

Processamento de Imagens

Resumo: Detecção de Arestas

Data: 28/11/2000

1 Detecção de Arestas

Estas notas apenas oferecem um resumo dos métodos de detecção de arestas apresentados em sala. Estas notas apenas resumem os métodos apresentados; quem quiser saber um pouco mais pode consultar as referências abaixo ([Nalwa] não entra em detalhes, mas dá um apanhado geral da teoria sem muita matemática).

A parte de discretização de derivadas esta em um outro conjunto de notas de aula e, portanto, não é discutida aqui.

Daqui por diante, seja $I(x, y) : \mathbb{R} \rightarrow \mathbb{R}^2$ uma imagem. Queremos encontrar arestas dentro desta imagem, isto é, (pedaços de) curvas através das quais há variações súbitas do nível de cinza (de I).

1.1 Gradient Thresholding

São métodos baseados em encontrar pontos (x, y) tais que

$$|\nabla I(x, y)| \geq C$$

onde o nível de corte C é uma constante (a ser determinada). Ao invés de produzir curvas, este método produz regiões de pontos que devem passar então por um processo de **thinning**. Por exemplo, pode-se usar a técnica de **nonmaximum supression**, isto é, eliminam-se pontos que não são máximos de $|\nabla I|$ na direção do gradiente ∇I .

A discretização deste processo envolve o cálculo discreto de ∇I em cada pixel da imagem; é comum usar a expressão $|\nabla I|^2 = I_x^2 + I_y^2$ para tanto; a discretização das derivadas parciais é feita através de uma convolução com uma máscara discreta (como descrito no outro conjunto de notas de aula).

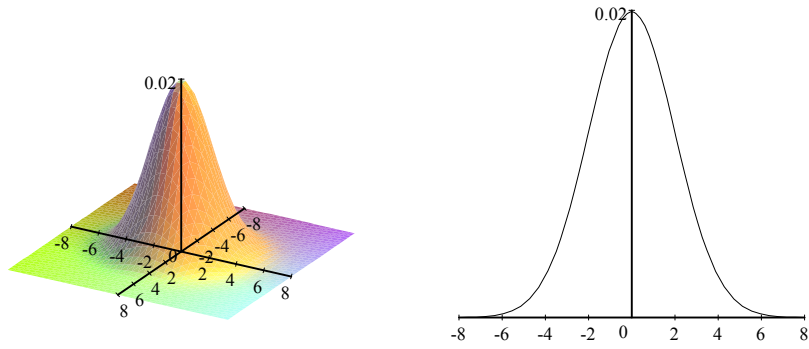
Canny apresentou um refinamento dessa idéia: o uso de dois cortes $C_1 > C_2$. O maior, C_1 , seria usado para começar a geração de arestas, enquanto o segundo seria usado para continuá-las em pixels vizinhos aos já declarados como pixels-de-aresta.

Como as derivadas parciais amplificam as altas freqüências da imagem, o gradiente é sensível a ruídos de alta freqüência. Tal sensibilidade pode ser amenizada com um filtro de passa-baixa (Gaussiano, por exemplo). Assim, é comum utilizar-se¹

$$|\nabla(G_t * I)| \geq C$$

¹Usamos G_t para uma Gaussiana de variância $\sigma^2 = 2t$; o uso desta notação faz com que G_t satisfaça a equação do calor: $\frac{\partial G_t}{\partial t} = \nabla^2 G_t$, como veremos logo a seguir.

ou versões discretas desta expressão para diminuir a sensibilidade ao ruído.



Gaussiana de variância 4 ($t = 2$)...

...e sua seção $x = 0$.

1.2 Laplaciano (operador de Marr-Hildreth)

O método de Marr-Hildreth propõe encontrar pontos onde o Laplaciano é nulo, isto é,

$$\nabla^2 I = \Delta I = I_{xx} + I_{yy} = 0$$

usando diversas possíveis notações. Para ser mais exato, apenas os **cruzamentos de zero** do Laplaciano seriam declarados como arestas. O Laplaciano é um operador linear e invariante por translações e portanto pode ser implementado como uma convolução. Ele também tem a vantagem de ser rotacionalmente simétrico (uma propriedade desejável na detecção de arestas) e conter a informação da direção da transição (cruzamentos positivo-para-negativo correspondem a arestas escuro-para-claro).

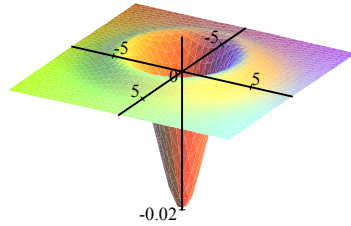
Novamente, o método é muito sensível a ruídos de alta frequência. Tal sensibilidade pode ser amenizada usando ao invés

$$\Delta (G_t * I) = \Delta G_t * I = 0$$

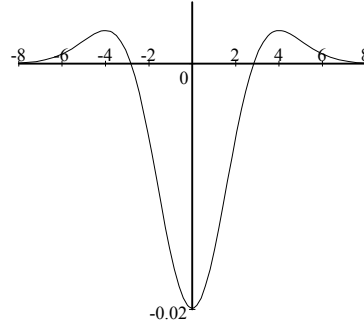
Note como isto pode ser calculado como uma simples convolução com o núcleo

$$\Delta G_t = \Delta \left(\frac{1}{4\pi t} e^{-\frac{1}{4t}(x^2+y^2)} \right) = \frac{x^2 + y^2 - 4t}{16\pi t^3} e^{-\frac{1}{4t}(x^2+y^2)}$$

cujo gráfico é exibido a seguir.



$\Delta G_2 (\sigma^2 = 4) \dots$



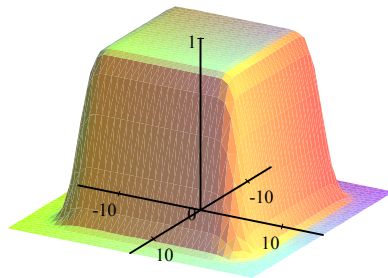
\dots e sua seção $x = 0$.

O grande problema deste método é a sua pobre localização, especialmente quando a aresta tem alta curvatura (e, portanto, perto de cantos e junções). O exemplo a seguir mostra isto

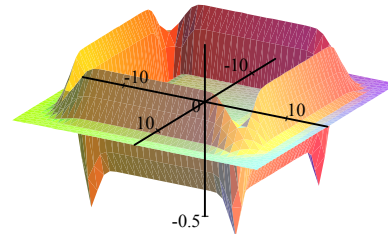
Exemplo 1 *Por exemplo, começando de um quadrado 10×10*

$$I(x, y) = \begin{cases} 1, & \text{se } |x| < 10 \text{ e } |y| < 10 \\ 0, & \text{caso contrário} \end{cases}$$

e convoluindo-o com uma Gaussiana de variância 1, obtemos a seguinte imagem e Laplaciano (exibidas aqui como gráficos de funções):

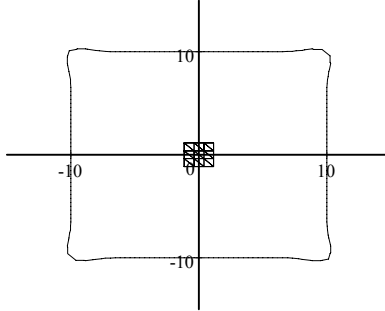


$G_{0.5} * I$



\dots e seu Laplaciano.

A seguir, mostramos as curvas de nível correspondentes a $\Delta(G_{0.5} * I) = 0$. Note que apenas a curva exterior corresponde de fato a um cruzamento de zero; como observado anteriormente, veja como é pobre a localização de arestas perto dos cantos do quadrado original – enquanto o arredondamento dos cantos é esperado devido à suavização pela Gaussiana, os cruzamentos de zero “escapam” do quadrado perto dos cantos:



Zeros de $\Delta(G_1 * I)$.

É bom ressaltar que a localização pobre não vem de ruído ou discretizações (os cálculos do exemplo acima foram realizados analiticamente; a discretização e a quantização só acontecem no momento da criação dos gráficos), mas de uma deficiência intrínseca do operador Laplaciano. Mesmo assim, o operador Laplaciano é muito usado para detecção de arestas.

Um outro método análogo ao do Laplaciano é o uso de núcleos de convolução que são diferenças de Gaussianas de variâncias próximas (filtros DOG ou *Difference Of Gaussians*). Em outras palavras, as arestas seriam os cruzamentos de zero de

$$(G_{t_1} - G_{t_2}) * I$$

onde t_1 e t_2 são em geral valores próximos um do outro. Para ver que este método é semelhante ao do Laplaciano, note que

$$\begin{aligned} G_\alpha &= \frac{1}{4\pi\alpha} e^{-\frac{1}{4\alpha}(x^2+y^2)} \Rightarrow \\ \Rightarrow \frac{\partial G_\alpha}{\partial \alpha} &= \frac{x^2 + y^2 - 4\alpha}{16\pi\alpha^3} e^{-\frac{1}{4\alpha}(x^2+y^2)} = \Delta G_\alpha \end{aligned}$$

Portanto, $G_{t_1} - G_{t_2} \approx (t_1 - t_2) \frac{\partial G_t}{\partial t} = (t_1 - t_2) (\Delta G_t)$ é um múltiplo de ΔG_t para $t_2 \approx t_1$, e seus cruzamentos de zero são os mesmos do método de Marr-Hildreth.

1.3 Método de Haralick

Uma maneira mais correta de realizar o método de Gradient Thresholding seguido de Nonmaximum Supression em uma operação consiste em encontrar diretamente máximos de $|\nabla I|$ na direção de ∇I . Em outras palavras, defina a função

$$g(x, y) = |\nabla I(x, y)|^2 = I_x^2 + I_y^2$$

(o quadrado aqui existe apenas para facilitar as contas). Queremos máximos de g na direção $\vec{v} = \frac{\nabla I}{|\nabla I|}$ do gradiente de I . Em outras palavras, precisamos

encontrar pontos onde $\partial_v g$ **cruxa zero**, onde $\partial_v g$ é a derivada direcional de g na direção de \vec{v} . Mas

$$\partial_v g = \nabla g \cdot \vec{v}$$

pois \vec{v} é unitário. Como

$$\begin{aligned} \nabla g &= \partial_x (I_x^2 + I_y^2) \vec{i} + \partial_y (I_x^2 + I_y^2) \vec{j} = \\ &= (2I_{xx}I_x + 2I_{xy}I_y) \vec{i} + (2I_{xy}I_x + 2I_{yy}I_y) \vec{j} \end{aligned}$$

e

$$\vec{v} = \frac{\nabla I}{|\nabla I|} = \frac{I_x \vec{i} + I_y \vec{j}}{\sqrt{I_x^2 + I_y^2}}$$

procuramos pontos onde

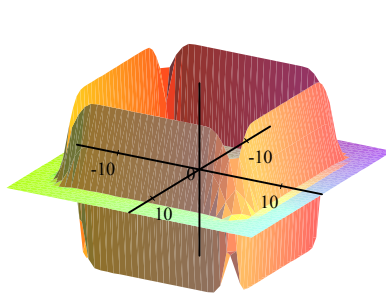
$$\nabla g \cdot \vec{v} = \frac{2}{\sqrt{I_x^2 + I_y^2}} (I_x^2 I_{xx} + 2I_x I_y I_{xy} + I_y^2 I_{yy}) = 0$$

isto é, o método de Haralick sugere que encontremos os cruzamentos de zero de

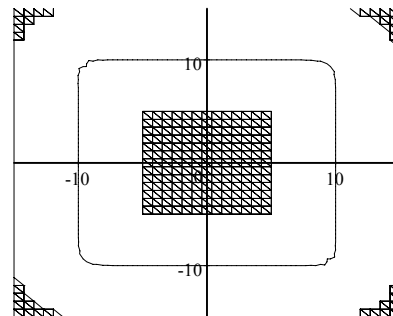
$$H \{I\} = I_x^2 I_{xx} + 2I_x I_y I_{xy} + I_y^2 I_{yy}$$

(desde que $|\nabla I| \neq 0$).

Exemplo 2 A figura abaixo à esquerda mostra o resultado deste operador aplicado a $G_{0,5} * I$, onde I é o quadrado 10×10 descrito na seção anterior. À direita, mostramos os zeros de $H \{G_{0,5} * I\}$; note que apenas a curva central corresponde de fato a **cruzamentos de zero** (os outros pontos satisfazem $|\nabla(G_{0,5} * I)| = 0$ e seriam removidos a seguir). Note como os cantos são arredondados mas não “escapam” do quadrado inicial.



$H \{G_{0,5} * I\} \dots$



\dots e seus zeros.

1.4 Métodos por Funcionais de Energia

Há uma gama de métodos que se baseam em minimização por funcionais de energia. Inicialmente, considere o seguinte exemplo de suavização de sinais.

Exemplo 3 Dada uma função $f : [a, b] \rightarrow \mathbb{R}$ considere os funcionais

$$E_1 \{g\} = \int_a^b (g(x) - f(x))^2 dx$$

$$E_2 \{g\} = \int_a^b (g'(x))^2 dx$$

para uma função diferenciável qualquer $g : [a, b] \rightarrow \mathbb{R}$. Note que E_1 mede quanto g se distancia de f e E_2 é uma medida de quanto g é “bem comportada”. Para obter uma aproximação suave de f , podemos considerar a função g que minimiza

$$E \{g\} = (E_1 + \alpha E_2) \{g\}$$

onde o parâmetro α controla qual propriedade é mais importante: obter uma função suave ou ficar perto de f . Assim, se $\alpha = 0$, uma solução possível é $g = f$ (se f for diferenciável); por outro lado, à medida que $\alpha \rightarrow \infty$, $g(x) = K$ (constante) torna-se uma das melhores soluções. Para encontrar g matematicamente, precisamos utilizar o cálculo das variações e encontrar uma equação diferencial para g , neste caso

$$g - \alpha g'' = f$$

Uma idéia semelhante pode ser usada para segmentação de sinais unidimensionais – o chamado “Weak String Model”.

Definio 4 Dada uma função $f : [a, b] \rightarrow \mathbb{R}$, o funcional de energia do **weak string model** é dado por

$$E \{g, S\} = \mu^2 \int_a^b (g(t) - f(t))^2 dt + \int_{[a,b]-S} (g'(t))^2 dt + \nu \cdot \#(S)$$

onde S é um conjunto de saltos $\{t_1, t_2, \dots, t_n\}$ tal que $a = t_0 < t_1 < t_2 < \dots < t_n < t_{n+1} = b$; g é uma função diferenciável em cada intervalo da forma $[t_i, t_{i+1}]$ (mas possivelmente descontínua em cada t_k); $\#(S) = n$ é simplesmente o número de elementos de S . Note que podemos ajustar a fidelidade ao sinal f (através do ajuste do parâmetro μ) e a quantidade de arestas desejados (através do ajuste de ν).

A idéia do modelo acima é encontrar g e S que minimizem $E \{g, S\}$; então, S será considerado o conjunto de arestas do sinal f original enquanto g será uma reconstrução suave-por-partes de f que atende a tal segmentação.

O caso bidimensional é semelhante, mas o funcional de energia se parece com

$$E \{g, S\} = \mu^2 \int_D (g - f)^2 dA + \int_{D-S} |\nabla g|^2 dA + \nu \cdot e(S)$$

onde S é em geral um conjunto de curvas e $e(S)$ pode ser tomado como a soma de seus comprimentos (há outras opções para $e(S)$).

O caso discreto unidimensional é mais simples. O funcional se torna

$$E\{g, S\} = \mu^2 \sum_j (g_j - f_j)^2 + \sum_{j \notin S} (g_{j+1} - g_j)^2 + \nu \cdot \#(S)$$

onde S é um conjunto de índices, cada um indicando a presença de um aresta entre f_j e f_{j+1} . Este funcional é normalmente re-escrito como

$$E\{g, l\} = \mu^2 \sum (g_j - f_j)^2 + \sum (g_{j+1} - g_j)^2 (1 - l_j) + \nu \sum l_j$$

onde l_j é a função característica de S , isto é, $l_j = 1$ se $j \in S$ e $l_j = 0$ quando $j \notin S$. O problema é então encontrar a seqüência g_j e a seqüência binária l_j que minimizem o funcional $E\{g, l\}$. Surpreendentemente, existem algoritmos de programação dinâmica que encontram g e l rapidamente neste caso discreto unidimensional (dada a seqüência inicial f).

Os casos contínuos e o caso discreto multidimensional são bem mais difíceis. Uma idéia é começar de $\{g, S\}$ quaisquer e mudá-los lentamente (localmente) para que a energia $E\{g, S\}$ caia. Tais métodos encontram mínimos locais do funcional E , mas não há garantia que os mínimos encontrados sejam próximos ao mínimo global. Resultados melhores são obtidos se mudanças bruscas (globais) são adicionadas esporadicamente, cada vez menos freqüentemente – a idéia básica do “simulated annealing”. Uma discussão mais profunda deste tópico foge ao escopo de nosso curso.

1.5 Difusão Linear e Difusão Não-Linear

Nesta seção, tocaremos levemente nas idéias que levam a modelos de difusão não-linear. De fato, estes modelos não são exatamente para detectar arestas diretamente, mas correspondem a um pré-processamento a ser aplicado na imagem que a suaviza (eliminando a sensibilidade a ruído) enquanto mantendo a nitidez de suas arestas.

Começemos notando que o processo de convolução com uma Gaussiana pode ser visto como um processo de difusão linear. De fato, dada uma imagem $I : \mathbb{R}^2 \rightarrow \mathbb{R}$, considere as várias imagens obtidas a partir de convoluções com Gaussianas de diferentes variâncias, isto é

$$\begin{aligned} L(x, y; t) &= G_t * I \\ G_t &= \frac{1}{4\pi t} e^{-\frac{1}{4t}(x^2+y^2)} \\ G_0 &= \delta(x, y) \text{ (por extensão)} \end{aligned}$$

(iamgine que $t \in \mathbb{R}^+$); note que

$$\frac{\partial L}{\partial t} = \frac{\partial (G * I)}{\partial t} = \frac{\partial G}{\partial t} * I = \Delta G * I = \Delta (G * I) = \Delta L$$

onde usamos novamente a propriedade citada no final da segunda seção (onde comentamos os filtros DOG). Em outras palavras, podemos pensar em L como a solução da seguinte equação diferencial linear com condições iniciais

$$\begin{aligned} L(x, y; 0) &= I(x, y) \\ \frac{\partial L}{\partial t} &= \Delta L = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) L \text{ (para } t \geq 0) \end{aligned}$$

Esta equação diferencial é conhecida como **equação do calor** ou **processo de difusão linear**. O nome vem do seguinte modelo físico: se uma placa metálica² uniforme (isolada do mundo externo) tem temperatura inicial (em $t = 0$) dada por $L(\cdot; 0) = I(\cdot)$, sua temperatura num determinado instante t é dada por $L(\cdot; t)$ satisfazendo a equação diferencial acima.

O mesmo modelo vale no caso unidimensional (para uma barra metálica):

$$\left. \begin{aligned} \frac{\partial L}{\partial t} &= \frac{\partial^2 L}{\partial x^2} \\ L(x, 0) &= I(x) \end{aligned} \right\} \Leftrightarrow L(x, t) = \frac{1}{\sqrt{4\pi t}} \left(e^{-\frac{1}{4t}x^2} \right) * I(x)$$

Como vimos nas seções anteriores, o processo de difusão linear suaviza a imagem por igual em todas as direções, inclusive através de seus arestas. Enquanto a suavização é necessária, alguns arestas que desejamos detectar podem ser deslocados durante esta suavização, ou mesmo eliminados.

Em geral, se a barra não é uniforme, podemos caracterizá-la termicamente pela sua condutância térmica $C(x)$. O modelo de difusão torna-se uma equação não-linear

$$\begin{aligned} \frac{\partial L}{\partial t} &= \frac{\partial}{\partial x} \left(C(x) \frac{\partial L}{\partial x} \right) = \frac{\partial C}{\partial x} \frac{\partial L}{\partial x} + C \frac{\partial^2 L}{\partial x^2} \\ L(x, 0) &= I(x) \end{aligned}$$

onde $I(x)$ é novamente a distribuição inicial de calor na barra. Note que $C(x)$ é alto nas partes da barra onde o calor se espalha rapidamente, e pequeno nas áreas mais “isolantes” com respeito ao calor. Se $C(x) = 0$ numa parte da barra então $\frac{\partial L}{\partial t} = 0 \Rightarrow L(x, t) = I(x)$ e a distribuição de calor, por mais desigual que seja, não muda com o tempo t .

A versão bidimensional é semelhante

$$\begin{aligned} \frac{\partial L}{\partial t} &= \frac{\partial}{\partial x} \left(C(x, y) \frac{\partial L}{\partial x} \right) + \frac{\partial}{\partial y} \left(C(x, y) \frac{\partial L}{\partial y} \right) = \nabla C \cdot \nabla L + C \cdot \Delta L \\ L(x, y; 0) &= I(x, y) \end{aligned}$$

(o termo à direita mostra como a equação acima é rotacionalmente simétrica).

Aqui está o centro da idéia de Perona e Malik: ao invés de tomarmos a “condutância térmica” como constante e chegarmos à difusão linear, por que não colocarmos uma condutância menor onde esperamos encontrar arestas, isto

²Exceto pelo apelo visual, não há necessidade de se pensar que a placa é de fato **metálica**.

é, onde o gradiente de L é grande? Assim, a imagem L se suavizaria apenas nos pontos onde a suavização é necessária, mantendo os arestas nítidos para uma análise via um detector de aresta. Por exemplo, podemos tentar

$$C(x, y) = \frac{1}{1 + |\nabla I|^2}$$

já evitando condutâncias infinitas. Infelizmente, pode-se mostrar que a equação diferencial que se obtém com esta expressão de $C(x, y)$ é mal comportada, isto é, ela amplifica instantaneamente qualquer ruído de alta frequência.

Para regularizar esta situação, pode-se tomar

$$C(x, y) = \frac{1}{1 + |\nabla(G_t * I)|^2}$$

e esta é a proposta de Perona e Malik. A discretização da equação diferencial é razoavelmente imediata (apenas precisamos dos operadores $\frac{\partial}{\partial x}$ e $\frac{\partial}{\partial y}$, e então precisamos decidir como discretizar $\frac{\partial}{\partial t}$) e oferece resultados ininteressantes. A presença da Gaussiana no denominador de $C(x, y)$ não deve assustar – afinal de contas, qualquer discretização de $|\nabla I|$ já envolveria uma suavização de qualquer forma.

Há outras alternativas para a escolha de $C(x, y)$; Nitzberg e Shiota, por exemplo, propuseram uma equação diferencial que não só preserva arestas mas também procura enfatizá-las espalhando níveis de cinza até os lados dos arestas em potencial. Novamente, uma análise mais profunda destes métodos escapa ao nosso escopo.

1.6 Referências

- [Nalwa] Vishvjit S. Nalwa, “*A Guided Tour of Computer Vision*”, Addison-Wesley, 1993.
- [Canny] J. Canny, “*A Computational Approach to Edge Detection*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No.6, November 1986, pp. 679–698
- [Marr] D. Marr and E. Hildreth, “*Theory of Edge Detection*”, Proceedings of the Royal Society of London, Series B, Vol. 207, 1980, pp.187–217
- [Haralick] R. M. Haralick, “*Digital Step Edges from Zero Crossing of Second Directional Derivatives*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, No.1, January 1984, pp. 58–68
- [Torre] Vincent Torre and Tomaso A. Poggio, “*On Edge Detction*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No.2, March 1986
- [Perona] P. Perona and J. Malik, “*Scale-Space and Edge Detection Using Anisotropic Diffusion*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No. 7, July 1990, pp. 629–639