned} P_1(T) &= \sum_{l=0}^T h(l) \\ P_2(T) &= 1 - P_1(T) \\ \mu_1(T) &= \frac{1}{P_1(T)} \sum_{l=0}^T lh(l) \\ \mu_2(T) &= \frac{1}{P_2(T)} \sum_{l=T+1}^{L-1} lh(l) \\ \sigma_1^2(T) &= \frac{1}{P_1(T)} \sum_{l=0}^T [l - \mu_1(T)]^2 h(l) \\ \sigma_2^2(T) &= \frac{1}{P_2(T)} \sum_{l=T+1}^{L-1} [l - \mu_2(T)]^2 h(l). \end{aligned}$$

Variando  $T$  de 0 a  $L - 1$ , o objetivo é encontrar o limiar  $T'$  para o qual  $J(T')$  é máxima. Otsu usa este limiar para separar as classes  $C_1 : [0, T']$  e  $C_2 : [T' + 1, L - 1]$ . No caso do modelo de

aproximação, podemos adotar  $v^{(0)} = \begin{bmatrix} P_1(T') \\ \mu_1(T') \\ \mu_2(T') \\ \sigma_1(T') \\ \sigma_2(T') \end{bmatrix}$ .

Observe que as técnicas acima foram aplicadas para toda imagem. Nada impede que a imagem seja dividida em blocos de pixels e uma regra de limiarização seja definida para cada bloco.

### 1.3 Limiarização por reconstrução morfológica

Em muitas situações o histograma da imagem não é bimodal, mas o objeto de interesse é mais claro (ou mais escuro) do que os valores dos pixels em uma vizinhança ao redor do objeto. A reconstrução morfológica pode ser usada para transformar o histograma em bimodal, sendo o objeto representado por uma elevação e o fundo por outra, e as técnicas acima são aplicadas para concluir a tarefa de segmentação.

Supondo que o objeto é mais escuro que sua vizinhança imediata, podemos aplicar uma IFT com uma semente no interior do objeto e função de custo de caminho  $f_{srec}$

$$\begin{aligned} f_{srec}(\langle q \rangle) &= I(q), \text{ se } q \in S, \text{ e } +\infty \text{ no caso contrário.} \\ f_{srec}(\pi \cdot \langle p, q \rangle) &= \max\{f_{srec}(\pi), I(q)\}. \end{aligned} \tag{14}$$

O mapa de custos da IFT é o resultado desta reconstrução morfológica superior, onde o fundo ficará com intensidade igual ou maior que a da vizinhança do objeto.