Introdução ao Processamento de Imagem Digital (MO443/MC920)

Prof. Alexandre Xavier Falcão

Primeiro semestre de 2005

1 Detecção de bordas

Esta seção apresenta duas outras abordagens para detecção do contorno fechado que descreve uma borda de objeto, como alternativa a técnica de perseguição de borda apresentada na aula 23. A primeira faz uma classificação binária dos pixels em borda ou fundo, e os segmentos de borda resultantes desta operação são ligados para formar o contorno final. A segunda parte de um contorno inicial que se deforma segundo um funcional de energia até aderir à borda do objeto na situação de energia mínima.

1.1 Classificação binária

Uma suavização para reduzir ruído seguida de um realce de bordas gera a imagem inicial para a segmentação. Uma alternativa é aplicar a técnica de segmentação por limiarização (aula 21) quando a imagem de realce é a magnitude de um vetor gradiente. Outra alternativa interessante, adotada por Marr & Hildret, usa como imagem de realce o resultado da convolução da imagem original pelo **negativo** do operador Laplaciano de uma Gaussiana $g(x,y)$ de média zero e desvio padrão σ.

$$g(x,y) = \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right). \tag{1}$$

Substituindo $r^2 = x^2 + y^2$, o Laplaciano $-\nabla^2 g(r)$ é o negativo da derivada segunda de $g(r)$ com relação a r.

$$-\nabla^2 g(r) = \left(\frac{\sigma^2 - r^2}{\sigma^4}\right)\exp\left(\frac{-r^2}{2\sigma^2}\right). \tag{2}$$

Por exemplo, para $\sigma = 1$, podemos gerar a máscara de convolução $\begin{bmatrix} -0.37 & 0 & -0.37 \\ 0 & 1 & 0 \\ -0.37 & 0 & -0.37 \end{bmatrix}$. Na imagem de realce, os pixels classificados como borda são aqueles em que ocorre uma transição de valor positivo para negativo, ou vice-versa, com relação ao valor de pelo menos um pixel adjacente— use adjacência circular de raio σ, e quanto maior for σ mais grossa será a borda.

Lembre-se que para evitar convolução com máscaras grandes podemos aplicar a filtragem em freqüência.

Uma vez classificados os pixels como borda ou fundo, a imagem binária resultante apresenta "buracos" nas bordas. Uma idéia interessante pode ser usar esses segmentos binários de borda como sementes da IFT e aplicar uma perseguição de bordas na imagem de realce. A seguir apresentamos duas outras técnicas de tratamento do problema.

1.1.1 Limiarização por histeresis

A limiarização por histeresis se baseia em dois limiares de classificação binária para a imagem de realce de bordas, um limiar inferior e outro superior. A regra para classificar um pixel como borda é ele ter realce ou maior que o limiar superior ou maior que o limiar inferior, desde que conexo a outro pixel cujo realce é maior que o limiar superior. Apesar de ser uma técnica local e simples, a limiarização por histeresis elimina várias bordas espúrias reduzindo buracos, e tem sido adotada em alguns operadores clássicos, como o operador de Canny.

1.1.2 Transformada de Hough

O princípio básico da transformada de Hough é mapear os pixels classificados como borda para um espaço de parâmetros de uma dada equação. Pixels que pertencem a segmentos de borda e que satisfazem a esta equação para um dado conjunto de parâmetros devem formar um aglomerado de pontos no espaço de parâmetros. A localização deste aglomerado de pontos permite identificar na imagem de bordas quais pixels pertencem à borda de interesse, eliminando os segmentos espúrios. A própria equação pode ser usada para fechar os buracos entre os segmentos selecionados.

O caso mais simples é a identificação de segmentos de reta. Todos os pixels que satisfazem a equação de uma reta $y = a'x + b'$ são mapeados no ponto (a', b') do espaço de parâmetros $a \times b$. Como não conhecemos a' e b', a idéia é quantizar o espaço de parâmetros para possíveis valores de a e b, depois acumular em cada posição (a, b) o número de pixels e quais pixels satisfizeram a equação $y = ax + b$. Os pixels da reta, portanto, formarão um aglomerado mais alto neste histograma bidimensional em torno de (a', b').

Um problema, porém, é que a e b assumem valores infinitos para retas verticais. Isto requer uma mudança de parâmetros para expressar a equação da reta como

$$\rho = x \cos\theta + y \sin\theta, \tag{3}$$

onde ρ é a distância da reta à origem do espaço $x \times y$ (isto é, $\rho = |\vec{n}|$, onde $\vec{n}$ é ortogonal à reta) e θ é o ângulo entre o vetor $\vec{n}$ e o eixo x. Neste caso, os valores de ρ e θ são quantizados de 0 até o valor D da diagonal da imagem e de $0°$ a $90°$, respectivamente. Uma matriz de acumulação $A(\rho, \theta)$ (histograma bidimensional) e uma lista de pixels $P(\rho, \theta)$ podem ser usadas para calcular a seguinte transformação.

Transformada de Hough para detecção de linhas:

Entrada: Imagem binária de bordas $\hat{I} = (D_I, I)$, onde $I(p) = 1$ indica borda e $I(p) = 0$ é fundo.
Saída: Matriz de acumulação $A(\rho, \theta)$ inicializada com zeros e lista de pixels $P(\rho, \theta)$ vazia.

1. Para cada valor de $\theta = 0, 1, \ldots, 90$ faça
2. Para todo pixel $p = (x_p, y_p) \in D_I$, tal que $I(p) = 1$, faça
3. $\rho \leftarrow round(x_p \cos\theta + y_p \sin\theta)$.
4. $A(\rho, \theta) \leftarrow A(\rho, \theta) + 1$.
5. insere p em $P(\rho, \theta)$.

Observe que as várias retas que passam por um mesmo pixel (x, y) viram senóides no espaço $\rho \times \theta$.

1.2 Contornos deformáveis

Contornos deformáveis são curvas paramétricas que se deslocam com o objetivo de minimizar uma função objetivo. O valor da função deve ser mínimo na borda do objeto a ser segmentado.

Entre os modelos de contornos deformáveis disponíveis na literatura, o modelo *Snakes* foi o primeiro proposto e é o mais citado.

1.2.1 Teoria

Uma *snake* é um contorno do tipo spline de continuidade controlada sob a influência de forças externas e forças internas. As forças internas atuam impondo suavidade ao contorno, enquanto as forças externas empurram a *snake* para características da imagem que se assemelhem a bordas ou segmentos de bordas. Pode ser definida no plano como um contorno paramétrico fechado composto de pontos, $\mathbf{v(s)} = (\mathbf{x(s)}, \mathbf{y(s)})$, onde $(\mathbf{x}, \mathbf{y}) \in \mathbf{R}^2$, $\mathbf{s} \in [0, 1]$ e $\mathbf{v(0)} = \mathbf{v(1)}$. A equação geral da energia deste modelo é escrita em termos das Energias Interna e Externa:

$$E_{snake}(v(s)) = \int_0^1 E_{int}(v(s)) + E_{ext}(v(s))\, ds \tag{4}$$

onde,

$$E_{int}(v(s)) = \frac{\alpha(s)\|\frac{\partial v}{\partial s}\|^2 + \beta(s)\|\frac{\partial^2 v}{\partial s^2}\|^2}{2} \tag{5}$$

e

$$E_{ext}(v(s)) = E_{image}(v(s)). \tag{6}$$

O objetivo é encontrar a solução que minimiza $\mathbf{E}_{snake}$. O termo $\mathbf{E}_{int}$ representa a Energia Interna do contorno cujo diferencial gera a Força Interna que impõe regularidade ao contorno.

Os termos $\frac{\partial v}{\partial s}$ e $\frac{\partial^2 v}{\partial s^2}$ representam as derivadas de primeira e segunda ordem, respectivamente. O coeficiente $\alpha(s)$ é o termo que controla a "tensão" entre os pontos do contorno, enquanto $\beta(s)$ é responsável pela "elasticidade" do contorno. A Energia Externa é calculada com base no gradiente da imagem, para que a *snake* seja atraída para pontos de borda. A Força Externa que executa esta ação pode ser escrita como:

$$F_{ext}(v(s)) = -\|\nabla(G_\sigma * I(v(s)))\|^2 \tag{7}$$

onde ∇ representa o operador gradiente, e $\mathbf{G}_\sigma * \mathbf{I}(\mathbf{v}(\mathbf{s}))$ denota a imagem trabalhada por um filtro de suavização gaussiano com desvio padrão σ.

A força interna, quando atua sozinha, pode levar a curva a tornar-se um único ponto, e por isto deve ser balanceada pela força externa, resultante do diferencial da Energia Externa.

1.2.2 Discretização

Com base em (4), (5) e (7), deve-se achar o contorno $v(\mathbf{s})$ que minimiza o funcional de energia. Para resolver este problema, é usado o cálculo variacional para derivar um sistema dinâmico (uma equação diferencial parcial) e chegar no estado de energia mínima no equilíbrio, usando programação dinâmica para achar a posição da *snake* que minimiza a energia discretizada. Para facilitar a solução, os coeficientes $\alpha(s)$ e $\beta(s)$ são considerados constantes.

O cálculo do extremo do funcional

$$E_{snake}(v(s)) = \int_0^1 \frac{\alpha\|\frac{\partial v}{\partial s}\|^2 + \beta\|\frac{\partial^2 v}{\partial s^2}\|^2}{2} + F_{ext}(v(s))\, ds \tag{8}$$

resulta na equação de Euler-Lagrange

$$F_{ext}(v(s)) - \frac{\partial}{\partial s}\alpha\frac{\partial v}{\partial s} + \frac{\partial^2}{\partial s^2}\beta\frac{\partial^2 v}{\partial s^2} = 0 \tag{9}$$

que é igual a

$$\alpha\frac{\partial^2 v}{\partial s^2} - \beta\frac{\partial^4 v}{\partial s^4} - F_{ext}(v(s)) = 0 \tag{10}$$

A curva que minimiza o funcional de energia deve satisfazer a equação de Euler-Lagrange (10). Para tornar a solução dinâmica, o contorno $\mathbf{v}$ é tratado como uma função de $\mathbf{s}$ e tempo $\mathbf{t}$:

$$\frac{\partial v(s,t)}{\partial t} = \alpha\frac{\partial^2 v(s,t)}{\partial s^2} - \beta\frac{\partial^4 v(s,t)}{\partial s^4} - F_{ext}(v(s,t)) \tag{11}$$

Quando a *snake* estabiliza, $\frac{\partial v(s,t)}{\partial t} = 0$ e temos a solução de (10).

Para a discretização, o método de diferenças finitas é usado, sendo h um passo pequeno ou diferença de s, e

- $v_i = v(ih) = (x(ih), y(ih))$
- $\mathbf{F_{ext}} = (F_x, F_y)$

Aplicando as diferenças finitas para frente e para trás alternadamente,

$$f'(x) = f(x+1) - f(x) \qquad (12)$$
$$f'(x) = f(x) - f(x-1)$$

a equação (10) pode ser discretizada em:

$$\begin{aligned} &\alpha(v_i - v_{i-1}) - \alpha(v_{i+1} - v_i) \\ &+\beta(v_{i-2} - 2v_{i-1} + v_i) - 2\beta(v_{i-1} - 2v_i + v_{i+1}) \\ &+\beta(v_i - 2v_{i+1} + v_{i+2}) + \mathbf{F_{ext}} = 0 \end{aligned} \qquad (13)$$

Esta equação pode ser escrita na forma matricial:

$$\mathbf{A}x + \mathbf{F}_x = 0 \qquad (14)$$
$$\mathbf{A}y + \mathbf{F}_y = 0$$

sendo

$$\mathbf{A} = \begin{pmatrix} A & B & C & 0 & 0 & 0 & \cdots & B \\ B & A & B & C & 0 & 0 & \cdots & C \\ C & B & A & B & C & 0 & \cdots & 0 \\ 0 & C & B & A & B & C & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & A & B & C \\ C & 0 & 0 & 0 & \cdots & B & A & B \\ B & C & 0 & 0 & \cdots & C & B & A \end{pmatrix} \qquad (15)$$

onde

$$\begin{aligned} A &= 2\alpha + 6\beta \\ B &= -\alpha - 4\beta \\ C &= \beta \end{aligned} \qquad (16)$$

A equação (14) pode ser resolvida iterativamente da seguinte forma (assumindo que $\mathbf{F}$ não muda em um passo de tempo):

$$\mathbf{A}x(t) + \mathbf{F}_x(t-1) = -\gamma(x(t) - x(t-1)) \qquad (17)$$
$$\mathbf{A}y(t) + \mathbf{F}_y(t-1) = -\gamma(y(t) - y(t-1)) \qquad (18)$$

onde γ é um intervalo de tempo pequeno. Após rearranjar os termos, temos:

$$x(t) = (\mathbf{A} + \gamma\mathbf{I})^{-1}(\gamma x(t-1) - \mathbf{F}_x(t-1)) \qquad (19)$$
$$y(t) = (\mathbf{A} + \gamma\mathbf{I})^{-1}(\gamma y(t-1) - \mathbf{F}_y(t-1)).$$

Usando as equações acima, $x(t)$ e $y(t)$ podem ser calculados iterativamente através de $x(t-1)$, $y(t-1)$, $\mathbf{F}_x(t-1)$, e $\mathbf{F}_y(t-1)$.

1.3 Algumas dificuldades de implementação

O modelo apresentado para contornos deformáveis apresenta algumas dificuldades de implementação.

1. Inicialização do contorno.

   O contorno inicializado em regiões homogêneas de brilho tende a se contrair, portanto o contorno deve ser inicializado fora do objeto de interesse. Mesmo assim, caso o contorno seja inicializado longe da borda de interesse, ele pode ficar preso em mínimos locais.

2. Estabilidade e convergência.

   O contorno pode oscilar em torno de bordas "fracas". A escolha dos parâmetros α e β é fundamental. Contornos com muita rigidez não aderem a identações e protusões da borda e contornos muito flexíveis tendem a passar por bordas "fracas".

3. Auto-cruzamento.

   Durante o movimento, o contorno pode formar "laços" cruzando a si próprio. Isto requer uma reamostragem dos pontos eliminando os laços.

4. Mudança de topologia.

   O modelo apresentado só trata o caso de uma única borda de interesse por vez.

A literatura de contornos deformáveis é extensa e consiste basicamente de trabalhos que procuram resolver as dificuldades supracitadas. Neste aspecto, métodos de segmentação baseados em região são mais simples de serem implementados.