Capítulo 2

Detecção de Pontos de Borda

Neste trabalho, estamos interessados na extração de bordas de objetos em imagens digitais. Neste sentido, o problema inicial é a determinação dos pontos que podem pertencer às bordas (genericamente denominados *pontos de bordas* (*edges*)), os quais são caracterizados por uma abrupta variação da imagem (Jain, 1989) nas suas vizinhanças.

Neste capítulo faremos uma abordagem conceitual e formal para o problema de detecção de pontos de borda em imagens bidimensionais sem o objetivo de desenvolver técnicas especificas no tema. A apresentação que segue tem por finalidade principal colocar o problema da detecção de pontos de borda dentro do contexto das técnicas de snakes.

Seja então a Figura 2.1 onde mostramos os modelos (unidimensionais) usuais para os tipos de pontos de borda comumente encontrados em processamento de imagem (Canny, 1983).

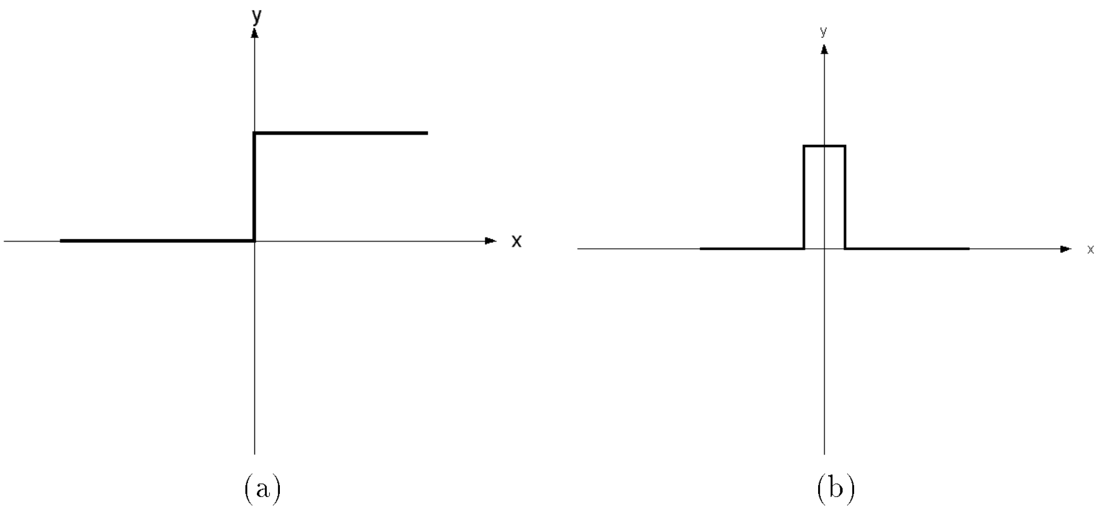


Figura 2.1: Modelos unidimensionais para pontos de borda.

Estas figuras representam modelos ideais. As funções (sinais) que obtemos na prática são em geral irregulares e apresentam ruídos.

A Figura 2.2 é um exemplo de uma situação mais próxima da prática. Nesta figura vemos um sinal com um degrau acentuado entre $x = 90$ e $x = 100$.

Uma possibilidade para caracterizar este comportamento é tomar os extremos locais de df/dx, ou seja, pontos de inflexão.

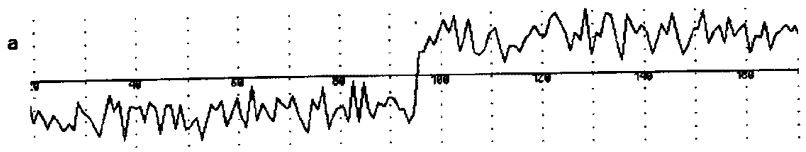


Figura 2.2: Sinal irregular f com ruido.

No entanto, na Figura 2.2 existem vários pontos de inflexão que não representam degraus. Assim, métodos que simplesmente localizam extremos locais da derivada fornecem falsos pontos de borda, inclusive nas proximidades do degrau procurado, o que implica em mais de uma resposta para um mesmo ponto.

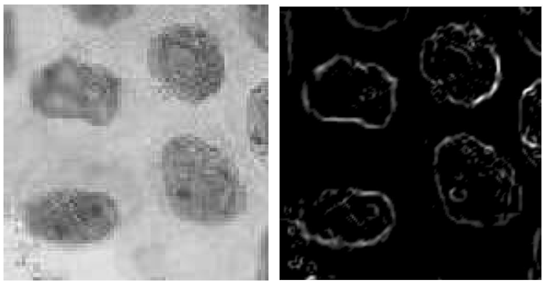
Estas considerações se estendem para o caso 2D. Neste caso, dada uma imagem, precisamos considerar também uma orientação além da taxa de variação da imagem em um dado ponto.

Uma referência fundamental no contexto de detecção dos pontos de borda é o trabalho de Canny (Canny, 1983). Partindo de considerações como as que fizemos acima, Canny estabelece inicialmente três critérios para avaliação do desempenho de um detetor: (a)baixa probabilidade de marcar $falsos$ pontos de borda ou de perder pontos de borda; (b)os pontos detectados devem estar próximos do centro da borda (ponto de inflexão no caso $1D$) correspondente; (c) fornecer resposta única para cada ponto de borda da imagem. Em seguida, estes critérios são desenvolvidos quantitativamente para enfim obter-se um detetor ótimo via métodos variacionais.

Segundo Canny (Canny, 1983), podemos classificar os métodos para detecção de pontos de bordas de acordo com as seguintes características: a) utilização de uma superfície que aproxima a imagem; b) métodos que estimam derivadas, usualmente primeira ou segunda derivadas da imagem; c) utilização de técnicas no domínio da freqüência.

Dentre estes métodos, aqueles do item (b) serão os enfocados nesta seção pela sua relação direta com o método de snakes (Black & Yuille, 1993).

Uma prática comum neste caso, é utilizar um limiar de corte para o valor da derivada (gradiente) com o objetivo apenas de *realçar* as regiões onde se encontram as bordas dos objetos de interesse (Jain, 1989), como mostra a figura 2.3. O mesmo efeito pode ser obtido via *zero-crossings* da segunda derivada (máximos locais da primeira derivada) (Jain, 1989).



(a) (b)

Figura 2.3: (a)Imagem original. (b)Realce das bordas via limiar para o gradiente.

A seguir, daremos uma formulação geral para a caracterização de pontos de borda baseada nas derivadas da imagem, ou mais precisamente, na Primeira Forma Fundamental associada à mesma.

2.1 Descrição Matemática

Partindo da idéia de que pontos de borda são pontos com alta variação da imagem (*alto contraste*), o primeiro passo será aplicar um modelo matemático que traduza este conceito quantitativamente.

Uma possibilidade neste sentido é usar a Primeira Forma Fundamental (Sapiro, 1997) associada à imagem e usar o fato de que os autovalores (autovetores) da mesma em um dado ponto fornecem as taxas (direções) de máxima e mínima variação da imagem no ponto, respectivamente (Dubrovin *et al.* , 1984).

Seja então

$$\Phi\left(u_1, u_2\right) : D \subset \Re^2 \rightarrow \Re^m, \tag{2.1}$$

uma imagem multi-variada com m componentes:

$$\Phi_i\left(u_1, u_2\right) : D \subset \Re^2 \rightarrow \Re, i = 1, ..., m. \tag{2.2}$$

Como exemplo de imagens multi-variadas podemos citar imagens coloridas codificadas num sistema usual como o RGB ou HLS e imagens multi-canal obtidas por satélites. Sejam agora dois pontos $P = (u_1^0, u_2^0)$ e $Q = (u_1^1, u_2^1)$ e a diferença:

$$\Delta\Phi = \Phi\left(P\right) - \Phi\left(Q\right).$$

da imagem calculada nestes pontos.

Se a distância $d(P,Q)$ for pequena, esta diferença pode ser aproximada pela diferencial

$$d\Phi = \sum_{i=1}^{i=2} \frac{\partial \Phi}{\partial u^i} du^i,$$

O vetor $d\Phi$ tem dimensão m e sua norma Euclideana quadrada é dada por:

$$d\Phi^2 = \sum_{i=1}^{i=2} \sum_{j=1}^{j=2} \left\langle \frac{\partial \Phi}{\partial u^i}, \frac{\partial \Phi}{\partial u^j} \right\rangle du_i du_j. \tag{2.3}$$

onde $<,>$ indica o produto interno usual. Esta forma quadrática é denominada a Primeira Forma Fundamental da superfície correspondente (Dubrovin *et al.* , 1984).

Relembrando conceitos fundamentais em geometria Riemaniana, ressaltemos primeiramente que os coeficientes g_{ij} dados por:

$$g_{ij} = \left\langle \frac{\partial \Phi}{\partial u^i}, \frac{\partial \Phi}{\partial u^j} \right\rangle,$$

formam uma matriz simétrica $[g_{ij}]$, definida-positiva, a qual fornece uma métrica (tensor simétrico de segunda ordem) sobre a superfície correspondente à imagem. Assim podemos calcular normas de vetores tangentes e distância entre pontos sobre a superfície sem qualquer referência ao espaço ambiente (Dubrovin *et al.* , 1984).

Usando a notação padrão em geometria Riemaniana, nós denotamos:

$$d\Phi^2 = \sum_{i=1}^{i=2} \sum_{j=1}^{j=2} g_{ij} du_i du_j = \begin{bmatrix} du_1 \\ du_2 \end{bmatrix}^T \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} du_1 \\ du_2 \end{bmatrix}. \tag{2.4}$$

Para um vetor unitário $\overrightarrow{v} = (\cos\theta, sen\theta)$, $d\Phi^2(\overrightarrow{v})$ indica a taxa de variação da imagem na direção de $\overrightarrow{v}$, a qual é dada por:

$$d\Phi^2(\overrightarrow{v}) = \begin{bmatrix} \cos\theta \\ sen\theta \end{bmatrix}^T \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} \cos\theta \\ sen\theta \end{bmatrix} \tag{2.5}$$

É um fato conhecido (Dubrovin *et al.* , 1984) que os extremos da forma quadrática (2.4) são obtidos nas direções dos autovetores do tensor métrico $[g_{ij}]$, e os valores máximo e mínimo correspondentes são os autovalores associados. Os autovalores são dados por:

$$\lambda_{\pm} = \frac{g_{11} + g_{22} \pm \sqrt{(g_{11} - g_{22})^2 + 4g_{12}^2}}{2}. \tag{2.6}$$

Os autovetores associados (ortogonais entre si) serão denotados por $(\cos\theta_{\pm}, sen\theta_{\pm})$ com as direções $\theta_{\pm}$ dados respectivamente pelas seguintes expressões:

$$\theta_{+} = \frac{1}{2}\arctan\left(\frac{2g_{12}}{g_{11} - g_{22}}\right), \tag{2.7}$$

$$\theta_{-} = \theta_{+} + \frac{\pi}{2}. \tag{2.8}$$

Observe que $\lambda_{\pm}$ são valores não negativos uma vez que $[g_{ij}]$ é positiva definida. O autovetor θ_{+} fornece a direção de máxima variação da imagem em um ponto e o autovetor θ_{-} fornece a direção de mínima variação da imagem neste mesmo ponto. Os autovalores correspondentes fornecem as respectivas taxas de máxima e mínima variação no ponto dado.

Por exemplo, para uma imagem em níveis de cinza $(m = 1)$ teremos:

$$\lambda_{+} = \|\nabla\Phi\|^2, \lambda_{-} = 0, onde\,(\cos\theta_{+}, sen\theta_{+}) \parallel \nabla\Phi, \tag{2.9}$$

recuperando assim a caracterização usual de pontos de borda para imagens deste tipo (Sapiro, 1997).

Contudo, se $m > 1$, o valor de λ_{-} pode não ser identicamente nulo. Uma primeira caracterização para pontos de borda neste caso, análoga a selecionar $\|\nabla\Phi\|^2$ no caso de $m = 1$, seria uma função do tipo $f = f(\lambda_{+} - \lambda_{-})$. Por exemplo, possíveis escolhas para f seriam (Sapiro, 1997):

$$\begin{aligned} f &= (\lambda_{+} - \lambda_{-})^{1/p},\ p > 0 \\ &\text{ou,} \\ f &= \sqrt{\lambda_{+}}. \end{aligned}$$

Analogamente ao caso $m = 1$, os pontos de borda seriam aqueles para os quais $|f(\lambda_{+} - \lambda_{-})|$, ultrapasse um limiar previamente estabelecido (Sapiro, 1997).

Esta caracterização de pontos de borda tem a vantagem de agregar as informações locais da imagem em duas únicas quantidades $\lambda_{\pm}$ dadas pela expressão (2.6) facilitando sua aplicação. Além disso, uma vez que os coeficientes g_{ij} constituem um tensor, esta caracterização de pontos de borda é independente do sistema de coordenadas (isto é, da representação de cor) usado (Sapiro, 1997; Sapiro, 1995).

Embora simples e geral, esta metodologia tem limitações decorrentes da instabilidade numérica que pode ocorrer no cálculo de derivadas para imagens com ruídos (Terzopoulos, 1986b; Poggio *et al.* , 1985; Snyder, 1991).

A seguir analisaremos o problema da instabilidade numérica. No próximo capítulo veremos o método de snakes como uma metodologia para a extração das informações geométricas de interesse do conjunto de pontos detectado.

2.2 Bordas no Espaço de Escalas

No texto que segue, diremos que uma função $f : \Re^m \rightarrow \Re$; $f = f\left(u^{i_1}, u^{i_2}, ..., u^{i_m}\right)$ é de classe C^n se suas derivadas parciais:

$$\frac{\partial^n}{\partial u^{i_1} \partial u^{i_2} ... \partial u^{i_m}};\ i_1 + i_2 + ... + i_m = n,\ i_1, i_2, ..., i_m \in N, \qquad (2.10)$$

são contínuas. Diremos ainda que uma função f é *suave* caso f seja de classe C^∞ e usaremos a notação:

$$N_n^m = \{(i_1, i_2, ..., i_m) \mid i_1 + i_2 + ... + i_m = n,\ i_1, i_2, ..., i_m \in N\}$$

O cálculo das derivadas que surgem em (2.3) sofre em geral de instabilidade numérica devido ruídos na imagem original (Terzopoulos, 1986b; Dolgopolova & Ivanov, 1966). Uma solução bem conhecida para este problema é filtrar (convoluir) a imagem original com um kernel $\mathbf{s}$, satisfazendo as seguintes propriedades (MA & Li, 1998):

1. $s \in C^n$,
2. $s\left(u_1, u_2\right)$ é uma função par de u_1 e u_2. Além disso:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} s\left(u_1, u_2\right) du_1 du_2 = 1,$$

3. Para qualquer $(i, j) \in N_k^2 (k = 1, ..., n)$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \left|u_1\right|^i \left|u_2\right|^j s\left(u_1, u_2\right) du_1 du_2$$

existe e é finita.

Por simplicidade, vamos supor $m = 1$ em (2.1), e assim teremos $s : \Re^2 \rightarrow \Re$ para o kernel acima.

Assim, dado um operador diferencial qualquer $\mathcal{D}$, aplicamos este operador a uma versão *suavizada* $\widetilde{\Phi}$ de Φ dada pela convolução $\Phi * s$:

$$\widetilde{\Phi} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Phi\left(v, w\right) s\left(u_1 - v, u_2 - w\right) dv dw, \tag{2.11}$$

obtendo-se assim:

$$\mathcal{D}\widetilde{\Phi} = \mathcal{D} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Phi\left(v, w\right) s\left(u_1 - v, u_2 - w\right) dv dw, \tag{2.12}$$

Fazendo uso das propriedades diferenciais do kernel s podemos derivar sob o sinal de integração obtendo finalmente:

$$\mathcal{D}\Phi = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Phi\left(v, w\right) \left[\mathcal{D} s\left(u_1 - v, u_2 - w\right)\right] dv dw. \tag{2.13}$$

Por exemplo, se:

$$\mathcal{D} = D_n = \sum_{(i,j) \in N_n^2} b_{ij} \frac{\partial^n}{\partial u_1^i \partial u_2^j}, \tag{2.14}$$

teremos a expressão (2.13) correspondente dada por:

$$D_n\Phi = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Phi\left(v, w\right) \left[\sum_{(i,j) \in N_n^2} b_{ij} \frac{\partial^n}{\partial u_1^i \partial u_2^j} s\left(u_1 - v, u_2 - w\right)\right] dv dw. \tag{2.15}$$

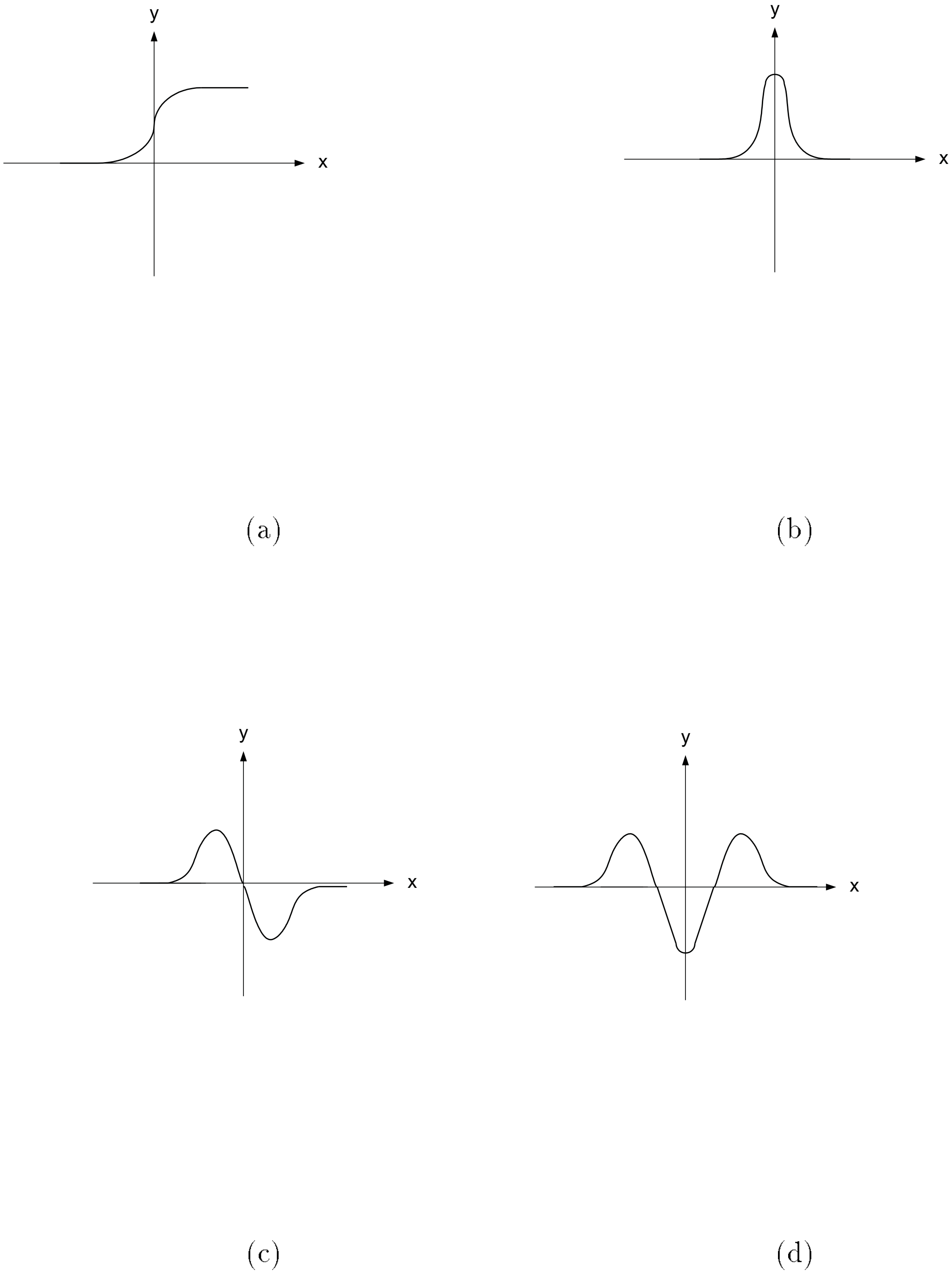
Para o caso 1D representado na Figura 2.1, mostramos a seguir as funções obtidas por (2.15) para $D_n = d^n/dx^n$, $n = 0, 1, 2, 3$, onde D_0 denota o operador identidade. O kernel s pode ser a gaussiana definida logo a seguir em (2.17).

É importante observar que o processo de filtragem dado pela convolução $\Phi * s$ tem implicitamente as seguintes propriedade:

(1) Linearidade: a convolução da soma é a soma das convoluções;

(2) Invariância por Deslocamento (*Spatial shift invariance*): o peso de s na convolução (2.11) depende apenas do vetor diferença $(u_1 - v, u_2 - w)$.

Se a estas propriedades acrescentarmos a imposição de que s tenha simetria circular (*isotropia*) bem como as Leis de Invariância por Escala da Análise Dimensional (Teorema dos π, Apêndice B) então chegamos ao resultado notável de que o *Kernel*



(a) (b)

(c) (d)

Figura 2.4: (a)Degrau da figura 2.1.a suavizado.(b)Primeira derivada.(c)Segunda derivada.(d)Terceira derivada

Gaussiano é o único operador no espaço de escalas (Florack *et al.* , 1992). Seu desvio padrão σ pode ser identificado com escala espacial (Florack *et al.* , 1992).

A importância do kernel Gaussiano é corroborada por indícios biológicos de que a funcionalidade da retina humana pode ser parcialmente modelada por filtros Gaussianos de várias larguras ou por suas derivadas (Florack *et al.* , 1992).

A utilização do Kernel Gaussiano foi o ponto de partida dos métodos de representação multi-escala de imagens (Witkin, 1983). A idéia central destes métodos (Koenderink, 1984) se resume em imergir a imagem original em uma família de imagens derivadas $I(x,y,t)$ obtidas pela convolução da imagem original $I_0(x,y)$ com o Kernel Gaussiano:

$$I\left(x,y,t\right)=G_t*I_0, \tag{2.16}$$

onde:

$$G_\sigma\left(x,y\right)=C\sigma^{-1}\exp\left(-\frac{1}{4\sigma}\left(x^2+y^2\right)\right) \tag{2.17}$$

é a função de Gauss.

Como indicado por Koenderink (Koenderink, 1984) e Hummel (Hummel, 1986) esta família de imagens fornece a solução fundamental da equação (do calor), tomando a imagem original como condição inicial; ou seja (Junior & Iorio, 1988):

$$\frac{\partial I\left(x,y,t\right)}{\partial t}=c\nabla^2 I\left(x,y,t\right) \tag{2.18}$$

$$I\left(x,y,0\right)=I_0\left(x,y\right)$$

onde c é uma constante (denominada constante de difusão). Koenderink desenvolveu sua formulação a partir de dois critérios (1)-(2) acima, bem como do principio de causalidade baixo:

Princípio de Causalidade: não podem ser criados "artefatos" quando aumentamos a escala.

Estes critérios levam naturalmente à equação de difusão.

Os métodos multi-escala procuram localizar as estruturas de interesse (pontos de borda, por exemplo) em uma escala mais grossa para em seguida fazer o *tracking* destas estruturas no espaço de escalas até a resolução desejada (Witten, 1993). Na Figura 2.5 representamos esquematicamente este processo.

Nesta figura, $\sigma_4 > \sigma_3 > ... > \sigma_0$. Identificada uma estrutura de interesse na escala σ_4 faz-se o tracking da mesma no espaço de escalas até a escala mais fina σ_0.

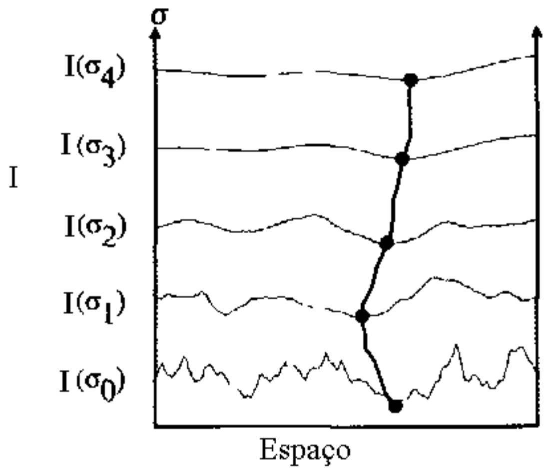
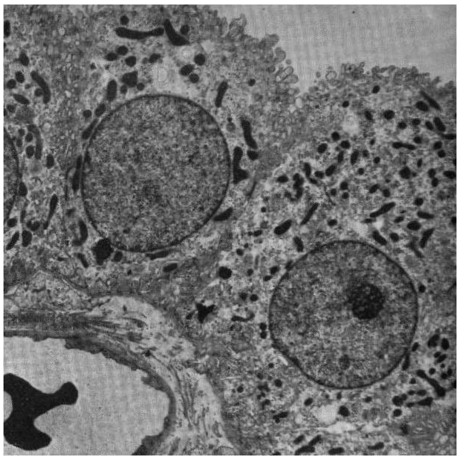
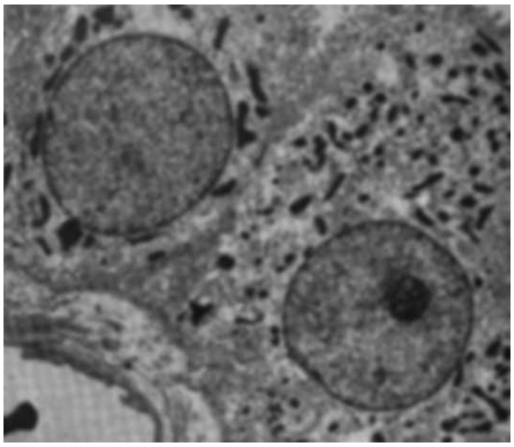


Figura 2.5: Tracking no espaço de escalas.

Esta abordagem para o espaço de escalas baseada na equação de difusão linear (2.18) apresenta o efeito mostrado na figura 2.6.a de "difundir" indiscriminadamente tanto a borda na direção do background da imagem como vice-versa. Este efeito é útil para o método de snakes discutido no próximo capítulo em função dos métodos numéricos a serem utilizados. Mas, no contexto de detecção de pontos de borda, é indesejado por aumentar a imprecisão na localização dos mesmos (Perona & Malik, 1990).



(a)



(b)

Figura 2.6: Célula de Gato: (a)Imagem de Células original.(b)Imagem suavizada por gaussiana.

Esta observação foi ponto de partida para aos métodos multi-escalas não-lineares introduzidos por (Perona & Malik, 1990). Nestes métodos, a idéia central é abandonar a linearidade pressupostas na convolucão (2.11), permitindo que a constante da equação de difusão em (2.18) possa variar no espaço e possivelmente no tempo (Alvarez *et al.* , 1992; Catte *et al.* , 1992). Tem-se agora um problema de valor inicial do tipo:

$$\frac{\partial I(x,y,t)}{\partial t} = div\left(c(x,y,t)\nabla I\right), \tag{2.19}$$

$$I(x,y,0) = I_0(x,y), \tag{2.20}$$

onde por simplicidade vamos supor a imagem I em tons de cinza ($m = 1$ em (2.1)).

A primeira questão nesta nova abordagem é como definir $c(x,y,t)$.

Seguindo a apresentação dada em Perona-Malik (Perona & Malik, 1990), vamos supor que os pontos de borda são alinhados com o eixo y ($I_y = 0$). Assim, sem perda de generalidade, a equação de difusão não-linear (2.19) toma numa vizinhança da borda a forma:

$$\frac{\partial I(x,y,t)}{\partial t} = \frac{\partial}{\partial x}\left(c(x,y,t)\, I_x(x,y,t)\right). \tag{2.21}$$

Vamos escolher c como sendo uma função do gradiente de I: $c(x,y,t) = g(I_x(x,y,t))$. Definindo $\phi(I_x) \equiv g(I_x) \cdot I_x$ podemos colocar a equação acima na forma:

$$I_t = \frac{\partial I}{\partial t} = \frac{\partial}{\partial x}\left(\phi(I_x)\right) = \phi'(I_x) \cdot I_{xx}. \tag{2.22}$$

Nós estamos interessados em analisar a variação no tempo da declividade da borda: $\frac{\partial I_x}{\partial t}$. Para isto, devemos observar primeiro que se $c(x,y,t) > 0$ então a solução I de (2.19) será suave (Perona & Malik, 1990). Logo, a ordem de diferenciação poderá ser invertida; ou seja:

$$\frac{\partial}{\partial t}(I_x) = \frac{\partial}{\partial x}(I_t) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\phi(I_x)\right) = \frac{\partial}{\partial x}\left(\phi'(I_x) \cdot I_{xx}\right). \tag{2.23}$$

$$\phi'' \cdot I_{xx}^2 + \phi' \cdot I_{xxx}.$$

Vamos supor que a borda seja orientada tal que $I_x > 0$ (ver Figura 2.4.a). No ponto de inflexão temos $I_{xx} = 0$, e $I_{xxx} << 0$ uma vez que o ponto de inflexão na Figura 2.4.a corresponde ao ponto com declividade máxima (Figura 2.4.b).

Assim, em uma vizinhança do ponto de inflexão a derivada $\partial I_x / \partial t$ tem sinal oposto a $\phi'(I_x)$. Se $\phi'(I_x) > 0$ a declividade do ponto de borda decresce com o tempo. Caso contrario, se $\phi'(I_x) < 0$ a declividade vai aumentar com o tempo. Assim, a borda torna-se *mais fina*; ou seja,a largura do pulso na Figura 2.4.b diminui. Temos portando mais precisão na localização da borda que é o efeito desejado!

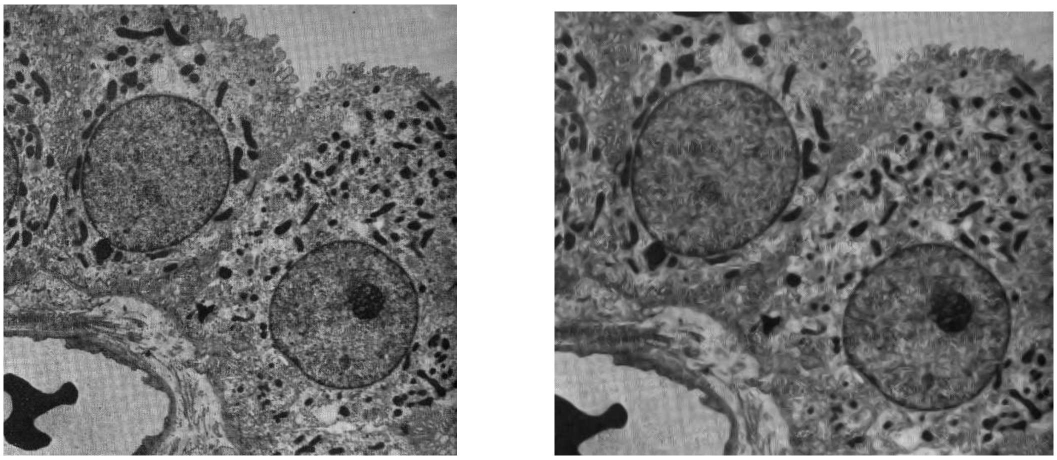
Uma extensão interessante da difusão não linear para imagens multi-variadas é usado em (Sapiro, 1995). Nesta referência usa-se a equação:

$$\frac{\partial \Phi}{\partial t} = g_{cor}\left(\lambda_+, \lambda_-\right) \frac{\partial^2 \Phi}{\partial \theta_-}, \tag{2.24}$$

onde $\lambda\pm$ e θ_- são dadas pelas expressões (2.6)-(2.8), respectivamente. A função $g_{cor}\left(\lambda_+, \lambda_-\right)$ é qualquer função decrescente da diferença $\left(\lambda_+ - \lambda_-\right)$ acrescentada com o objetivo de obter maior controle sobre a difusão local, a exemplo da função c em (2.19).

A equação (2.24) define uma difusão não linear e anisotrópica. Tal método tende a preservar as bordas uma vez que a difusão da imagem ocorre ortogonalmente à direção que as define (θ_+ em (2.7)) e, como g_{cor} é decrescente, a tendência é que a borda se torne mais fina com o tempo.

A Figura 2.7 esclarece o efeito esperado com esta definição para $c(x, y, t)$. Observe que neste caso a difusão da imagem em direção ao background (e vice-versa) é dificultada pois nos pontos de borda o autovetor θ_- em (2.8) é tangente à borda em questão. Isto dificulta a "mistura" entre objeto e background que é precisamente o efeito desejado nesta metodologia.



(a) (b)

Figura 2.7: (a)Célula de Gato: (a)Imagem de Células original.(b)Imagem obtida por difusão anisotrópica.

Mais recentemente, o trabalho de (Florack *et al.* , 1995) coloca os modelos dados por (2.18) e (2.19) em uma formulação covariante através da qual define os conceitos de espaços de escala linear e não linear mais precisamente.

Do ponto de vista das snakes, a suavização Gaussiana é o procedimento comumente usado para diminuir ruídos na imagem. A difusão entre background e objetos é neste caso um efeito desejável em função dos métodos de otimização empregados. No entanto, a formulação implícita proposta em Sapiro (Sapiro, 1997) é um belo exemplo de como a difusão anisotrópica pode aparecer naturalmente também no contexto das snakes.

No próximo capítulo, nos deteremos na formulação paramétrica das snakes e na sua relação com o que foi discutido neste capítulo.