Capítulo 3

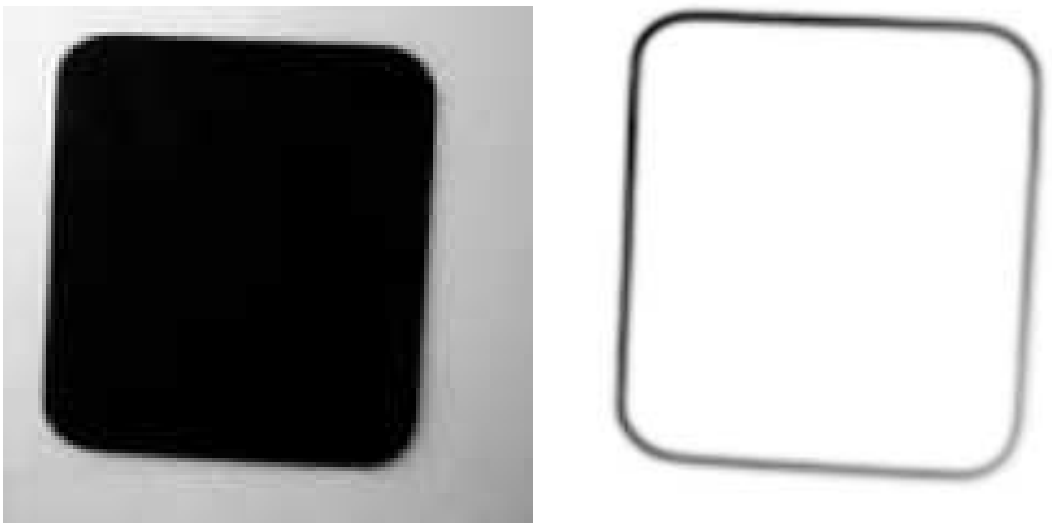
Modelo Original de Contorno Ativo

No capítulo anterior desenvolvemos uma caracterização para pontos de borda baseada na primeira forma fundamental associada à imagem.

No texto que segue estaremos sempre supondo imagens em tons de cinza e assim $m = 1$ na expressão (2.1).

Pelo que foi discutido no capítulo anterior podemos realçar as bordas selecionando pontos da imagem (suavizada por uma Gaussina) com alto gradiente. O resultado obtido nada mais é que uma filtragem passa-alta da imagem a qual realça as fronteiras dos "objetos" da cena.

A Figura 3.1.(b) a seguir mostra o resultado da aplicação desta caracterização para a imagem da Figura 3.1.(a).



(a) (b)

Figura 3.1: (a)Imagem original. (b)Imagem filtrada.

Até este momento, não temos informação precisa da geometria das bordas dos

objetos. Por exemplo, não temos ainda resposta para a seguinte pergunta: qual a curva plana que melhor representa a borda do objeto representado na Figura 3.1.(a).

É neste contexto que o método de snakes (ou de Contorno Ativo) se insere primeiramente (Black & Yuille, 1993). Vejamos alguns detalhes do modelo original de snakes.

3.1 Modelo Original de Snakes

Como já dito na introdução, snakes são um tipo particular de modelo deformável no qual um template inicial (curva) é deformado em direção á borda desejada pela ação de forças internas e externas a ele (Kass *et al.* , 1988; Black & Yuille, 1993). O que se deseja neste processo é que a snake atinja a posição de equilíbrio apenas quando o modelo aproximar a borda desejada.

O resultado do processo fornece assim informações da forma e da posição da borda na imagem. É neste sentido que diremos que o método de snakes é uma técnica de *extração de bordas* de objetos em uma imagem.

Basicamente, este método originalmente proposto por Kass at al. (Kass *et al.* , 1988), parte da idéia de formular o problema da extração de bordas via a otimização de um funcional montado a partir das características de interesse (bordas, no caso) (Gunn, 1996a).

Precisamente, retomemos a definição de pontos de borda para imagens em tons de cinza:

$$\|\nabla I\|^2 > T, \tag{3.1}$$

onde T é um limiar previamente estabelecido usando-se, por exemplo, um histograma da intensidade do campo gradiente na imagem.

Observemos a Figura 3.1. A curva que melhor representa a borda desejada será aquela que melhor se ajuste aos pontos onde $\|\nabla I\|^2$ é alto. Uma maneira precisa de formular isto é dizer que estamos procurando um mínimo (global) para o funcional:

$$E_{ext}(c) = \int_c P(c(s))\, ds, \tag{3.2}$$

onde:

$$P = -\|\nabla I\|^2, \tag{3.3}$$

e s é um parâmetro (não necessariamente comprimento de arco). Observemos, que o problema da extração da borda fica assim colocado como um problema de otimização

de um funcional definido em um espaço de curvas a ser estabelecido a priori (Black & Yuille, 1993).

Posto desta forma, teremos um problema numericamente instável uma vez que a otimização do funcional acima via métodos variacionais é um problema mal-condicionado (Terzopoulos, 1986b). Isto obriga a introdução de um termo de regularização (Poggio *et al.* , 1985; Snyder, 1991).

Nestas condições,o funcional de energia da snake ficará formado por um termo intrínseco, dependente apenas da geometria interna da mesma e por uma energia de imagem dada por um funcional do tipo (3.2).

Uma possibilidade para a energia da snake é dada por uma adaptação ao caso 2D to modelo de placa fina largamente usado em modelos deformáveis 3D (Black & Yuille, 1993). Foi esta a escolha feita por Kass at al. (Kass *et al.* , 1988) no modelo original o qual passamos a tratar agora.

3.1.1 Energia da Snake

Como dito acima, a energia da snake será composta por um termo diretamente relacionado à imagem que guarda as características de interesse da mesma e um termo de *regularização* ou de *suavização* (Poggio *et al.* , 1985; Snyder, 1991). Este último, além de garantir estabilidade numérica irá guardar características geométricas supostas a priori para a solução.

O primeiro passo para a definição precisa da energia é definir o espaço de curvas (Cohen, 1991). Seja então Ad um espaço de curvas parametrizadas por um parâmetro s, $c(s) = (x(s), y(s))$, tal que:

$$Ad = \left\{ c : [0,1] \rightarrow D \subset \Re^2;\ c \in C^4 \right\},$$

onde D é um domínio de interesse, podendo ser todo o domínio de definição da imagem ou apenas uma parte deste.

O grau de diferenciabilidade quatro, imposto aos elementos de Ad, é necessário uma vez que adota-se em (Kass *et al.* , 1988) a formulação variacional para o método de otimização (*Formulação Forte*). Uma outra alternativa seria utilizar a *Formulação Fraca* e então utilizar o Método dos Elementos Finitos (Cohen & Cohen, 1993) o que exigiria um menor grau de diferenciabilidade.

Em Kass at al. (Kass *et al.* , 1988) o funcional de energia da snake tem a forma geral dada por:

$$E(c) = E_{ext}(c) + E_{int}(c) + E_{vínculo},$$

onde E_{ext} é um funcional do tipo definido em (3.2), denominado a *energia externa* da snake:

$$E_{ext}(c) = \int_0^1 P(c(s))\, ds, \tag{3.4}$$

E_{int} é o temo de regularização, denominado *energia interna*, dado por:

$$E_{int}(c) = \int_0^1 \left(\omega_1 \|c'(s)\|^2 + \omega_2 \|c''(s)\|^2\right) ds, \tag{3.5}$$

onde ω_1 e ω_2 são parâmetros internos denominados parâmetros de elasticidade e rigidez, respectivamente, e $c'(s) = dc/ds$ e $c''(s) = d^2c/ds^2$.

O potencial externo P vai gerar um campo de forças externo à curva que deverá "puxar" a mesma em direção às bordas.

Outras possibilidades para a energia interna podem ser encontradas em (Snyder, 1991) onde encontramos uma discussão criteriosa para o problema da regularização em visão computacional.

O termo $E_{vínculo}$ corresponde a forças de vínculo que podem ser definidas a priori ou através de interação com o usuário. Não trataremos deste termo nesta discussão.

Assim, a energia da snake toma a forma:

$$E(c) = \int_0^1 \left(\omega_1 \|c'(s)\|^2 + \omega_2 \|c''(s)\|^2\right) ds + \int_0^1 P(c(s))\, ds. \tag{3.6}$$

Este funcional corresponde à energia potencial de uma corda fina imersa em um campo de forças externo definido pelo potencial P (uma variação do modelo de placa fina usado para superfícies (Black & Yuille, 1993)). Vejamos alguns detalhes deste modelo iniciando pelos parâmetros do mesmo.

O efeito dos parâmetros internos ω_1 e ω_2 na energia acima pode ser explicitado considerando-se as expressões para as derivadas $c'(s)$ e $c''(s)$ em (3.6). Seja então o comprimento de arco α da curva c. Então, as derivadas em (3.6) tomam a forma:

$$\frac{dc}{ds} = \frac{dc}{d\alpha}\frac{d\alpha}{ds}, \tag{3.7}$$

$$\frac{d^2c}{ds^2} = K\overrightarrow{N}\left(\frac{d\alpha}{ds}\right)^2 + \overrightarrow{T}\frac{d^2\alpha}{ds^2}, \tag{3.8}$$

onde $\overrightarrow{T}$ e $\overrightarrow{N}$ denotam o vetor tangente unitário e o vetor normal unitário, respectivamente, e K é a curvatura. Assim, os parâmetros ω_1 e ω_2 em (3.6) multiplicam, respectivamente, os seguintes termos:

$$\left\| \frac{dc}{ds} \right\|^2 = \left(\frac{d\alpha}{ds} \right)^2, \tag{3.9}$$

$$\left\| \frac{d^2c}{ds^2} \right\|^2 = \left\| K \overrightarrow{N} \left(\frac{d\alpha}{ds} \right)^2 + \overrightarrow{T} \frac{d^2\alpha}{ds^2} \right\|^2 \tag{3.10}$$

Vemos desta forma que o termo que multiplica ω_1 varia durante a deformação somente se o comprimento de arco da curva se alterar. Portanto, este termo está diretamente vinculado à elasticidade da curva (*corda*), o que justifica a denominação *parâmetro de elasticidade* para ω_1.

O termo (3.10) inclui a curvatura e o vetor normal e, portanto, está diretamente vinculado à forma da curva. Quanto maior o fator ω_2, mais dependente é a energia da corda de sua geometria, o que implica em maior resistência à deformação ou mais precisamente, em maior rigidez da corda. Em síntese: o coeficiente ω_1 pondera uma energia associada a uma compressão ou a um estiramento da snake enquanto que ω_2 pondera uma energia associada à curvatura.

No trabalho de Kass at al. (Kass *et al.* , 1988) usam-se as equações de Euler-Lagrange para extremizar o funcional de energia (Dubrovin *et al.* , 1984). Observemos que o este funcional é da forma:

$$E = \int_0^1 L\left(c, c', c''\right) ds. \tag{3.11}$$

onde:

$$c' = \frac{dc}{ds},\ c'' = \frac{d^2c}{ds^2},\ \text{e}\ L\left(c, c', c''\right) = \omega_1 \left\| c'\left(s\right) \right\|^2 + \omega_2 \left\| c''\left(s\right) \right\|^2 + P\left(c\right), \tag{3.12}$$

Portanto, as equações de Euler-Lagrange correspondentes têm a forma geral (Dubrovin *et al.* , 1984):

$$\frac{\partial L}{\partial x} - \frac{\partial}{\partial s}\left(\frac{\partial L}{\partial x'} \right) + \frac{\partial^2}{\partial s^2}\left(\frac{\partial^2 L}{\partial x''} \right) = 0, \tag{3.13}$$

$$\frac{\partial L}{\partial y} - \frac{\partial}{\partial s}\left(\frac{\partial L}{\partial y'} \right) + \frac{\partial^2}{\partial s^2}\left(\frac{\partial^2 L}{\partial y''} \right) = 0, \tag{3.14}$$

as quais são geralmente escritas na notação compacta:

$$\frac{\partial L}{\partial c} - \frac{\partial}{\partial s}\left(\frac{\partial L}{\partial c'}\right) + \frac{\partial^2}{\partial s^2}\left(\frac{\partial^2 L}{\partial c''}\right) = 0. \tag{3.15}$$

Para o caso da Lagrangeana L dada em (3.12) obtém-se o seguinte sistema de equações diferenciais ordinárias:

$$-2\frac{d}{ds}\left(\omega_1 \frac{dc}{ds}\right) + 2\frac{d^2}{ds^2}\left(\omega_2 \frac{d^2 c}{ds^2}\right) - \nabla P = 0. \tag{3.16}$$

Do ponto de vista mecânico, esta equação explicita o fato de que a *energia potencial* definida pelo funcional (3.11) terá um extremo quando as forças internas e de campo (força externa dada por ∇P) se anulam. Este equilíbrio é o ponto central da técnica e dele depende a eficiência da mesma.

Um detalhe que a expressão (3.16) acima deixa explícito é a alta dependência das forças internas com relação à forma da curva. Esta dependência pode ser crítica, particularmente nas proximidades da borda, o que motivou trabalhos como (Xu *et al.* , 1994) onde se procura substituir a componente normal das forças internas por um termo de mais fácil controle.

Outra observação importante é com relação ao potencial P dado por (3.3). Este potencial está vinculado à caracterização de pontos de borda apresentada no capítulo anterior. Outras possibilidades discutidas em Kass at al. (Kass *et al.* , 1988) são:

$$P_{linha} = \pm I(x,y), \text{ e } P_{zero} = \left(\nabla^2 (G_\sigma * I)\right)^2, \tag{3.17}$$

onde P_{linha} geraria um campo que atrairia a snake para linhas escuras (sinal $+$) ou claras (sinal $-$) e P_{zero} atrairia a snake para zero-crossings do Laplaciano, os quais localizam os pontos de borda segundo a teoria de Marr-Hildreth (Marr & Hildreth, 1980).

É conveniente neste momento ressaltar o fato de que o método de snakes não pressupõe que a imagem I, usada na definição do potencial P em (3.3), seja o resultado de uma filtragem passa alta da image inicial. De fato, podemos usar a definição (3.3) para P com a imagem I sendo apenas uma versão suavizada da imagem original.

Vejamos agora uma generalização do modelo apresentado nesta seção.

3.2 Generalização do Modelo Original via Mecânica Lagrangeana

A idéia básica neste caso é imergir o modelo dado por (3.16) em um meio viscoso a então usar elementos da formulação Lagrangeana da mecânica clássica (Goldstein, 1981) para obter as equações de movimento para a curva (snake).

A formulação Lagrangeana é mais conveniente que a Newtoniana por conduzir a equações (Equações de Euler-Lagrange do movimento) cuja forma não depende do sistema de coordenadas usado (Goldstein, 1981). Vejamos inicialmente os elementos básicos desta teoria para, em seguida, apresentar o modelo de snakes na sua formulação mais geral.

3.2.1 Princípio de Hamilton

Nesta formulação da Mecânica Clássica, a configuração instantânea de um sistema físico é descrita pelos valores de suas coordenadas *generalizadas* $q_1, q_2, ..., q_n$ e corresponde a um ponto num hiperespaço Cartesiano onde os $q's$ formam os eixos coordenados. Este espaço $2n - dimensional$ é denominado espaço de configuração e será denotado por $\overline{E}$.

À medida que o tempo t evolui, o estado do sistema muda, e portanto, o ponto que o caracteriza no hiperespaço se desloca traçando uma curva que define a evolução do sistema. Assim, a expressão *movimento do sistema* corresponde ao movimento do ponto no hiperespaço ao longo da curva correspondente a qual fica parametrizada na forma:

$$c(t) = (q_1(t), q_2(t), ..., q_n(t)). \tag{3.18}$$

Na formulação Lagrangeana da Mecânica o comportamento de um sistema físico pode ser obtido via o *Princípio de Hamilton* (Goldstein, 1981) que estabelece o seguinte: Para sistemas mecânicos para os quais todas as forças (exceto as de vínculo) são derivadas de um potencial (generalizado), o movimento do sistema de um instante t_1 a um instante posterior t_2 é tal que a integral de linha:

$$J = \int_{t_1}^{t_2} (T - V)\, dt \tag{3.19}$$

onde T é a energia cinética e V a energia potencial do sistema, tem um valor extremo (máximo ou mínimo local) para a curva correspondente.

A função $T - V$ em (3.19) é denominada Lagrangeana do sistema e será denotada por L.

A extremização do funcional J leva a equações de Euler-Lagrange da forma (Goldstein, 1981):

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i},\ i = 1, ..., n \tag{3.20}$$

onde $\dot{q}= dq/dt$.

Neste texto, denotaremos este sistema na forma compacta:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) = \frac{\partial L}{\partial q}, \tag{3.21}$$

onde $q = (q_1, ..., q_n)$ é o vetor das coordenadas generalizadas do sistema.

3.2.2 Formulação Geral das Snakes

A idéia central é submeter um contorno inicial (corda fina) a uma dinâmica para a qual os extremos do funcional (3.19) sejam trajetórias no hiperespaço convergindo para os contornos de interesse.

O primeiro passo é então montar a Lagrangeana do sistema, e para tal devemos escolher coordenadas convenientes e expressar as forças que definirão a dinâmica nestas coordenadas.

Não vamos restringir a dimensão do espaço de curvas admissíveis a um valor finito. Assim, teremos um espaço de configuração com dimensão infinita do tipo Ad na seção 3.1.1 acima. Uma curva do espaço Ad é um ponto deste espaço enquanto que uma curva deformando-se no tempo é representada por uma trajetória do tipo (3.18) em Ad.

Seja então uma curva que se deforma no tempo. Seja μ a densidade linear de massa ao longo da curva e γ o coeficiente de amortecimento correspondente à ação das forças de atrito. Assim, associaremos à curva uma energia potencial E_p do tipo dado pela expressão (3.11) e uma energia cinética T que depende da velocidade; ou seja (Black & Yuille, 1993):

$$E_p(c) = \int_c \left(\omega_1 \left\|c'(s)\right\|^2 + \omega_2 \left\|c''(s)\right\|^2\right) ds + \int_c P(c(s))\, ds, \tag{3.22}$$

$$T = \mu \int_c \left\|\frac{\partial c}{\partial t}\right\|^2 ds. \tag{3.23}$$

O movimento da curva estará sujeito também a forças de atrito dadas por um potencial generalizado dependente apenas da velocidade e de γ. Em geral este potencial (U_{at}) é definido por (Goldstein, 1981):

$$\frac{\partial U_{at}}{\partial t} = -\frac{1}{2}\lambda \left\|\frac{\partial c}{\partial t}\right\|^2, \tag{3.24}$$

Então a Lagrangeana do problema pode ser escrita na forma:

$$L\left(c,\dot{c},c'.c''\right)=T-E_p-U_{at}, \tag{3.25}$$

Uma vez que temos dois parâmetros independentes s e t, o funcional J em (3.19) toma a forma:

$$J=\int_{t_1}^{t_2}\int_{0}^{1} L ds dt \tag{3.26}$$

o que conduz às seguintes equações de Euler-Lagrange (Goldstein, 1981; Dubrovin *et al.* , 1984):

$$\frac{\partial L}{\partial x}-\frac{\partial}{\partial t}\left(\frac{\partial L}{\partial\,\dot{x}}\right)-\frac{\partial}{\partial s}\left(\frac{\partial L}{\partial x'}\right)+\frac{\partial}{\partial s^2}\left(\frac{\partial L}{\partial x''}\right)=0, \tag{3.27}$$

$$\frac{\partial L}{\partial y}-\frac{\partial}{\partial t}\left(\frac{\partial L}{\partial\,\dot{y}}\right)-\frac{\partial}{\partial s}\left(\frac{\partial L}{\partial y'}\right)+\frac{\partial}{\partial s^2}\left(\frac{\partial L}{\partial y''}\right)=0. \tag{3.28}$$

Usando (3.25) e retomando a notação compacta introduzida em (3.15) podemos escrever este sistema na forma:

$$\frac{\partial \widetilde{L}}{\partial c}-\frac{\partial}{\partial t}\left(\frac{\partial \widetilde{L}}{\partial\,\dot{c}}\right)-\frac{\partial}{\partial s}\left(\frac{\partial \widetilde{L}}{\partial c'}\right)+\frac{\partial}{\partial s^2}\left(\frac{\partial \widetilde{L}}{\partial c''}\right)=-\frac{\partial}{\partial\,\dot{c}}\left(\frac{\partial U_{at}}{\partial t}\right), \tag{3.29}$$

onde $\widetilde{L}$ corresponde à parte conservativa $(T-V)$ de L e U_{at} à parte dissipativa da Lagrangeana do sistema, como definido em (3.24).

Ao sistema (3.29) devemos agregar condições iniciais e de contorno que garantam a unicidade da solução. Podemos também agregar condições de vínculo decorrentes por exemplo de interação com o usuário.

As condições de contorno dependem das curvas admissíveis serem abertas ou fechadas. No caso de curvas fechadas o sistema a ser resolvido pode ser da forma:

$$\mu\,\ddot{x}+\lambda\,\dot{x}-\frac{\partial}{\partial s}\left(\omega_1\frac{\partial x}{\partial s}\right)+\frac{\partial^2}{\partial s^2}\left(\omega_2\frac{\partial^2 x}{\partial s^2}\right)=-\frac{1}{2}\frac{\partial P}{\partial x}, \tag{3.30}$$

$$\mu\,\ddot{y}+\lambda\,\dot{y}-\frac{\partial}{\partial s}\left(\omega_1\frac{\partial y}{\partial s}\right)+\frac{\partial^2}{\partial s^2}\left(\omega_2\frac{\partial^2 y}{\partial s^2}\right)=-\frac{1}{2}\frac{\partial P}{\partial y}, \tag{3.31}$$

sujeito a

$$c(0,t) = c(1,t), \quad \frac{\partial c(s,0)}{\partial t} = \overrightarrow{v_0}(s), \tag{3.32}$$

onde a primeira condição (de contorno) explicita o fato da curva ser fechada e a segunda condição (inicial) fornece a velocidade da curva no instante inicial.

Uma propriedade útil deste novo modelo é que a snake agora tem inércia. Assim, ela pode escapar de mínimos locais o que em princípio aumenta as suas chances de chegar ao mínimo global. Contudo, a real eficiência do método depende de um delicado *trade-off* entre os parâmetros do modelo o qual dificilmente é conseguido na prática.

Em (Giraldi & Oliveira, 1999) estudamos estas dificuldades tanto qualitativa quanto quantitativamente para modelos de snake definidos em um espaços de configuração com dimensão finita. No Apêndice D desenvolvemos os elementos básicos deste trabalho.

É oportuno destacar que a substituição do espaço admissível Ad por um espaço de dimensão finita, por exemplo, através do uso de splines (Black & Yuille, 1993) ou NURBS (Qin & Terzopoulos, 1996), pode diminuir a complexidade do método gerando uma formulação elegante e com um apelo geométrico proveniente do sistema dinâmico resultante (Giraldi & Oliveira, 1999).

O sistema (3.30)-(3.32) não tem em geral solução analítica uma vez que o potencial P pode ser absolutamente genérico. Assim, torna-se necessária a utilização de métodos numéricos para obter uma aproximação da solução desejada.

Vejamos a análise dos parâmetros.

Elasticidade e Rigidez

Do ponto de vista contínuo o efeito de ω_1e ω_2 permanece o mesmo já discutido na seção 3.1.1.

O efeito destes parâmetros no comportamento do modelo acima é analisado também em trabalhos como (Leymarie & Levine, 1993; Giraldi & Oliveira, 1999). Contudo, o desenvolvimento de procedimentos sistemáticos para estimar estes parâmetros tem recebido pouca atenção, como ressaltado por (R. Fisker, 1998).

A análise de convexidade apresentada em (Davatzikos & Prince, 1999) e reproduzida na seção D.1 é um caminho para limitar o intervalo de variação destes parâmetros numa dada aplicação.

Outra possibilidade seria permitir que estes parâmetros variem no tempo e no espaço o que pode aumentar o custo computacional do modelo (Davatzikos & Prince, 1999).

Em (Xu *et al.* , 1994) encontramos um outro método baseado na adição de um novo termo de energia tal que a força normal exercida sobre cada ponto tenha a mesma intensidade, independente da forma do contorno. O novo modelo torna-se menos sensível à escolha de parâmetros, mas por outro lado é pouco eficiente para extrair contornos subjetivos.

Neste trabalho, a metodologia assumida será considerar estes parâmetros constantes as quais podem ser atualizadas durante a evolução da snake (Leymarie & Levine, 1993).

Massa Linear e Coeficiente de Amortecimento

Estes parâmetros estão relacionados com a estabilidade numérica e a taxa de convergência, como descrito a seguir.

A situação ideal é escolher estes parâmetros de tal forma que a velocidade da snake decaia rapidamente somente quando a snake estiver próxima da borda desejada. Este comportamento é conhecido como *amortecimento crítico*, sendo bem conhecido no contexto da teoria de oscilações (Nayfeh & Mook, 1979), e significando que a parte transiente da solução contínua decai rapidamente.

No modelo descrito pelas equações (3.30)-(3.31) nós precisamos ser cuidadosos pois o mesmo é em geral não linear. Por exemplo, se simplificarmos a energia 3.4 de tal forma que a equação de movimento torne-se aquela de um oscilador harmônico forçado (Black & Yuille, 1993), nós podemos conseguir amortecimento crítico em alguns modos de vibração mas não em todos. O mesmo será verdadeiro para o modelo dado pela expressões (3.30)-(3.31), como demonstrado no Apêndice D.

Vejamos agora, o efeito destes parâmetros na convergência do método numérico.

3.2.3 Discretização do Modelo

Supondo a existência e unicidade da solução de (3.30)-(3.31), para o conjunto de condições iniciais e de contorno dadas por (3.32), vamos apresentar o método básico utilizado na literatura para discretizar o modelo de snakes (Leymarie & Levine, 1993).

A metodologia aqui é aquela padrão das técnicas para discretização de equações diferenciais por diferenças finitas. Escolhemos um passo τ no tempo e um passo h ao longo da snake. Com isto, a snake ficará discretizada em um certo número de pontos (*snaxels*) e o tempo em uma seqüência de instantes.

Dentre as possibilidades para aproximar as derivadas parciais costuma-se escolher diferenças centrais dadas por (Chapra & Canale, 1988):

$$\dot{x}\left(s,t-\tau\right) \approx \frac{1}{2\tau}\left[x\left(s,t\right)-x\left(s,t-2\tau\right)\right], \tag{3.33}$$

$$\ddot{x}\,(s,t-\tau) \approx \frac{1}{\tau^2}\left[x\,(s,t) - 2x\,(s,t-\tau) + x\,(s,t-2\tau)\right]. \tag{3.34}$$

$$\frac{\partial^2 x}{\partial s^2}\,(s,t) \approx \frac{1}{h^2}\left[x\,(s-h,t) - 2x\,(s,t) + x\,(s+h,t)\right]. \tag{3.35}$$

$$\frac{\partial^4 x}{\partial s^4}\,(s,t) \approx \frac{1}{h^4}\left[x\,(s-2h,t) - 4x\,(s-h,t) + 6x\,(s,t) - 4x\,(s+h,t) + x\,(s+2h,t)\right]. \tag{3.36}$$

Existem certos aspectos a serem considerados na escolha do tipo de diferenças finitas usado para discretizar o modelo. Por exemplo, é necessário que a matriz de rigidez correspondente seja simétrica e possua o mínimo de bandas possível (Leymarie & Levine, 1993).

Usando estas expressões, obtém-se a seguinte discretização do sistema (3.30)-(3.31):

$$\left[\left(\frac{\mu}{\tau^2} + \frac{\lambda}{2\tau}\right) I + K\right] X^n = R_X\left(\bullet, \overline{t} - \tau\right) + \left(\frac{2}{\tau^2}\right) X^{n-1} - \left(\frac{\mu}{\tau^2} I - \frac{1}{2\tau} C\right) X^{n-2}, \tag{3.37}$$

$$\left[\left(\frac{\mu}{\tau^2} + \frac{\lambda}{2\tau}\right) I + K\right] Y^n = R_Y\left(\bullet, \overline{t} - \tau\right) + \left(\frac{2}{\tau^2}\right) Y^{n-1} - \left(\frac{\mu}{\tau^2} I - \frac{1}{2\tau} C\right) Y^{n-2}, \tag{3.38}$$

onde K é a matriz simétrica proveniente da discretização dos termos envolvendo as derivadas espaciais, (X^n, Y^n) é o vetor de snaxels no instante $\overline{t} = n\tau$, e $R\left(\bullet, \overline{t} - \tau\right)$ é o campo externo sobre a snake discretizada (Leymarie & Levine, 1993).

As equações (3.37)-(3.38) podem ainda ser escritas na forma compacta abaixo:

$$\begin{aligned} AX^n &= B_X^{n-1,n-2}, \\ AY^n &= B_Y^{n-1,n-2}, \end{aligned} \tag{3.39}$$

onde :

$$A = \left[\left(\frac{\mu}{\tau^2} + \frac{\lambda}{2\tau}\right) I + K\right], \tag{3.40}$$

e

$$B_X^{n-1,n-2} = R\left(\bullet, \overline{t} - \tau\right) + \left(\frac{2}{\tau^2}\right) X^{n-1} - \left(\frac{\mu}{\tau^2} I - \frac{1}{2\tau} C\right) X^{n-2}, \tag{3.41}$$

$$B_Y^{n-1,n-2} = R\left(\bullet, \overline{t} - \tau\right) + \left(\frac{2}{\tau^2}\right) Y^{n-1} - \left(\frac{\mu}{\tau^2} I - \frac{1}{2\tau} C\right) Y^{n-2}. \tag{3.42}$$

3.2.4 Análise Numérica

Vamos agora estudar as propriedades do método de discretização acima. Para fornecer resultados coerentes, este esquema numérico deve ter as propriedades de consistência, estabilidade e convergência (Hirsch, 1988).

A consistência expressa o fato de que a equação discretizada (3.39) deve tender à equação diferencial correspondente quando $\tau \to 0$ e $h \to 0$. Esta condição é diretamente verificada pelas expressões (3.33)-(3.36).

Quanto a estabilidade do esquema numérico usado devemos ressaltar que estaremos supondo um limite superior finito para a intensidade do campo externo.

Assim, a análise da estabilidade do sistema (3.39) pode ser feita estudando-se o condicionamento da matriz A, seguindo a metodologia apresentada em (Leymarie & Levine, 1993). Se A for bem condicionada então teremos uma garantia de estabilidade para o esquema numérico da seção anterior (Golub & Loan, 1985; Hirsch, 1988).

O número de condição de uma matriz A, não singular, na norma 2, é dado por (Golub & Loan, 1985):

$$\kappa\left(A\right) = \frac{\lambda_{\max}\left(A\right)}{\lambda_{\min}\left(A\right)}, \tag{3.43}$$

onde $\lambda_{\min}\left(A\right)$ e $\lambda_{\max}\left(A\right)$ denotam os autovalores de A com maior e menor valor absoluto, respectivamente. Pela análise de sensibilidade de sistemas lineares sabemos que quanto mais próximo $\kappa\left(A\right)$ estiver da unidade mais bem condicionada será a matriz A. Conseqüentemente, o sistema linear (3.39) terá melhor estabilidade (Golub & Loan, 1985).

Lembrando que A tem a forma $(\beta I + K)$ nós podemos escrever:

$$\kappa\left(A\right) = \kappa\left(\beta I + K\right) = \frac{\lambda_{\max}\left(K\right) + \beta}{\lambda_{\min}\left(K\right) + \beta}. \tag{3.44}$$

Supondo snakes fechadas, é oportuno ressaltar que a matriz K possui neste caso um autovetor dado por $(1, 1, ..., 1)^T$ cujo autovalor é nulo (Leymarie & Levine, 1993). Este autovetor corresponde à translações da snake, as quais não alteram a

geometria intrínseca da mesma, deixando portando as forças internas inalteradas (daì o autovalor associado ser nulo). Com esta observação podemos simplificar esta expressão para:

$$\kappa(A) = \frac{\lambda_{\max}(K)}{\beta} + 1. \tag{3.45}$$

Por outro lado, $\lambda_{\max}(K)$ pode ser limitado pela relação:

$$\lambda_{\max}(K) \leq \max_j \left(\sum_i |K_{ij}| \right). \tag{3.46}$$

De (3.45)-(3.46) vem que valores altos de β podem melhorar o condicionamento da matriz A e conseqüentemente a estabilidade do sistema linear (3.39). Uma vez que:

$$\beta = \left(\frac{\mu}{\tau^2} + \frac{\lambda}{2\tau} \right), \tag{3.47}$$

fica claro que podemos aumentar esta constante aumentando os valores de μ e λ ou diminuindo o passo no tempo τ. Se diminuirmos muito o passo no tempo teremos um resultado mais preciso, mas isto implicará num esforço computacional extra que pode tornar a performance do algoritmo insatisfatória.

Por outro lado o aumento indiscriminado de λ implica em uma perda de energia cinética muito rápida o que pode levar a uma solução distante da borda procurada.

O resultado acima deixa claro também a necessidade de escolher o passo no tempo com cuidado pois o mau condicionamento do sistema linear decorrente de um passo alto pode implicar em uma solução muito distante da solução real do sistema linear.

De qualquer modo, o resultado acima é suficiente para garantir estabilidade pois este mostra que o erro entre a solução exata do sistema linear e aquela encontrada computacionalmente para cada interação pode ser controlado pela escolha de parâmetros adequados. Fica portanto demonstrada a estabilidade do esquema numérico.

Quanto ao parâmetro h, sua escolha é muito dependente da aplicação pois este deve ser pequeno suficiente para que a curva possa se ajustar ao contorno de interesse.

Quanto à convergência do método numérico; isto é, a garantia de que a solução numericamente computada para o sistema linear converge para a sua solução exata quando $\tau \to 0$ e $h \to 0$, basta usar o Teorema de Equivalência de Lax (Hirsch, 1988)

para obtê-la, como conseqüência direta da consistência e estabilidade discutidas acima.

Outro aspecto importante são os critérios de parada para o processo interativo. Encontramos duas propostas na literatura (Leymarie & Levine, 1993), ambas partindo do fato das bordas corresponderem a vales da energia potencial (3.22) nos quais a curva deve ficar confinada. Na primeira, interrompe-se o processo interativo quando:

$$\left\|(X^n, Y^n) - \left(X^{n-1}, Y^{n-1}\right)\right\| < \varepsilon, \tag{3.48}$$

para um ε previamente estabelecido.

No segundo critério, postula-se que a energia de campo média E_l sobre a curva teria taxa de variação pequena próximo da borda; ou seja, que:

$$E_l(t) = \frac{\int_0^1 E_{ext} ds}{\int_0^1 \left\|\frac{dc}{ds}\right\| ds}, \tag{3.49}$$

tenha pequena variação quando a curva c está próxima da borda procurada. Assim, um critério de parada seria:

$$|\Delta E_l| = |E_l(t) - E_l(t-\tau)| < \varepsilon, \tag{3.50}$$

para uma tolerância ε previamente estabelecida.

Ambos os critérios dependem de baixa velocidade da curva para serem satisfeitos. Suponha que o campo externo não varie de forma abrupta nas proximidades da borda, como resultado de uma suavização Gaussiana num esquema multi-escala (Leymarie & Levine, 1993). Então, o segundo critério seria menos sensível a oscilações da curva em torno do ponto de equilíbrio por envolver uma média com relação ao comprimento de arco da curva.

Entretanto, se a borda define um vale de potencial muito estreito, pequenas variações da posição da snake nas proximidades da borda podem levar a mudanças grandes no numerador de $E_l(t)$ não havendo vantagem em usar o critério (3.50).

Neste trabalho, escolhemos o critério (3.48) por ser mais simples. Os resultados que obtivemos até agora com este critério são satisfatórios

3.3 Limitações do Modelo Original

Apesar de suas vantagens para extração de contornos, o modelo original de snakes apresenta várias limitações discutidas em (McInerney, 1997; Gunn & Nixon, 1997;

Xu *et al.* , 1994).

Primeiramente, o funcional de energia (3.22) é em geral não convexo e assim, os métodos de otimização baseados em informações locais (tais como os métodos variacionais descritos nas seções 3.1.1 e 3.2) não oferecem garantia de que a solução encontrada seja o mínimo global (Gunn & Nixon, 1997). Este problema implica que a solução é sensível à inicialização do método, seja para o modelo original dado por (3.16) ou para sua generalização (3.29).

Ao problema da não-convexidade da energia acrescenta-se as dificuldades com a componente normal das forças internas a qual pode ser obtida a partir dos dois primeiros termos da expressão (3.16) (Gunn & Nixon, 1997). Esta força tem o efeito de suavizar a solução final mas tende a contrair a curva inicial (suposta fechada) fazendo com que a mesma colapse em um ponto caso as forças de campo não sejam suficientes para equilibrá-la. Tal comportamento traz dificuldades para a escolha dos parâmetros do modelo.

Esse comportamento da força normal interna está diretamente relacionado com a não invariância da energia (3.22) por transformações de escala (Gunn & Nixon, 1997; Ip & Shen, 1998). Isto fica explicito na equação (3.9) a qual mostra que a energia da snake varia com a mudança do comprimento de arco. Assim, por exemplo, a energia interna de uma circunferência ($(s \rightarrow r\,(\cos(s), sen(s)))$) diminui quando seu raio diminui.

Tal comportamento pode não ser desejável quando possuímos informações a priori sobre a geometria dos objetos de interesse. Por exemplo, quando sabemos que os objetos de uma cena são deformações de um protótipo, não é em geral interessante que escalas menores tenham prioridade. Em resumo, a energia interna da snake deve ser modificada para admitir invariância por outras transformações além da rotação e translação.

Uma outra limitação inerente ao modelo da seção 3.1 é a sua dificuldade de tratar com mudanças da topologia da curva inicial; isto é, splits da curva inicial (McInerney & Terzopoulos, 1995; McInerney, 1997). A menos que acrescentemos outros recursos ao modelo, a evolução da curva baseada na minimização do funcional (3.22) impõe que a variedade correspondente á curva permaneça conexa durante a sua evolução; isto é, a curva não pode sofrer subdivisões (*splits*).

Neste trabalho estamos particularmente interessados em soluções para os problemas dos mínimos locais e limitações topológicas do modelo original. As questões referentes à invariância são também discutidas pela sua relação com outros temas abordados neste trabalho.

A seguir desenvolvemos o modelo de Balloon que foi uma primeira abordagem para o problema da não convexidade da energia.

3.4 Modelo de Balloon

Se considerarmos $\mu = 0$ e $\gamma = 1$ na equação (3.23), encontraremos o seguinte sistema:

$$\dot{x} - \omega_1 \frac{\partial^2 x}{\partial s^2} + \omega_2 \frac{\partial^4 x}{\partial s^4} = -\frac{1}{2}\frac{\partial P}{\partial x}, \tag{3.51}$$

$$\dot{y} - \omega_1 \frac{\partial^2 y}{\partial s^2} + \omega_2 \frac{\partial^4 y}{\partial s^4} = -\frac{1}{2}\frac{\partial P}{\partial y}, \tag{3.52}$$

Se agora usarmos um método de discretização equivalente àquele usado na seção 3.2.3, teremos o seguinte sistema para resolver:

$$[I + \tau K] X^n = X^{n-1} + \tau R_X \left(\bullet, \overline{t} - \tau\right), \tag{3.53}$$

$$[I + \tau K] Y^n = Y^{n-1} + \tau R_Y \left(\bullet, \overline{t} - \tau\right), \tag{3.54}$$

Em (Cohen, 1991) esta formulação é discutida e suas dificuldades básicas analisadas. As soluções apresentadas deram origem ao modelo de *Balloon* , que apresentaremos abaixo por sua relação com o nosso trabalho.

3.4.1 Instabilidade devido a Forças de Imagem

Vamos examinar o efeito da força de imagem $R\left(\bullet, \overline{t} - \tau\right)$, como definido anteriormente.

Na seção (3.2.3) vimos que quanto menor o passo no tempo τ melhor a estabilidade numérica do modelo (3.39). Lembremos também que toda a análise de estabilidade estava dependente da suposição de que o vetor $R\left(\bullet, \overline{t} - \tau\right)$ é limitado. As componentes deste vetor, denotadas por $R_{\imath}\left(\bullet, \overline{t} - \tau\right)$, são dadas pelo gradiente:

$$R_{X_i}\left(\bullet, \overline{t} - \tau\right) = \frac{\partial P}{\partial x}\left(v_i^{n-1}\right), \tag{3.55}$$

$$R_{Y_i}\left(\bullet, \overline{t} - \tau\right) = \frac{\partial P}{\partial y}\left(v_i^{n-1}\right), \tag{3.56}$$

onde v_i^{n-1} é o snaxel i, da snake discretizada, no instante $\overline{t} - \tau$ e P é o potencial definido em (3.3).

Embora supondo $R\left(\bullet, \overline{t} - \tau\right)$ limitado, se a componente $R_i\left(\bullet, \overline{t} - \tau\right)$ é muito grande o ponto v_i^{t-1} pode ser deslocado de um valor muito alto fazendo com que a snake passe além da borda desejada em uma vizinhança de v_i^{t-1}. Este comportamento pode causar oscilações e instabilidade no movimento da snake diminuindo a performance do método ou mesmo fazendo com que a snake estabilize em um mínimo diferente daquele desejado.

Se nós escolhemos τ pequeno suficiente tal que o movimento $\tau F\left(v^{t-1}\right)$ nunca seja grande, por exemplo, nunca seja maior que a distância entre dois pixel vizinhos, então este problema seria evitado. Contudo, procedendo desta forma, gradientes com menor intensidade atrairiam pouco a curva devido a atenuação decorrente da multiplicação por τ.

Desta forma, em lugar de atuar no passo τ de discretização do tempo, a solução apresentada em (Cohen, 1991) é normalizar a força externa segundo:

$$\begin{aligned} F &= -\frac{\nabla P}{\|\nabla P\|}, \text{ se } \|\nabla P\| > T \\ F &= 0, \text{ caso contrário.} \end{aligned} \tag{3.57}$$

onde T é um limiar necessário para evitar problemas numéricos decorrentes da divisão por valores próximos de zero.

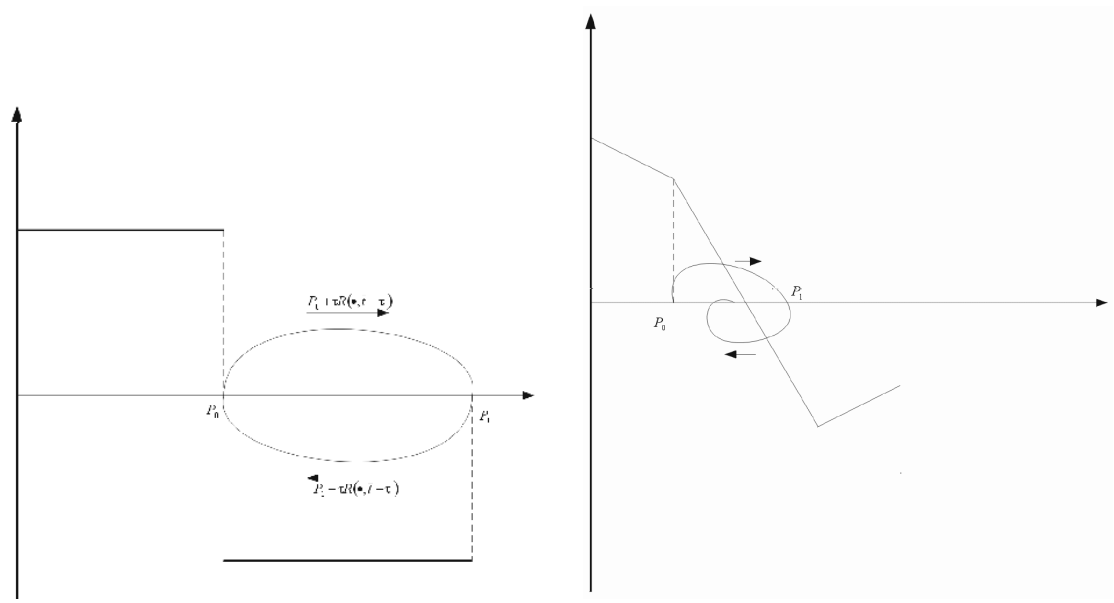
Assim, o deslocamento da snake fica controlado, evitando-se passos muito altos devido ao campo externo.

Um problema desta metodologia é que pequenos e grandes gradientes de imagem terão a mesma influência sobre a curva. Isto tornar o modelo mais sensível a mínimos locais, particularmente para imagens com muitos artefatos. Uma forma de diminuir este problema é aplicar previamente um detetor de bordas à imagem original e, em seguida, aplicar o modelo de snakes para extrair a borda desejada. Contudo, isto não resolve totalmente o problema, como será discutido a seguir.

3.4.2 Discretização no Espaço

A força $F = \nabla P$ é conhecida apenas nos nós da malha correspondente à imagem, e portanto, podem existir zero-crossings fora do grid correspondente. Isto conduz a oscilações entre pixels vizinhos a um ponto de borda (ver Figura 3.2) o que pode impedir que a snake atinja o equilíbrio.

Este problema é resolvido por interpolação bilinear para estimar F em pontos não pertencentes à malha original. Assim teremos uma definição continua de F evitando o problema citado (ver 3.2.b).



(a) (b)

Figura 3.2: (a)Força descontínua provocando oscilações. (b)Interpolação evita instabilidades melhorando convergência.

3.4.3 Inicialização

A escolha da curva inicial para o modelo correspondente ao sistema (3.51)-(3.52) deve levar em conta dois problemas básicos:

a) Se a curva não está próxima o suficiente de uma borda, ela não será atraída por ela,

b) Se a curva não está submetida a forças externas, ela colapsa porque a energia interna dada pela expressão (3.5) tem como mínimo global um segmento de reta (snake aberta) ou um ponto (snakes fechadas).

O comportamento do item (a) é facilmente compreendido observando-se que o alcance da força de imagem é em geral estritamente local. A Figura 3.3, que representa um modelo unidimensional para pontos de borda (Canny, 1983), deixa claro este problema.

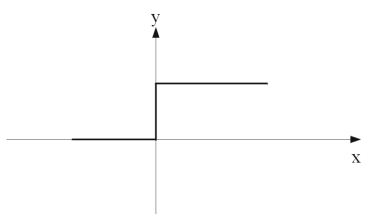


Figura 3.3: Gradiente não nulo apenas em uma vizinhança do ponto de borda.

Quanto ao item (b), uma observação simples esclarece tal comportamento. Su-

ponha que tomemos a seguinte circunferência para inicializar o modelo:

$$c(s,t) = r(\cos(s,t), sen(s,t)) \tag{3.58}$$

onde r é o raio da circunferência. Vamos supor também s seja o parâmetro de arco para simplificar o resultado. Então, substituindo esta expressão nas equações (3.51)-(3.52), encontraremos após um simples cálculo das derivadas o seguinte resultado:

$$\frac{\partial c}{\partial t} = -(\omega_1 + \omega_2)\{r(\cos(s,t), sen(s,t))\} - \frac{1}{2}\nabla P \tag{3.59}$$

ou seja, as forças internas possuem uma resultante normal à circunferência e dirigida para o centro da mesma (vetor entre chaves tem a direção do raio). Se a força externa tem pouca intensidade, esta resultante interna faz a snake colapsar em um ponto.

É interessante também analisarmos o significado geométrico das aproximações por diferenças finitas das derivadas espaciais de segunda e quarta ordem dadas em (3.35)-(3.36). Seja então uma snake $c = c(s)$ discretizada segundo a seqüencia $s_i = ih; i = 0, 1, ..., (N-1)$. As Figuras 3.4(a)-(b) mostram duas vizinhanças do ponto $c(s_i) = P_i$ em uma snake discretizada. Neste caso, as expressões (3.35)-(3.36) podem ser escritas na seguinte forma:

$$\frac{dc}{ds}(s_i) = \frac{2}{h^2}\left[\frac{1}{2}(P_{i+1} + P_{i-1}) + P_i\right] = \frac{2}{h^2}V_1, \tag{3.60}$$

$$\frac{d^4c}{ds^4}(s_i) = \frac{2}{h^4}\left\{\left[\frac{1}{2}(P_{i+2} + P_{i-2}) + P_i\right] - 4\left[\frac{1}{2}(P_{i+1} + P_{i-1}) + P_i\right]\right\} = \tag{3.61}$$

$$\frac{2}{h^4}(V_2 - 4V_1). \tag{3.62}$$

onde os vetores V_1 e V_2 estão representados nas Figuras 3.4.a,b, sendo dados pelas expressões entre colchetes das equações anteriores.

Estas expressões serão utilizadas durante este texto no momento oportuno.

3.4.4 Equações do Balloon

Para resolver as dificuldades (a),(b) da seção anterior, a solução proposta em (Cohen, 1991) é adicionar uma força normal extra ao modelo a fim de fazer a snake convergir para aborda em regiões onde o campo externo seja nulo.

O modelo de *Balloon* é dado por:

$$\frac{\partial c}{\partial t} - \omega_1\frac{\partial^2 c}{\partial s^2} + \omega_2\frac{\partial^4 c}{\partial s^4} = \kappa_1 F - \kappa_2 \vec{n}, \tag{3.63}$$

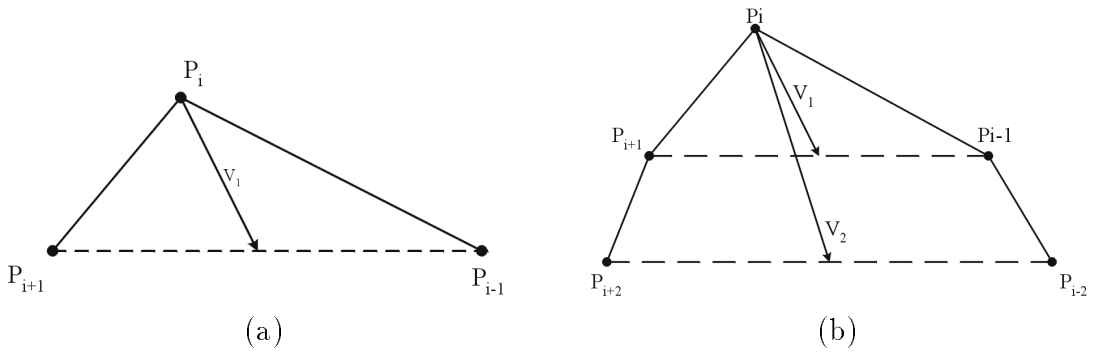


Figura 3.4: (a)Segunda derivada. (b)Derivada de quarta ordem .

onde F é o campo externo normalizado segundo a expressão (3.57), $\vec{n}$ é a normal (unitária) à snake e κ_1, κ_2 são fatores de escala. Agora a curva se comporta como um "balão", daí o nome do modelo, o qual se expande até atingir o equilíbrio pela ação das forças de imagem. O modelo discreto obtido por um esquema de discretização idêntico àquele da seção 3.2.3 é dado pelo seguinte sistema linear:

$$[I + \tau K] X^n = X^{n-1} + \tau \left[R_X \left(\bullet, \overline{t} - \tau \right) - N_X \left(\bullet, \overline{t} - \tau \right) \right], \tag{3.64}$$

$$[I + \tau K] Y^n = Y^{n-1} + \tau \left[R_Y \left(\bullet, \overline{t} - \tau \right) - N_Y \left(\bullet, \overline{t} - \tau \right) \right], \tag{3.65}$$

onde $N \left(\bullet, \overline{t} - \tau \right) = (N_X, N_Y)$ é a matriz correspondente à nova força normal sobre a snake.

O novo modelo tem quatro parâmetros. Em (Cohen, 1991) encontra-se a sugestão de que κ_1 e κ_2 sejam da mesma ordem, com κ_1 ligeiramente maior que κ_2. Assim, quando a snake atinge a borda, o campo externo poderá "segurar " a snake evitando que a mesma ultrapasse a borda.

Quanto à rigidez e elasticidade, é interessante que os coeficientes não nulos da matriz de rigidez K tenham a mesma ordem de magnitude. Em (Cohen, 1991) obteve-se bons resultados quando estes parâmetros são da ordem de h^2 para ω_1 e da ordem de h^4 para ω_2.