

Processamento digital de imagens

Agostinho Brito

Departamento de Engenharia da Computação e Automação
Universidade Federal do Rio Grande do Norte

14 de abril de 2016

Segmentação de imagens

- A segmentação subdivide uma imagem em suas partes constituintes ou objetos.
- É uma das tarefas mais complicadas em processamento de imagens.
- Algoritmos são baseados nas propriedades de descontinuidade e similaridade.
 - Descontinuidade: fronteiras das regiões são diferentes.
 - Similaridade: pixels em uma região pode ser aglomerados por critério semelhante.
- Exemplos:
 - Detecção de pontos e linhas
 - Detecção de arestas
 - Aglomeração de regiões

Segmentação de imagens

O processo de segmentação divide uma região R em n subregiões, $R_1, R_2, R_3, \dots, R_n$, de modo que:

- 1 $\bigcup_{i=1}^n R_i = R$
- 2 R_i é um conjunto conectado, para $i = 1, 2, \dots, n$.
- 3 $R_i \cap R_j = \emptyset$, para todo i e j , $i \neq j$.
- 4 Uma propriedade P aplicada a uma região R_i satisfaz à relação $P(R_i) = \text{VERDADEIRO}$ para $i = 1, 2, \dots, n$.
- 5 Para esta mesma propriedade, $P(R_i \cup R_j) = \text{FALSO}$ para duas regiões adjacentes R_i e R_j .

Detecção de pontos, linhas e bordas.

- Detecção de pontos: uso de derivadas de segunda ordem (laplaciano)

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

- Pode-se assumir que um ponto ocorre quando

$$g(x, y) = \begin{cases} 1 & \text{se } |R(x, y)| \geq T \\ 0 & \text{caso contrário} \end{cases}$$

- $R(x, y)$ é o resultado da resposta da máscara no ponto central (x, y) de uma região e T é um limiar não negativo que regula a intensidade da resposta que será considerada como ponto isolado. Ex: T igual a 90% do maior valor absoluto presente na imagem.

Detecção de linhas:

- Uso do laplaciano com valores positivos da resposta (evita linhas grossas)
- Detecção de linhas em direções específicas

-1	-1	-1
2	2	2
-1	-1	-1

Horizontal

2	-1	-1
-1	2	-1
-1	-1	2

$+45^\circ$

-1	-1	2
-1	2	-1
2	-1	-1

-45°

-1	2	-1
-1	2	-1
-1	2	-1

Vertical

- Avaliando-se a resposta destas quatro máscaras, é possível escolher a mais proeminente como candidata a detectar a linha existente em uma direção.
- Etapas
 - Suavização da imagem (atenuação de ruído)
 - Detecção de pontos de borda (extração de candidatos potenciais)
 - Localização da borda (selecionar apenas os que realmente pertencem a uma borda)

Detecção de bordas

- Uso do vetor gradiente

$$\begin{aligned}\nabla f &= \text{mag}(\nabla \mathbf{f}) \\ &= [G_x^2 + G_y^2]^{1/2} \\ &= \left[\left(\frac{\partial f}{\partial y} \right)^2 + \left(\frac{\partial f}{\partial x} \right)^2 \right]^{1/2} \\ &\simeq |G_x| + |G_y|\end{aligned}$$

- Uso dos operadores de Prewitt (um pouco mais rápido) e Sobel (melhor suavização).

Filtros avançados

- Detetor de borda de Marr-Hildreth

- Mudanças de intensidade são independentes de escala.
- Uso do operador laplaciano do gaussiano ($\nabla^2 G$), ou LoG.
- Encontrar os cruzamentos de zero. Para uma matrix 3x3, um pixel é de cruzamento se pelo menos os sinais de dois de seus vizinhos opostos (norte/sul, leste/oeste, diagonais) são diferentes.
- Bordas são de 1 pixel de espessura.
- Um limiar para o cruzamento de zero que seja diferente de “zero” pode ser escolhido. Ex: percentual do máximo valor do LoG.
- Cerca de 99,7% do volume de uma gaussiana encontra-se em $\pm 3\sigma$ ao redor da média. Um filtro de tamanho $n \times n$ deve ser escolhido tal que $n < 6\sigma$. Ex: $\sigma = 4 \rightarrow n = 25$.

$$\nabla^2 G = \frac{\partial^2 G(x, y)}{\partial x^2} + \frac{\partial^2 G(x, y)}{\partial y^2}$$

$$G(x, y) = \exp^{-\frac{x^2+y^2}{2\sigma^2}} \quad \nabla^2 G(x, y) = \left[\frac{x^2 + y^2 - \sigma^2}{\sigma^4} \right] \exp^{-\frac{x^2+y^2}{2\sigma^2}}$$

- Negativo do LoG para máscara 5×5

0	0	-1	0	0
0	-1	-2	-1	0
-1	-2	16	-2	-1
0	-1	-2	-1	0
0	0	-1	0	0

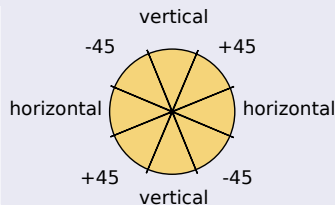
Detector de bordas de Canny

- Objetivos:
 - Baixa taxa de erros
 - Pontos de borda devem estar bem localizados
 - Resposta de um único ponto de borda

Etapas do detector

- Convolução com o filtro Gaussiano, cálculo da magnitude e ângulo do gradiente.
- Afinação das cristas largas do gradiente.
 - Classificação dos pontos quanto às orientações Horizontal, Vertical, $+45^\circ$, -45° (intervalos de $\pm 22.5^\circ$).
 - Para os vizinhos na orientação determinada para o pixel, verificar os seus gradientes.
 - Supressão de não máximos: se o valor da magnitude do gradiente $M(x, y)$ for inferior a pelo menos um de seus vizinhos, faça $g_N(x, y) = 0$; caso contrário, faça $g_N(x, y) = M(x, y)$. $g_N(x, y)$ é a imagem com supressão.

Direções dos ângulos



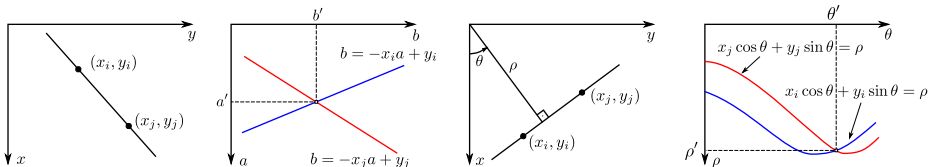
Detector de bordas de Canny (cont.)

- Limiarização com histerese é usada para a quebra do contorno (borda tracejada).
 - Dois limiares T_1 e T_2 . $T_1 > T_2$ são usados.
 - Se o pixel é tal que $g_N(x, y) \geq T_1$, é assumido como ponto de borda forte.
 - Para os pixels restantes, aqueles em que $g_N(x, y) \geq T_2$, são assumidos como ponto de borda fraco.
 - Para todos os vizinhos dos pontos de borda fraco, procurar nos seus 8-vizinhos se há algum ponto de borda forte. Caso haja, este é marcado como parte da fronteira.
 - Sugestão de Canny: $T_H/T_L = 3/1$ ou $T_H/T_L = 2/1$

Segmentação de imagens

Transformada de Hough

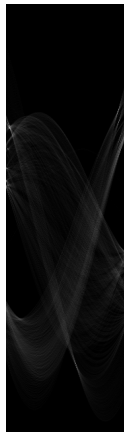
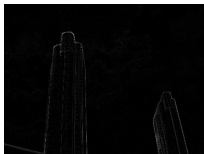
- Tem por objetivo encontrar as equações das retas presentes em uma imagem.
- Considere dois pontos (x_i, y_i) e (x_j, y_j) . Apenas uma reta da forma $y_i = ax_i + b$ passa por ambos os pontos.
- Se a equação da reta for reescrita da forma $b = -ax_i + y_i$, a representação no espaço de parâmetros mostra uma interseção das novas retas formadas por (x_i, y_i) e (x_j, y_j) .
- Problema: retas se aproximando da inclinação vertical ($a \rightarrow \infty$).
- Solução: representação em coordenadas polares.



Transformada de Hough

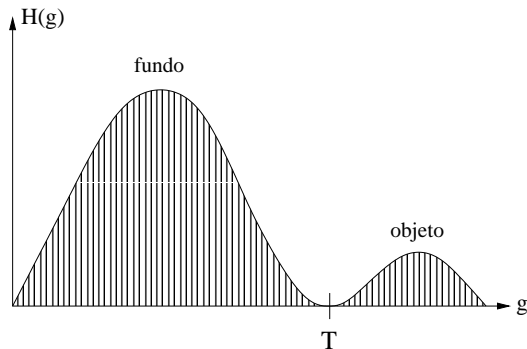
- Uso de matriz acumuladora $A(p, q)$ que subdivide o espaço de parâmetros $\rho\theta$.
- A matriz assume variações $-90^\circ \leq \theta \leq 90^\circ$ e $-D \leq \rho \leq D$, onde D é o diâmetro da imagem.
- Inicia-se o processo de cálculo da transformada com $A(p, q) = 0$.
- O processamento é realizado em uma imagem de borda (Ex: obtida pelo filtro de Canny).
- Para cada pixel de borda, uma função paramétrica é traçada. Para cada valor de θ incrementando, arredonda-se o valor de ρ resultante para a posição mais próxima na matriz, (p, q) . Acrescenta-se em 1 o valor de $A(p, q)$, ou seja $A(p, q) = A(p, q) + 1$.
- Examinar a matriz acumuladora em busca de máximos.
- Examinar a conectividade entre pixels de uma reta associada a uma célula escolhida.

Transformada de Hough



Limiarização

- Histogramas de intensidade



Limiarização global simples

- Selecionar uma estimativa global para o limiar, T .
- Segmentar a imagem em dois grupos de pixels usando T .
- Calcular as médias de tons de cinza de cada grupo, m_1 e m_2 .
- Calcular um novo valor de limiar

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

- Repetir a segmentação até que a diferença entre os limiares seja menor que um parâmetro pré-definido ΔT .

Método de Otsu

- Assume que o histograma pode ser representado por duas funções de densidade de probabilidade.
- Procura determinar um limiar ótimo de modo a maximizar a variância entre classes.
- Para um dado limiar k , a variância entre classes pode ser escrita como

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

ou

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

- $m(k)$ é a média acumulada até o nível k .
- m_G é a média geral dos pixels na imagem.
- Para n_i pixels no nível i , $P_1(k)$ é dado pela soma cumulativa

$$P_1(k) = \sum_{i=0}^k p_i = \sum_{i=0}^k \frac{n_i}{MN}$$

Método de Otsu: etapas

- Calcular o histograma normalizado da imagem de entrada, p_i .
- Calcular as somas acumuladas $P_i(k)$
- Calcular as médias acumuladas $m(k)$.
- Calcular a média global m_G .
- Calcular a variância entre classes $\sigma_B^2(k)$, para $k = 0, 1, 2, \dots, L - 1$ (L níveis de cinza).
- Procurar nos valores de $\sigma_B^2(k)$ aquele que manifesta valor máximo. Se mais de um máximo for encontrado, o limiar ótimo será a média dos valores de k que produziram os máximos.
- Caso contrário, escolha k como limiar.

Melhorando a segmentação

- Suavização prévia
- Uso de informação de bordas para processar o histograma
- Limiarização local
 - Particionando a imagem em blocos
 - Realizando a limiarização em uma vizinhança de tamanho fixo.
 - Usando médias de movimento (segmentação de documentos)
 - z_{k+1} : intensidade do ponto encontrado na etapa de digitalização $k + 1$.
 - média neste novo ponto

$$m(k+1) = \frac{1}{n} \sum_{i=k+2-n}^{k+1} z_i = m(k) + \frac{1}{n}(z_{k+1} - z_{k-n})$$

- Uso do limiar $T_{xy} = bm_{xy}$, com $b = 0,5$, por exemplo. O valor de n é, via de regra, 4 a 5 vezes a largura média do traço.