Estudo e Desenvolvimento de Técnicas para o cálculo de Curvaturas e Eixos de Simetria

LEANDRO FARIAS ESTROZI*

*Tese apresentada ao Instituto de Física de
São Carlos, da Universidade de São Paulo,
como parte dos requisitos para obtenção
do título de Doutor em Ciências, sub-área:
Física Computacional.*

Orientador:

Prof.Dr. Luciano da Fontoura Costa

São Carlos
2003

USP/IFSC/SBI

*Este trabalho contou com apoio financeiro da CAPES e do CNPq.

Estrozi, Leandro Farias
R a “Estudo e desenvolvimento de técnicas para o cálculo
de curvaturas e eixos de simetria”
Leandro Farias Estrozi – São Carlos, 2003.
160 p.
Tese(Doutorado)-Instituto de Física de São Carlos, 2003.
Orientador: Prof. Dr. Luciano da Fontoura Costa
1.Processamento de Imagens. 2.Curvatura. 3.Esqueletos.
I. Título.

UNIVERSIDADE
DE SÃO PAULO

Instituto de Física de São Carlos

Av. Trabalhador São-carlense, 400
Caixa Postal 369
13560-970 - São Carlos - SP - Brasil

MEMBROS DA COMISSÃO JULGADORA DA TESE DE DOUTORADO DE
LEANDRO FARIAS ESTROZI, APRESENTADA AO INSTITUTO DE FÍSICA DE
SÃO CARLOS, UNIVERSIDADE DE SÃO PAULO, EM 25-07-2003.

COMISSÃO JULGADORA:

Prof. Dr. Luciano da Fontoura Costa (Presidente) – IFSC/USP

Prof. Dr. Roberto de Alencar Lotufo – UNICAMP

Prof. Dr. Esmerindo de Sousa Bernardes – IFSC/USP

Prof. Dr. Roberto Marcondes César Júnior – IME/USP

Prof. Dr. Luís Gustavo Nonato – ICMC/ USP

Dedico este trabalho
aos meus pais e
à fonte de toda minha inspiração:
Bárbara.

Agradecimentos

A o Papai do céu, pelos desejos realizados, pela Saúde e pela Força.

A os meus pais e meus irmãos, por seu amor incondicional.

A Bárbara, por trilhar todos os caminhos ao meu lado (muitas vezes me guiando).

A o amigo Luiz Augusto Consularo, por todos os ensinamentos e por sua paciência quase infinita.

A os amigos Edson Koiti Nakamura e Andrea Gomes Campos Bianchi, pela divertida convivência, pela amizade e pelo apoio.

A o amigo Luiz Gonzaga Rios Filho, por seus conselhos, sua juventude, alegria e

serenidade.

Ao amigo Ricardo Fabbri, pela revisão e inteligência sempre surpreendente.

Ao amigo Marcos Gonçalves, pela fundamental ajuda nos momentos mais críticos.

Ao Prof. Dr. Peter Giblin, pelos inestimáveis ensinamentos, amabilidade e dedicação.

Ao Prof. Dr. Luciano da Fontoura Costa, pela dedicada orientação e pelas oportunidades.

A todos os meus amigos que direta ou indiretamente contribuíram para a manutenção de minha sanidade mental.

A todos os colegas que passaram pelo grupo de pesquisas em visão cibernética desde 1996, pelo ambiente de trabalho extremamente agradável, alegre e sincero que proporcionaram.

Aos docentes e funcionários, em especial às secretárias e bibliotecárias do Instituto de Física de São Carlos, pela boa vontade e empenho em suas funções.

A Donald E. Knuth por ter criado o T_EXe me livrado do MS-Word.

Resumo

Neste trabalho, estuda-se propriedades e implementações de duas importantes representações de formas denominadas curvatura e eixos de simetria (ou *esqueletos*), relacionando-as ao clássico modelo de processamento da informação visual proposto por Marr. Para tanto, fez-se necessária uma releitura dos critérios para a representação de formas já existentes, além da proposição e comparação de algumas abordagens numéricas através das quais curvaturas e esqueletos são calculados, com especial atenção dada a características como robustez e invariância. No caso da curvatura, que é uma medida clássica da geometria diferencial sem divergências quanto a sua definição e interpretação, foi dado maior enfoque à robustez dos métodos quanto a presença de ruídos de diversas naturezas e quanto a sensibilidade na escolha dos parâmetros que levaram aos resultados mais precisos. Além disso, propôs-se uma versão não-derivativa da curvatura (denominada circularidade local) a fim de contornar o problema advindo das instabilidades do cálculo de derivadas de dados reais amostrados. Ainda relativamente à curvatura, um método baseado na transformada de Fourier bidimensional foi proposto e comparado à abordagem unidimensional já estabelecida e a outro método padrão (*B-Splines*/Medioni) de cálculo de curvatura de contornos digitais. Já no caso dos esqueletos, a existência de diversas interpretações geométricas e matemáticas para o conceito e as inúmeras abordagens numéricas para o cálculo dos mesmos demandaram a revisão da literatura sob a luz do modelo inicialmente introduzido para que se pudesse impor uniformidade à terminologia e situar as diferentes abordagens e implementações dentro de um panorama comum, sendo que dois novos métodos para o cálculo de esqueletos foram criados quando do desenvolvimento deste trabalho (esqueletos por dilatações exatas e esqueletos afins baseados em áreas).

Abstract

In this work, properties and implementations of two important shape representations had been studied: curvature and symmetry axis (or *skeletons*) and they had been related to the Marr's classical visual information processing model. Thus, it was necessary a re-interpretation of previously existent criteria for shape representation in addition to the proposition and comparison of some numerical approaches for curvature and skeletons calculations with special attention given to robustness and invariance characteristics. In the curvature case, a classic measure from differential geometry with no divergence about its definition and interpretation, we focused mainly on the numerical methods robustness in presence of several noise categories and on the best parameter set choice sensibility. Additionally, a non-derivative version of curvature (also called local circularity) was proposed in order to circumvent problems related to the sampled real data derivatives numerical instabilities. Still talking about curvature, a 2D Fourier transform based method had been proposed and assessed in comparison to the 1D approach and other standard digital contour curvature calculation method (B-Splines/Medioni). In the skeleton's case, the existence of several geometrical and mathematical interpretations of the concept and the innumerable numeric approaches to their calculation demanded a literature survey regarding the initially introduced model to impose uniformity to the terminology and to situate the different approaches and implementations into a common framework where two new methods to skeleton's calculations had been created (exact dilations based skeletons and area based affine skeletons).

Conteúdo

1	Introdução e Objetivos	3
1.1	Visão humana, ciência e auto-conhecimento	3
1.2	Olhos, as janelas da alma	5
1.3	Informação, Imagem e Representação	7
1.4	Análise de Formas, Reconhecimento e Classificação	8
1.5	Objetivos	9
1.6	Organização e principais contribuições	10
1.7	O caminho percorrido	11
1.8	Contextualização frente a trabalhos anteriores	12
2	Modelo de Processamento da Informação Visual	13
2.1	Motivação	13
2.2	Finalidade	14
2.3	Os três níveis de processamento	15
2.4	Representações, Processos e Características	17
2.4.1	Elementos da Análise de Formas	19
2.4.2	Representações	21
2.4.3	O modelo através de um exemplo.	26
2.5	CrITÉRIOS para a avaliação das representações de formas	30
2.6	Avaliação dos processos	38
3	A Representação por Curvatura	40
3.1	Motivação	40
3.2	A curvatura frente aos critérios para a representação de formas	44
3.3	Comparação de <i>performance</i> entre três diferentes abordagens para a estimação numérica de curvatura de contornos.	46
3.3.1	Motivação	47

3.3.2	Estimação de curvatura baseada em Fourier	51
3.3.3	Cálculo de curvatura baseada na aproximação da curva por splines	55
3.3.4	O sistema de avaliação	56
3.3.5	Resultados	63
3.3.6	Discussão	68
3.3.7	Conclusões	74
3.4	Circularidade Local ou Curvatura não-derivativa multi-escala	75
3.4.1	A curvatura por área	76
3.4.2	Interpretação geométrica	78
3.4.3	Cálculo da Curvatura-Área	79
3.4.4	Uma visão comparativa	81
3.4.5	Conclusões	83
4	A Representação por Eixos de Simetria	86
4.1	Motivação	86
4.2	Definições, esqueletos e eixos de simetria	90
4.3	Esqueletos, abordagens e outras entidades geométricas	91
4.4	Os esqueletos frente aos critérios para a representação de formas	92
4.5	Esqueletização através de dilatações exatas	95
4.5.1	Introdução	96
4.5.2	Dilatações exatas e esqueletos multi-escala	96
4.5.3	Representação multi-escala de formas	99
4.5.4	Conclusões	99
4.6	Considerações sobre diagramas de Voronoi e Eixos Mediais	100
4.6.1	Introdução	100
4.6.2	Diagramas de Voronoi	101
4.6.3	Os eixos mediais	108
4.6.4	Relacionamento entre diagramas de Voronoi e Eixos Mediais	112
4.6.5	Conclusões	120
4.7	Conjuntos de Simetria e Paralelas Afins por Área	120
4.7.1	Transformações, Invariâncias e Geometrias	121
4.7.2	Introdução	123
4.7.3	Curvas Convexas	125
4.7.4	Curvas não convexas	131

4.7.5	Computando as paralelas afins	133
4.7.6	Conclusões	137
5	Conclusões e Desenvolvimentos Futuros	138
5.1	Considerações Finais	138
5.2	Uma visão comparativa entre as representações por curvatura e eixos de simetria	139
5.3	Contribuições	140
5.4	Desenvolvimentos Futuros	140
	Bibliografia	143

Lista de Figuras

1.1	(a) Imagem representando uma face. (b) Três setores circulares que induzem a presença de um triângulo.	7
2.1	Modelo geral para o processamento da informação visual. Representações vão sendo criadas a partir de uma representação inicial de baixo nível e processos vão implicitando ou explicitando características até que uma última representação de alto nível dê acesso a características que permitam a tomada de decisão.	18
2.2	Representação gráfica de um círculo.	21
2.3	(a) Representação fotográfica de um polegar. (b) Representação por linhas do mesmo polegar, sendo que a região nele enquadrada é representada em (c) por meio de um grafo valorado cujos nós correspondem aos diversos segmentos (marcados com seus respectivos comprimentos) e as arestas correspondem aos ângulos entre os segmentos.	23
2.4	Duas versões de um rosto e duas diferentes representações invariantes à presença da barba.	25
2.5	(a) Representação de baixo nível (fotografia) de um objeto. (b) Representação intermediária onde os pixels selecionados como pertencentes ao objeto foram escurecidos formado dois conjuntos de pixels (pretos e brancos) em (c). (d) Contorno do objeto. (e) Funções paramétricas horizontal e vertical do contorno. (f) Curvatura. (g) Círculos indicam pontos do contorno que correspondem aos picos da função curvatura.	27
2.6	(a) Contorno do objeto. (b) Esqueleto extraído do contorno. (c) Círculos que indicam os pontos do contorno que correspondem às extremidades do esqueleto.	29

2.7	(a) Imagem cujas regiões claras estão irregularmente ditribuidas porém com uniformidade de níveis de cinza. (b) Imagem com regiões em arranjo regular mas com intensidades variadas. (c) Limiarização da imagem (a). (d) Limiarização da imagem (b). (e) Transformada de Fourier 2D da imagem (a). (f) Transformada de Fourier 2D da imagem (b).	33
2.8	Sistemas de coordenadas locais bidimensionais sobre uma: (a) representação por contorno de uma forma suave, (b) representação por eixos mediais de uma forma suave, (c) representação por contorno de um retângulo e (d) uma representação por eixos mediais de um retângulo.	37
3.1	Três diferentes interpretações do conceito de curvatura. (a) Curvatura como a taxa de variação do campo tangente, (b) curvatura como a aceleração centrípeta de uma partícula que percorre a forma com velocidade constante e (c) o conjunto do inverso dos raios dos círculos osculantes à forma.	41
3.2	Conjunto de pontos (a) cujas posições estão espacialmente quantizadas e representando (b) um círculo e (c) um polígono.	42
3.3	O gato de Attneave	43
3.4	Círculos senoidais com variado número de ramos e raios internos.	58
3.5	Espirais com diferentes evoluções na largura.	58
3.6	Alguns dos contornos gerados por <i>B-splines</i> considerados.	59
3.7	Formas sintetizadas por componentes do espaço de Fourier com diferentes períodos e composições harmônicas.	59
3.8	Quantização por (a) caixas quadradas e (b) intersecção da grade de uma curva paramétrica específica e respectivas regiões ampliadas (c) e (d).	61
3.9	Pontos repetidos (a) e representação após a remoção de tais pontos (b).	61
3.10	Extensão da escala espacial considerada como uma medida de sensibilidade.	63
3.11	Histograma de tempos para a versão interpolada do método Fourier 1D. Valor médio: $0,008 \pm 0,003$ s.	64
3.12	Histograma de tempos do método por Fourier 2D. Valor médio: $3,5 \pm 0,2$ s.	65
3.13	Histograma de tempos para o método Medioni. Valor médio: $0,002 \pm 0,005$ s.	65

3.14	Curva de erro típica para os métodos por Fourier 1D (a) e 2D (b). Ambas as escalas são logarítmicas.	66
3.15	Erros típicos obtidos na estimação da curvatura considerando-se uma série de combinações dos parâmetros de regularização σ_1 e σ_2 . Situação onde o erro mínimo recai sobre a diagonal.	67
3.16	Erros típicos obtidos na estimação da curvatura considerando-se uma série de combinações dos parâmetros de regularização σ_1 e σ_2 . Situação onde o erro mínimo não recai sobre a diagonal.	68
3.17	Erros na estimação da curvatura para o método por Fourier 1D sem interpolação. Valor médio: $0,05 \pm 0,04$	69
3.18	Erros na estimação da curvatura para o método por Fourier 2D. Valor médio: $0,05 \pm 0,05$	69
3.19	Erros na estimação da curvatura para o método de Medioni e Yasumoto. Valor médio: $0,4 \pm 0,2$	70
3.20	Erros na estimação de curvatura da técnica de Medioni e Yasumoto para diferentes esquemas de quantização: (a) GIQ, (b) SBQ, (c) scanner e (d) câmera de vídeo.	70
3.21	Histograma da dispersão dos valores do parâmetro de escala espacial (desvio padrão) para o método Fourier 1D (sem interpolação).	71
3.22	Histograma da dispersão dos valores do parâmetro de escala espacial (desvio padrão) para o método Fourier 2D.	71
3.23	Histograma da extensão do parâmetro de escala que garante um erro não maior que 10% acima do erro mínimo para a técnica por Fourier 1D.	72
3.24	Histograma da extensão do parâmetro de escala que garante um erro não maior que 10% acima do erro mínimo para a técnica por Fourier 2D.	72
3.25	Ilustração da interferência espacial para a abordagem por Fourier 2D.	73
3.26	O quadrado (e qualquer outro polígono) é um contorno cujas posições dos pontos de alta curvatura são bem conhecidos porém o valor da curvatura naqueles pontos é indefinida.	76
3.27	(a) Segmento de curva genérico. (b) Segmentos de curva com máxima/mínima área para um par fixo de pontos P_1P_2 e três diferentes valores de p . (c) Segmentos de curva com máxima área para um p fixo e três diferentes pares P_1P_2	78

3.28	(a) Um contorno γ e quatro cordas diferentes c_1, c_2, c_3 e c_4 cortando γ em segmentos de comprimento p . O valor de “curvatura” associado ao ponto médio de cada segmento de contorno (o pequeno quadrado à direita) é a área A entre a corda e o correspondente segmento de contorno dividido por $\frac{p^2}{\pi}$. (b) Curvatura-área correspondente.	80
3.29	A área é incrementada, a cada passo, pela área (com sinal) dos triângulos hachurados.	81
3.30	Toda diagonal “ p -deslocada” de A ($A(i, i+p)$ com $1 \leq i \leq N$) contém os valores de área para todas as cordas unindo pontos em γ que estão separados por um perímetro p . Tais valores divididos por p^2/π são boas aproximações para a curvatura-área conforme mostrado nas Figuras 3.30(c), 3.30(e) e 3.30(g) para três diferentes escalas: $p=2$, 5 e 8, respectivamente.	82
3.31	Contorno (a) e respectivas curvaturas (b) por área (linha grossa) e clássica (linha fina). A pequena seta indica a orientação da parametrização.	83
3.32	Contorno (a) e respectivas curvaturas (b) por área (linha grossa) e clássica (linha fina). A pequena seta indica a orientação da parametrização.	84
3.33	Contorno (a) e respectivas curvaturas (b) por área (linha grossa) e clássica (linha fina). A pequena seta indica a orientação da parametrização.	84
3.34	Contorno (a) e respectivas curvaturas (b) por área (linha grossa) e clássica (linha fina). A pequena seta indica a orientação da parametrização.	84
4.1	Quatro diferentes interpretações do conceito de eixo medial. (a) O conjunto dos centros dos círculos maximais, (b) o conjunto das singularidades da transformada distância, (c) as posições dos choques entre frentes que se propagam na direção normal à forma e (d) o diagrama de Voronoi estendido de conjuntos conectados de pontos. . .	87
4.2	(a) Forma e esqueleto traçado “manualmente”. (b) Forma e respectivo eixo medial. (c) Forma e respectivo SLS. (d) Forma e respectivo PISA.	89

4.3	Forma original e suas dilatações exatas para diversos valores de d : (a) Forma original, (b) $d = 2$, (c) $d = 10$ e (d) $d = 20$	97
4.4	Pilha de esqueletos multi-escala da forma em 4.3(a) correspondentes diversos limiares.	98
4.5	Pilha de reconstruções dos esqueletos multi-escala internos da Figura 4.4.	99
4.6	Um simples diagrama de Voronoi definido por três pontos isolados. . .	102
4.7	Dois elementos de Voronoi e_1 e e_2 , e as correspondentes regiões de Voronoi.	103
4.8	Diagrama de Voronoi obtido para um ponto isolado e_1 e um conjunto de infinitos pontos ao longo de um segmento de reta.	104
4.9	Dois pontos ilustrando a propriedade de mínima equidistância dos pontos do diagrama de Voronoi. Somente P_1 é minimamente equidis- tante a pelo menos dois elementos.	105
4.10	Uma forma contínua retangular, considerada como sendo composta por infinitos pontos (i.e. todos os pontos sobre sua borda).	107
4.11	(a) Pequenos círculos dentro do retângulo correspondem a alguns pontos do diagrama de Voronoi. (b) Diagrama de Voronoi obtido quando todos os infinitos pontos sobre o retângulo são considerados. .	107
4.12	A distância entre um ponto e uma curva é medida ao longo da direção normal à curva SE tal direção pode ser definida.	108
4.13	Forma de corpo com sua representação “central”.	109
4.14	(a) Eixo medial de uma elipse com três bolas maximais dentro. (b) Es- trela e seus eixos mediais internos e externos.	109
4.15	Alguns objetos para os quais o eixo medial clássico não está definido. (a) Uma curva que se auto-intercepta; (b) uma forma aberta; (c) um conjunto de pontos desconectados.	110
4.16	O conceito proposto de bola maximal ilustrado sobre uma curva de Limaçon. B_1 é maximal mas, nem B_2 , nem B_3 o são.	111
4.17	Os eixos mediais estendidos de: (a) uma parábola; (b) um ponto e um segmento de reta; (c) Três pontos no plano; (d) um retângulo, que é exatamente o mesmo que para a definição clássica de EM. Repare a similaridade entre os EMs em b, c, d e as respectivas figuras da seção 4.6.2.	112

4.18	Forma poligonal e seu diagrama de Voronoi (seguindo a Definição 3) a partir das arestas e dos vértices considerados como elementos de Voronoi (a) , e seu respectivo eixo medial através da remoção de alguns ramos do DV (b)	113
4.19	Aproximando o Eixo Medial a partir do diagrama de Voronoi de pontos amostrado ao longo da borda de objetos.	114
4.20	Qualquer ponto do DV é o centro de uma bola maximal e portanto, um ponto do EM.	115
4.21	(a) Diagrama de Voronoi (linha tracejada) de um ponto (e_1) e um segmento de parábola (e_2). (b) Diagrama de Voronoi (linha tracejada) de e_1 e infinitos elementos (e_t) ao longo do mesmo segmento de parábola (note a linha vertical adicional no meio da parábola). (c) Eixo medial do mesmo conjunto de dados.	118
4.22	Relacionamento entre DV e EM.	118
4.23	(a) Imagem original. (b) Transformação de escala não-uniforme. (c) Transformação de inclinação horizontal. (d) Transformação afim.	122
4.24	(a) Forma original e respectivo CSAA. (b) Forma distorcida por transformação afim e respectivo CSAA. (c) Forma original e EM. (d) Forma distorcida por transformação afim e EM.	124
4.25	Esqueletos euclidianos de: Uma curva fechada simples (a) , a mesma curva na presença de ruído (b) e as versões transformadas afins de ambas (c) e (d)	125
4.26	Esqueletos afins de: Uma curva fechada simples (a) , a mesma curva na presença de ruídos (b) e as versões transformadas afins de ambas (c) e (d)	126
4.27	Curva γ com equação polar $r = r(\theta)$ relativa ao pólo X	126
4.28	Cinco paralelas-área de um triângulo.	129
4.29	Eixo medial por área (linhas grossas) e conjunto de simetria por área (linhas finas e grossas) de um triângulo.	130
4.30	Duas paralelas-área de um quadrilátero. O número de cúspides muda de quatro para seis conforme o valor da constante de área aumenta.	130
4.31	Uma curva não convexa. A área cortada por uma corda pode conter partes positivas (+) e partes negativas (-).	131
4.32	A paralela-área zero (afastada de γ) próxima a duas inflexões de γ	133

- 4.33 Esquerda: a paralela-área para uma área pequena e positiva. Direita: para uma área pequena e negativa. 133
- 4.34 (a) Uma elipse com dois conjuntos de cordas para diferentes valores de área. (b) Superfície de área correspondente e respectivas curvas de nível. (c) Uma curva não convexa com dois conjuntos de cordas para diferentes valores de área. (d) Superfície de área correspondente e respectivas curvas de nível. 135
- 4.35 Esquerda: Uma paralela ruidosa (ampliada da cúspide superior esquerda da Figura 4.30). Direita: Após suavização. 136
- 4.36 O eixo medial por área de uma curva não convexa. 137

“Todo homem, por natureza, deseja o conhecimento.”
Aristóteles (384 BC - 322 BC) em *Metaphysics*.

Capítulo 1

Introdução e Objetivos

1.1 Visão humana, ciência e auto-conhecimento

“Conhece-te a ti mesmo.” - *escritura no portal do templo de Delfos.*

Por natureza, o homem busca conhecer sua essência e desvendar os mistérios da própria existência. Seja no campo filosófico, seja nas questões sociais, na religião, ou mesmo no mais técnico estudo de sua própria fisiologia, o homem, ao questionar sobre si mesmo, procura respostas para todos os seus “*porquês*” tomando por base o que sabe dos “*comos*”, lutando assim para transformar o desconhecido em ciência sem atenuar, por maiores que sejam as suas descobertas, a sua natural sede pelo entendimento. E, se cogitar sobre o “penso, logo existo” ou descobrir as razões da apoptose podem ser vistos como formas diferentes de auto-descoberta, que dirá entender os mistérios que cercam o funcionamento do cérebro? Assim, não se pode deixar de interpretar os estudos sobre o cérebro, mesmo que restrito a apenas o sistema visual, como um ensejo de se conhecer melhor as nossas próprias características.

Este trabalho vai nesta direção, embora de uma forma extremamente rudimentar, indireta e simplificada, estaremos buscando entender o “porquê” da necessidade humana em *representar e descrever idéias através de imagens*, estudando “como” a natureza permite ao homem executar tal tarefa.

Das duas grandes necessidades exclusivamente humanas: o conhecimento acerca do universo e o conhecimento acerca de si próprio, nos concentraremos em tentar

satisfazer a segunda, embora a *física* concentre-se essencialmente na primeira e de esta tese apresentar-se como um trabalho em “*física-computacional*”.

Foi justamente a co-existência destas duas necessidades que possibilitou ao homem a criação da cibernética¹[64]. Ao adquirir conhecimento sobre o funcionamento das “*coisas do mundo*” através da física e da química, o homem pôde criar a eletrônica e a computação. Ao entender melhor as “*coisas do homem*” através da biologia e da filosofia, o homem pôde criar a medicina e a psicologia. E, no momento em que juntou tais “*coisas*”, surgiu a cibernética².

Apesar de comumente associada ao sonho da criação do “*ciborg*” ou “*ciber-organismo*”, a cibernética tem como objetivo maior ampliar as conexões entre “as coisas do mundo” como a física, a química, a eletrônica e a computação e “as coisas do homem” como a biologia, a filosofia, a medicina e a psicologia, criando com isso teorias, dispositivos ou seres que misturam o que de melhor se sabe sobre estes dois mundos. Pode-se dizer então que *cibernética* é a palavra pioneira para “*interdisciplinaridade*”.

Na cibernética, o estudo conjunto de sistemas naturais e artificiais catalisa o desenvolvimento de metodologias que fundem noções provenientes da medicina (incluindo aí também toda a biologia) e da eletrônica (incluindo a computação), permitindo ao homem criar dispositivos que, em certo grau, se comportam como entidades biológicas ou exercem atividades consideradas “inteligentes”.

Estando ciente de que o comportamento dos sistemas vivos é, ao mesmo tempo, complexo e eficiente, o pesquisador empenhado em progredir nesta área deve especializar-se ou concentrar-se em problemas bastante particulares, a fim de tornar possível uma contribuição efetiva, dada a diversidade, a complexidade e o número dos problemas que (ainda) permanecem em aberto.

A necessidade de especialização ou concentração sobre uma parcela dos problemas existentes nas atividades de pesquisa em cibernética acaba, por vezes, distanciando dos pesquisadores a idéia de que seu objetivo maior é: *criar sistemas artificiais cuja utilidade e eficiência se aproximem da do ser humano e, com isso, avançar na*

¹Cibernética: A ciência do controle e da comunicação no animal e na máquina - Norbert Wiener.

²Um terceiro ingrediente para o surgimento da cibernética foi o desejo humano de comportar-se como “ser criador” ou, mais tecnicamente, como “ser controlador”, de dominar o animal e a si mesmo, tal qual visto em lendas como Frankstein.

compreensão deste ser. Assim, neste trabalho que se concentra apenas na habilidade *visual* dos seres, ou seja, na área de pesquisa iniciada em 1993 pelo Prof. Luciano da Fontoura Costa dentro do Instituto de Física de São Carlos e denominada *visão cibernética*, tentaremos manter vivo, através deste objetivo, o caráter holístico desta pesquisa.

Portanto, devemos considerar as contribuições desta tese não apenas como “*um conjunto de algoritmos para a extração de medidas em imagens*”, mas também levar em conta que a motivação para a realização da mesma é a existência de indícios de que nosso cérebro trabalha da mesma forma, extraindo medidas das imagens captadas por nossos olhos. Assim, sempre que um trabalho estiver sendo inspirado e conduzido na direção dada por constatações biologicamente plausíveis, estaremos caminhando no sentido de compreender melhor o mais perfeito e misterioso processador de informações conhecido pelo homem, seja através dos sucessos, seja através dos fracassos das metodologias propostas.

1.2 Olhos, as janelas da alma

“Se podes olhar, vê. Se podes ver, repara.” - José Saramago em *Ensaio sobre a Cegueira*.

Animal racional é a definição dada por Aristóteles para o “bicho homem”. Costuma-se diferenciar o ser humano dos demais animais da natureza apelando-se para um conceito abstrato denominado *raciocínio*. Embora não haja o objetivo de se fazer uma análise profunda sobre o que é raciocínio, ou mesmo a sobre a polêmica definição de “inteligência”, é importante relacionar tais conceitos com a idéia de imagem. Fisicamente, uma imagem é o resultado da convergência (real ou virtual) dos raios luminosos para a formação de uma cena. Psicologicamente, uma imagem

é a representação (subjettiva ou objetiua) de uma idéia ou conceito. Dado que raciocínio é a capacidade de se criar e inter-relacionar idéias ou conceitos a partir de informações preuiamente conhecidas, e que nossas únicas fontes de informação são nossos órgãos sensoriais, podemos concluir (no sentido empirista de George Berkeley (1685-1753): “*esse est percipi*”) que “somos apenas aquilo que sentimos”³.

Tão grande é a importância dos dados sensoriais para o desenvolvimento mental e intelectual que, por não ter a capacidade de ouuir, um surdo, dotado de perfeitas cordas vocais, é incapaz de falar. Outro fato que ilustra a importância daquilo que vemos para com aquilo que pensamos é que, ao discursarmos sobre um dado assunto ou ao ouirmos um discurso alheio, sempre estamos automaticamente associando imagens ao que está sendo dito. Apenas como exercício, tente ensinar uma receita culinária a alguém, sem que apareçam em seus pensamentos imagens de travessas, colheres e xícaras ou do próprio prato pronto, mesmo que para desfrutar do mesmo, os olhos sejam absolutamente desnecessários. Outro fato curioso é que as pessoas, muitas vezes, na tentativa de se recordarem de algo que alguém lhes falou, tentam lembrar da imagem da pessoa dizendo aquilo que se quer lembrar, mesmo que esta imagem não guarde em si a informação relevante, ela ajuda o cérebro a “*indexar*” aquele momento em que a informação foi absorvida [125].

Dos órgãos sensoriais, os olhos são os que nos fornecem o maior volume de informação e, não obstante, é ao processamento de informações visuais que nosso cérebro dedica a maior parte de sua capacidade em comparação com os demais sentidos.

Portanto, representar e descrever idéias através de imagens é uma necessidade e uma capacidade intrínseca do homem e, por essa razão, teremos como objetivo principal deste trabalho, o estudo de representações e descrições matemáticas que permitam ao computador realizar tarefas semelhantes às realizadas pelo homem no que se refere à seleção das características visuais que melhor representem as cenas ou as informações nelas contidas.

³Neil Young em “On the way home” e Thich Nhat Hanh em “O sol meu coração”.

1.3 Informação, Imagem e Representação

“Imagem (não) é tudo.”

A visão é o principal sentido dos seres humanos e da maioria dos animais. É através da visão que obtemos a maior parte das informações sobre o ambiente em que estamos e é com base nas informações visuais que tomamos a maioria de nossas decisões.

A sensação de que obtemos, através de nossos olhos, uma descrição precisa do mundo a nossa volta, acaba por surpreender aqueles que se propõem a criar um sistema artificial (robô ou programa de computador) capaz de tomar decisões com base em informações visuais.

Isto ocorre porque, na tentativa de se criar tal sistema, descobre-se que muitas das informações que acreditávamos estarem sendo captadas pelos nossos olhos estão, na verdade, sendo deduzidas pelo nosso cérebro. Por exemplo: não se vê o lado direito da face mostrada na Figura 1.1(a) mas, mesmo assim, nos sentimos plenamente capazes de reconhecê-la. Ou ainda melhor, não se vê o triângulo em 1.1(b) mas nosso cérebro deduz a sua presença.

Figura 1.1: (a) Imagem representando uma face. (b) Três setores circulares que induzem a presença de um triângulo.

Busca-se mostrar com isso que imagens são representações da informação, ou seja, uma imagem representa uma parcela especial da informação total presente no ambiente já que nem sempre *toda* a informação é necessária. Sempre que se está representando, estão ocorrendo simplificações ou filtragens da informação. No caso

de uma mera fotografia como a da Figura 1.1(a), perde-se não somente a outra metade da face mas também informações como o perfil (informação tridimensional) e a cor (informação cromática).

Outro ponto importante é que, para lidar com tais simplificações inerentes à necessidade de se representar a informação através de imagens, passamos a ter uma necessidade intrínseca de processamento. Na prática não se percebe isso mas, o cérebro encontra-se permanentemente ocupado com a tarefa de processar as imagens captadas pelos olhos e delas extrair, selecionar ou até mesmo deduzir informações.

Tal processamento tem se mostrado extremamente complexo e, em linhas gerais, este é o problema de que trata a visão computacional. Trata-se de COMO fazer com que sistemas artificiais sejam capazes de processar as informações contidas em imagens e, com base nelas, tomar decisões.

Veremos, no próximo capítulo, que *representações* e *processos* compõem a base de toda uma teoria que busca modelar o funcionamento dos sistemas que visam a tomada de decisões com base em informações visuais.

1.4 Análise de Formas, Reconhecimento e Classificação

“Pode um cego guiar outro cego? Não cairão os dois num buraco?” -
Lucas 6(39).

Ao buscar reconhecer ou classificar um objeto com base em informações visuais, executa-se um processo de tomada de decisão. Dado que o processamento de imagem fornece medidas sobre as características das imagens analisadas, deve-se, com base nestas informações (medidas), decidir se um objeto é ou não reconhecido como sendo um dos objetos previamente analisados, se é um objeto novo ou se podemos classificá-lo como pertencente a determinada classe de objetos.

Neste contexto, as informações extraídas de uma imagem a fim de se conhecer suas características e o uso de tais informações dentro do processo de tomada de decisão constitui o que é chamado de *análise de formas*.

Assim como o físico que, para tentar descobrir as características de um sistema desconhecido, realiza medidas (sejam elas mecânicas, térmicas, ópticas, elétricas ou

magnéticas) sobre o mesmo e então, com base nestas medidas, tira conclusões que o levarão a uma melhor compreensão do sistema, o analista de formas⁴ procura reconhecer ou classificar imagens baseando-se em um conjunto de medidas criteriosamente escolhido.

Um importante aspecto deste processo é que, seja para compreender um sistema, seja para reconhecer ou classificar uma imagem, não é necessário um número infinito de medidas, ou seja, com apenas algumas medidas cuidadosamente escolhidas pode-se chegar à conclusão desejada, o que significa, no caso do físico, ser capaz de prever o comportamento do sistema ou, no caso do analista de formas, reconhecer ou classificar a imagem.

Neste sentido, existe uma linha de pesquisa denominada *seleção de atributos* (ou “*feature selection*”) bem situada tanto frente ao processamento de imagens quanto ao estudo dos mecanismos da aprendizagem e à mineração de dados [120, 71, 107, 25, 111], mostrando a importância da escolha de tais medidas e que esta escolha implica numa seleção ou filtragem da informação contida nas imagens. Além disso, a seleção de atributos indica que, para fins de reconhecimento e classificação, nem toda informação é relevante e que escolher representações que enfatizem os aspectos (informações) mais importantes é essencial para o êxito desta tarefa.

1.5 Objetivos

Tecnicamente, os principais objetivos deste trabalho estão relacionados a proposição e a avaliação de novos métodos numéricos para o cálculo de curvaturas e esqueletos de contornos, considerando os diversos problemas envolvidos e que afetam não somente a qualidade dos resultados obtidos mas que também impõem dificuldades à seleção dos parâmetros envolvidos, especialmente no caso da curvatura. Já conceitualmente, os objetivos são os de uniformizar a nomenclatura geralmente adotada na literatura que é bastante confusa, especialmente no caso dos esqueletos, dadas as origens diversas das principais abordagens. Pretende-se também situar entidades até então puramente geométricas e matemáticas dentro de um modelo de processamento de informação visual como o que foi proposto por Marr [63].

⁴pessoa, robô ou programa de computador

1.6 Organização e principais contribuições

No capítulo 2, motivamos e apresentamos o modelo de processamento da informação visual que será usado para situar as representações de formas (especialmente as curvaturas e os esqueletos) junto às tarefas de seleção, extração e composição de características⁵, aqui denominado processamento. Em seguida, as tentativas anteriores de se formular critérios gerais para a avaliação das representações são discutidas e reformuladas.

O capítulo 3 trata de maneira mais próxima da representação por curvatura, suas vantagens, desvantagens e conexões com outras representações, além de situá-la perante os critérios discutidos no capítulo 2 com destaque dado à influência de ruídos e de fatores de escala e regularização. Posteriormente, uma versão não-diferencial desta medida é proposta além de duas novas técnicas para a sua estimação numérica ao lado de um método de avaliação sistemático e comparativo.

Os esqueletos, que neste trabalho denotam os *eixos mediais* (ou *medial axes*) e os *conjuntos de simetria* (ou *symmetry sets*), são abordados no capítulo 4, onde são feitas considerações sobre a questão da subjetividade deste conceito no que se refere às suas definições e sua posição frente aos critérios para representações de formas propostos no capítulo 2. Além disso, a sua relação com outras representações e conceitos é discutida e as diferentes abordagens através das quais eles são calculados são avaliadas segundo o modelo de processamento de informações visuais.

Por fim, no capítulo 5, fazemos as considerações finais sobre o problema da escolha de diferentes representações para o processamento de imagens e as consequências destas escolhas no desenvolvimento de novas técnicas relativamente aos problemas de desenvolvimento de algoritmos e implementações e à qualidade e utilidade dos resultados obtidos. Na sequência, as principais contribuições deste trabalho são levantadas e as linhas mais atraentes e promissoras de desenvolvimentos futuros são discutidas.

⁵O conceito de *característica* é introduzido na seção 2.4.1.

1.7 O caminho percorrido

Da mesma forma que a reconstituição de um crime auxilia nas investigações, esclarecendo fatos que poderiam não vir à tona de outra maneira, descrever (ainda que sucintamente) a ordem com que os caminhos foram trilhados para a realização deste trabalho pode vir a facilitar a compreensão da tese com um todo, conectando partes que de outra forma não ficariam explicitamente relacionadas.

Ao iniciar o estudo de medidas diferenciais para a análise de formas que já vinham sendo desenvolvidas dentro do grupo de pesquisas onde este trabalho teve sua gênese [103], viu-se na curvatura de contornos e superfícies um promissor objeto de estudos e que interessantes descobertas poderiam nascer da investigação de alguns dos métodos numéricos para o cálculo de curvatura já existentes, dado que a bem estabelecida teoria por trás do conceito de curvatura (geometria diferencial) não deixa margem a dúvidas sobre o que se esperar desta medida no contexto da análise de formas, antevendo e justificando suas boas e más propriedades. Naquele momento, a visita do Prof. James Sethian ao nosso grupo de pesquisas e o estudo dos seus algoritmos de propagação de frentes [184] motivado por ocasião de sua visita proporcionaram ao aluno o primeiro contato com a idéia de transformada distância. Este contato, em conjunto com a recente criação do método das dilatações exatas e eixos mediais por propagação de rótulos [46] (mostrado na seção 4.5) pelo orientador, conduziram o aluno ao estudo dos esqueletos, suas propriedades, algoritmos e aplicações. Posteriormente, a constatação da necessidade comum de aboragens multi-escala tanto para a curvatura quanto para os esqueletos e o conhecimento dos trabalhos de Leyton relacionando matematicamente estas duas entidades [123], acabaram por catalizar o estudo conjunto destas representações sob o ponto de vista de um modelo que pudesse relacionar com mais rigor as inúmeras semelhanças observadas entre elas. E finalmente, a oportunidade de um estágio de seis meses junto do Prof. Peter Giblin (Liverpool) e do contato com sua abordagem mais teórica permitiu a ampliação de horizontes sobre as representações por eixos de simetria levando ao estudo e desenvolvimento dos esqueletos invariantes à transformações afins apresentados na seção 4.7. Todos esses fatos influenciaram na adaptação do modelo de processamento da informação visual de Marr [135], o que acabou possibilitando o tratamento comparativo aqui dado entre tais classes de representações.

1.8 Contextualização frente a trabalhos anteriores

Dentro do grupo de pesquisas em visão cibernética onde este trabalho foi produzido, são conduzidas pesquisas sobre os mais diversos aspectos da visão computacional, biológica, análise de formas e reconhecimento de padrões. Do enfoque essencialmente multidisciplinar dado pelo coordenador do grupo e orientador deste trabalho, nasceram linhas de pesquisa integradas e que vão desde estudos eletrofisiológicos sobre a visão dos invertebrados [45, 194], passando por experimentos psicofísicos sobre atenção seletiva [174], paralelismo em visão [37], análise de formas [103] e cores [154], visão estéreo [72], indo até estudos entre forma e função de estruturas neurais por meios de técnicas de processamento de imagens [51, 53]. Todos estes trabalhos tiveram como inspiração um sistema denominado Cyvis-1 [52] e cuja filosofia é a de utilizar-se dos conhecimentos advindos de sistemas biológicos para a construção e integração de diversos módulos de processamento de informações visuais especializados, sendo que um dos conjuntos de módulos que compõem o Cyvis-1 compreende a análise de formas e é neste contexto que o trabalho aqui apresentado se situa. Não são meros acasos as semelhanças entre o que é proposto em [52] e o modelo de processamento que será exposto no próximo capítulo, sendo tais semelhanças as melhores contextualizações deste trabalho frente aos trabalhos que já vinham e que continuam sendo desenvolvidos dentro do referido grupo de pesquisas.

Capítulo 2

Modelo de Processamento da Informação Visual

“Você não pode depender dos seus olhos quando sua imaginação está fora de foco.” - Mark Twain.

2.1 Motivação

A necessidade do uso de um modelo de processamento da informação visual como o que será adotado neste trabalho nasceu da dificuldade intrínseca e notória na avaliação diferentes metodologias e algoritmos em visão computacional [29, 76, 54, 77], especialmente na avaliação das numerosas abordagens para o cálculo dos esqueletos (eixos mediais e conjuntos de simetria) dada a ampla conexão destes com diversos outros conceitos como, por exemplo, curvatura [123], transformada distância, diagramas de Voronoi, frentes propagantes, etc. Por concordarmos com o fato de que muitos dos algoritmos descritos na literatura são testados em um número muito pequeno de imagens e também que isto se deve a dois principais fatores: o inumerável conjunto de imagens-teste necessárias a uma adequada caracterização da *performance* de um dado algoritmo e a dificuldade em se formar um conjunto verdadeiramente representativo dos dados (mesmo ainda que restrito a um único e específico problema ou aplicação), optamos (assim como em [54]) pela tentativa de modelagem e caracterização dos *algoritmos* (processos e representações) mais do que dos dados sobre os quais eles atuam. Porém, diferentemente da análise es-

tatística proposta em [54], buscamos fundamentar nossa análise na suposta validade de num modelo de processamento da informação visual semelhante¹ ao proposto por Marr e num conjunto de critérios posteriormente formulados para a avaliação das representações de formas e para os processos que atuam sobre elas.

2.2 Finalidade

Embora, pela teoria evolutiva, o surgimento do primeiro ser dotado da primeira estrutura foto-sensível tenha ocorrido por mero fruto do acaso², o processamento da informação visual pelos seres vivos tem necessariamente uma finalidade: a de melhorar suas possibilidades de sobrevivência ou, como quer a teoria, de melhorar sua adaptação ao meio. Assim, não é cabível conceber a existência de um sistema que processe informações visuais sem ao menos uma finalidade específica. Seja para a procura de alimento, seja na busca de parceiros para a reprodução ou até para a complexa comunicação simbólica entre seres de um mesmo bando, existe sempre uma finalidade que acaba por direcionar (ou até mesmo determinar) que tipo de processamento visual será realizado por um dado ser.

Ao se estudar e propor mecanismos artificiais que venham a processar imagens, não se pode agir diferentemente. Não é possível a criação de nenhum algoritmo sem que para este haja uma finalidade. Tal fato, aparentemente óbvio, é destacado aqui pois, muitas vezes, ao nos concentrarmos no estudo e na modelagem de sistemas de visão, mergulhando em abstrações ou formalismos mais teóricos, nos esquecemos que estes estão sempre definidos por seus próprios objetivos e que tais objetivos devem ser sempre considerados quando da proposição e uso de modelos. Em contraste, por exemplo, com o estudo das interações entre a matéria, onde os modelos são propostos sem o conhecimento da finalidade de tais interações³, o estudo do processamento da informação visual não pode ser desconectado de uma finalidade, ou seja, não é possível modelar ou propor sistemas de processamento da informação visual *per si*.

¹A diferença está no fato de aceitarmos a presença adicional de mecanismos “*top-down*”, ao passo que Marr restringiu-se aos mecanismos “*bottom-up*”.

²Sendo esta uma das causas de descrença nesta teoria [171].

³Damos graças à gravidade que nos prende ao solo, onde nossa vida é possível, mas não sabemos exatamente o porquê das coisas serem assim.

Este paradigma proposto por Aloimonos [1] e conhecido como visão intencional (ou “*purposive vision*”) mostra-se especialmente útil na avaliação de diferentes representações, pois se comparadas isoladamente de uma aplicação ou finalidade, aqui denominada “*finalidade global do processamento*”, a objetividade de qualquer um dos critérios levantados na seção 2.5 fica seriamente comprometida.

Embora estejamos de acordo com as idéias de Aloimonos sobre a importância de se incluir no modelo de Marr aspectos da visão ativa, ou seja, a participação dos sistemas no processo espaço-temporal de aquisição das imagens, neste trabalho, nos limitaremos a uma visão estacionária do mundo onde sua descrição, ainda que incompleta, se dá a partir de imagens e representações estáticas.

2.3 Os três níveis de processamento

Em [135], Marr propôs uma divisão da análise do processamento da informação visual em três partes ou níveis de processamento, úteis para a distinção dos papéis dos diversos conceitos envolvidos na solução de problemas de decisão baseada em imagens. Tais níveis são brevemente sintetizados em [85, 200, 189] e re-avaliados por [141]. Os três níveis listados e discutidos a seguir servirão para relacionar a questão da finalidade com o modelo de processamento e suas partes, posteriormente detalhadas através de exemplos.

Nível da Teoria Computacional Descreve o que o sistema deve fazer ou, em outras palavras, que tipo de informação deve ser obtida a partir dos dados de entrada através da determinação da lógica de uma estratégia que venha a executar tal tarefa.

Sendo este o nível mais abstrato, ele relaciona-se claramente com a discussão anterior sobre a necessidade de se ter um objetivo bem definido, sendo este a base para a formulação adequada do segundo nível. Além disso, de acordo com a interpretação de Wright [200]: “*Este nível não apenas define o que está sendo computado mas também o porquê e quais os vínculos da tarefa que podem ser explorados durante a sua execução. (...) A ausência da distinção entre “o que” o “como” minou tentativas anteriores de se*

formar uma teoria mais rigorosa da visão, o que levou muitos trabalhos a se perderem em argumentos subjetivos sobre detalhes de implementação.”

Nível algorítmico (modelo representacional e processual) Descreve como deverá ser efetuada a computação propriamente dita. No caso do modelo proposto por Marr, isto é feito definindo-se as representações (estruturas de dados) e os processos (algoritmos) que as manipularão.

Este nível será mais profundamente discutido na próxima seção pois estudá-lo e compreender o relacionamento entre as partes que o compõe é de fundamental importância na criação e na avaliação das diferentes implementações possíveis.

Nível de Implementação Descreve a realização física do algoritmo, ou seja, os *meios* através dos quais os itens anteriores serão concretizados.

Embora seja essencial, o terceiro nível expressa apenas o fato de que diferentes implementações de um mesmo algoritmo são possíveis, conforme o exemplo dado por Wright em [200]: *“A detecção de bordas usando-se o operador de Marr-Hildreth [137] poderia ser implementado num computador, num circuito eletrônico dedicado ou pelas células da retina de um animal”*.

Nos capítulos seguintes, algumas implementações distintas para o cálculo de curvaturas e de esqueletos estarão sendo avaliadas de acordo com estes três níveis.

Outro ponto a ser esclarecido é que o modelo exposto a seguir não tem a pretensão de corresponder exatamente à forma como os sistemas biológicos processam a informação visual, embora seja este um sonho de muitos pesquisadores da visão e de sua formulação ser fortemente baseada em evidências coletadas acerca dos sistemas biológicos. Este modelo é apenas um dentre outros [81] que poderiam ter surgido (infelizmente não surgiram muitos) na tentativa de se descobrir como uma máquina de estados (discretos ou não) pode representar e descrever idéias através de imagens sendo, por fim, capaz de tomar decisões e alterar seu comportamento com base neste tipo de informação. Nossas principais preocupações são, no entanto, as de ratificar o poder deste modelo através da análise de algumas importantes representações amplamente estudadas na literatura, no caso a curvatura e os eixos de

simetria, seja com respeito às suas propriedades, seja relacionado-as aos processos (algoritmos) que as criam, evitando o uso de argumentações e critérios subjetivos comumente encontrados em outros trabalhos.

Não é nosso objetivo fazer aqui uma revisão sobre o que já se conhece do funcionamento do sistema visual humano. Porém, para que se justifique ao menos uma das características do modelo adotado, faz-se necessária a observação de que, no cérebro, verifica-se a presença de uma estrutura hierárquica de processos e representações que partem da retina, onde a organização dos fotorreceptores em conjunto com algumas camadas de células nervosas atuam na ênfase das bordas e contrastes presentes nas imagens captadas [133, 134, 137, 136], seguindo para o córtex visual, onde outras camadas de neurônios mostram-se capazes de medir a orientação angular de tais bordas [11, 127]. Embora ainda saiba-se relativamente pouco sobre o funcionamento completo desse sistema, toda a informação que se dispõe aponta na direção de um sistema que integra as consagradas abordagens “bottom-up” e “top-down” numa cadeia de processos que supomos estarem ligados por representações expressas na forma de pulsos nervosos que, numa certa altura (muito além do conhecimento experimental atual do homem sobre o cérebro), permitem o reconhecimento dos objetos presentes nas cenas. Esse é o embasamento biológico que é dado ao modelo abaixo descrito e que será usado posteriormente na avaliação das representações e processos propostos neste trabalho.

2.4 Representações, Processos e Características

Um dos mais importantes modelos para o processamento da informação visual (e talvez o único modelo formal até então) foi proposto por Marr [135] poucos anos antes de sua morte. Nele, três elementos básicos são usados como pilares: As *características*, as *representações* e os *processos*. Resumidamente, segundo este modelo, os processos atuam sobre as representações extraindo características para a formação de novas representações ou para a tomada de decisão. Dessa forma, pode-se pensar no processamento da informação visual como uma cadeia de representações e processos (cada vez mais complexos e de nível mais elevado) onde diversas características vão se tornando explícitas ou implícitas até que se consiga uma representação que

explícite a(s) característica(s) que permita(m) a tomada de decisão. A Figura 2.1 busca ilustrar o modelo.

Figura 2.1: Modelo geral para o processamento da informação visual. Representações vão sendo criadas a partir de uma representação inicial de baixo nível e processos vão implicitando ou explicitando características até que uma última representação de alto nível dê acesso a características que permitam a tomada de decisão.

Faremos aqui uma re-leitura das idéias inicialmente propostas por Marr, a fim de complementar alguns dos pontos ainda passíveis de discussão à luz de alguns dos trabalhos publicados nos últimos vinte anos, onde percebe-se que muito ainda está por ser feito para que se construa uma teoria computacional definitiva para o problema de se representar e descrever idéias através de imagens.

Para isso, retornaremos aos elementos mais básicos que compõem a análise de formas através de representações como um processo matemático-computacional que, apesar de já amplamente utilizado e difundido, ainda carece de certa formalização (ou, segundo Marr, de uma teoria computacional) que proporcione uniformidade à terminologia e permita que as diferentes abordagens e implementações sejam comparadas com base nos referidos elementos básicos aqui apresentados e discutidos.

2.4.1 Elementos da Análise de Formas

Assim como na geometria, onde para se compreender o que é uma *reta* faz-se necessário o conceito de *ponto*, em análise de formas, para se compreender o que é *forma* é necessário o conhecimento do que é *informação*. E embora informação não seja um conceito puramente axiomático e de ainda existirem controvérsias acerca de sua definição, pode-se dizer que informação está intimamente ligada ao registro e a transmissão de idéias ou, como preferem os matemáticos, a *ordem* de uma seqüência de símbolos [185]. Fato importante é que a informação pode ser armazenada e transmitida através de diversos meios e quando estes são materiais ou geométricos, a informação se manifesta através da forma.

Como curiosidade, etimologicamente, a palavra informação deriva do latim *informare* que significa “dar forma para”, o que conota uma *imposição de estrutura* sobre uma determinada massa ou sobre um conjunto de entidades geométricas [39, 38].

Apenas para fins didáticos, imaginemos que uma certa *informação* que possuímos seja a velocidade de uma partícula num dado instante de tempo, representada através de um vetor no espaço tridimensional.

O segundo elemento básico da análise de formas são as *características* da informação. Uma característica é parte específica da informação total que, em si, encerra um conceito ou idéia. Uma característica é um subconjunto da informação total que, independentemente desta, é capaz de classificá-la, mas não necessariamente é capaz de distingui-la de outras informações com a mesma característica.

Como exemplo, podemos dizer que as componentes do vetor velocidade citado acima são, cada uma delas, características do vetor ou da informação nele contida. Porém, uma única característica não é suficiente para determinar toda a informação, assim como uma componente não é capaz de definir um único vetor.

Forma é a expressão material ou geométrica da informação. Ou seja, toda forma é fruto de algum tipo de informação. Por exemplo, o contorno serrilhado de uma chave guarda a informação que permite a abertura da fechadura e o contorno da chave como um todo guarda a idéia de *chave* que, independentemente do seu serrilhado, faz com que nós a guardemos no bolso e não a coloquemos na boca pensando tratar-se de um chiclete.

Mesmo quando não há uma realidade física envolvida, a forma ainda funciona como um recipiente da informação. Um círculo, por exemplo, é uma forma ideal advinda de conceitos como “eqüidistância a um ponto”, “curvatura positiva e constante” ou mesmo $e^{i\theta}$. Essas três maneiras equivalentes de representar a idéia de círculo competem entre si sendo que cada uma, por sua própria natureza, evidencia certas características daquela forma.

Em “eqüidistância a um ponto”, salienta-se a existência de um ponto central e de uma *distância fixa* que separa todos os pontos do círculo de seu referido centro.

Em “curvatura positiva e constante”, faz-se a relação de um ponto do círculo com os outros pontos do próprio círculo, sem nenhum apelo a existência de um centro que, embora saibamos existir, é desnecessário dentro desta segunda representação. Um objeto de “curvatura positiva e constante” significa um objeto que tem seu campo tangente⁴ constante em magnitude mas uniformemente variado em sua direção, formando um caminho circular, da mesma forma que uma partícula eletricamente carregada, viajando em uma região de campo magnético constante, descreve círculos⁵, sem que o centro de tais círculos tenha, por isso, nenhuma importância ou propriedade física especial. Quer-se destacar com isso fato de que, ao passar de uma representação para outra, a informação e a forma são mantidas, porém as características evidenciadas pelas diferentes representações mudam.

Em $e^{i\theta}$ ou, equivalentemente, $\cos(\theta) + i\sin(\theta)$, temos um terceiro tipo de representação para a idéia de círculo. Porém nesta, não mais se evidencia a existência de um centro e de uma distância fixa ou mesmo do “alinhamento curvo” dos pontos que o compõem, mas sim o fato de que um círculo pode ser construído pela simples composição perpendicular de dois movimentos harmônicos de mesma amplitude defasados em 90 graus.

Talvez a quantidade de informação contida em um simples círculo e expressa em suas diversas características através das diferentes representações venha a surpreender quem se limitava a vê-lo apenas através de sua representação mais natural, a representação gráfica, mostrada na Figura 2.2.

O fato é que cada representação destaca determinadas características ao mesmo

⁴Formado pela relação de um ponto com sua vizinhança através de uma ligação diferencial.

⁵Desconsiderando-se perdas por emissão.

Figura 2.2: Representação gráfica de um círculo.

tempo em que oculta outras e não se pode julgar as representações como sendo boas ou ruins sem que antes se defina claramente *para quê* tais representações serão empregadas. Isso, por consequência, conecta a existência de uma representação à existência de um *processo* que a emprega (se não fôssemos *processá-la*, por que defini-la ou mesmo avaliá-la?).

Eis então o “quinto elemento” da análise de formas: o processo. Podemos dizer que o processo consiste na seqüência e no tipo das operações que serão efetuadas sobre a representação com a finalidade de extrair as características que serão usadas no passo seguinte, que pode ser a montagem de uma outra representação de nível mais elevado ou o uso efetivo das informações na tomada de decisão, na classificação ou no reconhecimento dos objetos da cena.

A conexão entre os processos e a finalidade global do processamento é comum a muitos trabalhos em visão computacional [10, 193] que visam comparar sistematicamente diferentes métodos e algoritmos, muitas vezes empiricamente concebidos. Porém, ainda é pouco explorada a conexão entre as representações que tais métodos utilizam (implícita ou explicitamente) e a finalidade global do processamento, da forma como será realizada nos próximos capítulos.

2.4.2 Representações

A importância dada por Marr em seus trabalhos às representações como parte de um modelo que ele mesmo denominava “representacional” apoiado pela a idéia de Brady de que “a principal característica do entendimento de imagens (IU) é a construção de ricas descrições” [30] demonstra como tais estruturas que organizam a

informação podem influenciar os resultados gerados por sistemas de visão que seguem este modelo.

No intuito de sistematizar o estudo das diferentes representações existentes, iniciaremos um trabalho de categorização das mesmas separando-as em diferentes níveis de abstração. Mesmo que as maneiras, através das quais pode-se representar a informação visual (ou imagem), sejam extremamente diversas, ainda assim podemos dividi-las em três principais grupos:

Representações de baixo nível. Tais representações compreendem as fotografias, os hologramas, as tomografias, as imagens digitais (matrizes) 2 e 3D e todas as demais representações onde não são efetuadas divisões, agrupamentos ou associações entre diferentes partes da informação visual. Denominaremos as informações representadas nesta forma de dados crus ou *dados não-processados*.

Representações de alto nível. Neste grupo encontram-se os grafos orientados ou as redes⁶, os gráficos vetoriais e as descrições matemáticas e/ou geométricas que agrupam elementos como pontos, retas, curvas e superfícies em objetos mais complexos e de maior significância tornando evidente ao menos uma relação entre as diferentes características salientadas. As informações, quando assim representadas, serão denominadas *dados processados*.

Representações de nível intermediário. Este grupo de representações é o primeiro passo no sentido de se criar uma representação de alto nível a partir de dados crus ou não-processados. Caracterizam-se pela divisão em partes ou segmentação da informação, no entanto, ainda sem a realização explícita de associações que relacionem partes ou características distintas da imagem. Como exemplo, podemos citar as bordas, as texturas, o espectro de Fourier, a curvatura, os eixos de simetria, etc.

Para ilustrar melhor a diferença entre estes três grupos de representações, apresentamos a Figura 2.3.

À esquerda, na Figura 2.3(a), temos a imagem de um polegar representada por meio de fotografia (representação de baixo nível). À direita, na Figura 2.3(b) o

⁶Grafos com custos ou pesos associados às arestas, também chamados de grafos valorados.

Figura 2.3: (a) Representação fotográfica de um polegar. (b) Representação por linhas do mesmo polegar, sendo que a região nele enquadrada é representada em (c) por meio de um grafo valorado cujos nós correspondem aos diversos segmentos (marcados com seus respectivos comprimentos) e as arestas correspondem aos ângulos entre os segmentos.

mesmo polegar está representado por segmentos que reproduzem o traço das linhas geralmente utilizadas no processo de identificação de impressões digitais (representação de nível intermediário) e, em baixo (2.3(c)), um grafo (representação de alto nível) cujos nós correspondem aos segmentos de reta encontrados em 2.3(b) marcados com o comprimento do respectivo segmento e cujas arestas indicam o ângulo entre os pares de segmentos conectados.

Embora a fotografia seja uma representação mais fiel do polegar⁷ em comparação ao conjunto de linhas, nos defrontamos com o fato de que, na Figura 2.3(a) existem informações não relevantes para o processo de identificação se comparado com os traços da Figura 2.3(b). Queremos mostrar com este exemplo que, numa representação, nem sempre faz-se necessária a preservação de todas as informações e que selecioná-las (ou filtrá-las), na maioria das vezes, facilita bastante o trabalho.

Uma última constatação relacionada à Figura 2.3 é que somente a partir de uma representação de alto nível (como em 2.3(c)) torna-se possível a automatização do processo de tomada de decisão, ou seja, o computador não é capaz de reconhecer ou classificar um objeto (ou forma) tomando por base apenas uma representação intermediária. Sempre é preciso que informações vindas de diferentes partes do objeto sejam comparadas a fim de que se possa obter uma característica que possibilite a decisão.

Vale observar que este agrupamento das representações em diferentes níveis de abstração constitui apenas uma dentre as inúmeras possibilidades de agrupamento, sendo esta a mais conveniente para a aplicação sistemática do modelo de Marr conforme será melhor ilustrado na seção 2.4.3. Apenas como exemplo de outros tipos de divisões encontrados na literatura [189], podemos citar agrupamentos mais simples como os que dividem as representações entre “baseadas em contornos” e “baseadas em regiões”, ou divisões relacionadas ao caráter “local” ou “global” destas estruturas.

Segundo Loncaric [128], “*métodos para a **representação** de formas resultam em uma representação não-numérica da forma original (por exemplo, um grafo) de forma que as características importantes da forma sejam preservadas. Sendo que a palavra importantes na sentença acima tipicamente possui diferentes significados para diferentes aplicações.*” Já os métodos para **descrição** de formas resultam em dados numéricos ou medidas que podem ser utilizadas tanto como um auxiliar para enriquecer a representação associada quanto para comporem um vetor de características da imagem útil no reconhecimento e na classificação da mesma em comparação com outras imagens através do uso de uma métrica comparativa.

⁷Apesar das perdas de tridimensionalidade e cor.

“A habilidade de focar atenção nas coisas importantes é uma característica definitiva de inteligência.” - Robert J. Shiller em *Irrational Exuberance*.

Outro aspecto importante relativo à seleção de informações relevantes é que se pode, com isso, obter invariância ou “robustez” no processo, ou seja, ao considerar apenas as informações tidas como relevantes, evita-se que leves alterações na imagem acarretem em grandes mudanças nas informações representadas. Por exemplo, na Figura 2.4 vemos duas fotos de um mesmo rosto mas que, nas representações por *pontos de referência* do lado direito (comumente empregadas em sistemas de reconhecimento de faces), o fato das duas fotografias diferirem quanto à presença da barba não é considerado.

Figura 2.4: Duas versões de um rosto e duas diferentes representações invariantes à presença da barba.

Desta forma, introduzimos a idéia de invariância que é a propriedade que uma dada representação tem de não ser afetada por determinados tipos de transformações. Na seção 4.7.1 trataremos dessa mesma idéia de um ponto de vista mais formal.

2.4.3 O modelo através de um exemplo.

A fim de ilustrar tanto o modelo do processamento da informação visual quanto os diferentes grupos de representações e processos, imaginemos a realização da seguinte tarefa (lembrando que a finalidade atua na determinação da solução): Detectar os pixels correspondentes aos cantos do objeto mostrado na Figura 2.5(a). Esta figura, em si, já é a primeira representação (a de baixo nível) que contém a fotografia de uma tampinha de garrafa com N^2 pixels. Aqui, conforme a descrição deste grupo de representações, ainda não foi realizado nenhum tipo de agrupamento ou associação entre tais pixels⁸.

Sabendo-se que as características “cantos” estão localizadas na borda da tampinha, mais especificamente nos pontos de alta curvatura desta borda, o primeiro processo realizado no intuito de se *decidir* quais pixels serão detectados como sendo cantos é o de segmentar a imagem em dois conjuntos de pixels, obviamente, os pixels do objeto e os pixels do fundo (*objetivo* do primeiro processo).

Cientes de que o fundo é predominantemente cinza, executamos sobre a representação inicial o processo de marcar como sendo fundo os pixels cinza (*definição* do primeiro processo). Isso pode ser feito selecionando-se os pixels cujos componentes de cor (RGB) são iguais (dentro de uma determinada tolerância), excluindo-se os valores extremos de intensidade, ou seja, o preto e o branco (*implementação* do primeiro processo).

Como resultado deste processo, obtemos a primeira representação de nível intermediário mostrada nas figuras 2.5(b) e 2.5(c). De acordo com a especificação deste grupo de representações, embora tenha sido feita uma segmentação de dois conjuntos de pixels, estes ainda não foram relacionados e nem postos em pé de comparação. Todavia, a *característica* da imagem denominada **borda** foi tornada explícita enquanto outras características como a cor ficaram ocultas⁹.

O segundo processo fará uma nova subdivisão da imagem, agora para selecionar

⁸Apenas para deixar claro que, à toda representação, está associado um processo que a cria, no caso da representação fotográfica, o processo de projeção óptica junto com a sensibilização de uma superfície foto-sensível compõem o respectivo processo.

⁹Lembre-se de que nós apenas **marcamos** os pixels como sendo ou não pixels de fundo e que todas as demais informações ainda não processadas continuam presentes dentro da representação, ou seja, não houve perda de informação.

Figura 2.5: (a) Representação de baixo nível (fotografia) de um objeto. (b) Representação intermediária onde os pixels seleccionados como pertencentes ao objeto foram escurecidos formando dois conjuntos de pixels (pretos e brancos) em (c). (d) Contorno do objeto. (e) Funções paramétricas horizontal e vertical do contorno. (f) Curvatura. (g) Círculos indicam pontos do contorno que correspondem aos picos da função curvatura.

os pixels da borda do objeto de uma maneira ordenada ou paramétrica (objetivo do processo). Pixels de borda são pixels do objeto com pelo menos um vizinho pertencente ao fundo (definição do processo) e, através de uma máscara digital 3x3 que percorre a imagem, define-se uma vizinhança e uma regra de deslocamento da máscara que pode então seleccionar ordenadamente os pixels desejados¹⁰ (implementação do processo).

¹⁰Processo também conhecido como algoritmo de perseguição de contorno [49].

Mais uma vez, o resultado do processo é uma nova representação com novas características em destaque. Neste caso, o contorno paramétrico, que é mostrado na Figura 2.5(d) e cujas coordenadas X e Y (horizontal e vertical, respectivamente) são apresentadas em 2.5(e).

O terceiro processo extrai uma medida diferencial deste contorno chamada curvatura que é mostrada na Figura 2.5(f) (veja o capítulo 3 para maiores detalhes sobre esta medida). Esta função que ainda constitui uma representação de nível intermediário somente passa a ser uma representação de alto nível no momento em que efetuamos o quarto e último processo que levará à uma tomada de decisão e não a uma nova representação. Este último processo, a localização dos máximos locais da curvatura (objetivo do processo) consiste em procurar por pontos da função curvatura cujo valor seja maior que os valores em suas adjacências (definição do processo). A procura por tais máximos locais pode ser feita através de um simples teste comparativo do valor da função curvatura num ponto com os seus vizinhos (implementação) que pode ser visto, em última análise, como uma relação entre informações de diferentes partes da representação, o que levaria a uma representação de alto nível se, neste ponto, não estivéssemos atingido nosso objetivo que era o de detectar os cantos do objeto, mostrados na Figura 2.5(g).

Com isso, constata-se que toda representação (de qualquer nível) está sempre associada a um processo que a cria e, a fim de evitar a confusão entre características e representações, dizemos que uma representação, juntamente com qualquer um dos processos que possam tê-la criado, deve permitir o retorno à(s) representação(ões) anterior(es) que poderia(m) ter sido usada(s) em sua criação, fato este que não se observa com as características. Em outras palavras, a preservação da informação permitindo o processo inverso é o que diferencia uma representação de uma característica.

É importante salientar que, nos processos, a formulação do objetivo, da definição e da implementação é fundamental e é essencialmente com esta atividade que os analistas de imagens, em geral, têm se preocupado nos últimos anos, ou seja, muita atenção tem sido dada à formulação e desenvolvimento de processos, porém, pouco se discute acerca das diversas representações utilizadas, bem como suas correspondentes características. Mesmo quando se fala da escassez de trabalhos que

avaliam e comparam a *performance* dos inúmeros métodos propostos na literatura (processos) [100], constata-se uma escassez ainda maior de trabalhos que avaliam e comparam as representações mais comumente utilizadas, ainda que estas sejam bem menos numerosas.

Neste sentido, lembrando que *o propósito das representações é prover descrições úteis dos aspectos do mundo real e que as estruturas do mundo real, portanto, têm grande importância na determinação da natureza das representações que são usadas* [135] (p.13), poderemos então tentar enumerar alguns critérios de qualidade para as representações de formas ratificando e/ou retificando algumas das tentativas anteriores [135, 31] a fim de, posteriormente, situar as representações abordadas neste trabalho segundo tais critérios.

Um último aspecto desta modelagem do processamento de informações visuais é que diversos *caminhos* (cadeias de representações e processos) podem ser seguidos para a realização de uma mesma tarefa. Assim, se ao invés da curvatura do contorno optássemos por calcular os eixos de simetria do mesmo, um outro caminho estaria sendo seguido e um resultado semelhante poderia ser alcançado conforme mostra a Figura 2.6(c).

Figura 2.6: (a) Contorno do objeto. (b) Esqueleto extraído do contorno. (c) Círculos que indicam os pontos do contorno que correspondem às extremidades do esqueleto.

A utilização desta segunda cadeia de processos nos leva a um resultado bastante semelhante ao obtido através do cálculo da curvatura. Porém, repare que devido à *implementação* que extraiu o esqueleto do contorno (descrita na seção 4.5), temos

a presença de uma extremidade a mais, o que levou a detecção de um “canto” extra (na parte superior direita do objeto). Isto não significa, no entanto, que a *representação* por eixos de simetria (esqueleto) escolhida neste segundo caminho seja pior que a representação dada pela curvatura, e sim que a *implementação* que extraiu a curvatura do contorno era melhor que a *implementação* que extraiu o seu respectivo esqueleto.

Eis aqui uma questão delicada pois, em muitos casos, não é fácil distinguir se o problema está na representação utilizada ou na implementação do processo que a criou, principalmente se não conhecemos bem todas as características e propriedades dos processos e das representações envolvidas ou mesmo se não formos capazes de relacioná-los com outras representações afim de descobrir as causas do bom ou mau desempenho verificado. No exemplo das Figuras 2.5 e 2.6, o conhecimento da existência de uma conexão direta e objetiva (matematicamente definida) entre os máximos de curvatura e as extremidades dos eixos de simetria [123] é o que nos leva à conclusão do parágrafo anterior, a partir da qual poder-se-ia buscar outra solução com base em outro tipo de implementação para o cálculo dos eixos de simetria.

Um fato que estará sempre sendo considerado, conforme já salientado no início deste capítulo, é que a aplicação (ou finalidade) da cadeia de processos e representações determinará qual o caminho mais adequado à tomada de decisão baseada em informações visuais.

2.5 Critérios para a avaliação das representações de formas

A existência de infinitas representações de formas e a necessidade de se criar algoritmos que resolvam problemas práticos em visão computacional levou pesquisadores como Marr e Brady [135, 31] a formularem um conjunto de critérios segundo os quais as representações poderiam vir a ser situadas na intenção de facilitar a escolha de uma ou outra representação para a criação de algoritmos. Diz Loncaric em [128] que “nem todos os métodos são apropriados para todo tipo de forma e para toda aplicação, isto é, a escolha do método depende das propriedades da forma a ser descrita e da particular aplicação.” e que “critérios de avaliação consistentes ainda

não existem para métodos descritores de formas”. Apesar da grande diversidade das possíveis representações a serem julgadas perante tais critérios, o fato de todas elas estarem direta ou indiretamente relacionadas a conceitos matemáticos e geométricos previamente estabelecidos (em especial da geometria diferencial) e da existência de um modelo dentro do qual tais representações são utilizadas, permitiu que os critérios abaixo transcritos fossem criados.

Dado que muitos destes critérios, na sua formulação original, carecem de maior objetividade a fim de facilitar (ou até mesmo de possibilitar) sua aplicação a um conjunto mais amplo de casos reais além daqueles usados como exemplos pelos seus autores, nos daremos a liberdade de, posteriormente a alguma análise, reformular, alterar, incluir ou mesmo excluir critérios no intuito de tornar factível sua aplicação sobre as representações aqui estudadas ou, eventualmente, sobre quaisquer outras representações.

A referida falta de objetividade dos critérios propostos na literatura, traduzida pelo uso de expressões do tipo: “*custos razoáveis*”, “*representações apropriadas*” [200], “*relativamente eficiente*” e “*sub-partes perceptuais*” [31], talvez seja uma das razões pelas quais estes ainda não tenham se estabelecido efetivamente como uma ferramenta para aqueles que “arquitetam” ou avaliam algoritmos de processamento de imagens.

Acessibilidade [135] Diz respeito a “ser possível” extrair uma representação a partir de uma imagem a custa de quantidades “viáveis” de uso de memória e tempo de processamento.

Este critério, por ser carregado de subjetividade e estar mais relacionado aos processos do que às representações propriamente ditas, acaba, por estas razões, sendo de difícil aplicação no momento em que diferentes representações são comparadas. Além disso, dado que preocupações com tempo de processamento e uso de memória são típicas da avaliação de *processos*, este critério, se interpretado superficialmente, pode acabar assumindo um caráter óbvio e, portanto, negligenciável.

Na tentativa de solucionar tais problemas, propomos uma nova versão onde a acessibilidade passa a estar mais proximamente ligada ao modelo acima exposto.

Acessibilidade Diz respeito a ordem de complexidade do algoritmo mais simples que se conheça, ou seja, ao algoritmo conhecido de menor ordem de complexidade, capaz de extrair as características necessárias à construção de uma determinada representação com precisão suficiente para que se alcance a finalidade global do processamento.

Apesar de, nesta outra forma, a acessibilidade ainda estar atrelada ao processo e não somente à representação e de a ordem de complexidade de algoritmos ser uma medida nem sempre facilmente quantificada, a medida de acessibilidade torna-se mais objetiva ainda que características como a complexidade de *implementação* não seja levada em conta.

Escopo [135] Relaciona-se à “adequação” da representação a determinadas classes de formas e/ou problemas no sentido da representação ser uma descrição canônica (ou compacta) da informação de interesse contida da forma.

Embora neste caso não julguemos necessária uma reformulação do critério, precisamos fazer alguns esclarecimentos através de um exemplo. Imaginemos o seguinte problema: calcular a separação média entre as regiões claras das imagens 2.7(a) e 2.7(b). Temos aí duas classes de imagens onde, na primeira, existe uniformidade nos níveis de cinza das regiões claras porém o posicionamento destas regiões não é regular no plano. Já no segundo caso, embora o posicionamento seja visivelmente regular, as intensidades luminosas variam consideravelmente de região para região. Duas possíveis representações que poderiam vir a tornar explícita a separação média entre tais regiões são obtidas através de (I) um processo de limiarização mostrado nas imagens 2.7(c) e 2.7(d) e (II) a representação obtida através da transformada de Fourier bidimensional mostrada nas imagens 2.7(e) e 2.7(f).

Enquanto, para a primeira classe de imagens, o processo de limiarização e a respectiva “representação limiar” destaca precisamente as posições das regiões que podem então ter sua separação média facilmente calculada, a representação por espectro de Fourier não dá indícios claros do valor desta separação. Já para a segunda classe de imagens, o panorama se inverte. Dada a regularidade do posicionamento das regiões no plano, o espectro de Fourier mostra claramente o espaçamento médio

Figura 2.7: (a) Imagem cujas regiões claras estão irregularmente distribuídas porém com uniformidade de níveis de cinza. (b) Imagem com regiões em arranjo regular mas com intensidades variadas. (c) Limiarização da imagem (a). (d) Limiarização da imagem (b). (e) Transformada de Fourier 2D da imagem (a). (f) Transformada de Fourier 2D da imagem (b).

entre as regiões através do inverso da máxima frequência espacial observada (pontos claros formando uma cruz na Figura 2.7(f)), enquanto que a representação limiar sofre com a variação dos níveis de intensidade luminosa daquela imagem.

Ainda sobre os mesmos dois tipos de imagens, se imaginarmos um segundo problema como, por exemplo, **contar** o número de regiões claras, é simples constatar que somente a representação limiar provê uma solução direta, ainda assim, apenas para a imagem 2.7(a), já na representação pelo espectro de Fourier, em nenhum dos

casos a informação sobre o número de regiões claras torna-se evidente.

Com este exemplo em mente, podemos dizer que o *escopo* de uma representação está relacionado à variedade de classes de imagens e de problemas em que ela pode ser empregada de maneira a prover uma solução (ou parte dela).

Em resumo, dado um problema a ser resolvido sobre um certo conjunto de classes de imagens, apenas um conjunto restrito representações têm escopo capaz de englobar possíveis soluções. Complementarmente, dada uma representação, apenas um conjunto restrito de problemas e classes de imagens são cobertas pelo seu escopo. Por fim, enquanto a relação de pertinência de um problema/imagem ao escopo de uma representação é dada pela possibilidade de solução do referido problema, o grau desta pertinência está ligado à simplicidade com que se chega a tal solução.

Unicidade [135] Refere-se ao mapeamento biunivocamente definido entre as formas e as representações, considerando-se uma classe de representações¹¹ e todas as possíveis classes de formas.

Por razões óbvias deseja-se que formas distintas possuam representações distintas. Porém, visto que este critério é equivalente ao critério da *sensibilidade* descrito abaixo, a unicidade, embora fundamental, não será explicitamente considerada.

Estabilidade [135] Significa que leves variações/alterações na forma devem acarretar em pequenas mudanças na respectiva representação.

Este critério indica que formas semelhantes devem possuir representações semelhantes e que perturbações de pequena escala (ruídos) não devem afetar a representação de maneira drástica, no sentido de continuidade, assim como na matemática onde, num mapa regular, variações infinitesimais no domínio devem levar a variações infinitesimais na imagem.

É interessante lembrar que estes critérios são concomitantes ao surgimento da teoria de escala-espço e que este é o primeiro critério a conectar a representação com a noção de *escala* sugerindo, conforme o trabalho de Marr e Nishihara [138],

¹¹Definimos uma *classe de representações* como o conjunto das representações geradas pelo mesmo tipo de processo.

um desacoplamento entre características estáveis e detalhes de pequena escala. Mais uma vez, grandezas subjetivas como “leves variações” ou “pequenas escalas” podem inviabilizar a aplicação deste tipo de critério a casos onde a diferença entre “características estáveis” e “detalhes” ou “ruídos” são sutis. Adicionalmente, o critério de estabilidade mostra-se estritamente relacionado ao conceito de escala-espaco [126] pois impõe uma ordem às escalas ao fazer distinção entre detalhes (informações de pequena escala) e características estáveis (informações de grande escala).

Sensibilidade [135] Diz respeito à capacidade da representação de capturar pequenas diferenças entre formas semelhantes.

O critério da sensibilidade está relacionado ao da estabilidade devido à ligação com a idéia de diferentes escalas espaciais em que uma forma pode ser representada. Além disso, a sensibilidade engloba a unicidade acima descrita, uma vez que para uma representação ser “sensível” está pressuposta a unicidade. A inter-conexão entre a estabilidade e a sensibilidade através da idéia de diferentes escalas espaciais abre espaço para a criação de um novo critério que substitua e englobe as características que se deseja numa representação expressas por meio destes dois critérios.

Caráter escala-espaco Refere-se a capacidade da representação de permitir (por si só) a seleção de características em diferentes escalas espaciais a partir de parâmetros que preservem a causalidade escala-espaco [197, 110].

O caráter escala-espaco de uma representação, portanto pressupõe simultaneamente estabilidade e sensibilidade, pois traz para junto do modelo as características da teoria de escala-espaco [126]. Vale observar que operações de suavização como, por exemplo, as governadas por equações de difusão, podem ser, em princípio, combinadas com a maioria das representações a fim de permitir a seleção de escalas. No entanto, deve-se diferenciar as representações que têm caráter intrinsecamente escala-espaco daquelas que *precisam ser compostas* com tais suavizações.

Em [31] Brady propôs a inclusão de mais três importantes critérios que serão discutidos a seguir. No entanto, é importante salientar o direcionamento que a proposição destes critérios teve em direção às representações axiais como os cilindros generalizados e as simetrias locais suavizadas (ou “*Smoothed local symmetries*”).

(SLS)”) [32], por isso, faremos uma re-interpretação destes critérios a fim de facilitar sua aplicação a classes mais amplas de representações. Outro fato digno de nota é que estes critérios, em sua forma original, têm uma dependência hierárquica, ou seja, para que uma representação supra os requisitos do segundo critério, deve antes suprir os requisitos do primeiro e assim por diante.

Suporte local rico [31] Diz-se da representação que pode ser obtida com base em informações de caráter local e que permita a reconstrução da forma, ou seja, que preserve a informação.

Quanto a preservar a informação, o próprio modelo aqui adotado impõe que a representação permita o retorno ou a reconstrução da representação original, caso contrário, seria uma característica e não uma representação. Já com relação ao caráter local, sua principal motivação vem da necessidade de se poder representar imagens que contenham oclusões, isto é, ser capaz de montar a representação mesmo na ausência de algumas das partes. Além disso, o caráter local pressupõe a possibilidade de se determinar o que Brady chamou de “*local frames*” que são *sistemas de coordenadas locais* relativos a cada ponto (ou parte) da representação semelhantes ao que são, por exemplo, os triedros de Frenet para a representação de curvas espaciais por curvatura e torção (veja [61],p.19).

Extensão suave e enquadramento [31] O enquadramento está relacionado à seleção, dentre todos os possíveis sistemas de coordenadas locais, daqueles que englobam ou equivalem aos sistemas não selecionados. Em outras palavras, o enquadramento consiste em **desconsiderar** todos os sistemas de coordenadas locais que podem ser descritos através de outros sistemas de coordenadas. A estes “*local frames*” mais importantes dá-se o nome de apenas “*frames*”. A extensão suave traduz a necessidade de que o enquadramento preserve a “consistência” dos sistemas de coordenadas locais selecionados de forma que, ao se passar de um frame para outro, não ocorram inversões ou perda de orientação.

Portanto, para que uma representação possua extensão suave e enquadramento, existe o pré-requisito de haver um suporte local rico e este deve permitir uma espécie de “compactação” expressa através da redundância entre os diferentes sistemas de

coordenadas locais. A Figura 2.8 mostra como a idéia de enquadramento pode ser aplicada a duas diferentes representações dependendo das características da forma representada. Por possuir lados retos, os retângulos (na parte inferior da figura) permitem que poucos sistemas locais de coordenadas (*frames*) sejam necessários para descrevê-los (maior enquadramento) enquanto que as formas suaves (na parte superior da figura) necessitam de infinitos sistemas locais de coordenadas para uma descrição precisa.

Figura 2.8: Sistemas de coordenadas locais bidimensionais sobre uma: (a) representação por contorno de uma forma suave, (b) representação por eixos mediais de uma forma suave, (c) representação por contorno de um retângulo e (d) uma representação por eixos mediais de um retângulo.

Hierarquização propagável [31] É a capacidade da representação organizar a forma em sub-partes perceptuais dentro de uma hierarquia de relevância perceptual onde diferentes frames são “conectados” de forma consistente.

As representações axiais como o eixos mediais de Blum [22], os cilindros e cones generalizados [156], as simetrias locais suavizadas [32] e os eixos de simetria para a inferência de processos (“*Process-Infering Symmetry Axis (PISA)*”) [124] criaram expectativas quanto a sua aplicação na descrição estrutural das formas e este critério sintetiza o que se espera de uma representação capaz de captar a organização geral da forma [8]. Este é, sem dúvida, o mais subjetivo dos critérios até aqui expostos,

podendo sempre estar relacionado a questões perceptuais [21] da maneira como, por exemplo, Attneave relacionou os cantos (extremos de curvatura) de uma silhueta com a relevância perceptual de suas partes. Todavia, na intenção de reformulá-lo dotando-o de um caráter mais objetivo, utilizaremos a divisão em grupos de representações feita no início da seção 2.4.2 de forma a substituí-lo pelo seguinte critério:

Hierarquia global Diz-se da representação de baixo nível ou de nível intermediário que permite a criação de uma representação de alto nível cuja relação entre as diferentes características salientadas possa ser estruturada hierarquicamente.

Assim, evitamos o apelo à percepção e criamos um critério um pouco mais abrangente que a hierarquização propagável, já que não há mais a necessidade da representação ter um suporte local rico e de permitir o enquadramento e a extensão suave para que organize a forma em sub-partes dentro de alguma hierarquia. Apesar da teoria de escala-espaco implicar numa hierarquização das escalas espaciais, podem existir outros aspectos das representações que levem a outros tipos de hierarquia global.

Invariância É a capacidade que uma representação tem em não ser afetada por certos tipos de transformações.

Por se tratar de um conceito já bem estabelecido, fica bastante simples a verificação da presença desta propriedade em uma determinada representação e em certos momentos, este critério virá a opor-se ao critério da sensibilidade anteriormente exposto porém fica a cargo da finalidade do processamento decidir se é melhor a representação ser mais sensível ou mais invariante.

Nos próximos capítulos, as representações-alvo deste trabalho são avaliadas perante os critérios aqui discutidos.

2.6 Avaliação dos processos

Conforme dito anteriormente, muita atenção tem sido dada à formulação e ao aperfeiçoamento dos processos em detrimento das representações. Porém, seria injusti-

ficável negligenciar a importância dos processos e de toda a complexidade de relações que estes guardam com os demais elementos do modelo de processamento da informação visual aqui exposto. Assim, dado que a estrutura do presente trabalho busca destacar inicialmente os elementos menos estudados do modelo, deixaremos as contribuições deste trabalho quanto ao desenvolvimento de novos processos para as posteriores seções 3.3, 3.4, 4.5 e 4.7 onde são detalhados alguns métodos particularmente úteis na obtenção de curvaturas e eixos mediais.

Capítulo 3

A Representação por Curvatura

3.1 Motivação

A geometria diferencial dá ao conceito de curvatura 2D, pelo menos três diferentes interpretações, sendo elas: (i) a taxa de variação angular do campo tangente à forma, (ii) a magnitude da aceleração centrípeta de uma partícula que percorre a forma com velocidade constante e (iii) o inverso dos raios dos círculos que melhor se ajustam localmente à forma¹. A Figura 3.1 ilustra essas interpretações sobre quatro pontos de um mesmo contorno.

Para um contorno contínuo, cuja forma analítica é conhecida, apesar das diferentes interpretações, não há ambigüidade quanto ao cálculo da função curvatura, podendo-se dizer que qualquer equação obtida por meio de qualquer uma das interpretações é correta e deve, portanto, levar ao mesmo resultado. Porém, no caso de contornos obtidos através da amostragem de dados reais (imagens), ou seja, quando *não existem* as equações que descrevem tais formas, o cálculo da curvatura passa a ser um problema “mal posto” e a solução deste problema passa a depender de conhecimentos sobre os dados antes da amostragem e de critérios ligados à finalidade do cálculo. Por exemplo, se sabemos que o conjunto de pontos da Figura 3.2(a) foi produzido pela quantização² de uma forma suave, é válido considerá-lo como sendo a representação discreta de um contorno aproximadamente circular, cabendo, nesta condição, o uso de métodos como em [105] onde se efetua a suavização da

¹tambem chamados *círculos osculantes*

²a quantização espacial geralmente introduz ruídos de alta frequência no sinal.

Figura 3.1: Três diferentes interpretações do conceito de curvatura. (a) Curvatura como a taxa de variação do campo tangente, (b) curvatura como a aceleração centrípeta de uma partícula que percorre a forma com velocidade constante e (c) o conjunto do inverso dos raios dos círculos osculantes à forma.

forma para a atenuação de ruídos (Figura 3.2(b)). Por outro lado, se soubermos que o referido conjunto de pontos é fruto da quantização de uma forma poligonal, também passa a ser válido um resultado advindo de cálculos muito mais simples como em [177, 180, 14] onde simplesmente são calculados os ângulos entre lados consecutivos da representação (poligonal) na grade (Figura 3.2(c)).

O fato é que, para os casos reais, a curvatura (no sentido matemático rigoroso) “*não pode ser calculada exatamente, podendo apenas ser estimada*” [199]. Entretanto, a principal razão que nos leva, ainda assim, a considerar a curvatura como uma representação de formas no contexto da modelagem do processamento da informação visual é a existência de indícios de que o sistema visual do homem e de outros animais seja especialmente sensível à orientação das linhas contidas nas ima-

Figura 3.2: Conjunto de pontos (a) cujas posições estão espacialmente quantizadas e representando (b) um círculo e (c) um polígono.

gens captadas pelos olhos³ [94, 95, 127, 75]. Em outras palavras, há indícios de que o cérebro monte uma representação das cenas que destaca certas características muito semelhantes às características salientadas por este conceito da geometria diferencial chamado curvatura [4], indicando ser esta uma boa aproximação de uma das possíveis representações construídas pelo cérebro para processar imagens.

Attneave e Arnoult [6, 7] foram os primeiros a estudar a importância da curvatura para a percepção visual humana através de experimentos psicofísicos onde figuras contendo apenas os pontos de mais alta curvatura de um dado objeto mostraram se excepcionalmente capazes de representá-lo. Ou seja, supondo que a quantidade total de informação contida em uma imagem não está uniformemente distribuída sobre a mesma, os resultados de Attneave e Arnoult indicam que esta informação está concentrada sobre os pontos de alta curvatura (ou cantos) das bordas da imagem, o que ratifica as idéias de Marr e Hildreth sobre a importância das bordas [134] e

³Não se esquecendo de que os sistemas naturais de visão também são afetados pelo problema da quantização espacial das imagens (ainda que, de alguma forma, solucionem o problema) pois eles geralmente são compostos por fotoreceptores discretos que, em última análise, são semelhantes aos fotodetectores de uma câmera digital comum.

mais, diz quais partes destas bordas guardam a maior parte desta informação. O clássico desenho de Attneave mostrado na Figura 3.3 ilustra como, em muitos casos, bastam os pontos de alta curvatura para que uma idéia, no caso a de um gato, seja transmitida.

Figura 3.3: O gato de Attneave

A curvatura, que é a primeira representação de formas abordada neste trabalho, mede localmente o quão “curvado” ou “torto” é o objeto num determinado ponto, traduzindo essa medida em termos do raio do círculo (ou esfera, no caso 3D) que melhor de ajusta ao objeto naquele ponto⁴.

O uso da curvatura como uma representação de formas já é bastante difundido sendo que inúmeros trabalhos sobre o assunto podem ser encontrados na literatura, desde a aplicação do conceito na solução de problemas práticos [12, 104, 50], passando pela proposição [177, 47] e avaliação [15, 199, 68, 65, 112] de técnicas para o cálculo da representação, até o relacionamento da curvatura com outros conceitos como os eixos de simetria [123] (ou *esqueletos*) e teorias como a teoria de escala-espço [147, 105], além de trabalhos que buscam sua extensão para geometrias não-euclidianas [181, 70, 148].

Devido às várias interpretações possíveis para o conceito de curvatura ilustradas na Figura 3.1, podemos encontrar, na literatura, diversos métodos que buscam aproximar esta função baseando-se: (i) no cálculo da orientação do campo tangente [2, 4, 62, 159, 178, 201, 14], (ii) na aproximação da derivada segunda da curva considerada como um caminho [129, 142, 151] e (iii) na busca pelos círculos osculantes [192, 44]. Tal diversidade nos métodos encontrados é uma característica comum às representações cujas características salientadas têm grande relevância na

⁴Em 3D, devido ao maior grau de liberdade, além da curvatura, existe a chamada *torção* que também tem um importante papel na representação das formas tridimensionais

descrição das formas e, conseqüentemente, para a aplicação na solução de problemas práticos.

3.2 A curvatura frente aos critérios para a representação de formas

Faremos agora um posicionamento da curvatura frente aos critérios para representação de formas propostos na seção 2.5.

Acessibilidade Conforme especificado por este critério, a precisão suficiente para que se alcance a finalidade global do processamento define um conjunto de processos (algoritmos) dentre os quais o mais simples definirá a acessibilidade da representação por curvatura. Em termos práticos, se a precisão na estimação da curvatura não é importante, ou seja, se as características da forma que devem ser salientadas são apenas *qualitativas*, a acessibilidade desta representação fica definida por algoritmos bastante simples como os primeiros que foram propostos na literatura [177, 180, 14] onde o simples cálculo do cosseno dos ângulos entre as retas unindo pontos consecutivos do contorno são utilizados para se classificar pontos como sendo “dominantes”⁵ ou não. Porém, se deseja-se levantar características *quantitativas* sobre os pontos de um contorno, então a acessibilidade da curvatura é menor pois, nestes casos, é necessária a estimação de derivadas de primeira e segunda (em 3D, até de terceira) ordem de funções amostradas que, dependendo da precisão necessária, pode ser realizada através de métodos simples baseados em diferenças finitas [102, 167, 195] ou mesmo através de métodos mais sofisticados como os baseados em interpolações [142], morfologia matemática [121] ou em propriedades da transformada de Fourier [104].

Escopo Uma maneira de se avaliar o escopo de uma determinada representação R é fazer um levantamento das diversas aplicações que a adotam, avaliando-se a diversidade das classes de demais representações R' a partir das quais R é calculada (aqui denominadas representações-fonte). No caso da curvatura, as

⁵Os pontos dominantes de um contorno são os cantos ou extremos de curvatura.

representações-fonte estão limitadas a (i) contornos paramétricos amostrados (ou “*chain-codes*” de Freeman), (ii) interpolações (*B-splines* [142], séries de Fourier [104], etc.) e (iii) imagens em níveis de cinza [109]. Embora relativamente limitada quanto às suas representações-fonte, a curvatura, no que se refere às aplicações, esta entre as representações mais versáteis sendo utilizada desde a remoção de ruídos em imagens [132] até a caracterização morfométrica de neurônios [50], passando pelo reconhecimento de caracteres [79], pela indexação de bases de imagens [144, 149] e detecção de movimento [165].

Caráter escala-espço Conforme já amplamente discutido na literatura, a curvatura permite a análise multi-escala de formas se composta com esquemas de difusão [147, 104, 152], se calculada através de transformadas intrinsecamente multi-resolução como as *wavelets* [103] ou através do uso de diferentes máscaras de diferenças finitas [74].

Suporte local rico Devido a sua natureza local, a curvatura tem um suporte local rico, permitindo a reconstrução da forma representada (a menos de transformações de corpo-rígido) com seus sistemas de coordenadas locais (ver seção 2.5) correspondendo aos pares tangente-normal (orientados) a cada ponto da forma. No entanto, sendo a reconstrução a partir da curvatura um processo integrativo, este está sujeito as dificuldades e limitações comuns à resolução de integrais, sejam estas solúveis por meios analíticos ou numéricos.

Extensão suave e enquadramento Por permitir a seleção dos chamados “pontos dominantes” a curvatura pode levar a extensão suave e ao enquadramento dos sistemas de coordenadas locais correspondentes aos extremos de curvatura, uma vez que seu cálculo pressupõe a definição prévia de uma orientação.

Hierarquia global No caso de curvaturas obtidas através de métodos que a dotem de caráter escala-espço, pode-se formar uma organização hierárquica dos pontos dominantes através da dinâmica destes com relação a seleção de escalas espaciais crescentes.

Invariância Os tipos de transformações para as quais a curvatura clássica apresenta invariância são apenas as transformações de corpo rígido (translações

e rotações). No entanto, já existem trabalhos no sentido de ampliar essa invariância a fim de englobar as transformações afins [70, 148] mas algumas dificuldades relacionadas à estimação de derivadas de ordem elevada ainda não foram resolvidas.

3.3 Comparação de *performance* entre três diferentes abordagens para a estimação numérica de curvatura de contornos.

Nesta seção, uma cuidadosa comparação entre três técnicas para a estimação da curvatura de contornos espacialmente quantizados é apresentada. Duas das técnicas aqui consideradas são baseadas na transformada de Fourier (e serão portanto denominadas técnicas espectrais), sendo que uma opera sobre contornos representados através de sinais unidimensionais (contornos paramétricos) e a outra em sinais bidimensionais, ambas combinadas com regularizações Gaussianas que visam atenuar o ruído de quantização espacial. Enquanto a técnica 1D não representa novidade, já tendo sido alvo de estudos anteriores [103], a abordagem 2D será aqui descrita pela primeira vez⁶. A terceira técnica, baseada em splines, representa uma alternativa mais tradicional. Três classes de curvas são investigadas: formas sintéticas geradas por representações analíticas (equações), *B-splines* e curvas sintéticas geradas através da seleção de componentes no espaço de Fourier. Quatro esquemas de quantização foram considerados: Quantização por intersecção com a grade (Grid intersection quantization - GIQ), quantização em caixas quadradas (Square box quantization - SBQ), quantização pelo uso de um scanner de mesa e quantização pelo uso de uma câmera de vídeo. As *performances* dos métodos foram avaliadas em termos do tempo de execução, erro no valor das curvaturas obtidas e sensibilidade quanto aos parâmetros envolvidos. A terceira técnica mostrou-se a mais rápida, porém a que implicou em maiores erros e as técnicas espectrais apresentaram maior precisão e relativa robustez às configurações dos parâmetros. Uma última observação sobre a técnica espectral 2D revela a possibilidade extra de se obter valores de curvatura em todos os pontos da imagem e a presença de interferências conforme será descrito

⁶Esta seção já foi publicada na forma de artigo em revista [67] e citada por Mokhtarian [150].

adiante. Tais resultados são importantes não somente pela caracterização relativa de *performance* dos métodos considerados como também por compor um guia prático para aqueles interessados na aplicação de tais técnicas a problemas reais.

3.3.1 Motivação

A análise de formas bidimensionais é um dos mais clássicos, importantes e amplamente estudados problemas em visão computacional, encontrando aplicações em inúmeros problemas práticos. Porém, apesar da aparente simplicidade, não existe uma abordagem definitiva para a classificação e análise de formas 2D e mais: muito esforço ainda é necessário principalmente para que as técnicas já existentes sejam adequadamente avaliadas em termos comparativos. Dentre os diferentes métodos para a análise de formas 2D, aqueles baseados na curvatura da borda dos objetos constituem alguns dos mais abrangentes e promissores conforme será discutido detalhadamente em breve. Os dois principais objetivos desta seção são: introduzir um novo método baseado em transformada de Fourier 2D para a estimativa de curvatura e avaliar comparativamente a *performance* deste junto de dois outros métodos comumente utilizados para o mesmo fim. Os resultados decorrentes não somente produziram valiosas indicações sobre as vantagens e desvantagens específicas de cada método, como também esclareceram alguns detalhes de importância prática para aqueles que desejarem aplicar tais técnicas na solução de problemas reais. Por fim, o esquema de avaliação de *performance* adotado pode ser estendido a outras técnicas para o cálculo de curvatura não cobertos neste trabalho.

Primeiramente, até o ponto em que a preservação da informação interessa, curvatura é uma representação completa, no sentido em que o contorno original pode ser reconstruído (a menos de isometrias). Além disso, dado o fato de que nem todos os pontos da forma são igualmente relevantes do ponto de vista perceptual [6], a localização de pontos salientes (ou críticos) é uma tarefa extremamente importante para a extração de características e segmentação de contornos. Curvatura também é de fundamental importância no seguinte contexto: existem vários métodos que procuram localizar tais pontos críticos em termos dos pontos máximos e mínimos locais de curvatura, bem como regiões de curvatura zero (linhas retas e inflexões). Por último, interessantes analogias físicas baseadas na idéia de curvatura têm levado ao

surgimento de poderosas técnicas para análise de formas. Como exemplo, a teoria de elasticidade fornece o conceito da *energia de dobramento*, uma característica global da forma calculada em função da integral do módulo da curvatura ao quadrado ao longo do contorno [201, 28] que pode ser usada na caracterização da complexidade das formas [106].

Aqui, descreveremos e discutiremos um esquema para a avaliação de *performance* de métodos para a estimação digital de curvatura, bem como os resultados da comparação entre duas diferentes abordagens baseadas em propriedades da transformada de Fourier e um terceiro método que utiliza interpolação por *B-splines* [142]. A ênfase dada nesta seção à abordagem por Fourier é motivada pela inerentemente alta precisão apresentada por esta abordagem [104]. Em outras palavras, a técnica por Fourier permite a consideração de toda uma vizinhança em torno de cada ponto em que a curvatura está sendo calculada, e não apenas numa vizinhança limitada no tamanho e na resolução conforme acontece com diversos outros métodos alternativos com em [142]. A transformada de Fourier discreta pode também ser eficientemente calculada usando-se a FFT. Bônus adicionais permitidos por esta abordagem incluem o fato de que a forma pode ser suavizada no domínio de Fourier e que os descritores de Fourier, que são valiosos recursos para a caracterização e classificação de formas, são obtidos como subprodutos.

Basicamente, existem dois modos distintos através dos quais se pode estimar a curvatura de uma forma digital. A primeira e mais tradicional delas se dá a partir da borda parametrizada do objeto e a outra obtém a curvatura diretamente da imagem bidimensional (sem a necessidade de parametrização) através do conceito de curvas isopotenciais. Técnicas que usam contornos parametrizados são mais populares e amplamente usadas em diferentes problemas de visão computacional como em biomedicina, reconhecimento de impressões digitais e reconhecimento óptico de caracteres [172, 42, 131], apresentando diversas vantagens interessantes tais como o fato que lidar inerentemente com dados unidimensionais é computacionalmente menos custoso e a existência de uma literatura técnica substancial em análise de contornos [49].

No entanto, uma das principais desvantagens desta abordagem é o fato de que o objeto precisa ter o seu contorno extraído de forma parametrizada. Por outro

lado, a técnica que é aplicada diretamente nos dados bidimensionais não depende da extração do contorno e permite que hardware dedicado seja usado para implementar as rotinas de processamento. E ainda, o tamanho dos dados de entrada torna-se independente do perímetro do objeto, o que não acontece com a abordagem baseada em contorno. Uma boa discussão sobre a robustez da computação de derivadas sobre curvas de nível pode ser encontrada em [184].

Ao leitor são sugeridos [68, 168, 128, 139] para uma revisão detalhada de diversas técnicas de estimação de curvatura em imagens digitais. O principal problema que deve ser resolvido para a estimação de curvatura digital é que a expressão da curvatura envolve a derivada de dados discretos, o que é um problema mal formulado que deve ser contornado pela introdução de um procedimento de regularização. A abordagem padrão é baseada na suavização dos dados, por exemplo, pela convolução do contorno com funções Gaussianas, que é o caso, por exemplo, do trabalho clássico de Mokhtarian and Mackworth [147]. Abordagens alternativas incluem a interpolação local dos dados [142] e o uso de diferentes medidas discretas de curvatura, tal qual a *C-curvature* de Davis [58]. Por outro lado, as técnicas baseadas em 2D também têm recebido a atenção da comunidade de processamento de imagens. Alguns exemplos de trabalhos dedicados a esta abordagem podem ser encontrados em [40, 109].

Outro aspecto importante levantado nesta seção relaciona-se à validação e comparação das técnicas 1 e 2D baseadas em Fourier. De fato, pesquisadores da ciência da visão têm sido criticados não somente por não dar suficiente atenção à caracterização da *performance* de novas abordagens respectivamente a um conjunto representativo de dados, mas também por não comparar tais novas técnicas com alternativas mais tradicionais [99, 114]. Apenas recentemente alguns trabalhos abordaram o tópico da avaliação de *performance* de uma forma sistemática [13, 41, 59, 91] e uma das principais características desta seção é não servir de alvo para tal tipo de crítica. Primeiramente, uma atenção especial foi dada à definição de um conjunto representativo de curvas fechadas. A razão pela qual nos concentramos em curvas fechadas é que estas são inerentemente compatíveis com a estimação de derivadas por Fourier e porque estas correspondem às fronteiras dos objetos. Em outras palavras, o fato da extensão periódica dessas curvas ser contínua permite boa precisão na interpolação da curva e na estimação de suas derivadas, o que não é geralmente

verdade para curvas abertas.

Embora saibamos que é virtualmente impossível considerar um conjunto de dados de teste completamente representativo, acredita-se que é melhor considerar um conjunto amplo de formas, mesmo que não seja completamente representativo, do que considerar apenas um conjunto restrito como ocorre às vezes na literatura. Mais especificamente, três tipos de curvas foram consideradas: (a) curvas definidas por expressões analíticas, tais como círculos senoidais e espirais; (b) curvas definidas por *B-splines* [155]; e (c) curvas sintetizadas em termos de componentes espectrais [202]. É importante notar que, em cada caso, é possível obter uma quantificação analítica precisa da curvatura ao longo destes contornos para ser usada como um padrão de comparação. Cada curva é espacialmente amostrada antes da estimação numérica segundo quatro esquemas de quantização: (i) quantização por intersecção com a grade, (ii) quantização por caixas quadradas, (iii) usando um scanner de mesa e (iv) usando uma câmera de vídeo. Observe que (iii) e (iv) envolvem ainda a impressão dos contornos por uma impressora laser convencional. Estes dois últimos esquemas provêm uma caracterização mais global dos métodos considerados dado que correspondem a situações práticas típicas.

Grande atenção foi também dada na definição de figuras de mérito abrangentes e imparciais, permitindo a caracterização adequada das técnicas consideradas. As características mais importantes que foram levadas em conta foram a precisão no cálculo da curvatura definida em termos do erro RMS (*Root mean square*) entre a curvatura analítica e a curvatura obtida numericamente, a robustez quanto ao ajuste dos parâmetros, expressa em termos da distribuição dos parâmetros ótimos (os parâmetros de escala dados pelo desvio padrão das Gaussianas de suavização) e o tempo de execução.

Além disso, na tentativa de caracterizar a *performance* das técnicas consideradas de maneira formal, abrangente e comparativa, emergiram algumas questões particularmente interessantes, definidas inerentemente pelas características específicas das técnicas baseadas na transformada de Fourier que deveriam, pelo menos, serem respondidas pela metodologia de avaliação proposta. Como exemplo, já que derivadas de primeira e segunda ordem são necessárias nas técnicas por Fourier, seria interessante verificar o uso de desvios padrões distintos (para primeiras e segundas

derivadas) nas Gaussianas de suavização. Em outras palavras, dado que a derivada de segunda ordem implica numa maior “amplificação” das altas frequências do que a de primeira ordem, é interessante checar se o uso de uma suavização maior para a derivada segunda acarretaria em melhoras na precisão dos resultados. Outro ponto importante a ser investigado refere-se à robustez das técnicas dados diferentes conjuntos de parâmetros.

Esta seção inicia-se descrevendo as três abordagens numéricas consideradas para o cálculo da curvatura, incluindo a abordagem baseada na transformada de Fourier 2D (introduzida em [65]) que é descrita em maiores detalhes. Depois, o esquema que avaliação de *performance* é apresentado em detalhes, seguindo pelos resultados obtidos e respectiva discussão. A seção termina apresentando as conclusões gerais bem como as possibilidades de futuros desenvolvimentos.

3.3.2 Estimação de curvatura baseada em Fourier

Conforme é sabido da álgebra linear, transformações lineares unitárias provêm uma representação alternativa completa (no sentido de ser inversível) de sinais nos domínios do “tempo” (sinais 1D) ou do “espaço” (sinais 2D). De fato, o real efeito de tais transformações consiste em expressar o sinal em uma nova base que é obtida através de uma rotação do sistema de coordenadas. A vantagem de transformações como a transformada de Fourier, que é uma transformação quase-unitária, vêm do fato de que a base exponencial complexa usada nesta transformação, devido a sua natureza altamente correlacionada, normalmente permite que os sinais sejam representados numa forma muito mais compacta, ou seja, com apenas alguns componentes espectrais. No entanto, a transformada de Fourier também possui uma série de outras propriedades interessantes, tal como a propriedade de diferenciação, dada nas equações (3.1) e (3.2), respectivamente para sinais 1D e 2D, onde $H(f) = \mathfrak{F}(h(t))$ e $H(u, v) = \mathfrak{F}(h(x, y))$. Tais propriedades permitem um modo interessante de se estimar derivadas de funções amostradas, já que a transformada discreta de Fourier pode ser calculada rápida e eficientemente.

$$\frac{d^n}{dt^n} h(t) = \mathfrak{F}^{-1}\{(i2\pi f)^n H(f)\} \quad (3.1)$$

$$\frac{\partial^n}{\partial x^n} \frac{\partial^m}{\partial y^m} h(x, y) = \mathfrak{F}^{-1} \{ (i2\pi u)^n (i2\pi v)^m H(u, v) \} \quad (3.2)$$

onde $i = \sqrt{-1}$. Dada a natureza diferencial da curvatura, claramente observada na Eq.(3.3), as propriedades (3.1) e (3.2) podem servir de meios para que a curvatura seja numericamente estimada. Note que as variáveis x e y na Eq.(3.3) referem-se às funções paramétricas $x(t)$ e $y(t)$ definidas pelo contorno. Inicialmente proposto em [12] em termos da série de Fourier em conjunto com janelas de regularização do tipo Kaiser, esta possibilidade tem sido desenvolvida e aplicada numa série mais recente de desenvolvimentos [49, 104, 105] que se valem da transformada de Fourier e da suavização Gaussiana, sendo esta última essencial na regularização dos ruídos de alta frequência introduzidos pela quantização espacial. De fato, o desvio padrão da Gaussiana de suavização define um parâmetro de escala permitindo a representação da curvatura em multi-resolução [104, 105]. Enquanto tais trabalhos se concentraram na estimação da curvatura de contornos fechados bidimensionais, feito em termos da transformada de Fourier 1D, também é interessante investigar a possibilidade de uso da transformada de Fourier 2D. A principal vantagem desta abordagem é que a mesma pode ser usada para estimar a curvatura pelo uso da Eq. (3.4), ao longo de todas as curvas isopotenciais definidas pela superfície $\phi(x, y)$ contendo a forma do contorno original como uma de suas curvas de nível. Uma vantagem secundária é que o tamanho dos dados de entrada não varia necessariamente com o perímetro das formas conforme acontece na abordagem 1D.

$$k = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{(\dot{x}^2 + \dot{y}^2)^{3/2}} \quad (3.3)$$

$$k = \nabla \cdot \frac{\nabla \phi}{\|\nabla \phi\|} = \frac{\phi_{xx}\phi_y^2 - 2\phi_x\phi_y\phi_{xy} + \phi_{yy}\phi_x^2}{(\phi_x^2 + \phi_y^2)^{3/2}} \quad (3.4)$$

As próximas subseções apresentam o método baseado em splines, Fourier 1D e 2D, respectivamente, com atenção especial dada ao último já que este é apresentado aqui pela primeira vez.

Estimativa da curvatura baseada na propriedade de diferenciação da transformada de Fourier 1D

O método de estimação da curvatura discutida nesta sub-seção se originou do conceito de curvograma discutido em [104, 105] e tem sido considerado em várias aplicações [106]. Sejam $c(n) = (x(n), y(n))$ o contorno paramétrico de interesse com $n = 0, \dots, N-1$, e N o número de pontos ao longo do mesmo. O contorno pode ser representado como um sinal complexo $u(n) = x(n) + iy(n)$. Uma ferramenta fundamental para esta abordagem é o par de transformadas de Fourier 1D de $u(n)$ dado por:

$$U(s) = \mathfrak{F}\{u(n)\} = \sum_{n=0}^{N-1} u(n) e^{-i2\pi(sn/N)}, s = 0, \dots, N-1, \quad (3.5)$$

$$u(n) = \mathfrak{F}^{-1}\{U(s)\} = \frac{1}{N} \sum_{s=0}^{N-1} U(s) e^{i2\pi(sn/N)}, n = 0, \dots, N-1, \quad (3.6)$$

A função auxiliar $\chi(s)$ é uma ferramenta útil para a estimação das derivadas discretas de $u(n)$,

$$\chi(s) = \begin{cases} s & \text{se } s = 0, 1, \dots, (N - \lfloor N/2 \rfloor - 1), \\ N - s & \text{se } s = (N - \lfloor N/2 \rfloor), (N - \lfloor N/2 \rfloor + 1), \dots, (N - 1). \end{cases} \quad (3.7)$$

Sendo $\lfloor N/2 \rfloor$ a função que trunca o valor $N/2$. A função $\chi(s)$ implementa o alinhamento necessário à representação usualmente produzida pela DFT (i.e, a representação da frequência formada pela segunda parte seguida pela primeira parte do próximo período). Os sinais precisam ser suavizados por causa do efeito de amplificação das altas frequências produzido pela diferenciação numérica. A versão suavizada $u(n, \sigma)$ de $u(n)$ é dada por:

$$u(n, \sigma) = \mathfrak{F}^{-1}\{U(s)G_\sigma(s)\}, \quad (3.8)$$

onde $G_\sigma(s) = e^{-(\sigma\chi(s))^2}$. O deslocamento do contorno, neste caso, é evitado pela escolha de um parâmetro de escala espacial σ cujo valor seja o menor possível a fim

de filtrar o ruído de quantização espacial, mas de não distorcer demais o contorno. As derivadas suavizadas de primeira e segunda ordem de $u(n)$ ficam definidas como:

$$\dot{u}(n, \sigma) = \mathfrak{F}^{-1}\{i2\pi\chi(s)U(s)G_\sigma(s)\}, \quad (3.9)$$

$$\ddot{u}(n, \sigma) = \mathfrak{F}^{-1}\{-(2\pi\chi(s))^2U(s)G_\sigma(s)\}. \quad (3.10)$$

A curvatura multi-escala de $u(n)$ fica dada por:

$$k(n, a) = \frac{-\dot{u}(n, a)\ddot{u}^\dagger(n, a)}{|\dot{u}(n, a)|^3}, \quad (3.11)$$

onde $z^\dagger$ denota o complexo conjugado e $|z|$ denota o modulo de z .

Embora seja possível aplicar o método por Fourier 1D a contornos com qualquer número de pontos, é desejável que o número de pontos amostrados seja uma potência de dois a fim de que a eficiência da transformada rápida de Fourier (FFT) possa ser utilizada. Tal necessidade pode ser facilmente satisfeita através da simples interpolação linear ao longo da curva original a fim de produzir uma nova representação com o número de pontos desejados, ou seja, a menor potência de dois maior ou igual ao número de pontos original.

Método baseado em Fourier 2D

Dado um contorno c não necessariamente parametrizado, o primeiro passo para que se possa aplicar o método aqui proposto para calcular a curvatura de c consiste em criar uma superfície que contenha c como sendo uma de suas curvas de nível. Por se tratar de um contorno simples e fechado e apelando para a simplicidade, formamos uma superfície $\phi(x, y)$ onde

$$\phi(x, y) = \begin{cases} 0 & \text{se o ponto } (x, y) \text{ é externo ao contorno,} \\ 1/2 & \text{se } (x, y) \text{ está sobre contorno,} \\ 1 & \text{se } (x, y) \text{ está no interior do contorno.} \end{cases} \quad (3.12)$$

É claro que outras superfícies como a transformada distância dotada de sinal (positivo dentro e negativo fora) poderiam ter sido utilizadas, porém como a es-

colha pela superfície seria sempre arbitrária, preferimos optar pela simplicidade da função ϕ acima definida a contar pela rapidez na sua formação (o cálculo de uma transformada distância, mesmo que aproximada, demandaria muito mais tempo de processamento) e pelo fato de ϕ não ser uma superfície contínua. Assim, se a curvatura calculada sobre a superfície ϕ por este método gozasse de boa precisão (o que realmente aconteceu conforme será visto nas próximas sub-seções), não seria estranho imaginar que o mesmo método aplicado a uma superfície mais “bem comportada” como a transformada distância poderia levar a resultados ainda melhores.

Dessa forma, usando a propriedade (3.2) para calcular a equação (3.4), obtém-se uma aproximação numérica da curvatura de $\phi(x, y) = 1/2$. Da mesma forma que no caso 1D descrito na seção anterior, é também necessário regularizar $\phi(x, y)$ dado que esta superfície está espacialmente amostrada. Isto é feito através da convolução de $\phi(x, y)$ com uma Gaussiana 2D circularmente simétrica (Eq. (3.13)) que é eficientemente realizada no domínio de Fourier.

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (3.13)$$

Devemos observar que, embora inicialmente todos os pontos originais do contorno estejam sobre a mesma curva de nível, isto não é mais verdade após a regularização. No entanto, as curvas não se deslocam muito se apenas um pequeno grau de suavização for aplicado (suficiente apenas para filtrar as altas frequências decorrentes da amostragem). Assim, para extrair a curvatura, assumiremos que o contorno se manterá sobre as suas coordenadas (x, y) originais e delas tomaremos os valores da curvatura.

3.3.3 Cálculo de curvatura baseada na aproximação da curva por splines

Esta primeira abordagem assume que as derivadas da expressão da curvatura (Eq. (3.3)) são calculadas com base em uma aproximação local por partes dos pontos do contorno. Mais especificamente, suponha que a curva que será ajustada é um polinômio cúbico em t , com $t \in [0, 1]$ que aproxima o segmento da curva entre dois pontos do contorno $A(t = 0)$ e $B(t = 1)$. Portanto, para este segmento de curva, a aproximação

é definida como:

$$x(t) = a_1 t^3 + b_1 t^2 + c_1 t + d_1 \quad (3.14)$$

$$y(t) = a_2 t^3 + b_2 t^2 + c_2 t + d_2 \quad (3.15)$$

Calculando as respectivas derivadas e substituindo-as na expressão da curvatura, Eq. (3.3), pode ser facilmente verificado que a curvatura no ponto $A(t = 0)$ do contorno é:

$$k = 2 \frac{c_1 b_2 - c_2 b_1}{(c_1^2 + c_2^2)^{3/2}} \quad (3.16)$$

Medioni e Yasumoto [142] usaram *B-splines* cúbicas com nós igualmente espaçados para ajustar por partes os segmentos paramétricos da curva. Os coeficientes b_1 , b_2 , c_1 e c_2 da Eq. (3.16) acima podem ser calculados, a partir das expressões acima, como (veja [142] para maiores detalhes):

$$b_1 = \frac{1}{12}((x_{n-2} + x_{n+2}) + 2(x_{n-1} + x_{n+1}) - 6x_n), \quad (3.17)$$

$$b_2 = \frac{1}{12}((y_{n-2} + y_{n+2}) + 2(y_{n-1} + y_{n+1}) - y_n), \quad (3.18)$$

$$c_1 = \frac{1}{12}((x_{n+2} - x_{n-2}) + 4(x_{n+1} + x_{n-1})), \quad (3.19)$$

$$c_2 = \frac{1}{12}((y_{n+2} - y_{n-2}) + 4(y_{n+1} + y_{n-1})) \quad (3.20)$$

A curvatura é calculada pela substituição dos coeficientes acima na Eq. (3.16).

3.3.4 O sistema de avaliação

A fim de que diferentes métodos numéricos sejam adequadamente comparados, é importante definir um sistema de avaliação computacional completo que seja tão

justo e abrangente quanto possível. Esta tarefa vincula três questões principais: a definição de um conjunto adequado de contornos de teste, a modelagem dos esquemas de quantização espacial e a identificação de figuras de mérito capazes de expressar quão bem os métodos correspondem às expectativas que se têm sobre eles. As próximas subseções apresentam e discutem cada uma destas questões.

Os contornos considerados

Antes que as *performances* de diferentes técnicas para a estimação da curvatura sejam adequadamente avaliadas, é necessário definir o padrão de dados de entrada que serão fornecidos aos métodos. Três importantes propriedades são esperadas de tais formas. Primeiro, elas devem permitir uma descrição analítica da curvatura, já que isto é essencial para a quantificação da precisão de cada método considerado. Segundo, o conjunto de formas deve ser tão geral quanto possível a fim de proporcionar uma amostra representativa das formas que podem ser encontradas em problemas típicos de análise de formas. Enquanto a primeira destas propriedades, na prática, não significa um problema, é virtualmente impossível considerar um conjunto completamente representativo de formas. Aqui, o fator limitante é o recurso computacional necessário ao processamento e análise da *performance* de um conjunto muito grande de formas. Uma terceira propriedade requerida relaciona-se ao fato de que a curvatura analítica definida pelas curvas não deve ser muito alta ou variar de maneira demasiadamente abrupta, a fim de serem adequadamente representadas nas grades ortogonais utilizadas na quantização espacial das curvas, já que não é justo esperar que uma técnica numérica estime uma curvatura que não está adequadamente representada na imagem. Tal propriedade foi satisfeita pela exclusão de curvas que possuísem raios de curvatura menores que $1/3$ de um pixel (a unidade espacial), exceto para um número pequeno de regiões isoladas.

Para que sejam satisfeitos os requisitos acima, consideramos três diferentes classes de curvas simples e fechadas. Primeiro, conforme ilustrado nas Figuras 3.4 e 3.5, temos curvas analíticas tais como círculos senoidais e espirais, respectivamente (embora não sejam fechadas em suas versões analíticas, as espirais foram fechadas “manualmente”), que provêm uma ampla variedade de formas, apesar de limitadas pela natureza das respectivas funções (por exemplo, funções harmônicas tenderão a

Figura 3.4: Círculos senoidais com variado número de ramos e raios internos.

Figura 3.5: Espirais com diferentes evoluções na largura.

formar curvas suaves e assim por diante). Segundo, as splines cúbicas geradas em termos dos pontos de controle interativamente fornecidos pelo operador e mostradas na Figura 3.6 foram consideradas na tentativa de representar formas mais gerais tais quais aquelas encontradas em situações reais. Porém, a generalidade destas splines também é limitada por suas características inerentes como, por exemplo, sua continuidade paramétrica de ordem C^n . A terceira classe de formas apresentadas

Figura 3.6: Alguns dos contornos gerados por *B-splines* considerados.

na Figura 3.7 inclui curvas definidas no domínio de Fourier [202] e foram sintetizadas pela transformada inversa de Fourier. Novamente, tais curvas implicam em propriedades específicas tais como sendo C^n .

Figura 3.7: Formas sintetizadas por componentes do espaço de Fourier com diferentes períodos e composições harmônicas.

Esquemas de quantização espacial

Considerando que as formas selecionadas devem ser contínuas e suaves a fim de permitir a quantificação exata da curvatura que será usada como padrão de comparação, é necessário utilizar esquemas de quantização adequados que mapeiem estas curvas contínuas numa grade ortogonal comum a qualquer imagem digital. Quatro tradicionais esquemas de quantização foram considerados aqui a fim de proporcionar uma visão geral da *performance* das técnicas numéricas quanto à variação das condições de quantização. Dois dos quatro métodos são definidos precisamente em termos matemáticos: a quantização por interseção da grade [78] (*grid intersection quantization*) ou GIQ, e a quantização por caixas quadradas (*square box quantization*) ou SBQ. Os outros dois esquemas de quantização consistem no uso de um scanner de mesa (HP Scanjet 4L) e uma câmera de vídeo (Sony CCD Iris) na aquisição de imagens de impressões de alta qualidade das formas (HP LaserJet 4L, 300 dpi). A Figura 3.8 ilustra as representações GIQ 3.8(a) e SBQ 3.8(b) de um mesmo objeto junto de uma mesma região ampliada 3.8(c) e 3.8(d). Como o SBQ acarreta tipicamente numa quantização espacial mais densa da curva analítica original, poderíamos esperar que erros menores sejam obtidos por uma técnica numérica de estimação de curvatura operando sobre tais representações.

Deveria ser também observado que os esquemas GIQ e SBQ podem produzir representações com pontos duplos tais como o ponto marcado com um asterisco na Figura 3.9(a). Tais pontos podem transformar uma curva originalmente regular em uma curva não regular e por isso é importante removê-los. Isto é feito incorporando-se condições adicionais nos esquemas de quantização (remove-se arbitrariamente um dos dois pontos com mesmas coordenadas).

Algumas observações sobre a detecção de bordas incluem o fato de que, enquanto esta operação não é necessária nos esquemas GIQ e SBQ, a simples aplicação de um limiar seguido da detecção de bordas binárias foi o que permitiu a obtenção das formas adquiridas pelo scanner e pela câmera. Em situações práticas gerais, esquemas tradicionais de detecção de bordas (veja, por exemplo, [49, 84, 182]) podem ser aplicados. As leves diferenças entre as bordas obtidas por diferentes métodos são uniformizadas pela suavização Gaussiana inerente à aplicação de diferentes algoritmos para extração de contornos que geralmente afetam a estimação da curvatura

[15].

Figura 3.8: Quantização por (a) caixas quadradas e (b) intersecção da grade de uma curva paramétrica específica e respectivas regiões ampliadas (c) e (d).

Figura 3.9: Pontos repetidos (a) e representação após a remoção de tais pontos (b).

Figuras de mérito

O sistema de avaliação aqui adotado consiste em apresentar quatro versões de cada uma das curvas quantizadas pelos métodos acima citados a cada uma das

técnicas numéricas e comparar a curvatura numericamente estimada com a curvatura analítica ao longo do contorno. Conforme já foi observado, três características de *performance* foram especialmente levadas em conta: (a) os tempos de execução, (b) a precisão numérica e (c) a sensibilidade da técnica à configuração dos parâmetros, aqui quantificados em termos da dispersão dos parâmetros ótimos e a extensão das escadas espaciais para as quais os erros não excedam 10% do erro mínimo.

Uma das mais diretas e intuitivas formas de se quantificar o erro dá-se através de métricas tradicionais tais como o erro RMS, que é a abordagem adotada no presente trabalho. Mais especificamente, para cada forma e método considerado, um erro global ϵ é obtido através da equação (3.21), onde k_0 é a curvatura calculada usando-se as técnicas propostas, k_a é a curvatura analítica quantizada e N é o número de pontos do contorno.

$$\epsilon_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N (k_0(i) - k_a(i))^2} \quad (3.21)$$

No entanto, enquanto a distância euclidiana provê uma medida global de similaridade entre todas as curvaturas analíticas e estimadas ao longo das curvas, ela também apresenta algumas desvantagens. Por exemplo, uma grande diferença em apenas um único ponto pode influenciar genericamente o erro total. Uma vez que tal tipo de problema ocorrerá com qualquer métrica, nós tentamos controlar tais efeitos equalizando as curvaturas (ambas analítica e estimada) por meio de uma função sigmóide, limitando assim os valores de curvatura no intervalo $[-3, +3]$ antes que a distância euclidiana fosse calculada. Este processo envolveu a Equação (3.22), onde x denota o valor da curvatura e a é o máximo valor absoluto permitido.

$$S(x) = a \left[\frac{e^{x/a} - e^{-x/a}}{e^{x/a} + e^{-x/a}} \right] = a \left[1 - \frac{2}{e^{2x/a} + 1} \right] \quad (3.22)$$

Em busca de uma caracterização mais completa do erro estimado, as distâncias euclidianas foram organizadas em termos de histogramas.

Com relação à terceira medida de *performance* — a sensibilidade de cada método a configurações específicas de parâmetros — duas medidas foram consideradas. Primeiro, é interessante quantificar a dispersão dos parâmetros ótimos. Para cada

técnica, o melhor conjunto de parâmetros (i.e., desvio padrão) é identificado, em termos do mínimo erro pela abordagem da força-bruta (i.e., é feita uma varredura no espaço de parâmetros) e a respectiva distribuição é expressa através de histogramas. Valores altos de dispersão indicam que o valor de desvio padrão ótimo varia consideravelmente de uma curva para outra, o que torna a escolha de um conjunto adequado de parâmetros algo particularmente difícil. Valores baixos indicam que os parâmetros ótimos estarão situados em um intervalo razoavelmente pequeno. A segunda medida de sensibilidade expressa a extensão da região onde os respectivos parâmetros não conduziram a um erro superior a 10% do erro mínimo em cada caso (Figura 3.10), aqui considerados na forma de histogramas. Conseqüentemente, enquanto melhores *performances* são caracterizadas por grandes extensões, pequenas extensões indicam alta sensibilidade à configuração dos parâmetros.

Figura 3.10: Extensão da escala espacial considerada como uma medida de sensibilidade.

3.3.5 Resultados

Esta seção apresenta as medidas de *performance* das técnicas em questão com respeito aos três aspectos considerados: (a) tempo de execução, (b) erros no valor estimado e (c) sensibilidade aos parâmetros. Os resultados apresentados foram obtidos usando-se um computador Pentium(r) II 333MHz com 256Mb de memória RAM rodando implementações feitas em Delphi (versões 3 e 4) sobre a plataforma Windows(r) 95.

Tempos de execução

Histogramas dos tempos de execução, obtidos pelo uso do relógio interno do sistema operacional, são apresentados nas Figuras 3.11-3.13 para Fourier 1D, Fourier 2D e Medioni, respectivamente. Deve-se observar que os tempos para o Fourier 1D correspondem às versões interpoladas das curvas conforme descrito na seção 3.3.2, já que isto permitiria o uso da transformada rápida de Fourier (FFT). Observa-se também que os histogramas apresentados incluem todas as classes de curvas e esquemas de quantização. O tempo necessário à interpolação, no caso do Fourier 1D, não foi computado como parte dos tempos de execução. Quando interpolados, a maioria dos contornos ficou com 1024 pontos.

Figura 3.11: Histograma de tempos para a versão interpolada do método Fourier 1D. Valor médio: $0,008 \pm 0,003$ s.

Histogramas dos erros na estimação da curvatura

A Figura 3.14 apresenta o comportamento típico observado para o erro estimado em termos do parâmetro de escala (desvio padrão da Gaussiana σ) com respeito aos métodos por Fourier 1D e 2D. Deve-se observar que as escalas são logarítmicas.

A fim de verificar as possíveis melhorias advindas do uso de dois valores distintos para o desvio padrão σ_1 e σ_2 , respectivamente como parâmetros de regularização para as primeiras e segundas derivadas (veja Seção 3.3.1), os métodos baseados na transformada de Fourier foram executados para diversas combinações de tais

Figura 3.12: Histograma de tempos do método por Fourier 2D. Valor médio: $3,5 \pm 0,2$ s.

Figura 3.13: Histograma de tempos para o método Medioni. Valor médio: $0,002 \pm 0,005$ s.

parâmetros. As Figuras 3.15 e 3.16 apresentam resultados típicos com respeito à B-spline da Figura 3.6(b) e ao círculo senoidal da Figura 3.4(c), respectivamente. Estas figuras ilustram as situações onde o erro mínimo recai sobre e fora da diagonal principal ($\sigma_1 = \sigma_2$).

A partir da verificação em diversos casos de que valores distintos para o desvio padrão não contribuíram significativamente para minimizar os erros, todos os resultados subseqüentes foram apresentados com relação a um mesmo valor de desvio padrão para ambas primeiras e segundas derivadas. As Figuras 3.17-3.19 apresentam os histogramas de erros com respeito às três técnicas de estimação de curvatura

Figura 3.14: Curva de erro típica para os métodos por Fourier 1D (a) e 2D (b). Ambas as escalas são logarítmicas.

consideradas, incluindo todas as curvas e esquemas de quantização. A Figura 3.17 foi obtida para os erros produzidos pela técnica por Fourier 1D sem o procedimento de interpolação exposto na seção 3.3.2, a fim de permitir uma melhor comparação entre as diferentes quantizações. As Figuras 3.18 e 3.19 correspondem ao Fourier 2D e à abordagem de Medioni. A Figura 3.20 apresenta os erros para a abordagem de Medioni com respeito aos diversos esquemas de quantização.

Sensibilidade aos parâmetros

Esta seção apresenta os resultados obtidos quanto a quantificação da sensibilidade dos parâmetros necessários a aplicação da técnicas por Fourier 1 e 2D (a técnica de

Figura 3.15: Erros típicos obtidos na estimação da curvatura considerando-se uma série de combinações dos parâmetros de regularização σ_1 e σ_2 . Situação onde o erro mínimo recai sobre a diagonal.

Medioni envolve nenhum parâmetro). As Figuras 3.21 e 3.22 apresentam histogramas caracterizando a dispersão dos parâmetros de escala correspondentes ao mínimo erro no cálculo da curvatura respectivamente para os métodos por Fourier 1D e 2D. As Figuras 3.23 e 3.24 apresentam os histogramas da extensão do parâmetro de escala que garante um erro, no máximo, 10% maior que o erro mínimo obtido para uma dada curva, respectivamente para os métodos por Fourier 1D e 2D. E finalmente, a Figura 3.25 ilustra um curioso fenômeno observado para algumas curvas (especialmente aquelas apresentando estreitamentos) onde alguns picos de erro aparecem em decorrência da interferência entre porções da curva que, embora estejam perimetralmente distantes, estão próximas no espaço bidimensional.

Figura 3.16: Erros típicos obtidos na estimação da curvatura considerando-se uma série de combinações dos parâmetros de regularização σ_1 e σ_2 . Situação onde o erro mínimo não recai sobre a diagonal.

3.3.6 Discussão

Considerando as técnicas baseadas na transformada de Fourier, o primeiro resultado experimental a ser verificado (referente às Figuras 3.15 e 3.16) indicam que não há vantagem substancial no uso de diferentes parâmetros de regularização no cálculo das derivadas de primeira e segunda ordem, o que nos motivou a considerar $\sigma = \sigma_1 = \sigma_2$. Um resultado particularmente importante quanto às técnicas por Fourier está no fato de que o valor ótimo de σ (implicando no mínimo erro) deve ser grande o suficiente para filtrar os ruídos causados pela quantização espacial das curvas, mas deve também ser o menor possível a fim de não remover os detalhes de pequena escala que compõem a curva. Os valores ótimos de σ mostraram-se não

Figura 3.17: Erros na estimação da curvatura para o método por Fourier 1D sem interpolação. Valor médio: $0,05 \pm 0,04$.

Figura 3.18: Erros na estimação da curvatura para o método por Fourier 2D. Valor médio: $0,05 \pm 0,05$.

variar substancialmente entre diferentes classes de curvas, o que é uma propriedade desejável que depõe a favor da robustez dos métodos. Conforme indicado pelas Figuras 3.21 e 3.22, os valores ótimos de σ de mais de 70% das curvas estiveram contidos no intervalo $[1,5 \text{--} 4]$. Sobre a outra medida de sensibilidade, a extensão (ver Figura 3.10), cerca de 80% de todos os casos processados pelas técnicas Fourier 1D e 2D (ver Figuras 3.23 e 3.24) apresentaram uma extensão não maior que 10 pixels, indicando robustez quanto a escolha do σ .

Conforme indicado pelas Figuras 3.17, 3.18 e 3.19, a precisão das técnicas baseadas em Fourier proporcionam erros consistentemente menores (por um fator de

Figura 3.19: Erros na estimação da curvatura para o método de Medioni e Yasumoto. Valor médio: $0,4 \pm 0,2$.

Figura 3.20: Erros na estimação de curvatura da técnica de Medioni e Yasumoto para diferentes esquemas de quantização: (a) GIQ, (b) SBQ, (c) scanner e (d) câmera de vídeo.

aproximadamente 10) que os obtidos pelo método de Medioni. Por outro lado, esta última técnica não envolve a escolha de parâmetros e é cerca de 4 vezes mais rápida que a técnica Fourier 1D (ver Figuras 3.11, 3.12 e 3.13). A precisão supe-

Figura 3.21: Histograma da dispersão dos valores do parâmetro de escala espacial (desvio padrão) para o método Fourier 1D (sem interpolação).

Figura 3.22: Histograma da dispersão dos valores do parâmetro de escala espacial (desvio padrão) para o método Fourier 2D.

rior alcançada pelos métodos baseados na transformada de Fourier foi identificada como consequência do uso de vizinhanças maiores em tornos dos pontos da curva no cálculo das derivadas. Outra tendência geral observada inclui o fato de que a abordagem por Fourier 1D considerada em conjunto com a interpolação linear descrita na Seção 3.3.2 provê tempos de execução muito menores que os observados pela técnica por Fourier 2D ou mesmo pelo Fourier 1D sem interpolação, o que caracteriza-se como consequência direta do uso da FFT. Por outro lado, os erros na estimação da curvatura e a sensibilidade às configurações dos parâmetros mostraram-se simi-

Figura 3.23: Histograma da extensão do parâmetro de escala que garante um erro não maior que 10% acima do erro mínimo para a técnica por Fourier 1D.

Figura 3.24: Histograma da extensão do parâmetro de escala que garante um erro não maior que 10% acima do erro mínimo para a técnica por Fourier 2D.

lares para estas duas técnicas (Fourier 1 e 2D, ver Figura 3.14). Deve-se observar também que enquanto a velocidade do Fourier 2D depende do tamanho do mínimo retângulo que envolve a curva, o Fourier 1D depende diretamente do perímetro da curva. O fenômeno particularmente interessante de *interferência* foi observado no Fourier 2D. Conforme fica claro na Figura 3.25, curvas contendo estreitamentos ou “gargalos”, i.e, porções que estão próximas no espaço 2D mas distantes ao longo do perímetro, tendem a produzir picos de erro. Isto tem sido verificado como consequência do fato de que a máscara de convolução abrange, em grandes escalas,

Figura 3.25: Ilustração da interferência espacial para a abordagem por Fourier 2D.

não apenas a vizinhança do ponto onde os operadores diferenciais estão sendo cal-

culados, mas também pontos de outras porções da curva. As posições de tais picos de erro são conseqüentemente determinadas pela escala espacial do estreitamento. Deve-se observar que tal fenômeno não implica em desvantagem à estimação da curvatura, já que ele ocorre em escalas espaciais muito maiores do que aquelas em que se observam os mínimos erros, porém o fenômeno pode deixar indeterminada a representação escala-espaco derivadas da abordagem 2D.

Pouca variação de *performance* foi observada em relação às classes de curvas consideradas e aos esquemas de quantização. O que indica que os esquemas GIQ e SBQ são bons modelos para o processo de amostragem realizados por scanners e câmeras de vídeo (ver Figura 3.14), ao menos no que tange o cálculo de medidas como a curvatura. Verificou-se uma exceção no caso do método de Medioni onde o SBQ implicou em erros muito maiores que as demais quantizações (Figura 3.20) já que o esquema SBQ produz contornos mais “serrilhados” (Figura 3.8). No entanto, no caso dos métodos por Fourier, o SBQ freqüentemente acarretou em maior precisão como uma possível conseqüência de uma representação mais densa das curvas dada por tal esquema (ao menos no que se refere às grades ortogonais).

3.3.7 Conclusões

Esta seção apresentou uma abordagem comparativa abrangente e formal à estimação numérica de curvatura usando duas técnicas baseadas na transformada de Fourier (1 e 2D) e de um método clássico proposto por Medioni e Yasumoto [142] usado aqui como um padrão de comparação. Diversos tipos de curvas, incluindo curvas analíticas paramétricas, *B-splines* e curvas sintetizadas no domínio da freqüência [202] foram consideradas juntamente de quatro diferentes esquemas de quantização destas curvas: interseção na grade (GIQ), caixas quadradas (SBQ), um scanner de mesa e uma câmera de vídeo. Um cuidado especial foi dado à escolha das curvas a fim de garantir a sua representação adequada na grade ortogonal, a fim de garantir que suficiente informação estivesse disponível às técnicas de estimação numérica. A *performance* das técnicas foi comparada de acordo com três principais critérios: tempos de execução, erro no valor calculado e sensibilidade aos parâmetros (expressos em termos da dispersão e da extensão dos parâmetros ótimos), e dados experimentais abrangentes foram obtidos para que se pudesse caracterizar cada uma das técnicas

consideradas.

Genericamente falando, a técnica de Medioni é particularmente adequada para situações onde o tempo de execução requerido é o menor, mas onde também não há grande preocupação em termos da precisão dos valores estimados. Isto é, em grande parte, consequência do fato de que esta técnica não inclui nenhum tipo de filtragem de ruídos. Por outro lado, as técnicas baseadas na transformada de Fourier mostram-se como alternativas particularmente interessantes para aplicações que requerem maior precisão e mais, a técnica por Fourier 1D é mais rápida que o Fourier 2D, mas este último tem a vantagem de permitir o cálculo da curvatura não somente sobre o contorno original, mas também de todo o espaço bidimensional. O que pode ser bastante útil em situações que envolvem o cálculo da curvatura para a solução de equações diferenciais parciais [184]. O interessante fenômeno de interferência foi observado para a técnica por Fourier 2D.

Além disso, os valores ótimos de σ se mostraram robustos em relação à diferentes formas, permitindo, por exemplo, estimar um intervalo fixo de busca para o ajuste automático de parâmetros para uma dada aplicação. Por caracterizar em termos quantitativos as vantagens e desvantagens de cada um dos métodos considerados, além de indicar escolhas adequadas para os respectivos parâmetros envolvidos, os resultados obtidos também proporcionam valiosas informações para aqueles interessados na aplicação das técnicas consideradas em problemas reais.

3.4 Circularidade Local ou Curvatura não-derivativa multi-escala

Conforme visto anteriormente, a curvatura de contornos abertos [47] e fechados é um importante descritor de formas cuja utilidade é bastante reconhecida em processos tanto de reconhecimento quanto de classificação. É também um fato clássico oriundo dos trabalhos de Attneave [6, 7] que os pontos de alta curvatura concentram a maioria da informação perceptualmente relevante [96] conforme é ilustrado pela não menos clássica Figura 3.3 onde pode-se reconhecer um gato dormindo apenas pelos pontos de alta curvatura conectados por segmentos de retas.

Embora Hubel et al [94, 95, 127] tenham identificado evidências biológicas sobre

a detecção de orientação nos primeiros estágios do sistema visual humano que vieram a confirmar a importância de medidas de orientação tais como a curvatura, não é necessariamente verdade que a curvatura *euclidiana* (advinda da geometria diferencial clássica) é a medida de orientação implementada por nossos neurônios dentro do córtex visual. A robustez a ruídos e distorções tanto projetivas quanto afins (ver seção 4.7.1) apresentada pelos sistemas de visão naturais que podem facilmente reconhecer objetos mesmo na presença de tais distorções é o principal argumento para que se pesquise outras medidas não-derivativas de curvatura como a que é proposta aqui.

Um segundo e mais técnico argumento é que dada a sua natureza diferencial, a estimação da curvatura clássica de contornos reais (amostrados e, conseqüentemente ruidosos) é um problema cuja solução não é trivial e procedimentos de suavização são freqüentemente necessários por grande parte dos métodos encontrados na literatura [170, 199, 173, 68, 105] (veja também a seção 3.3). Isto tem motivado um relacionamento de maior proximidade entre curvatura e teorias de escala-espço [105] como uma potencial solução mas o fato é que a curvatura clássica sofre de instabilidades nos cantos onde as derivadas não são bem definidas conforme é mostrado na Figura 3.26.

Figura 3.26: O quadrado (e qualquer outro polígono) é um contorno cujas posições dos pontos de alta curvatura são bem conhecidos porém o valor da curvatura naqueles pontos é indefinida.

3.4.1 A curvatura por área

A idéia de uma medida de curvatura baseada em áreas foi descrita pela primeira vez em [186], porém naquele trabalho não foram realizados maiores desenvolvimentos, não foi dada uma interpretação geométrica à medida, não foi feita uma comparação, mesmo que qualitativa, com a curvatura clássica e nem mesmo as características não derivativas e de sensibilidade a ruídos foram levantadas. Devido ao método aqui

proposto ser baseado em áreas, a “curvatura” resultante será chamada de *curvatura-área* e ela apresenta as seguintes vantagens:

1. É intrinsecamente multi-escala, já que a escolha de um parâmetro de escala constitui o primeiro passo em direção ao conceito de curvatura-área conforme será mostrado na seção 3.4.2 e, pelo menos para os casos verificados, não são criadas novas estruturas (extremos de curvatura) quando se vai da menor escala para escalas maiores até o limite em que o fator de escala atinge a ordem de grandeza da separação mínima entre dois extremos de curvatura-área na menor escala (um requisito para que tal espaço de representações multi-escala forme um escala-espaço, de acordo com [126]). A garantia de que tais novas estruturas não aparecem para o caso geral ainda não foi realizada, constando por isso na lista de desenvolvimentos futuros da seção 5.4 . Além disso, o método descrito na seção 3.4.3 produz a curvatura-área para todas as possíveis escalas e, a qualquer momento, qualquer escala particular pode ser selecionada de acordo com o atual nível de ruído ou de acordo com o interesse numa escala específica. Assim, o problema de a que nível de detalhes este descritor de formas deve capturar características relevantes da forma é eficientemente tratado.
2. Não depende de derivadas de nenhuma ordem. Ao contrário da curvatura euclidiana tradicional que é um conceito intrinsecamente local que depende de derivadas de primeira e segunda ordem, ou da curvatura diferencial afim [181, 70, 148] que depende de derivadas de até quarta ordem, a curvatura-área aqui proposta é uma medida integral que pode ser precisamente aproximada através do simples cálculo de áreas, sendo pouco sensível a ruídos (no mesmo sentido em [18]) conforme mostra a seção 3.4.4.

Um trabalho análogo e bastante interessante que também propõe uma medida multi-escala não derivativa de curvatura pode ser encontrado em [87]. Lá, também são utilizadas as propriedades das cordas que secam um contorno porém, ao invés da área entre o contorno e suas cordas, é utilizada a soma das distâncias entre uma corda e os pontos do contorno. No entanto, ficará como parte de nossos desenvolvimentos

futuros (descritos na seção 5.4) um estudo mais aprofundado das semelhanças e diferenças entre o método aqui proposto e o trabalho de Han [87].

3.4.2 Interpretação geométrica

A fim de compreender melhor o “significado” do descritor de formas apresentado aqui, é necessário inicialmente um olhar mais matemático sobre ele.

Dada um segmento de curva S com comprimento p unindo dois pontos P_1 e P_2 no plano (veja Figura 3.27(a)) e supondo S parametrizado por comprimento de arco (por simplicidade) como $S(t) = (x(t), y(t))$, a área A entre S e a reta l unido P_1 e P_2 pode ser calculada através da seguinte integral:

$$A = \frac{1}{2} \int_0^p (xy' - yx') dt$$

Figura 3.27: (a) Segmento de curva genérico. (b) Segmentos de curva com máxima/mínima área para um par fixo de pontos P_1P_2 e três diferentes valores de p . (c) Segmentos de curva com máxima área para um p fixo e três diferentes pares P_1P_2 .

Agora, se procurarmos pelo segmento de curva S que maximiza/minimiza A mantendo-se p , P_1 e P_2 fixos, é possível demonstrar, a partir desigualdade isoperimétrica ([61], p.31), que a curva correspondente é dada por um **segmento circular** com extremidades em P_1 e P_2 conforme mostrado nas Figuras 3.27(b) e 3.27(c). Qualquer outro segmento de curva de comprimento p entre P_1 e P_2 terá área menor. Além disso, se apenas fixarmos p e deixarmos P_1 e P_2 livres, a forma que maximiza/minimiza A será um semicírculo com raio p/π e área $\pm p^2/\pi$. O caso

especial ($A = 0$) ocorre quando S é um segmento de reta ou quando S está “área-simetricamente” oscilante em torno de l (quando as áreas positivas e negativas se anulam).

Utilizando o fato de que $|A| \leq \frac{p^2}{\pi}$, definimos a curvatura-área assim:

Definição 1. *Dada uma curva γ e um valor de perímetro p (escala), a curvatura-área de um ponto $P \in \gamma$ é a área entre o segmento de γ centrado em P com comprimento p e a corda unindo suas extremidades dividido por $\frac{p^2}{\pi}$.*

De fato, de acordo com tal raciocínio, a curvatura-área mede⁷ quão circular é γ , dados os vínculos P_1 , P_2 e p . Curvatura-área é claramente diferente da curvatura clássica k , pois enquanto k mede o raio (com sinal) do círculo que melhor se ajusta a um dado ponto da curva, a curvatura-área mede a “circularidade”⁸ de um segmento de curva em torno de uma vizinhança p ao redor de cada ponto da curva.

3.4.3 Cálculo da Curvatura-Área

Apresentaremos, nesta seção, um algoritmo incremental para calcular a curvatura-área para todas as possíveis escalas (perímetros) cujos resultados (e robustez a ruído) serão mostrados na próxima seção.

Dado um contorno γ , a idéia básica é definir um valor de perímetro p e procurar por todas as cordas que cortam o contorno em segmentos com aquele perímetro e associar a área algébrica normalizada entre a corda e o segmento de contorno ao ponto médio deste conforme mostrado na Figura 3.28.

Uma vez que não estamos interessados, *a priori*, em nenhum valor de perímetro específico (escala), será apresentado um algoritmo capaz de calcular a curvatura-área para todas as possíveis escalas de forma incremental.

Algoritmo

Partindo de uma curva amostrada bidimensional simples γ , $\gamma(i)$ representa o i -ésimo ponto sobre γ , com i entre 1 e N (aproximando uma parametrização sobre

⁷Globalmente se p é relativamente grande ou localmente se p tende a zero.

⁸No mesmo sentido que em [176, 88, 183].

Figura 3.28: (a) Um contorno γ e quatro cordas diferentes c_1, c_2, c_3 e c_4 cortando γ em segmentos de comprimento p . O valor de “curvatura” associado ao ponto médio de cada segmento de contorno (o pequeno quadrado à direita) é a área A entre a corda e o correspondente segmento de contorno dividido por $\frac{p^2}{\pi}$. (b) Curvatura-área correspondente.

γ), descrevemos um algoritmo para calcular as áreas de todas as cordas unindo quaisquer dois pontos $\gamma(i) = P_1$ e $\gamma(j) = P_2$ e armazená-las numa matriz auxiliar $A(i, j)$. Isto é feito pelo incremento da variável *area* no pseudocódigo abaixo com a área do triângulo resultante de metade da terceira componente do produto vetorial entre v_1 e v_2 conforme ilustrado pela Figura 3.29.

Pseudocódigo:

```

A = zeros(N,N)
for i = 1..N
    j = (i mod N) + 1
    ant_j = j
    P1 = C(i)
    P2 = C(j)
    ant_P2 = P2
    area = 0
    while (j != i)
        v1 = ant_P2 - P1
        v2 = P2 - P1
        area = area + 0.5 * (v1 x v2)
        A(i,j) = area
        ant_j = j
        ant_P2 = P2
        j = (j mod N) + 1

```

```

    P2 = C(j)
  end while
end for

```

No pseudocódigo acima, *ant_* denota valores da iteração anterior, *x* denota a terceira componente do produto vetorial e *mod* é o resto da divisão inteira.

Figura 3.29: A área é incrementada, a cada passo, pela área (com sinal) dos triângulos hachurados.

Da definição de curvatura-área, dado um valor de perímetro p , a curvatura-área de um ponto P é a área entre o segmento de contorno centrado em P com comprimento p e a corda unindo suas extremidades (ver Figura 3.28).

Usando a matriz A produzida pelo algoritmo acima que contém valores de área para todas as possíveis cordas ligando pares de pontos (P_1, P_2) de γ codificados através de seus índices como $A(i, j)$ e supondo que tais índices possam ser usados como uma aproximação para o perímetro da curva⁹, a curvatura-área é obtida de A simplesmente extraindo-se fatias diagonais (paralelas à diagonal principal da matriz) e dividindo seus valores por p^2/π conforme mostrado nas Figuras 3.30(b), 3.30(d) e 3.30(f).

3.4.4 Uma visão comparativa

As duas principais razões que fazem da curvatura-área uma medida interessante é que (i) por ser uma medida integral, ela é pouco sensível a ruídos da mesma forma que os esqueletos invariantes a transformações afins propostos na seção 4.7 e (ii) por não ser uma medida estritamente local (por definição), diferentemente da curvatura

⁹Em grades ortogonais, quando se usa algoritmos padrão para a extração de contornos, a distância entre quaisquer dois pontos consecutivos do contorno pode ser 1 ou $\sqrt{2}$.

Figura 3.30: Toda diagonal “ p -deslocada” de A ($A(i, i + p)$ com $1 \leq i \leq N$) contém os valores de área para todas as cordas unindo pontos em γ que estão separados por um perímetro p . Tais valores divididos por p^2/π são boas aproximações para a curvatura-área conforme mostrado nas Figuras 3.30(c), 3.30(e) e 3.30(g) para três diferentes escalas: $p=2$, 5 e 8, respectivamente.

clássica, a curvatura-área pressupõe a influência simultânea de diferentes partes da forma sem a necessidade de preocupação com técnicas externas de regularização.

A Figura 3.31 mostra um contorno com três folhas (à esquerda) cujas funções curvatura clássica (linha fina) e curvatura-área (linha grossa) são mostradas à direita. Podemos observar que embora ambas as curvaturas tenham seus máximos e mínimos nas mesmas posições, a curvatura clássica faz melhor distinção entre máximos e mínimos uma vez que as concavidades da forma são mais agudas que as suas partes convexas, em outras palavras, a curvatura-área não faz a mesma dis-

tinção quantitativa entre as regiões côncavas e as regiões convexas da curva pois ela não mede variações no campo tangente da forma. Na Figura 3.32, uma versão rui-

Figura 3.31: Contorno (a) e respectivas curvaturas (b) por área (linha grossa) e clássica (linha fina). A pequena seta indica a orientação da parametrização.

dosa e distorcida (por uma transformação afim) do mesmo contorno é apresentada e as respectivas funções curvatura (à direita) ainda mantêm uma boa correspondência quanto às posições dos seus extremos. Observamos mais uma vez a relativa insensibilidade da curvatura-área quanto à taxa de variação do campo tangente que pode ser notada especialmente com relação à leve concavidade da parte direita daquele contorno que, na curvatura clássica (linha fina), causa um ponto de mínimo pouco pronunciado (no centro da função) enquanto que, na curvatura-área, causa um mínimo cuja profundidade é comparável aos demais extremos da função.

Tal característica da curvatura-área pode vir a ser útil em problemas onde partes da curva com diferentes graus de concavidade/convexidade são igualmente importantes e precisam ser detectados como nos casos das Figuras 3.33 e 3.34 onde os contornos apresentam diferentes graus de concavidade/convexidade e cujas curvaturas-área (linhas grossas) variam menos que a respectiva curvatura clássica.

3.4.5 Conclusões

Embora a curvatura-área apresentada nesta seção tenha semelhanças e diferenças fundamentais com relação à curvatura clássica e, por isso, apresente-se como uma promissora ferramenta para a caracterização de formas, ainda permanecem em aberto algumas questões que deverão ser respondidas em trabalhos futuros conforme será discutido na seção 5.4. Entre tais questões está o fato da curvatura-área

Figura 3.32: Contorno (a) e respectivas curvaturas (b) por área (linha grossa) e clássica (linha fina). A pequena seta indica a orientação da parametrização.

Figura 3.33: Contorno (a) e respectivas curvaturas (b) por área (linha grossa) e clássica (linha fina). A pequena seta indica a orientação da parametrização.

Figura 3.34: Contorno (a) e respectivas curvaturas (b) por área (linha grossa) e clássica (linha fina). A pequena seta indica a orientação da parametrização.

não compor uma representação legítima de formas uma vez que não se sabe como se poderia realizar o processo inverso que permita a reconstrução da forma dada a sua curvatura-área. Outro ponto seria a modificação de sua definição a fim de trocar o vínculo com o perímetro por outra entidade que fosse, por exemplo, invariante a transformações afins, permitindo assim a criação de uma curvatura invariante tais

transformações, semelhante aos esqueletos afins mostrados na seção 4.7. E, por fim, seria bastante interessante verificar as propriedades da extensão deste conceito para 3D onde, ao invés da área das cordas definidas por pares de pontos de um contorno seria calculado o volume de regiões delimitadas dos planos secantes definidos por trios de pontos de uma superfície. Trabalhos neste sentido já foram iniciados [17], porém muito ainda precisa ser desenvolvido.

Capítulo 4

A Representação por Eixos de Simetria

4.1 Motivação

Simetria é uma característica presente em quase todas as formas que estão a nossa volta [89] e suas manifestações podem ser observadas na anatomia de vírus, animais e plantas, na arquitetura e nas artes dos mais diversos povos [196], na ciência e até na música [117]. Na área da percepção visual, estudos indicam que rostos simétricos geralmente são mais atraentes [86, 113] havendo uma infinidade de outros casos onde a simetria tem papel fundamental.

Tal fato não poderia ter sido ignorado no contexto das representações de formas e, na década de 60, Blum, através de seus trabalhos sobre a caracterização de formas biológicas [22, 23, 24] propôs e analisou o primeiro conceito matemático que visava a captura dos eixos de simetria das formas. O “*medial axis*”, aqui denominado *eixo medial (EM)* surgiu como um novo descritor de formas através do qual formas biológicas puderam ser caracterizadas e estudadas de maneira objetiva. No entanto, as aplicações desta idéia não ficaram restritas apenas à análise de formas biológicas, mas mostraram-se úteis dentro dos mais variados problemas e aplicações.

O eixo medial é definido através da definição 2 e a associação de seus pontos com o valor dos raios das bolas maximais ficou também conhecida como “*medial axis transform*” ou *transformada do eixo medial*, pois são os raios das bolas maximais que permitem a reconstrução da forma a partir de seu respectivo eixo medial.

Definição 2. Considere S como um subconjunto do espaço euclidiano n -dimensional e

sua fronteira B . Uma bola aberta dentro de B que não é contida por nenhuma outra bola aberta dentro de B é denominada **bola maximal**.

O eixo medial (EM) de S é definido como o lugar geométrico formado pelos centros de suas bolas maximais e a transformada do eixo medial (TEM) de S é composta pelo EM e pelo conjunto dos respectivos raios das bolas maximais que o formam.

Semelhantemente à curvatura (ver capítulo 3), os EMs possuem várias (e equivalentes) interpretações sendo elas: (i) o conjunto dos centros dos círculos maximais, (ii) o conjunto das singularidades da transformada distância¹, (iii) as posições dos choques entre frentes que se propagam na direção normal à forma e (iv) o diagrama de Voronoi estendido de conjuntos conectados de pontos (ver seção 4.6.2). A Figura 4.1 ilustra essas quatro interpretações para um mesmo contorno indicando o estreito relacionamento entre os eixos mediais e outras entidades geométricas.

Figura 4.1: Quatro diferentes interpretações do conceito de eixo medial. (a) O conjunto dos centros dos círculos maximais, (b) o conjunto das singularidades da transformada distância, (c) as posições dos choques entre frentes que se propagam na direção normal à forma e (d) o diagrama de Voronoi estendido de conjuntos conectados de pontos.

Mais uma vez temos aqui uma representação de formas que, se aplicada a dados cuja forma analítica é conhecida, não apresenta ambigüidades e pode-se dizer que,

¹A transformada distância de um subconjunto S do espaço euclidiano é a função escalar não negativa $d(S, x)$ que associa, a cada ponto x daquele espaço, o valor da mínima distância entre S e x .

para cada forma, há uma única transformada do eixo medial a ela associada. No entanto, diferentemente da curvatura, para a idéia de “*esqueleto da forma*” (nome através do qual os EMs ficaram conhecidos) não existe apenas uma única definição, uma vez que esta é uma noção intuitiva e os EMs são apenas uma das possíveis maneiras de se definir objetivamente uma “representação fina”, uma “linha central” ou uma “estrutura de suporte” que represente o “traçado interno” da forma. Por essa razão, os *esqueletos* também receberam definições alternativas àquela originalmente dada por Blum, dentre as quais podemos citar os “*symmetry sets*” [35], aqui denominados *conjuntos de simetria (CS)*, as “*smoothed local symmetries*” [32], aqui denominadas *simetrias locais suavizadas (SLS)* e o “*process-inferring symmetry axis (PISA)*” [124], aqui denominado *eixo de simetria para inferência de processo*.

A Figura 4.2 ilustra o dilema entre a “imprecisa noção intuitiva” e as “arbitrárias versões objetivas” dos esqueletos através de uma letra ‘T’ cujo esqueleto é mostrado em 4 diferentes versões. Em 4.2(a), embora a linha interna seja uma representação mais “intuitiva” de uma letra ‘T’ se comparada às linhas internas das Figuras 4.2(b), 4.2(c) e 4.2(d) (que contêm alguns ramos extras e até desconectados), não existe na literatura uma definição objetiva que corresponda sempre ao esqueleto “perceptualmente mais fiel” como o esqueleto apresentado em 4.2(a).

Por ser subjetiva a decisão (isolada de contexto) por qualquer uma das definições acima citadas como sendo a melhor definição para o conceito de esqueleto, optamos por situar as diferentes definições frente aos critérios para a representação de formas (ver seção 2.5) a fim de constituir um cenário comparativo e uma terminologia comum a todas as abordagens que serão avaliadas em seguida e que se ocupam do cálculo deste tipo de representação.

A ausência de trabalhos de integração conceitual em análise de formas, especialmente com relação às representações axiais como o que será realizado aqui, tem como consequência o fato de, não raro, encontrarmos trabalhos que, para se definir o que é esqueleto, apelam para algum conjunto de características que um “bom esqueleto” deveria ter. Tais características geralmente incluem: (i) permitir a reconstrução da forma, (ii) ser o mais fino possível, (iii) ser homotópico, (iv) ser robusto às distorções causadas por ruídos de pequena escala, translações e rotações. Em contrapartida, outros trabalhos assumem implicitamente (e sem qualquer justificativa) que esque-

Figura 4.2: (a) Forma e esqueleto traçado “manualmente”. (b) Forma e respectivo eixo medial. (c) Forma e respectivo SLS. (d) Forma e respectivo PISA.

letos *são* os eixos mediais de Blum. Neste caso, provavelmente a única justificativa seja simplesmente que esta *é*, de fato, a primeira definição criada e a mais amplamente adotada na literatura.

E foi na tentativa de responder à questão “*o que é o esqueleto de uma forma?*” que nos vimos obrigados a retornar aos fundamentos do processamento da informação visual e a reformular os critérios para representação de formas (vide capítulo 2) a fim de então poder situar os esqueletos e suas variantes perante outras representações e relacioná-los a outros conceitos como, por exemplo, os diagramas de Voronoi (ver seção 4.6).

4.2 Definições, esqueletos e eixos de simetria

As quatro definições de esqueleto citadas até aqui (EM, CS, SLS e PISA) e que serão posicionadas perante os critérios para a representação de formas são bastante semelhantes entre si pois todas baseiam-se no conjunto de bolas/círculos maximais ou *bitangentes* da forma (ver definição 2), com exceção dos conjuntos de simetria que, ao contrário das demais definições, não fazem distinção se as bolas/círculos estão ou não contidos na forma. A principal diferença entre tais definições, conforme pode ser verificado em [35, 32, 124] está no ponto do círculo maximal que é tomado para a composição dos eixos de simetria. No caso do EM e do CS, os *centros* dos círculos maximais formam os eixos de simetria. Já no caso do SLS, o ponto médio da *corda* que une os pontos de tangência entre a forma e o círculo é usado para traçar os eixos de simetria, enquanto que para o PISA, usa-se o ponto médio do *arco* de círculo que une os mesmos pontos de tangência.

Apesar de aparentemente sutil, a diferença entre estas definições tem implicações que vão desde a não satisfação de alguns dos critérios para representação de formas (conforme mostra a seção 4.4) até a introdução de grandes dificuldades para a criação de algoritmos capazes de calcular aproximações destas representações, passando pela inadequação à solução de problemas práticos. De fato, apesar de todas estas definições terem sido propostas há vários anos, apenas os EMs foram extensivamente utilizados na prática e ainda não existem exemplos onde CS's, SLS's e PISA's tenham se mostrado úteis. Façamos aqui uma ligeira ressalva quanto ao PISA cuja proposição foi principalmente dirigida ao estudo e à modelagem do crescimento ou evolução espaço-temporal das formas e não para ser uma representação nos moldes do modelo proposto por Marr.

Como uma última observação, embora estejamos nos restringindo apenas às representações axiais, é interessante o conhecimento de uma outra classe de representações denominadas “*local rotational symmetries*” ou *simetrias rotacionais locais* cujas características (até hoje inexploradas na solução de problemas práticos) podem ser encontradas em [73]. Esse trabalho (uma tese de doutorado), além de propor um tipo de representação que visa a caracterização de formas arredondadas (situação onde as representações axiais não têm bom desempenho), apresenta uma

interessante discussão sobre os critérios para uma “boa” representação de formas.

4.3 Esqueletos, abordagens e outras entidades geométricas

Existem atualmente inúmeras abordagens para a extração de esqueletos de imagens na literatura. Entre as principais e mais comumente empregadas estão as que se baseiam no cálculo da *transformada distância* (TD) e posterior detecção de suas singularidades [179, 26, 200], as que calculam o diagrama de Voronoi de pontos amostrados sobre o contorno seguidos de alguma metodologia de poda (ou *pruning*) [158, 82], técnicas baseadas em operações morfológicas de erosão/dilatação acompanhadas por regras para a preservação da conectividade, também chamados de *thinning* e esquemas de propagação de frentes [69, 188] onde são detectados os choques entre tais frentes.

Nos próximos parágrafos, os termos *precisão geométrica*, *precisão topológica* e *precisão estrutural* estão definidos no item *acessibilidade* da seção 4.4.

Os métodos de esquelitização baseados em transformada distância [153, 3, 56, 200] surgiram motivados pela existência prévia de algoritmos eficazes (porém aproximados) para o cálculo dessa transformada (útil também em diversas outras aplicações [161]) e pela relativa simplicidade na implementação desta abordagem. No entanto, por depender deste cálculo consideravelmente custoso² e pelas dificuldades na localização precisa de suas singularidades (o que leva geralmente a produção de esqueletos não conectados), este tipo de metodologia não aparece como uma alternativa interessante, embora sua simplicidade a transforme numa das primeiras opções na extensão para três ou mais [26] dimensões.

As abordagens advindas da morfologia matemática [115], apesar de padecerem de problemas como a baixa precisão geométrica e estrutural, por ainda assim garantirem uma boa precisão topológica e por apresentarem um bom desempenho, são alternativas bastante empregadas em diversas aplicações.

Os resultados obtidos por esquelitizações com base nos diagramas de Voronoi são os que apresentam um dos melhores balanços considerando-se as diferentes pro-

²no caso da transformada distância **euclidiana** exata cuja isotropia dá maior precisão geométrica aos esqueletos gerados.

priedades desejadas nos esqueletos obtidos. Porém, por ser relativamente complexa a implementação de bons algoritmos para o cálculo dos diagramas de Voronoi e pela freqüente necessidade de aplicação de heurísticas de poda para eliminação de ramos indesejados, esta metodologia, apesar de promissora, ainda não teve um numeroso uso em aplicações práticas. Com relação a extensão tridimensional, muitos dos problemas já encontrados em 2D ficam bastante agravados no caso 3D. É importante ainda salientar que através deste tipo de abordagem surgiram os primeiros esqueletos intrinsecamente³ escala-espço [157] cujas características próprias desta teoria acabaram sendo posteriormente levadas a outras metodologias como em [46, 69].

Outra importante abordagem que se dá através de métodos de propagação de frentes que buscam aproximar uma solução (não viscosa) da equação eikonal (propagação na direção normal à curva) também está dentre as mais promissoras forma de cálculo de representações axiais, uma vez que apresentam um ótimo balanço entre desempenho, precisão e facilidade de implementação. Essa classe de algoritmos surge também como a principal alternativa para o cálculo de esqueletos de objetos tridimensionais devido a esse mesmo balanço em comparação às demais abordagens.

Existem ainda algumas outras metodologias bastante recentes ou mesmo pouco avaliadas como a baseada em redes neurais [57, 162], *snakes* [122] e lógica *fuzzy* [119, 166, 163, 164] e que devido à baixa repercução não foram utilizadas em muitas aplicações práticas.

4.4 Os esqueletos frente aos critérios para a representação de formas

Tentaremos mostrar nesta seção as qualidades e defeitos das representações por eixos de simetria e elucidar suas principais características utilizando os critérios propostos na seção 2.5.

Acessibilidade Diferentemente da curvatura onde a questão da precisão no cálculo da representação está ligada somente ao valor da curvatura correspondente a cada ponto do contorno, nas representações por eixos de simetria temos três

³Sem apelo a suavização gaussiana das formas.

diferentes “tipos” de precisão que podem vir a ser separadamente priorizados. São eles: (i) *precisão topológica* que refere-se ao fato da representação ter ou não a mesma topologia (número de buracos e de componentes conectados) da forma representada; (ii) *precisão estrutural* que refere-se a ordem das conexões entre os diversos ramos que compõem a representação axial e (iii) *precisão geométrica* que está relacionada à posição dos pontos do eixos de simetria. Diferentes problemas podem carecer de diferentes precisões dependendo de quais aspectos da representação são mais ou menos importantes para que se atinja a finalidade global do processamento (ver seção 2.2). Por exemplo, se num problema de reconhecimento de caracteres precisamos apenas fazer a distinção entre letras que sabemos estarem restritas a apenas c’s e d’s, então basta a precisão topológica para que c’s sejam diferenciados de d’s. Porém, se precisamos distinguir entre u’s e v’s, a precisão geométrica passa então a ser a mais importante. Portanto, há um balanço entre estes três tipos de precisão para cada específico problema. Definido este balanço, a acessibilidade corresponde então ao algoritmo mais simples capaz de atingir a(s) precisão(ões) mínima(s) para a realização de determinada tarefa.

Com relação às precisões topológica e estrutural, a acessibilidade das representações axiais fica restrita aos algoritmos que visam o cálculo dos eixos mediais, pois as demais definições nem sempre preservam a topologia e a estrutura da forma original. Já no caso da precisão geométrica, há uma disputa entre métodos baseados em frentes propagantes como os propostos por [69, 188] e métodos baseados no cálculo otimizado dos diagramas de Voronoi [158, 60] cujas aproximações dos eixos mediais maximizam os três tipos de precisão tendo bom desempenho tanto nos casos 2D quanto 3D se comparadas com técnicas baseadas, por exemplo, no cálculo da transformada distância.

Escopo Valendo-se do mesmo raciocínio utilizado para o levantamento do escopo da representação por curvatura, ao verificarmos a diversidade das representações-fonte (representações a partir das quais uma representação é calculada) podemos constatar que, dentre as definições candidatas a esqueleto da forma, os EMs são os que apresentam maior escopo pois podem ser calculados a partir da: (a) representação por bordas, (b) representação por transformada

distância, (c) representação por diagramas de Voronoi e (d) representação por choques [188]. Já as demais definições, por não corresponderem a um número tão grande de analogias e interpretações (ver Figura 4.1) acabam por ficar restritas ao cálculo a partir da representação por bordas. Já em comparação a classe de representações por curvatura, as representações axiais apresentam um escopo de abrangência semelhante, porém com pequena intersecção já que a única representação-fonte comum a estas duas classes é a representação por bordas. Isto significa que não são muitos os casos onde são obtidas soluções de igual acessibilidade através de representações axiais ou por curvatura.

Caráter escala-espço Por terem sido pouco desenvolvidas, as raras implementações que calculam os CS's, SLS's e PISA's não incluem a possibilidade de acesso a diferentes escalas espaciais destas representações e são relativamente poucas e recentes as implementações de eixos mediais que dão a esta representação um caráter escala-espço próprio não se utilizando de suavizações aplicadas sobre a própria forma como acontece em [27]. Poucas abordagens como as de Ogniewicz [158], Costa&Estrozi [48] e Falcão et al. [69] produzem eixos mediais cuja simplificação progressiva está embutida na própria maneira de se calcular a representação.

Suporte local rico De acordo com o próprio Brady [30], propositore deste critério, “medidas de simetria, apesar de serem menos locais, não podem vir a ser consideradas globais uma vez que não se utilizam simultaneamente de todas as partes da imagem”. Isso deixa todas as definições candidatas a esqueleto em pé de igualdade pois a necessidade de uso de dois ou mais pontos da forma para a constatação de simetria é comum a todas as abordagens conhecidas para construção de representações axiais. Já em comparação à curvatura (que é uma medida estritamente local), os esqueletos mostram-se em desvantagem, principalmente se o problema envolve a possibilidade de oclusão de partes da forma.

Extensão suave e enquadramento Nos casos das definições não homotópicas (SLS e PISA), não é possível garantir a extensão suave pois não há como se fazer uma transição gradual entre sistemas de coordenadas situados em diferentes componentes conectados, fato que já não ocorre com as definições que preservam

a topologia da forma original (EM e CS). Já o enquadramento é garantido por todas as representações axiais devido à característica de redução dimensional apresentada.

Hierarquia global Os eixos mediais, por poderem ser representados através de grafos como, por exemplo, em [83] (onde ramos e pontos de ramificação definem uma hierarquia das partes que compõem a forma [8, 118]) e por cada ponto do eixo relacionar pelo menos dois pontos distintos da forma representada, eles acabam por satisfazer este critério. No entanto, as demais definições que não têm correspondência direta com grafos, não permitem a definição de uma estrutura hierárquica.

Invariância Os eixos de simetria, de maneira geral, apresentam as mesmas características de invariância apresentadas pela curvatura no capítulo anterior: invariância a translações e a rotações. Quanto às transformações afins, alguns trabalhos já foram realizados [181, 146, 18, 17, 19, 20] no sentido de se criar uma definição alternativa de esqueleto que apresente invariância a tal transformação e na seção 4.7 são apresentadas algumas contribuições justamente nessa direção.

4.5 Esqueletização através de dilatações exatas

Mostraremos nesta seção como o conceito das dilatações exatas introduzido em [46] podem ser usadas na obtenção de esqueletos multi-escala e como uma nova representação multi-escala de formas pode ser obtida através de tais esqueletos. Além disso, por permitir a reconstrução precisa de versões simplificadas da forma original, a abordagem aqui proposta fica caracterizada por sua simplicidade e pela importante propriedade de não mover as bordas dos objetos enquanto estes vão sendo simplificados.

Esta seção já foi publicada na forma de artigo em revista [48] e a descrição de sua implementação em hardware pode ser encontrada em [130].

4.5.1 Introdução

Esqueletização e representações de formas multi-escala constituem dois dos mais promissores e desafiadores temas em análise de formas e em visão computacional. Estes temas, objetos de crescente atenção da comunidade, ainda permanecem em aberto dada a persistência de alguns problemas fundamentais. Enquanto um dos grandes desafios das técnicas de esqueletização é garantir a que o esqueleto resultante tenha a largura mínima (1 pixel) e seja conectado, a análise de formas multi-escala tem enfrentado o problema do deslocamento das bordas em função do grau de suavização aplicado a um dado objeto, um problema que é típico das abordagens multi-escala lineares tais como a suavização Gaussiana. Apesar do esforço na criação de sofisticadas metodologias multi-escala não-lineares (por exemplo, [191]), constata-se que o problema do deslocamento é somente controlado a um custo muito elevado e, de fato, não é completamente resolvido.

Aqui, apresentamos uma nova abordagem para a representação de formas que baseia-se no uso do conceito recém introduzido das dilatações exatas [50] para propagar rótulos associados ao contorno original para a produção de esqueletos multi-escala. Dado que as posições dos elementos (pontos) que compõem tais esqueletos não se deslocam quando se vai de uma escala menor para uma escala maior, uma família multi-escala de reconstruções da forma original pode ser obtida com elementos que também não se deslocam ao se passar de uma escala a outra.

Começamos por descrever como as dilatações exatas podem ser calculadas e usadas para propagar rótulos e calcular a transformada distância exata. Prosseguimos mostrando como esqueletos multi-escala podem ser obtidos das diferenças entre os rótulos propagados e a representação multi-escala de formas é então introduzida em termos de dilatações dos esqueletos multi-escala.

4.5.2 Dilatações exatas e esqueletos multi-escala

A principal característica das dilatações exatas é que elas representam todas as possíveis dilatações de qualquer forma num espaço espacialmente quantizado [50]. Seja B_R uma bola digital de raio R , i.e., todos os pixels contidos num círculo digital de raio R . Dilatações exatas podem ser computacionalmente implementadas

usando-se uma estrutura de dados especial contendo, em ordem ascendente, todas as possíveis distâncias em B_R . As posições relativas com respeito ao centro da bola são também armazenadas nesta estrutura junto de cada distância. Uma maneira simples de se obter tal estrutura de dados é calcular todas as distâncias dentro de B_R e então ordená-las. Uma vez que a estrutura está disponível, o processo das dilatações exatas pode ser realizado percorrendo-se os N pontos do contorno de um objeto para cada valor subsequente de distância. Em cada passo, i.e., para o atual pixel do contorno p e para a atual distância d , os pixels correspondentes às posições relativas associadas a d são marcados ao redor de p com o valor '1' (assumindo que inicialmente todos os pixels possuíam o valor zero). A Figura 4.3 ilustra o conceito das dilatações exatas.

Figura 4.3: Forma original e suas dilatações exatas para diversos valores de d : (a) Forma original, (b) $d = 2$, (c) $d = 10$ e (d) $d = 20$.

Dilatações exatas podem ser usadas no cálculo da transformada distância e para propagar rótulos com precisão (dentro de um esquema de propagação na direção normal, também conhecido como propagação *eikonal*). Enquanto a primeira pode ser imediatamente obtida pela atualização com o valor da distância atual ao invés de simplesmente '1', a propagação de rótulos é realizada em dois passos: (i) os pontos

do contorno são rotulados perimetralmente por números inteiros subseqüentes (que podem ser obtidos usando-se algoritmos clássicos de perseguição de bordas [169]) e (ii) durante a dilatação exata, atualiza-se com valor p no lugar de '1'. Esqueletos multi-escala podem então ser obtidos pelas diferenças entre os rótulos propagados. Mais especificamente, dado que os rótulos foram propagados, o valor máximo das diferenças entre cada um dos pixels $R_{i,j}$ e seus quatro vizinhos mais próximos $R_{i,j+}$ é determinado, criando-se assim uma imagem de diferenças D . Uma vez que uma rotulação unidimensional é usada sobre um contorno bidimensional, surgirão descon- tinuidades se rótulos oriundos de regiões perimetralmente distantes se encontrarem. A solução deste problema dá-se pelo simples complemento:

$$D(i, j) = \begin{cases} \max(R_{i,j} - R_{i,j+}) & \text{se } \max(R_{i,j} - R_{i,j+}) \leq \frac{N}{2}, \\ N - \max(R_{i,j} - R_{i,j+}) & \text{se } \max(R_{i,j} - R_{i,j+}) > \frac{N}{2}. \end{cases}$$

Uma família de esqueletos multi-escala pode então ser facilmente obtida simplesmente aplicando-se limiares sobre D , limiares estes que definem o parâmetro de escala. Observa-se que embora ambos os esqueletos internos e externos sejam ge- rados desta forma, é trivial separá-los por uma simples combinação "AND" com a forma original. A Figura 4.4 dá os esqueletos multi-escala da forma mostrada em 4.3(a) na forma de uma pilha (esqueletos mais simples sobre esqueletos mais complexos).

Figura 4.4: Pilha de esqueletos multi-escala da forma em 4.3(a) correspondentes diversos limiares.

Uma característica interessante de tais esqueletos é que pode ser verificado que eles produzem uma representação escala-espço válida, já que não são criados novos

elementos conforme se eleva o parâmetro de escala. Além disso, a posição dos pixels dos esqueletos não muda quando se vai de uma escala maior para uma escala menor, exceto pelo fato de que mais detalhes são adicionados. Aos pontos destes esqueletos podemos então adicionar o respectivo valor da transformada distância o que nos permitirá reconstruir versões da forma original.

4.5.3 Representação multi-escala de formas

Uma outra propriedade interessante dos esqueletos é que a forma original pode ser recuperada se, para cada escala e cada esqueleto correspondente, seus pixels forem dilatados por uma bola digital cujo raio seja igual ao valor da transformada distância associada àquele ponto. Conforme acima mencionado, já que as posições dos pixels do esqueleto não mudam com o parâmetro de escala, a mesma propriedade é estendida às reconstruções, ou seja, o contorno das formas reconstruídas não se desloca quando se passa de uma escala à outra. A Figura 4.5 apresenta as reconstruções da forma original a partir dos esqueletos internos também organizados numa pilha.

Figura 4.5: Pilha de reconstruções dos esqueletos multi-escala internos da Figura 4.4.

4.5.4 Conclusões

Apresentamos aqui um novo e simples meio de se representar formas multi-escala baseado no conceito de dilatações exatas, propagação de rótulos e esqueletos multi-escala. A metodologia apresentada tem uma série de propriedades interessantes, sendo a principal relacionada ao não deslocamento das bordas quanto da simplificação dos objetos dentro de um esquema escala-espaco. No entanto, ainda permanece a necessidade de uma avaliação mais profunda de outras possíveis vantagens e

desvantagens dessa abordagem em comparação ao que é encontrado na literatura. Em especial, podemos citar os trabalhos de Ogniewicz [157, 158] pois apresentam algumas semelhanças quanto à maneira de selecionar diferentes escalas e que por estarem baseados nos diagramas de Voronoi, poderão ser melhor compreendidos sob a luz das considerações feitas na próxima seção.

4.6 Considerações sobre diagramas de Voronoi e Eixos Mediais

Os eixos mediais e os diagramas de Voronoi estão entre as idéias mais influentes da visão computacional e análise de imagens. Relacionamentos entre eles com respeito a polígonos foram notados décadas atrás e, recentemente, tal relacionamento tem sido expandido para uma classe maior de formas. Mais especificamente, o diagrama de Voronoi tem sido considerado como um meio através do qual algoritmos ótimos de geometria computacional podem ser aplicados para calcular eixos mediais. Esta seção visa estabelecer um relacionamento teórico mais próximo entre diagramas de Voronoi e eixos mediais. Para isso, são propostas extensões das definições destes conceitos e as vantagens destas definições estendidas com respeito a alguns casos específicos, porém relevantes, são destacadas. Além disso, os eixos mediais acabam sendo caracterizados como um caso particular dos diagramas de Voronoi e as implicações deste fato são analisadas e discutidas.

4.6.1 Introdução

Diagramas de Voronoi e eixos mediais são importantes estruturas geométricas úteis em um amplo espectro de aplicações em computação gráfica e visão computacional. O principal objetivo desta seção é estudar e reavaliar o relacionamento entre estas duas estruturas a fim de congregá-las num panorama conceitual comum que poderá ser usado para re-organizar diversas abordagens alternativas encontradas na literatura. Ademais, tem-se a intenção de criar fundamentos para que as técnicas de esqueletização baseadas nos diagramas de Voronoi possam ser mais bem situadas, incluindo aqui a questão dos critérios de poda.

Apesar dos relacionamentos entre diagramas de Voronoi e eixos mediais terem sido discutidos amplamente na literatura (ver seção 4.6.4), as definições destes conceitos carecem de revisão, assim como é realizado nesta seção. Pequenas modificações na definição dos diagramas de Voronoi são apresentadas para estendê-lo adequadamente a formas contínuas/conectadas enquanto preservando a propriedade de equidistância (discutida na seção 4.6.2). Os eixos mediais, originalmente criados por Blum [22, 23, 24], por sua vez, são estendidos para objetos mais gerais como curvas abertas, eliminando a necessidade de se ter regiões internas e externas bem definidas, permitindo assim que a idéia de Blum seja aplicada a conjuntos de dados mais amplos. Várias implicações de tais resultados são aqui discutidas e ilustradas.

Na seção 4.6.2, a noção de diagrama de Voronoi (DV) é introduzida juntamente com sua definição usual. Uma versão estendida desta definição é então apresentada juntamente de suas vantagens. Alguns casos especiais onde a extensão do DV proposta nos dá resultados mais coerentes são também ilustrados. A seção 4.6.3 trata do conceito básico dos eixos mediais e da definição normalmente usada na literatura, seguido por uma definição estendida para conjuntos de dados mais gerais. Vantagens desta extensão são apresentadas junto de alguns casos ilustrativos. Na seção 4.6.4, as principais contribuições de ambas as definições propostas são revisadas considerando-se a integração conceitual dos diagramas de Voronoi e eixos mediais.

Casos onde os diagramas de Voronoi não são tipicamente aplicados, i.e., em objetos conectados, e casos onde os eixos mediais nunca haviam sido usados, i.e., em objetos desconectados, com auto-intersecções ou abertos, são caracterizados ajudando a organizar os conceitos dentro de um todo coerente, seguido por teoremas relacionando os DVs e os eixos mediais e respectivas provas. Finalmente, é feita uma discussão sobre como estas idéias poderiam ajudar a definir novas técnicas de esqueletização baseadas nos diagramas de Voronoi através de regras de poda mais consistentes.

4.6.2 Diagramas de Voronoi

O diagrama de Voronoi [9, 175, 160] é um conceito geométrico que baseia-se na proximidade a conjuntos de dados discretos/desconectados. A maneira mais fácil (e

tradicional) de explicar este conceito se dá através da Figura 4.6:

Figura 4.6: Um simples diagrama de Voronoi definido por três pontos isolados.

Considere um conjunto E de elementos no espaço euclidiano n -dimensional onde cada elemento pode ser qualquer subconjunto deste espaço, tipicamente um ponto ou um segmento de reta (considerando ambos DV clássico e generalizado). Isto define o contexto para o resto desta seção. É normal dizer, baseado em figuras como a 4.6, que o diagrama de Voronoi é a união das fronteiras entre as regiões de Voronoi definidas como o conjunto de pontos ao menos tão próximos de um dado elemento quanto de outro qualquer. De fato, a seguinte definição matemática de um DV é comumente empregada na literatura, de acordo com a pesquisa de Aurenhammer [9] e, mais recentemente, no trabalho de Brandt [34]:

Definição 3. A região de Voronoi de um elemento $e \in E$, denotado por $RV(e)$, é o conjunto de pontos ao menos tão próximos de e quanto de qualquer outro elemento de E :

$$RV(e) \doteq \{ p \in \mathbb{R}^n \mid d(p, e) \leq d(p, e'), \quad \forall e' \neq e, e' \in E \}$$

onde d é a distância euclidiana usual entre dois conjuntos. O diagrama de Voronoi de E , denotado como $DV(E)$, é normalmente definido como a união das fronteiras das regiões de Voronoi:

$$DV(E) \doteq \bigcup_i \partial RV(e_i)$$

onde o símbolo ∂ denota a fronteira de um conjunto.

Retornando ao exemplo da Figura 4.6, o DV definido são as linhas naquela figura. Neste caso, esta definição parece razoável e, de fato, as inúmeras propriedades elegantes e desejáveis do DV encontradas na literatura estão baseadas nela [9, 175]. Observe, no entanto, que não existem restrições sobre os elementos, de tal forma que qualquer conjunto de pontos pode ser considerado como sendo um elemento. Por exemplo, um segmento de reta poderia ser visto como um único elemento ou, de uma forma menos usual, como um conjunto de infinitos, distintos e conectados elementos subsequentes infinitamente próximos uns dos outros. Na verdade, uma importante propriedade dos diagramas de Voronoi está em permitir o agrupamento arbitrário de pontos para a formação de elementos válidos⁴. Isto dota o DV de características não-locais, porque quando as distâncias são avaliadas a fim de definir as regiões de Voronoi, a proximidade a elementos extensos leva em conta toda a sua extensão e não apenas porções específicas.

Em ambos os casos (com ou sem agrupamento de pontos na composição de elementos), a Definição 3 é perfeitamente aplicável e é desejável que ambos os resultados tenham as mesmas propriedades. O próximo exemplo ilustra estas duas possibilidades.

Figura 4.7: Dois elementos de Voronoi e_1 e e_2 , e as correspondentes regiões de Voronoi.

A Figura 4.7 mostra dois elementos: um ponto (e_1) e um segmento de reta (e_2). Repare que este último é considerado como um único objeto. As correspondentes regiões de Voronoi para os elementos e_1 e e_2 estão situadas respectivamente acima e abaixo da curva na Figura 4.7. Entre as linhas verticais pontilhadas, a união das fronteiras das RVs é um segmento parabólico que é, por esta razão, parte do DV. No restante do plano, a união das fronteiras das regiões de Voronoi define duas semi-retas

⁴Nas próximas seções o termo *agrupamento* será usado para elementos consistindo de mais de um único ponto.

que começam em cada um dos pontos indicados pelos pequenos quadrados. Estas linhas coincidem com as bissetrizes entre e_1 e ambas as extremidades de e_2 conforme indicado pelas linhas tracejadas. Note que o diagrama de Voronoi resultante é o lugar geométrico dos pontos equidistantes a e_1 e e_2 . Observemos agora o que acontece quando o segmento de reta e_2 é tratado como sendo composto por infinitos elementos (pontos), como mostra a Figura 4.8

Figura 4.8: Diagrama de Voronoi obtido para um ponto isolado e_1 e um conjunto de infinitos pontos ao longo de um segmento de reta.

Considere uma parametrização um-a-um sobre o segmento de reta $\overline{e_2e_3}$ onde o parâmetro real t o cobre completamente da esquerda para a direita. Cada elemento e_t do segmento $\overline{e_2e_3}$, exceto por suas extremidades, tem como sua região de Voronoi $RV(e_t)$, de acordo com a definição, uma linha perpendicular a $\overline{e_2e_3}$ passando por e_t . Qualquer ponto sobre essa linha está mais próximo de e_t do que de qualquer outro elemento. As RVs correspondentes às extremidades são as áreas indicadas na Figura por $RV(e_2)$ e $RV(e_3)$. Assim, se tentarmos encontrar a união de todas as fronteiras de todas as regiões de Voronoi, seguindo a definição, nós encontraremos o mesmo diagrama de Voronoi da Figura 4.7, só que agora acrescido da área cinza mostrada na Figura 4.8 (um DV com área não nula). O diagrama de Voronoi contém esta área por que as regiões de Voronoi dos elementos em $\overline{e_2e_3}$ (linhas verticais) cobrem toda a área cinza e, como a fronteira de qualquer linha reta (em $\mathbb{R}^n, n > 1$) é a própria linha reta [43], aquela área, de acordo com a Definição 3, deve ser também parte do DV. Portanto, dois diagramas de Voronoi diferentes (i.e., aqueles das Figuras 4.7 e 4.8) são obtidos para cada uma das possibilidades de agrupamento consideradas sobre o mesmo conjunto de pontos.

Adicionalmente, observemos na Figura 4.6 e em qualquer outro DV (clássico ou generalizado) encontrado na literatura que qualquer ponto do diagrama de Voronoi tem a propriedade de ser “minimamente eqüidistante” a, pelo menos, dois elementos. Isto significa que, se a distância de um ponto p do DV àqueles elementos é D , não existe outro elemento s tal que $d(s, p) < D$. A Figura 4.9 mostra dois pontos P_1 e P_2 sobre o plano do DV da Figura 4.6 que são eqüidistantes aos elementos de Voronoi, porém somente P_1 é minimamente eqüidistante.

Figura 4.9: Dois pontos ilustrando a propriedade de mínima eqüidistância dos pontos do diagrama de Voronoi. Somente P_1 é minimamente eqüidistante a pelo menos dois elementos.

Embora seja aparentemente sutil, esta é precisamente a razão que motivou uma redefinição do diagrama de Voronoi a uma versão mais geral, permitindo o tratamento coerente de conjunto de pontos conectados. A fim de contornar o problema ilustrado na Figura 4.8, onde nenhum dos pontos sobre a área cinza é minimamente eqüidistante a quaisquer elementos, a seguinte definição alternativa é proposta:

Definição 4. A região de Voronoi de um elemento $e \in E$, denotada por $RV(e)$, é o conjunto de pontos estritamente mais próximos de e do que de qualquer outro elemento em E :

$$RV(e) \doteq \{ p \in \mathbb{R}^n \mid d(p, e) < d(p, e'), \forall e' \neq e, e' \in E \}$$

Dessa forma, é garantido, por construção, que qualquer ponto que não está em nenhuma região de Voronoi, i.e., qualquer ponto no complemento da união de todas as regiões de

Voronoi, será eqüidistante a, pelo menos, dois elementos de Voronoi. Assim, enunciamos:

$$DV(E) \doteq \left[\bigcup_i RV(e_i) \right]^c$$

Embora [187] adote a mesma definição para o diagrama de Voronoi, ele não enuncia explicitamente suas vantagens, nem mesmo que esta é diferente daquela que é classicamente utilizada. Além disso, ambas as definições 3 e 4 nunca foram aplicadas diretamente a objetos contínuos como foi feito aqui no caso das Figuras 4.7 e 4.8.

Considerando esta definição alternativa, o caso ilustrado na Figura 4.8 torna-se idêntico ao da Figura 4.7. Isto acontece porque os pontos na linha em forma de parábola são os únicos pontos que não são estritamente mais próximos a um elemento do que a qualquer outro. Portanto, estes são os únicos pontos que não pertencem a nenhuma das regiões de Voronoi determinadas pela Definição 4. Os resultados produzidos pela Definição 4 diferem da Definição 3 apenas quando elementos conectados são considerados. Esta é a mais importante característica desta nova definição. Ela permite um tratamento teórico coerente de conjuntos de dados gerais tais como curvas contínuas (formas) sem apelar para qualquer tipo de amostragem de bordas (ver seção 4.6.5). É também notável que a Definição 4 mantém todas as bem estabelecidas propriedades dos DVs para elementos discretos e desconectados.

Um outro caso ilustrativo é tratado a seguir visando dar alguma noção sobre conjuntos de dados menos abstratos. A Figura 4.10 mostra uma forma retangular C . Considere C como sendo composto por um conjunto de infinitos, distintos, porém conectados, elementos. Para cada elemento de Voronoi distinto, por exemplo, $p_1 \in C$, pode ser verificado que a região de Voronoi a ele associada é uma semi-reta perpendicular ao retângulo em p_1 partindo de Q mas não contendo este ponto (veja p_1 na Figura 4.10). Este ponto especial Q , de acordo com a definição 4, não pertence a nenhuma RV pois não é um ponto exclusivamente mais próximo a nenhum elemento e , portanto, é um ponto do DV. Os quatro cantos do retângulo constituem exceções porque suas regiões de Voronoi são áreas (e não semi-retas), conforme indicado em cinza na Figura 4.10 para p_2 . Pontos nesta área são estritamente mais próximos de p_2 que de qualquer outro elemento em C .

Depois que todas as (infinitas) regiões de Voronoi são consideradas, o diagrama

Figura 4.10: Uma forma contínua retangular, considerada como sendo composta por infinitos pontos (i.e. todos os pontos sobre sua borda).

de Voronoi é obtido tomando-se os pontos correspondentes ao complemento da união de todas as regiões de Voronoi, seguindo-se a Definição 4. A Figura 4.11(b) mostra o diagrama de Voronoi obtido desta forma para o retângulo representado pela união de todos os pequenos círculos em 4.11(a) para formar a estrutura contínua em 4.11(b).

(a)

(b)

Figura 4.11: (a) Pequenos círculos dentro do retângulo correspondem a alguns pontos do diagrama de Voronoi. (b) Diagrama de Voronoi obtido quando todos os infinitos pontos sobre o retângulo são considerados.

É interessante observar que nas porções da forma onde a direção normal não é definida (tais como cantos e extremidades) as correspondentes regiões de Voronoi têm áreas não-nulas. A razão para tal fato é que a distância medida entre um ponto e uma curva (ou forma) se dá ao longo da direção normal à curva, por exemplo, p_1 , p_2 e p_4 na Figura 4.12. A única exceção ocorre quanto tal direção não é definida, i.e., nas extremidades ou vértices da curva (veja p_3 ou qualquer outro ponto entre as linhas pontilhadas na Figura 4.12).

É importante notar que esta nova definição para o diagrama de Voronoi e para as

Figura 4.12: A distância entre um ponto e uma curva é medida ao longo da direção normal à curva SE tal direção pode ser definida.

regiões de Voronoi não invalida as definições clássicas, mas apenas as estende convenientemente para conjuntos de dados contínuos/conectados. Quando aplicadas aos casos discretos comuns (Figura 4.6), o mesmo resultado é obtido. Além disso, será mostrado na seção 4.6.4 que a redefinição proposta também permite que diagramas de Voronoi e eixos mediais sejam adequadamente relacionados.

4.6.3 Os eixos mediais

Os eixos mediais [22, 23, 24] freqüentemente chamados de esqueletos, são um conceito geométrico de grande influência em análise de imagens amplamente utilizado para análise de formas e extração de características. Eles correspondem aos pontos localmente simétricos à forma, no sentido de ser uma estrutura central (ver Figura 4.13).

O eixo medial é normalmente definido na literatura como:

Definição 5. Considere S como um subconjunto do espaço euclidiano n -dimensional e sua borda B . Uma bola aberta dentro de B que não for contida por nenhuma outra bola aberta dentro de B é chamada de bola maximal. O eixo medial (EM) de S é definido como o lócus dos centros de suas bolas maximais.

A Figura 4.14(a) mostra o eixo medial de uma elipse e algumas das bolas ma-

Figura 4.13: Forma de corpo com sua representação “central”.

ximais que compõem seu eixo medial (i.e., a linha horizontal). É também de senso comum a existência do EM externo para formas côncavas, sendo este definido pela simples troca do termo dentro pelo termo fora na definição acima.

Figura 4.14: (a) Eixo medial de uma elipse com três bolas maximais dentro. (b) Estrela e seus eixos mediais internos e externos.

De qualquer modo, a definição clássica do EM ainda depende desta necessidade de se identificar regiões internas e externas e o EM da Definição 5 não pode, portanto, ser aplicado a objetos tais como curvas abertas, curvas que se auto-interceptam ou

mesmo conjuntos de pontos discretos/desconectados (Figura 4.15). Porém, também para tais objetos, é perfeitamente concebível (e desejável) que suas simetrias locais sejam representadas.

Figura 4.15: Alguns objetos para os quais o eixo medial clássico não está definido. (a) Uma curva que se auto-intercepta; (b) uma forma aberta; (c) um conjunto de pontos desconectados.

Por exemplo, considere a curva de Limaçon C da Figura 4.16. A bola aberta B_1 é “maximal” no sentido de que ela é um tipo de bola que não contém pontos de C e nenhuma outra bola deste tipo a contém estritamente. Já B_2 e B_3 não são maximais, visto que a primeira é contida por B'_2 e a última contém pontos de C . Este conceito de bola maximal é a chave para se estender os EMs a formas gerais. Os centros de todas as bolas “maximais” relativas a C na Figura 4.16 formam um eixo localmente simétrico (linha cinza) que poderia ser considerado como o EM de um Limaçon. Encorajados por este exemplo, propomos uma extensão da definição de eixo medial clássico a fim de que o conceito forjado por Blum seja aplicado a uma classe maior de objetos que apenas a curvas fechadas.

Definição 6. Considere S como sendo um subconjunto do espaço euclidiano n -dimensional. Uma bola aberta B é dita ser uma bola maximal de S se ela satisfaz as seguintes propriedades:

1. B pertence ao conjunto G das bolas abertas que não contém nenhum ponto de S .
2. B não é um subconjunto próprio de qualquer outra bola em G .

Figura 4.16: O conceito proposto de bola maximal ilustrado sobre uma curva de Limaçon. B_1 é maximal mas, nem B_2 , nem B_3 o são.

O eixo medial de S é então definido como o locus dos centros de todas as bolas maximais de S que a tocam em, pelo menos, dois pontos.

A razão para tomar apenas as bolas que tocam S em, pelo menos, dois pontos vem do fato de que existe uma quantidade numerável de bolas maximais que tocam S em somente um ponto [140, 198]. Estas bolas coincidem com as bolas osculantes naqueles pontos de S . No entanto, nestes casos, não há simetria local já que não existem dois pontos correspondentes que possam ser identificados como localmente simétricos.

As linhas cinzas na Figura 4.16 e 4.17 indicam o eixo medial estendido de vários objetos anteriormente intratáveis sob a definição clássica. É fato importante que a Definição 6 aplica-se também aos objetos fechados anteriores. Note, por exemplo, que o EM estendido do retângulo na Figura 4.17(d) coincide exatamente com o EM clássico da mesma figura.

A extensão do EM aqui proposta permite, desta forma, a extração das simetrias locais de objetos muito mais gerais. Esta generalidade é o que fundamentará a relação entre o EM e o VD estendidos, conforme explicado em detalhe na próxima seção.

Figura 4.17: Os eixos mediais estendidos de: (a) uma parábola; (b) um ponto e um segmento de reta; (c) Três pontos no plano; (d) um retângulo, que é exatamente o mesmo que para a definição clássica de EM. Repare a similaridade entre os EMs em b, c, d e as respectivas figuras da seção 4.6.2.

4.6.4 Relacionamento entre diagramas de Voronoi e Eixos Mediais

Revisão de trabalhos relacionados

Diagramas de Voronoi e eixos mediais são conceitos profundamente inter-relacionados. Por exemplo, não é por mera coincidência que as Figuras 4.6 e 4.17(c); 4.7 e 4.17(b); 4.11(b) e 4.17(d) são similares.

Décadas atrás, Kirkpatrick [108] foi o primeiro a descobrir uma forma de encontrar o EM a partir de um DV para formas poligonais convexas simples. Mais tarde, Lee [116] estendeu isto para formas poligonais não convexas e sem buracos (Figura 4.18). Nesta abordagem, Lee notou que o EM exato de uma forma poligonal pode ser obtido de seu respectivo DV quando os elementos de Voronoi considerados fossem cada um dos lados (como segmentos de reta abertos) e cada um dos vértices do polígono. Naquele trabalho, as arestas do diagrama de Voronoi incidentes nos vértices côncavos do polígono deveriam ser removidas. Esta ação justifica-se pelo fato de que tais arestas não contêm qualquer centro de qualquer bola maximal.

Neste ponto, é importante ressaltar que a definição usada para construção do DV foi a Definição 3. De acordo com este trabalho, o DV é um superconjunto do EM para formas poligonais. Outros trabalhos ampliaram a aquisição do EM de polígonos mais gerais a partir de seus DVs tais como polígonos com buracos [190, 143]. Porém, a idéia permanecia essencialmente a mesma.

Figura 4.18: Forma poligonal e seu diagrama de Voronoi (seguindo a Definição 3) a partir das arestas e dos vértices considerados como elementos de Voronoi (a), e seu respectivo eixo medial através da remoção de alguns ramos do DV (b).

Uma nova abordagem sobre este tema apareceu em 1991 com o trabalho de Brandt e Algazi [34, 33]. Basicamente, eles mostraram que o EM de uma forma contínua e regular pode ser aproximado por um subconjunto do DV de pontos amostrados sobre a borda da forma. Eles também mostraram que esta aproximação melhora conforme a densidade da amostragem aumenta. Isto é ilustrado na Figura 4.19(a), onde pode-se ver claramente que quando a amostragem aumenta, o EM do retângulo também melhora em qualidade. Na Figura 4.19(d), uma versão aproximada do EM da figura do menino é claramente visível.

Trabalhos subseqüentes na mesma direção, em sua maioria, contribuíram com técnicas de poda dos DVs para a obtenção de subconjuntos destes que aproximassem os EMs. Frequentemente, como em [158], o critério de poda proposto cria parâmetros de ajuste multi-escala para a escolha do nível de detalhes captados pelo esqueleto. Embora estes trabalhos relacionem DV e EM, em essência, eles o fazem não através da busca por uma relação definitiva entre EM e DV, mas através da visão de que o DV é meio através do qual algoritmos de geometria computacional pode ser aplicados no cálculo do EM.

Figura 4.19: Aproximando o Eixo Medial a partir do diagrama de Voronoi de pontos amostrado ao longo da borda de objetos.

Talvez o trabalho mais importante sobre o relacionamento entre EM e DV (e também outras entidades como o *cut locus*) foi apresentado por Wolter em [198] e, mais recentemente, numa publicação em colaboração com Sherbrooke e Patrikalakis [187]. Wolter definiu uma estrutura geométrica chamada *cut locus* que, de acordo com suas próprias palavras, “oferece um panorama comum à unificação clara de diferentes conceitos como o DV, EM e outros conjuntos de pontos equidistanciais”. Apesar das similaridades entre os objetivos deste e de nosso trabalho, algumas diferenças merecem destaque. Por exemplo, Wolter restringiu o DV a conjuntos de pontos discretos e não o relacionou explicitamente com o EM. Na seção seguinte, nós estabelecemos um relacionamento próximo entre DV e EM com base nas redefinições destes dois conceitos apresentadas nas seções anteriores.

O relacionamento proposto

Nesta seção, mostraremos e provaremos que o diagrama de Voronoi e o eixo medial são conceitos muito próximos. Começaremos pelo seguinte teorema (seguindo as Definições 4 e 6) ilustrado pela Figura 4.20:

Teorema 1. *Todo ponto do diagrama de Voronoi é também um ponto do eixo medial.*

Figura 4.20: Qualquer ponto do DV é o centro de uma bola maximal e portanto, um ponto do EM.

Visão geral da prova: Esta prova de equivalência consiste em mostrar que qualquer ponto do DV é o centro de alguma bola maximal. Para tanto, construímos uma bola centrada em um ponto qualquer do DV e mostramos que ela obedece as duas condições necessárias para ser uma bola maximal. Isto é feito na forma de duas sub-provas separadas nas afirmações 1 e 2.

Demonstração. Nós inicialmente lembramos que qualquer ponto do DV é equidistante a, pelo menos, dois elementos de Voronoi. Estes elementos equidistantes serão chamados VS_P . Considere a bola aberta $B(P, R)$ centrada em P e com raio R igual à distância entre P e qualquer um dos elementos em VS_P . Será mostrado que $B(P, R)$ é uma bola maximal.

Afirmação 1: $B(P, R)$ não contém nenhum ponto de qualquer elemento de Voronoi.

Suponha que B contém um ou mais pontos de um elemento S . Assim, P seria mais próximo a S e, por esta razão, deveria ser parte da região de Voronoi de S . Portanto, de acordo com a Definição 4, P não poderia ser um ponto do DV, fato que está em contradição com a suposição inicial.

Até agora, provamos que B não contém pontos de qualquer elemento de Voronoi.

Afirmção 2: Qualquer bola aberta $B'(P', R')$ que seja um super-conjunto completo de $B(P, R)$ deve possuir, pelo menos, um ponto de um elemento de VS_P .

Suponhamos $B \subsetneq B'$ e que B' não contém qualquer ponto de nenhum elemento. Em particular, e sem perda de generalidade, um ponto de um elemento $S \in VS_P$ não está em B' . Dado que $B \subsetneq B'$, e usando uma propriedade bem conhecida das bolas euclidianas⁵, temos que:

$$d(P, P') + R \leq R' \quad (4.1)$$

Segue imediatamente da desigualdade triangular que:

$$d(P', S) \leq d(P, P') + d(P, S)$$

De (4.1):

$$d(P', S) \leq R'$$

Mas nenhum ponto de S pertence a B' , o que significa:

$$d(P', S) \geq R'$$

Com isso, $d(P', S) = R'$.

Já que é garantida a existência de, pelo menos, dois elementos em VS_P , seja $S' \neq S$ um segundo elemento em VS_P . Seguindo exatamente o mesmo raciocínio feito sobre S , chegamos a $d(P', S') = R'$. Agora, neste momento da prova, temos:

S e S' são dois elementos distintos na borda de ambas B e B' .

⁵Brandt fornece uma prova simples desta propriedade em [34].

Isto implica⁶ que B' não pode conter B e, ao mesmo tempo, ser diferente de B , uma contradição à suposição inicial (B é um subconjunto próprio de B'). Assim, B' deve conter ao menos um ponto de um elemento de VS_P , caso contrário, ele coincidiria com B .

Com isso, B é uma bola máxima e, conseqüentemente, P é um ponto do eixo medial.

□

Teorema 2. *Quando não houver nenhum agrupamento de pontos para a formação dos elementos de Voronoi, um ponto do eixo medial será também um ponto do diagrama de Voronoi.*

Demonstração. Primeiramente, lembramos que para um ponto P do eixo medial, existe uma bola maximal B associada e, de acordo com a definição de bola maximal, B não contém pontos do objeto. Além disso, da Definição 6, B tem ao menos dois pontos do objeto em sua borda, evidentemente equidistantes de P . Assim, se os pontos do objeto não estão compondo o mesmo elemento de Voronoi (i.e., os elementos considerados estão todos desagrupados), P não está mais próximo de algum ponto específico do objeto do que qualquer outro ponto. Portanto, como uma RV é o conjunto de pontos mais próximos de um elemento do que de qualquer outro, P não está em nenhuma RV. Finalmente, de acordo com a Definição 4, P é um ponto do DV.

□

Devemos restringir esta equivalência a elementos não-agrupados porque, conforme previamente discutido, a possibilidade de agrupar pontos para compor elementos de Voronoi extensos dá ao diagrama de Voronoi resultante (conhecido como DV generalizado) características não-locais. Diferentemente, o eixo medial é sempre local, ou seja, depende apenas do conjunto de pontos e não de possíveis agrupamentos como é o caso do DV. Este fato é ilustrado na Figura 4.21.

⁶Como B e B' são bolas abertas distintas, tratemos C e C' , respectivamente, como suas bordas. Dessa forma, C e C' têm dois pontos em comum, P_1 no elemento S e P_2 no elemento S' . Primeiramente, considerando P_1 , C pode tocar C' em P_1 de somente duas formas: tangencialmente ou secantemente (com tangentes diferentes ou iguais em P_1). Considerando a primeira possibilidade (tangente), a única maneira de se garantir a existência de um segundo ponto em comum P_2 é que C seja igual a C' , o que não é permitido pela suposição inicial. Portanto, apenas a segunda possibilidade (secante) é válida. Assim, dado que C e C' são secantes em P_1 (elas têm diferentes tangentes naquele ponto), C necessariamente tem pontos dentro e fora de C' (dada a continuidade da borda da bola). Fato que faz C não-contido em C' .

Figura 4.21: (a) Diagrama de Voronoi (linha tracejada) de um ponto (e_1) e um segmento de parábola (e_2). (b) Diagrama de Voronoi (linha tracejada) de e_1 e infinitos elementos (e_t) ao longo do mesmo segmento de parábola (note a linha vertical adicional no meio da parábola). (c) Eixo medial do mesmo conjunto de dados.

A Figura 4.21 mostra dois DVs (4.21(a) e 4.21(b)) e um EM (4.21(c)). Em 4.21(a), os pontos agrupados compondo o elemento e_2 causam a ausência da linha do diagrama de Voronoi no meio da parábola. Em 4.21(b), onde não há agrupamento, todos os distintos elementos pontuais contribuem para a presença daquele segmento, levando a um DV que coincide com o EM estendido mostrado em 4.21(c).

Figura 4.22: Relacionamento entre DV e EM.

Podemos analisar os diagramas de Voronoi e os eixos mediais para ver quem contém quem. De acordo com as Definições 4 e 6, o diagrama de Voronoi é conceitualmente mais abrangente porque permite o tratamento de conjuntos de dados

mais gerais que podem vir a ser tratados como agrupamentos ou pontos isolados. O eixo medial, dado sua natureza estritamente local, pode ser reduzido ao diagrama de Voronoi no caso de não existirem agrupamentos.

Graficamente, no entanto, o esquema de quem contém quem se inverte. O eixo medial é mais amplo, sendo um superconjunto do diagrama de Voronoi. A Figura 4.22 ilustra esta colocação conceitual e gráfica.

Esta conexão bem definida elucidada a questão acerca da extração do eixo medial a partir do diagrama de Voronoi. Ela mostra claramente, como proceder a fim criar regras não-empíricas para *extrair* o EM do DV e não apenas “aproximar” um usando o outro.

Extraindo o Eixo Medial do diagrama de Voronoi

Existem diversos trabalhos sobre a esqueletização de formas baseadas em diagramas de Voronoi [34, 116, 33, 5]. Estes métodos têm sido aplicados em situações onde esqueletos com características como estabilidade a rotações, ruídos e rapidez no processamento são necessárias. No entanto, porque eles estão baseados na Definição 3, onde o DV aparece como um super-conjunto do EM, eles precisam apelar para certas regras de poda afim de obter boas aproximações do EM. De qualquer forma, é um problema escolher a melhor ou a mais correta técnica para se podar o DV, considerando-se que todas as técnicas propostas para esse fim tem características positivas e negativas [33].

Tais métodos poderiam vir a se beneficiar dos relacionamentos aqui estabelecidos da seguinte forma. A Definição 4, se aliada a adequada informação de conectividade dos elementos de Voronoi, produz um DV que é igual ao eixo medial, conforme afirmam os teoremas 1 e 2. Para elementos conectados, não aparecem ramos do DV entre eles. Assim, a correta aproximação do EM fica prontamente disponível sem a remoção de ramos “externos” como é feito em [34]. Além disso, este critério de conectividade é independente de regiões internas ou externas, assim, aplica-se naturalmente também a curvas abertas e/ou com auto-intersecções. Isso proporciona uma forma de aproximar o EM destes objetos a partir dos algoritmos de DV existentes.

Os métodos para a extração do EM do DV de polígonos também poderiam se beneficiar das mesmas considerações. O DV da Figura 4.18(a) está correto de acordo com a Definição 3. Em contraste, a Definição 4 produziria diretamente a Figura 4.18(b), sem as arestas incidentes nos vértices côncavos. Aquelas arestas “extras”, presentes na Figura 4 estão dentro das regiões de Voronoi porque estão mais próximas dos vértices côncavos do que de qualquer outro elemento, fato que pela Definição 4 o excluiria automaticamente do DV.

4.6.5 Conclusões

Esta seção apresentou uma nova e mais coerente abordagem aos diagramas de Voronoi e aos eixos mediais, com atenção especial dada ao seu inter-relacionamento. Ao redefinir ambos os conceitos numa forma mais abrangente, tornou-se possível esclarecer e generalizar vários conceitos. Por exemplo, diagramas de Voronoi foram devidamente estendidos para atuar sobre formas contínuas e o eixo medial foi definido para conjuntos de pontos isolados e curvas com auto-intersecções. Ademais, os eixos mediais mostraram-se ser um caso particular dos diagramas de Voronoi e foi também mostrado que o ato de se podar os DVs afim de se obter os respectivos EMs é desnecessário, considerando-se as novas definições. Tais resultados apontam para uma potencial reavaliação e unificação de várias das abordagens de poda descritas na literatura. De fato, os resultados obtidos também indicam que a “convergência” dos DVs para os EMs representa um conceito que pode ser naturalmente melhorado em termos da adjacência entre os pontos da borda dos objetos.

Finalmente, a abordagem descrita também apresenta potencial para sugerir que novos métodos sejam criados. Por exemplo, a consideração da conectividade na grade ortogonal (conectividade 4 ou 8) pode ser explicitamente levada em conta afim de eliminar naturalmente as arestas de Voronoi que, de outra forma, seriam produzidas entre pontos que são adjacentes na grade (como, por exemplo, em [34]).

4.7 Conjuntos de Simetria e Paralelas Afins por Área

Os eixos mediais e de simetria baseados em área, também conhecidos como esqueletos afins resistentes a ruídos [18] são potenciais descritores de forma cujas características

de invariância e robustez tornam-os adequados para a aplicação em tarefas de visão computacional. Nesta seção, motivaremos o uso destas entidades, estudaremos algumas de suas propriedades matemáticas e proporemos uma definição aplicável a curvas não convexas junto com um algoritmo incremental rápido para computá-los⁷.

4.7.1 Transformações, Invariâncias e Geometrias

As transformações geométricas mais simples que conhecemos são as translações e rotações. Estes dois tipos de transformações, quando compostos dão nome a uma categoria de transformações chamada *transformações de corpo rígido* ou *isometrias*. A maioria dos métodos utilizados em visão computacional para a análise de formas apresenta invariância a este tipo de transformação. É o caso, por exemplo, da curvatura e do CS de um dado contorno que, independentemente da posição ou orientação deste, mostram-se inalterados.

Um terceiro tipo de transformação é a transformação de escala uniforme onde o tamanho de um dado objeto é alterado, porém a proporcionalidade entre as diferentes partes do mesmo não muda. A composição de uma transformação de corpo rígido com uma transformação de escala uniforme recebe o nome de *transformação de similaridade* pois, apesar de alterar a posição, a orientação e o tamanho da forma, as razões entre os comprimentos e os ângulos entre diferentes partes da forma se mantêm, bem como as linhas que eram paralelas antes e que continuam paralelas após aplicado este tipo de transformação⁸.

Na seqüência, podemos considerar a transformação de escala *não*-uniforme como um quarto tipo de transformação onde aplica-se, nas direções horizontal e vertical, diferentes fatores de escala. Neste caso, o objeto transformado perde a relação de similaridade com objeto original, sendo esta a forma mais simples de se distorcer um objeto. A Figura 4.23(b) ilustra este tipo de distorção.

Existe um quinto tipo de transformação chamada inclinação (em inglês, *shearing*) cujo resultado de sua atuação é ilustrada na Figura 4.23(c), onde percebe-se a inclinação na direção horizontal.

⁷Existe um artigo correspondente a esta seção sendo preparado[66].

⁸Denominaremos tal característica de *preservação do paralelismo*.

Figura 4.23: (a) Imagem original. (b) Transformação de escala não-uniforme. (c) Transformação de inclinação horizontal. (d) Transformação afim.

A expressão para este tipo de transformação, levando um ponto de coordenadas (x, y) em outro ponto de coordenadas (u, v) é:

$$x = u + a * v$$

$$y = v$$

onde a é chamado *fator de inclinação*.

Embora seja evidente que uma inclinação também distorce a imagem original (por exemplo, os ângulos são alterados), pode-se verificar que o paralelismo é preservado bem como a área da imagem.

A composição de qualquer um destes dois tipos de distorção (escala não-uniforme ou inclinação) com uma transformação de similaridade é o que dá origem ao conjunto de transformações conhecidas por *transformações afins*. Na Figura 4.23(d), mostramos uma imagem distorcida por uma transformação afim. Esta categoria de transformações tem por característica preservar o paralelismo e ter na razão entre as áreas dos objetos uma entidade invariante.

Neste ponto, é importante fazer uma observação sobre a diferença entre entidade preservada e entidade invariante. Ao falarmos das transformações geométricas, dis-

semos, por exemplo, que nas transformações de similaridade, os ângulos e as razões entre os comprimentos são preservados, significando que, independentemente da transformação de similaridade aplicada sobre o objeto, estas medidas jamais serão alteradas. Porém, quando dizemos que, numa transformação afim, a área dos objetos é invariante, isto significa que a área do objeto transformado é alterada na mesma proporção da transformação realizada.

O principal fato que nos motiva a criar e estudar entidades geométricas que sejam invariantes às transformações afins é que estas servem de *aproximação* para o que acontece quando um objeto do mundo tridimensional é projetado num plano. Dizemos aproximação, pois a transformação geométrica que modela exatamente o processo de projeção óptica é um grupo ainda mais geral de transformações chamadas de *transformações projetivas* [145, 90]. No entanto, por constituírem um grupo muito geral de transformações geométricas e, conseqüentemente, por preservarem pouquíssimas entidades (apenas a razão cruzada ou razão da razão do comprimento das linhas de um objeto), a criação e o estudo de elementos projetivamente invariantes está além de nossas pretensões.

Apenas para dar uma idéia do que é um conjunto de simetria invariante às transformações afins e de suas vantagens, apresentamos na Figura 4.24(a) uma forma e, junto dela, o seu respectivo conjunto de simetria afim por área (o CSAA, uma das contribuições deste trabalho que será descrita adiante). Em 4.24(b), a mesma forma distorcida por uma transformação afim apresenta um CSAA que, apesar de não ser exatamente o mesmo de 4.24(a), está alterado na mesma proporção da distorção afim aplicada. Para salientar as vantagens do CSAA sobre o eixo medial euclidiano tradicional, em 4.24(c) e 4.24(d), a mesma distorção foi aplicada sobre a forma, porém vê-se que o EM é alterado de uma maneira onde a hierarquia de suas ramificações muda, o que já não ocorre com o CSAA.

4.7.2 Introdução

A construção de conjuntos de simetria e de eixos mediais de curvas planas e superfícies tridimensionais tem se mostrado útil para a análise de formas. O conjunto de simetria de uma curva plana fechada e simples é o lugar geométrico dos centros dos círculos tangentes à curva em mais de um lugar ('círculos bitangentes'), jun-

Figura 4.24: (a) Forma original e respectivo CSAA. (b) Forma distorcida por transformação afim e respectivo CSAA. (c) Forma original e EM. (d) Forma distorcida por transformação afim e EM.

tamente com seus pontos limite (extremidades). Já os eixos mediais usam apenas aqueles círculos bitangentes cujos raios são iguais à mínima distância do centro do círculo até a curva ('círculos maximais'). Estas construções, por serem baseadas na distância expressa na forma do raio de um círculo, são invariantes à transformações euclidianas no plano: rotações, translações e reflexões. Algumas tentativas foram feitas com o intuito de modificar esta definição para uma que também seja invariante à transformações *afins*, ou seja, transformações lineares não-singulares compostas com translações. Um delas, inspirada no trabalho de Moisan [146] foi explorada em [18] e uma investigação do caso 3D foi iniciada em [17]. Uma aplicação em imagens médicas tridimensionais da versão 2D destes chamados *esqueletos afins* foi publicada em [101]. A idéia básica está no uso da *área*, ao invés da *distância*, como medida. Isso tem o duplo efeito de tornar a construção invariante à transformações afins e de dotá-la de maior robustez à presença de ruídos.

A Figura 4.25 mostra uma curva fechada simples, seu respectivo eixo medial 4.25(a) e os efeitos do ruído sobre ele 4.25(b). A versão transformada (afim) destas figuras é mostrada em 4.25(c) e 4.25(c). Enquanto os eixos euclidianos sofrem uma forte distorção como consequência da transformação afim, os eixos afins mostrados em 4.26(c) correspondem aos eixos originais de 4.26(a) se sujeitos à mesma transformação. A grande robustez advinda do uso da área ao invés da distância é também ilustrada

neste exemplo pelo fato de que o ruído adicionado à curva implicou na formação de novos ramos em 4.25(b), enquanto que praticamente não afetou os respectivos eixos afins. Tal invariância às transformações geométricas e robustez a ruídos fazem dos eixos mediais afins uma ferramenta particularmente adequada para a representação de formas e conseqüente caracterização.

Figura 4.25: Esqueletos euclidianos de: Uma curva fechada simples (a), a mesma curva na presença de ruído (b) e as versões transformadas afins de ambas (c) e (d).

Nesta seção, continuaremos a exploração do caso 2D dos eixos mediais e dos conjuntos de simetria afins adotando uma definição ligeiramente diferente daquela adotada em [146, 18], a fim de prover uma construção mais satisfatória para o caso de curvas não convexas. Ao final, apresentamos um algoritmo para calcular as ‘paralelas afins’ de uma dada curva—este tipo de ‘evolução afim’ das curvas se dá através destas paralelas—e algumas propriedades matemáticas dos eixos mediais e dos conjuntos de simetria resultantes baseados na área.

4.7.3 Curvas Convexas

Seja γ uma curva plana fechada simples. Usaremos também γ para uma parametrização da curva, por exemplo, $\gamma(t) = (a \cos t, b \sin t)$ é (a parametrização de) uma

Figura 4.26: Esqueletos afins de: Uma curva fechada simples (a), a mesma curva na presença de ruídos (b) e as versões transformadas afins de ambas (c) e (d).

elipse quando a e b são constantes. Considere as cordas de γ . Estaremos interessados na *área entre as cordas e a curva*. Esta é uma idéia trivial quando γ é convexa se definirmos a orientação da corda para determinar de que ‘lado’ da corda a área será contabilizada. Porém, para curvas não-convexas, necessitamos de uma convenção para dar à área uma interpretação única. Veja §4.7.4.

Considere inicialmente uma curva plana que pode ser definida pela equação polar $r = r(\theta)$, $r > 0$, com o pólo em X dentro da curva e uma corda fazendo um ângulo α com a horizontal conforme mostrado na Figura 4.27. Por exemplo, uma curva convexa tem esta propriedade para qualquer ponto X dentro de γ .

Figura 4.27: Curva γ com equação polar $r = r(\theta)$ relativa ao pólo X .

A área da região da curva ‘acima’ da corda é

$$A(\alpha) = \frac{1}{2} \int_{\alpha}^{\alpha+\pi} r^2 d\theta$$

de forma que

$$A'(\alpha) = \frac{1}{2} r(\alpha + \pi)^2 - \frac{1}{2} r(\alpha)^2$$

e

$$A''(\alpha) = r(\alpha + \pi)r'(\alpha + \pi) - r(\alpha)r'(\alpha),$$

onde ' significa a diferenciação com relação à α . Isto mostra o seguinte

- Proposição 1.** 1. $A' = 0$ se e somente se $r(\alpha) = r(\alpha + \pi)$, i.e., X é o ponto médio da corda;
2. $A' = A'' = 0$ se e somente se, adicionalmente, as tangentes à curva, na duas extremidades da corda, forem paralelas.

De fato, observando mais cuidadosamente os sinais, podemos mostrar que

3. A tem um mínimo não degenerado se e somente se X é o ponto médio da corda e as tangentes a γ nas extremidades da corda se encontrarem ‘acima’ da corda. Mais precisamente, dá-se à corda uma orientação e toma-se a área A à esquerda da corda, sendo que a última condição requer que as tangentes também se encontrem à esquerda da corda.

Podemos imaginar A como uma família de funções $A : U \times S^1 \rightarrow \mathbf{R}$, onde U é o interior de γ e A associa a (X, α) a área A como descrito acima. Para o tipo de curva considerada, o círculo S^1 pode ser identificado com a própria curva γ , por exemplo, através de uma das extremidades da corda que passa por X . Assim, estaremos considerando uma família de funções sobre a curva parametrizada por $X \in \mathbf{R}^2$. Isto é análogo à família de funções distância (ao quadrado) que está intimamente associada com as paralelas (euclidianas), evolutas, conjuntos de simetria e eixos mediais de uma curva γ (vide [80, 36]). Mais precisamente, as paralelas euclidianas são *frentes de onda*, a evoluta é o *conjunto de bifurcação* e a evoluta mais o conjunto de simetria formam o *conjunto de bifurcação completo* da família de funções distância euclidianas (ao quadrado).

Pode-se aplicar as mesmas idéias à família de funções área, pelo menos quando a curva é *convexa*, de forma que ela fica definida pela equação polar com pólo em qualquer ponto do interior de γ . Dado que ainda estamos trabalhando com uma família de funções, muitas das definições e propriedades seguem exatamente o mesmo raciocínio que no caso euclidiano.

- Proposição 2.** 1. Uma *paralela-área* será definida como o conjunto de pontos X para os quais A tem um valor fixo e, para algum α , $\frac{\partial A}{\partial \alpha} = 0$. Pela Proposição 1(1), este é o conjunto dos pontos médios das cordas que cortam a curva num valor fixo de área. Estas cordas, por consequência, são tangentes à paralela-área que, portanto, também pode ser vista como o *envelope* das cordas.
2. A *evoluta-área* é o conjunto de pontos X tais que, para algum α , $\frac{\partial A}{\partial \alpha} = \frac{\partial^2 A}{\partial \alpha^2} = 0$. Usando a Proposição 1(2), este é o conjunto de pontos médios das cordas para as quais as tangentes a γ em suas extremidades são paralelas. A evoluta-área foi estudada anteriormente sob o nome de ‘lugar geométrico dos pontos médios de tangentes paralelas’ (ou ‘midpoint parallel tangent locus’ (MPTL)) por Holtom [93]. Ela tem relacionamento muito próximo de outra construção invariante às transformações afins que é o ‘conjunto de simetria por envelope afim’ (o lugar geométrico dos centros das cônicas que têm ao menos dois pontos de contato de ordem 3 com γ).
3. Uma paralela-área tem cúspides onde ela encontra a evoluta-área (da mesma forma que no caso euclidiano). Já a evoluta-área tem cúspides onde γ tem a mesma curvatura nos pontos de contato das tangentes paralelas correspondentes. Observe que, apesar das aparências, esta é uma condição invariante às transformações afins.
4. Os conjuntos de simetria por área de γ podem então ser definidos como o lugar geométrico das auto-intersecções das paralelas-área. Um ponto X deste conjunto pertence ao *eixo medial por área* se a função área correspondente a X tem um mínimo absoluto com relação à rotação da corda em torno de X . O conjunto de simetria por área terá extremidades nas cúspides da evoluta-área, bifurcações e cúspides, sendo que estas ocorrerão na evoluta-área. Os eixos mediais por área não possuirão cúspides.
5. Pode ser mostrado também que todo ponto de γ ‘contribui para o conjunto de simetria por área’. Ou seja, dado P sobre γ , sempre é possível encontrar uma corda com

extremidade em P cujo ponto médio está sobre o conjunto de simetria por área. Está é uma propriedade importante, uma vez que se deseja que todas as partes de γ sejam usadas no cálculo do conjunto de simetria por área.

6. Outra questão importante, mas que nós não abordamos neste trabalho, é até que ponto o conjunto de simetria afim (ou o eixo medial afim) determina a curva γ ? Ou seja, que informação adicional seria necessária para re-construir γ ? Para o conjunto de simetria ou eixo medial euclidiano, sabemos da necessidade de se conhecer a função r que determina o raio dos círculos bitangentes. Qual seria o resultado correspondente para as versões baseadas na área?

Um triângulo proporciona um exemplo simples de paralelas e evoluta por área para uma curva convexa. Qualquer triângulo é equivalente a um triângulo equilátero sob ação de uma transformação afim. Para torná-lo suave e estritamente convexo, precisamos suavizar levemente os seus três cantos e ‘entortar’ levemente os lados na direção externa. Não é difícil ver que as paralelas-área têm seis cúspides—antes da suavização dos cantos, estas eram os pontos de encontro entre dois ramos de hipérboles tangenciais—nos pontos onde a corda passa por um dos vértices do triângulo. Suavizar os cantos não remove as cúspides, de fato, apenas as tornam cúspides normais (como $y^2 = x^3$), já que, conforme dito acima, estas ocorrem sobre a paralela-área quando as tangentes à curva nas extremidades da corda correspondente são paralelas. A Figura 4.28 mostra várias paralelas-área de um triângulo.

Figura 4.28: Cinco paralelas-área de um triângulo.

O correspondente eixo medial consiste de três segmentos das três medianas desenhados a partir do ponto médio de um lado até um terço da distância entre este ponto e o vértice oposto, ou seja, até o centróide do triângulo. Isto é dado pelas auto-inteseções das paralelas-área até a terceira das paralelas da Figura 4.28. O conjunto de simetria, por sua vez, continua até a metade da distância ao longo da mediana, que é quando a paralela torna-se ‘dupla’, como mostrado na quinta paralela da Figura 4.28. Isso corresponde à área constante igual a metade da área

do triângulo, ou seja, as cordas que tangenciam aquela paralela bissecta a área do triângulo. Veja a Figura 4.29.

Figura 4.29: Eixo medial por área (linhas grossas) e conjunto de simetria por área (linhas finas e grossas) de um triângulo.

Também é interessante considerar o caso onde γ é um quadrilátero. Podemos assumir, usando uma transformação afim, que dois lados adjacentes têm o mesmo comprimento e são perpendiculares, como mostra a Figura 4.30. Encontramos agora que o número de cúspides na paralela-área pode mudar com o aumento do valor da constante de área.

Figura 4.30: Duas paralelas-área de um quadrilátero. O número de cúspides muda de quatro para seis conforme o valor da constante de área aumenta.

Isto acontece porque, como mostra a Figura 4.30, uma corda que corta a quantidade de área requerida e tem uma de suas extremidades sobre um vértice do quadrilátero pode interceptar um lado para um pequeno valor de área ou outro lado

para um valor grande de área. Na figura da esquerda não há ‘tangente ao quadrilátero no vértice’ (imagine o vértice como sendo levemente arredondado) paralela ao lado onde repousa a outra extremidade da corda enquanto que, na figura da direita existe tal tangente, o que resulta numa cúspide na paralela-área. Obviamente, a transição entre estes dois casos se dá quando a área requerida alcança aquela do triângulo formado por três vértices adjacentes do quadrilátero. Este ramo do conjunto de simetria por área terminará no ponto onde ele encontra uma das diagonais do quadrilátero.

4.7.4 Curvas não convexas

Figura 4.31: Uma curva não convexa. A área cortada por uma corda pode conter partes positivas (+) e partes negativas (-).

A Figura 4.31 mostra uma curva que não é convexa. Da mesma forma que antes, consideramos todas as cordas para esta curva, uma corda sendo definida por suas duas extremidades *ordenadas* de forma que a corda é orientada como em AB ou CD na figura. A área ‘algébrica’ cortada pela corda orientada é então contabilizada positivamente se a área fica à esquerda da corda e negativamente se a área fica à direita. Podemos então continuar procurando por todas as cordas que, desta forma, cortam uma certa quantidade de área e tomar o envelope destas para definir a paralela-área. As auto-interseções destas paralelas definirão o conjunto de simetria afim.

Da mesma forma que no caso convexo, os pontos do envelope estão geralmente sobre os pontos médios das cordas, porém existe alguns casos críticos de paralelas que precisam ser examinados.

Caso 1 Considere uma corda cortando a área requerida mas sendo tangente a γ como a corda CD da Figura 4.31. Se movermos C para baixo, esperamos que D seja capaz de se mover para a esquerda ou para a direita sobre γ a fim de manter a mesma área (algébrica). Se, por outro lado, movermos C para cima, não haverá lugar para onde D possa ir a fim de manter a área. Um cálculo direto mostra que, neste caso, a paralela-área tem uma *inflexão* (como em $y = x^3$) no ponto médio de CD , sendo a tangente inflexional o próprio segmento CD . Na situação da Figura 4.31, a paralela-área tem pontos abaixo de CD que estão à esquerda do ponto médio e acima de CD que estão à direita. Um caso especial ocorre quando a própria curva γ tem uma inflexão em D . Neste caso, cálculos mostram que o conjunto de simetria é novamente tangente à corda CD no ponto médio mas, desta vez, este assemelha-se à $y = \pm x^4$ nas proximidades deste ponto.

Caso 2 A corda que corta a área requerida pode ser uma *linha bitangente* à γ . Neste caso, a corda é paralela às tangentes de γ nas duas extremidades da corda. O resultado sobre a paralela-área é um *ponto isolado* quando ambas as tangentes estão do mesmo lado da corda e um *cruzamento* quando as tangentes estão em lados opostos.

Caso 3 *A paralela de área nula* Observe que a própria curva γ constitui uma parte da paralela de área nula. Entretanto, considere a tangente sobre uma inflexão de γ . Ela pode ser considerada (no caso limite) como uma corda cujas extremidades coincidem e que corta área zero (lembrando que antes de obter a área devemos conhecer ambas as extremidades da corda). Podemos então deslocar a corda levemente mantendo a condição de área nula. Tomando uma inflexão simples $y = x^3 - ax^4 + \dots$ ($a \neq 0$) a equação local da paralela-área é $y = \frac{5}{4a}x^2 + \dots$ com $ax \geq 0$. Isto é uma ‘meia-parábola’ que termina na inflexão de γ que pode ser vista na Figura 4.32 (uma ampliação da parte inferior da Figura 4.33) onde a paralela-área zero próxima a duas inflexões é mostrada. (Nesta implementação, a paralela-área acompanha γ depois que ela tem tangência com γ nas inflexões.) Veja que a paralela tem duas cúspides.

Pode-se acompanhar o progresso de uma paralela-área conforme o valor da área passa por zero. À esquerda da Figura 4.33 vê-se a paralela-área para uma pequena

Figura 4.32: A paralela-área zero (afastada de γ) próxima a duas inflexões de γ .

área positiva. Ao passar pela área zero, a paralela engloba toda a curva γ junto com duas partes mostradas na Figura 4.32. Com uma pequena área negativa, a paralela é mostrada à direita na Figura 4.33. Na medida em que a área torna-se mais e mais negativa, diminui e desaparece primeiro um, depois o outro dos dois pedaços fechados. De acordo com o Caso 2, isto ocorre quando a corda é bitangente a γ .

Figura 4.33: Esquerda: a paralela-área para uma área pequena e positiva. Direita: para uma área pequena e negativa.

4.7.5 Computando as paralelas afins

Partindo de uma curva amostrada bidimensional simples γ , $\gamma(i)$ representa o i -ésimo ponto sobre γ , com i entre 1 e N , descreveremos um algoritmo para calcular o conjunto de paralelas-área de γ que permitirá posteriormente aproximar o conjunto de simetria afim através da detecção das auto-intersecções destas paralelas usando algoritmos padrões de geometria computacional [16].

Da definição das paralelas-área, dado um valor de área A_0 , os pontos da paralela estão sobre os pontos médios das cordas que unem pontos distintos P_1 e P_2 de γ onde a área algébrica entre estas cordas e a curva γ é igual a A_0 . Dado que é necessário calcular as paralelas para todos os diferentes valores de área a fim de aproximar o conjunto de simetria por área, calcularemos, para cada par de pontos P_1, P_2 de γ , a respectiva área e a armazenaremos numa matriz auxiliar $A(i, j)$ onde $P_1 = \gamma(i)$ e $P_2 = \gamma(j)$.

Na forma de pseudo-código:

```
A = zeros(N,N)
for idx1 = 1..N
    idx2 = (idx1 mod N) + 1
    prev_idx2 = idx2
    P1 = C(idx1); P2 = C(idx2)
    prev_P2 = P2; area = 0
    while (idx2 != idx1),
        v1 = prev_P2 - P1
        v2 = P2 - P1
        area = area + 0.5 * (v1 x v2)
        A(idx1,idx2) = area
        prev_idx2 = idx2
        prev_P2 = P2
        idx2 = (idx2 mod N) + 1
        P2 = C(idx2)
    end while
end for
```

Aqui, idx denota um índice, $prev_$ denota valores da iteração anterior, x denota a terceira componente do produto vetorial e mod é o resto da divisão inteira.

O algoritmo acima produz uma matriz A com valores de área para todos os possíveis pares de pontos P_1 e P_2 sobre γ , codificados através dos índices i e j . Consideremos A como uma superfície parametrizada em i e j .

Na Figura 4.34(b), a superfície A é mostrada para a elipse representada na

Figura 4.34(a) onde dois conjuntos de cordas compondo duas diferentes paralelas (por meio de seus envelopes) podem ser vistas. Repare que para cada conjunto de cordas há uma respectiva curva de nível sobre A . Uma figura similar correspondente a uma curva não convexa é apresentada nas Figuras 4.34(c) e 4.34(d).

Figura 4.34: (a) Uma elipse com dois conjuntos de cordas para diferentes valores de área. (b) Superfície de área correspondente e respectivas curvas de nível. (c) Uma curva não convexa com dois conjuntos de cordas para diferentes valores de área. (d) Superfície de área correspondente e respectivas curvas de nível.

Assim, para desenhar uma paralela-área específica com área A_0 , é necessário detectar em A todas as posições (correspondentes às cordas) cujos valores sejam iguais a A_0 (ou suficientemente próximos, dado que A é uma aproximação numérica da área). Plotando os pontos médios destas cordas, teremos uma boa aproximação da particular paralela-área.

Por ser necessária, posteriormente, a detecção das auto-interseções de cada paralela a fim de desenhar um aproximação do conjunto de simetria por área, precisamos extrair aquelas posições de A (com mesmos valores de área correspondentes a uma curva de nível de A) de uma forma ordenada, ou seja, precisamos parametrizar tal curva de nível.

Para a maioria dos valores de área, a correspondente curva de nível tem apenas um componente conectado mas, no caso geral, esta pode ser composta por vários

componentes desconectados. No entanto, individualmente, cada componente ainda é uma curva que pode ser facilmente parametrizada.

Uma possível abordagem para detectar tais posições de uma forma parametrizada é usar A_0 como um limiar seguido por um algoritmo de perseguição de bordas que produzirá curvas de nível de A parametrizadas.

Finalmente, dado que A está definida num espaço de parâmetros discretos e que para cada curva de nível de A parametrizada existe uma paralela-área associada (e também parametrizada), pode-se considerar tal paralela como uma curva poligonal ou um conjunto de segmentos de reta. Assim, pela busca de todas as auto-intersecções deste polígono, é possível aproximar e desenhar o conjunto de simetria por área.

Dado que curvas representadas em imagens digitais têm necessariamente que ser espacialmente quantizadas sobre uma grade ortogonal, anisotropias implicadas por tal representação causa ruídos nas paralelas obtidas, tornando-as serrilhadas (veja o lado esquerdo da Figura 4.7.5) ao ponto que várias auto-intersecções falsas são detectadas. Para minimizar este problema, uma suavização Gaussiana (ver [49] p.170) é aplicada sobre as paralelas. Conforme ilustrado no lado direito da Figura 4.7.5, este esquema leva a uma redução substancial do serrilhado, acarretando nas corretas auto-intersecções das quais os eixos mediais por área podem ser obtidos.

Figura 4.35: Esquerda: Uma paralela ruidosa (ampliada da cúspide superior esquerda da Figura 4.30). Direita: Após suavização.

Um exemplo de um eixo medial por área de uma curva não convexa γ é mostrado na Figura 4.36. Enquanto os eixos mediais euclidianos para curvas fechadas são sempre conectados, o mesmo não vale para a versão afim. Tal fato é ilustrado na Figura 4.36 que inclui eixos desconectados partindo das inflexões do contorno.

Figura 4.36: O eixo medial por área de uma curva não convexa.

4.7.6 Conclusões

Esta seção abordou a questão dos eixos mediais por área de ambos os pontos de vista teórico e prático, incluindo um algoritmo numérico para a estimação de tais eixos envolvendo a parametrização das cordas e respectiva suavização Gaussiana.

O caso 3D do conjunto de simetria por área—o conjunto de simetria por volume—apresenta desafios matemáticos consideravelmente maiores que o caso 2D. Alguns trabalhos têm sido feitos neste sentido por I.A.Bogaevsky mas aqui, deixaremos o 3D para os desenvolvimentos futuros.

Capítulo 5

Conclusões e Desenvolvimentos Futuros

5.1 Considerações Finais

Após toda a discussão sobre as representações de formas, sobre a modelagem do processamento da informação visual, sobre os critérios para as representações e do detalhamento de alguns dos pontos relacionados aos processos (algoritmos) envolvidos no cálculo de representações por curvatura e eixos de simetria, podemos concluir que a estruturação do desenvolvimento de programas de computador que visam a tomada de decisão com base em informações visuais, se baseada num modelo como o que foi utilizado neste trabalho, permite uma re-organização dos conceitos envolvidos de forma que uma imagem fotográfica passa a diferir de uma imagem de bordas não somente na questão do significado dos pixels que as compõem, mas também quanto ao nível da respectiva representação associada. Assim, ao encadear processos no sentido de enfatizar a informação-chave para a tomada de decisão, deixamos o caráter puramente heurístico e passamos a poder conectar, de uma maneira mais objetiva e estruturada, as diversas representações conhecidas e estudar os diferentes processos que as conectam com olhos voltados à preservação das características fundamentais à finalidade do processamento. Tendo sido todos estes termos definidos e exemplificados no decorrer desta tese, acreditamos estarmos contribuindo para a construção de uma linguagem comum à muitas das abordagens que solucionam problemas de visão computacional.

Esperamos ainda que aperfeiçoamentos sejam agregados ao cenário conceitual que foi desenvolvido e que definições e critérios possam vir a ser reformulados a

fim de que um número cada vez maior de representações e processos seja tratável dentro deste cenário. Numa visão de futuro, conforme será mostrado na seção 5.4, imaginamos que seja então possível o desenvolvimento de um novo tipo de software de processamento de informações visuais nos moldes, por exemplo, da programação visual permitida pelo Khoros[97], porém valendo-se de uma orientação a objetos semelhante à adotada pelo VTK [98] para dar às diferentes classes um posicionamento (tanto hierárquico quanto funcional) de acordo com o modelo de Marr. Isso poderia contribuir na validação deste modelo, pois problemas não solucionáveis através do esquema (representação-processo)-decisão poderiam levar a modificações nas suposições iniciais como, por exemplo, a da presença simultânea de abordagens “bottom-up” e “top-down”.

5.2 Uma visão comparativa entre as representações por curvatura e eixos de simetria

Comparativamente, as representações por curvatura e por eixos de simetria têm agora, além das ligações já conhecidas e que as relacionam do ponto de vista conceitual, um elo através de seus respectivos papéis dentro do modelo de processamento da informação visual¹. Enquanto as características salientadas são distintas, através do relacionamento conceitual pode-se perceber quais caminhos (ou processos) são capazes de levar os extremos de curvatura a extremidades dos esqueletos e como trechos de alta e baixa curvatura podem estar relacionados as peculiaridades da função raio associada aos eixos mediais.

Do ponto de vista dos critérios adotados e discutidos para cada representação, as representações axiais satisfizeram um conjunto maior de critérios, em especial os eixos mediais uma vez que explicitam diversas características, uma delas sendo o conjunto de suas extremidades que correspondem diretamente à principal característica destacada pela curvatura cuja acessibilidade mostrou-se maior e menos sujeita a divergências metodológicas. Os eixos mediais também estão entre as poucas representações para as quais foram criadas técnicas multi-escala sem apelo a regras de

¹Esperamos que este elo possa ser estendido a outras representações e que as relações entre os diversos conceitos tradicionalmente utilizados em processamento de imagens possam ser melhor compreendidos através das características do modelo adotado.

difusão e que, portanto proporcionam meios alternativos de filtragem e seleção de nível de detalhes.

5.3 Contribuições

Como contribuições deste trabalho podemos citar o próprio reavivamento da aplicação e desenvolvimento de modelos para o estudo do processamento de imagens, a reavaliação dos critérios para a representação de formas e respectivo posicionamento das representações por curvatura e eixos de simetria.

Com relação à curvatura, foi criado e validado um esquema para a comparação de diferentes processos (algoritmos) que visam a sua aproximação numérica onde interessantes resultados puderam ser analisados (ver seção 3.3.6) ao mesmo tempo em que propusemos um novo processo (baseado na transformada de Fourier bidimensional) para o cálculo da curvatura. Como uma alternativa aos métodos diferenciais, foi apresentado o conceito de circularidade local cujas características levantadas sinalizam o seu potencial como uma nova medida para caracterização de contornos.

A respeito dos esqueletos, o posicionamento frente aos critérios das diferentes definições (eixos mediais, conjuntos de simetria, simetrias locais suavizadas e cixo de simetria para a inferência de processo) e a discussão sobre o caráter intuitivo das representações axiais trouxe para um mesmo cenário, conceitos que até então só estavam relacionados pela busca da captura dos tipos mais essenciais de informação contida nas formas como a topologia e a estrutura de ramificação. Destaque deve então ser dado aos eixos mediais devido a sua ampla interconexão com diversos conceitos provenientes de diferentes áreas como a transformada distância e os diagramas de Voronoi.

5.4 Desenvolvimentos Futuros

Por sabermos que a atividade de pesquisa vive mais do surgimento de novas perguntas do que da elaboração de novas respostas e que, no desenrolar deste trabalho, várias questões surgiram e ficaram em aberto, enumeraremos nesta seção os pontos passíveis de maiores desenvolvimentos e/ou que carecem de mais aprofundado

estudo na intenção de tornar este trabalho mais completo, útil e abrangente.

1. Submeter outras representações de formas aos critérios apresentados em 2.5 formando um “*Handbook*” de representações, processos e características adequadas para cada particular problema de análise de imagens.
2. Aplicar a técnica para o cálculo de curvatura baseada na transformada de Fourier bidimensional (vide seção 3.3.2) na implementação de algoritmos de dinâmica de formas governadas por fluxo de curvatura baseando em “*Level Set Methods*” [184].
3. Avaliar a possibilidade ou não do surgimento de novas estruturas quando da simplificação da curvatura-área (seção 3.4) para que se possa afirmar se esta medida forma realmente um modelo de escala-espaco.
4. Levantamento das semelhanças, vantagens e desvantagens da circularidade local aqui proposta em comparação com a medida de “acúmulo de distância corda-ponto” apresentada em [87].
5. Criação de uma metodologia que permita a reconstrução da forma a partir de sua circularidade local, promovendo-a assim à posição de representação de formas.
6. Substituição do perímetro na definição da circularidade local por uma entidade invariante a transformações afins criando, desta forma, um método não derivativo de cálculo de curvatura afim.
7. Extensão da circularidade local para 3D formando assim a “esfericidade local” e posterior avaliação de suas potencialidades na análise de objetos tridimensionais.
8. Integrar a técnica de esqueletização por propagação de rótulos (vide seção 4.5) com técnicas rápidas para o cálculo da transformada distância como o que é proposto por Cuisenaire em [55], substituindo assim as dilatações exatas por algoritmos menos custosos visando sua aplicação em problemas 3D.
9. Aplicar os conjuntos de simetria afins por área (vide seção 4.7) na solução de problemas práticos onde o tipo de invariância apresentado mostrar-se útil.

10. Prototipagem de um software de processamento de imagens baseado no modelo de (representação-processo)-decisão onde o trabalho nos níveis algorítmicos e de implementação (ver seção 2.3) fosse “encapsulado” em objetos cujas entradas e saídas permitissem a criação de soluções com base apenas no nível da teoria computacional e na finalidade global do processamento.
11. Estudar a possibilidade de uso de diferentes modelos através da leitura dos trabalhos metarepresentacionais de Francis Heylighen [92].

Bibliografia

- [1] Y. Aloimonos. *Active Vision Revisited*, pages 1–18. Active Perception. Lawrence Erlbaum Associates, 1993.
- [2] I. M. Anderson and J. C. Bezdek. Curvature and deflection of discrete arcs: A theory based on the commutator of scatter matrix pairs and its application to vertex detection in planar shape data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:27–40, 1984.
- [3] C. Arcelli and G.S. di Baja. Finding local maxima in a pseudo-euclidean distance transform. *Computer Vision, Graphics, and Image Processing*, 43(3):361–367, 1988.
- [4] H. Asada and M. Brady. The curvature primal sketch. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:2–14, 1986.
- [5] D. Attali and A. Montanvert. Computing and simplifying 2d and 3d continuous skeletons. *Computing Vision and Image Understanding*, 67(3):261–273, 1997.
- [6] F. Attneave. Some informational aspects of visual perception. *Psychological Review*, 61:183–193, 1954.
- [7] F. Attneave and M.D. Arnoult. The quantitative study of shape and pattern perception. *Psychological Bulletin*, 53(6):452–471, 1956.
- [8] J. August, K. Siddiqi, and S.W. Zucker. Ligature instabilities in the perceptual organization of shape. *Computer Vision and Image Understanding*, 76(3):231–243, 1999.
- [9] F. Aurenhammer. Voronoi diagrams - a survey of a fundamental geometric data structure. *ACM Comput. Surveys*, 23(3):345–405, 1991.

- [10] S. Bailey, M. Forshaw, and M. Hodgets. A methodology for goal oriented image processing performance evaluation. In *Proceedings of the Workshop on Performance Characterisation and Benchmarking of Vision systems*, page 54, Los Palamos de Gran Canaria, jan 1999.
- [11] H. B. Barlow. Pattern recognition and the responses of sensory neurons. *Ann. N.Y. Acad. Sci.*, 156:872–881, 1969.
- [12] M. Baroni and G. Barletta. Digital curvature estimation for left ventricular shape analysis. *Image and Vision Computing*, 10(7):485–494, sep 1992.
- [13] J.L. Barron, D.J. Fleet, and S.S Beauchemin. Performance of optical flow techniques. *International Journal of Comuter Vision*, 12(1):43–77, 1994.
- [14] J. R. Bennet and J. S. Mac Donald. On the measurement of curvature in a quantized environment. *IEEE Transactions on Computers*, 24(8):803–820, 1975.
- [15] J.R. Bennet and J.S. McDonald. On the measurement of curvature in a quantized environment. *IEEE Trans. Comput.*, 24:803–820, 1975.
- [16] J.L Bentley and T. Ottmann. Algorithms for reporting and counting geometric intersections. *IEEE Transactions on Computers*, 28(9):643–647, sep 1979.
- [17] S. Betelu, G. Sapiro, and A. Tannenbaum. Affine invariant erosion of 3d shapes. In *Proceedings of the Eighth International Conference On Computer Vision*, pages 174–180, Vancouver, British Columbia, Canada, jul 2001.
- [18] S. Betelu, G. Sapiro, A. Tannenbaum, and P.J. Giblin. Noise-resistant affine skeletons of planar curves. In *6th European Conference on Computer Vision*, volume 1842 of *Lecture Notes in Computer Science*, pages 742–754. Springer-verlag, jun 2000.
- [19] S. Betelu, G. Shapiro, A. Tanenbaum, and P.J. Giblin. On the computation of the affine skeletons of planar curves and the detection of skew symmetry. *Pattern Recognition*, 34(5):943–952, 2001.

- [20] S. Betelu, G. Shapiro, A. Tanenbaum, and P.J. Giblin. On the computation of the affine skeletons of planar curves and the detection of skew symmetry. *Pattern Recognition*, 34(5):943–952, 2001.
- [21] I. Biederman. Recognition-by-components: A theory of human image understanding. *Psychol. Rev.*, 94:115–147, 1987.
- [22] H. Blum. *A transformation for extracting new descriptors of shape*, pages 362–380. *Models for the Perception of Speech and Visual Forms*. Mit Press, 1967.
- [23] H. Blum. Biological shape and visual science. *J. Theoret. Biol.*, 38:205–287, 1973.
- [24] H. Blum and R.N. Nagel. Shape description using weighted symmetric axis features. *Pattern Recognition*, 10(3):167–180, 1978.
- [25] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. Warmuth. Occam's razor. *Information Processing Letters*, 24(6):377–380, 1987.
- [26] G. Borgefors. Distance transformations in arbitrary dimensions. *Computer Vision, Graphics, and Image Processing*, 27:321–345, 1984.
- [27] G. Borgefors, G. Ramella, and G.S. di Baja. Hierarchical decomposition of multiscale skeletons. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1296–1312, nov 2001.
- [28] J.E. Bowie and I.T. Young. An analysis technique for biological shapes ii. *Acta Cytologica*, 21(3):455–464, 1977.
- [29] K. W. Bowyer and P. J. Phillips. *Empirical Evaluation Techniques in Computer Vision*. IEEE Computer Society Press, 1998.
- [30] M. Brady. Computational approaches to image understanding. *Computing Surveys*, 14(1):3–71, 1982.
- [31] M. Brady. *Criteria for Representation of Shape*, pages 39–84. *Human and Machine Vision*. Academic Press, 1983.

- [32] M. Brady and H. Asada. Smoothed local symmetries and their implementation. *International Journal of Robotics Research*, 3:36–61, 1984.
- [33] J.W. Brandt. Convergence and continuity criteria for discrete approximations of the continuous planar skeleton. *CVGIP: Image Understanding*, 59(1):116–124, 1994.
- [34] J.W. Brandt and V.R. Algazi. Continuous skeleton computation by voronoi diagram. *CVGIP: Image Understanding*, 55(3):329–338, 1992.
- [35] J.W. Bruce, P.J. Giblin, and C.G. Gibson. Symmetry sets. *Proceedings of the Royal Society of Edinburgh Section A- Mathematics*, 101:163–186, 1985.
- [36] J.W. Bruce, P.J. Giblin, and C.G. Gibson. Symmetry sets. *Proceedings of the Royal Society of Edinburgh Section A- Mathematics*, 101:163–186, 1985.
- [37] O. M. Bruno. *Paralelismo em visão natural e artificial*. PhD thesis, Instituto de Física de São Carlos, 2000.
- [38] R. Capurro. *On the Genealogy of Information*, pages 259–271. Information - New Questions to a Multidisciplinary Concept. Akademie Verlag, 1996.
- [39] R. Capurro and B. Hjørland. *The concept of Information*, volume 37 of *Annual Review of Information Science and Technology*, chapter 8, pages 343–411. Information Today, Inc., 2003.
- [40] C.H. Chen, J.S. Lee, and Y.N. Sun. Wavelet transformation for gray-level corner detection. *Pattern Recog.*, 28(6):853–861, 1995.
- [41] K. Cho, P. Meer, and J. Cabrera. Performance assessment through bootstrap. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(11):1185–1198, 1997.
- [42] M.M.S. Chong, R.K.L. Gay, H.N. Tan, and J. Liu. Automatic representation of fingerprints for data compression by b-spline functions. *Pattern Recognition*, 25(10):1199–1210, 1992.
- [43] G. Choquet. *Topology*. Academic Press, 1966.

- [44] D. Cocurjolly, S. Miguet, and L. Tougne. Discrete curvature based on osculating circle estimation. In *4th International Workshop on Visual Form, Capri, Italy, LNCS 2059*, pages 303–312, 2001.
- [45] L. F. Costa. On how jumping spiders can teach us to construct real time hough transformers. In *Workshop sobre Visão Cibernética*, pages 70–73, 1994.
- [46] L.F. Costa. Multidimensional scale space shape analysis. In *Proc. International Workshop on Synthetic-Natural Hybrid Coding and Three Dimensional Imaging, Santorini, Greece*, pages 15–17, 1999.
- [47] L.F. Costa. Estimating derivatives and curvature of open curves. *Pattern Recognition*, 35(11):2445–2451, nov 2002.
- [48] L.F. Costa and L.F. Estrozi. Multiresolution shape representation without border shifting. *Electronics Letters*, 35(21):1829–1830, 1999.
- [49] L.F. Costa and R.M. Cesar Jr. *Shape Analysis and Classification, Theory and Practice*. CRC Press, 2001.
- [50] L.F. Costa, E.D.M. Manoel, and J.S. Tanaka. A new model for neural curvature detection and its investigations in terms of neuromorphometric measures. In *III Workshop on Cybernetic Vision*, pages 18–24, 1999.
- [51] L.F. Costa and R.R. Poznanski. Methods in neuronal modeling. *Trends in Neuroscience*, 22(7):329–330, 1998.
- [52] L.F. Costa, V.O. Roda, and R. Köberle. A biologically-inspired system for visual pattern recognition. In *IEEE International Symposium on Industrial Electronics*, pages 25–30, Santiago-Chile, 1994.
- [53] L.F. Costa and T.J. Velte. Automatic characterization and classification of ganglion cells from the salamander retina. *Journal of Comparative Neurology*, 404(1):33–51, 1999.
- [54] P. Courtney, N. Thacker, and A. Clark. Algorithmic modeling for performance evaluation. In *Proceedings of the ECCV Workshop on Performance Characteristics of Vision Algorithms*, 1996.

- [55] O. Cuisenaire and B. Macq. Fast euclidean distance transformation by propagation using multiple neighborhoods. *Computer Vision and Image Understanding*, 76(2):163–172, 1999.
- [56] P-E. Danielsson. Euclidean distance mapping. *Computer Graphics and Image Processing*, 14:227–248, 1980.
- [57] A. Datta, S. Pal, and S. Chakraborti. Image thinning by neural networks. *Neural Computing & Applications*, 11(2):122–128, oct 2002.
- [58] S.L. Davis. Understanding shapes: Angles and sides. *IEEE Transactions on Computers*, C-26(3):236–242, mar 1977.
- [59] D. Demigny and T. Kamlé. A discrete expression of canny’s criteria for step edge detector performances evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(11):1199–1211, 1997.
- [60] T.K. Dey and W. Zhao. Approximate medial axis as a voronoi subcomplex. In *Solid Modeling and Applications*, pages 356–366, Saarbrücken-Germany, 2002.
- [61] M.P. doCarmo. *Differential Geometry of Curves and Surfaces*. Prentice Hall, 1976.
- [62] J. S. Duncan, F. Lee, A. W. M. Smeulders, and B. L. Zaret. A bending energy model for measurement of cardiac shape deformity. *IEEE Trans. Med. Imaging*, 10(3):307–320, 1991.
- [63] S. Edelman and L.M. Vaina. *Marr, David (1945-80)*. International Encyclopedia of the Social & Behavioral Sciences, 2001.
- [64] I. Epstein. *Cibernética e Comunicação*. Editora da Universidade de São Paulo, 1926.
- [65] L.F Estrozi, A.G Campos, L.G Rios, R.M. Cesar Jr, and L.F. Costa. Comparing curvature estimation techniques. In *IV Simpósio Brasileiro de Automação Inteligente*, pages 58–63, 1999.
- [66] L.F. Estrozi, L.F. Costa, and P.J. Giblin. Affine area parallels and symmetry sets. (*em preparação*), 2003.

- [67] L.F. Estrozi, L.G. Rios-filho, A.G.C. Bianchi, R.M. Cesar Jr, and L.F. Costa. 1d and 2d fourier-based approaches to numeric curvature estimation and their comparative performance assessment. *Digital Signal Processing*, 13:172–197, 2003.
- [68] D.P. Fairney and P.T. Fairney. On the accuracy of point curvature estimators in a discrete environment. *Image and Vision Computing*, 12(5):259–265, jun 1994.
- [69] A. X. Falcao, L. F. Costa, and B. S. da Cunha. Multiscale skeletons by image foresting transform and its application to neuromorphometry. *Pattern Recognition*, 35(7):1571–1582, jul 2002.
- [70] O. Faugeras and R. Keriven. Scale spaces and affine curvature. In R. Mohr and C. Wu, editors, *Workshop on Geometrical Modeling & Invariants for Computer Vision (GMICV'95 - Xian)*, pages 17–24, 1995.
- [71] J.T. Feddema, C.S.G. Lee, and O.R. Mitchell. Weighted selection of image features for resolved rate visual feedback-control. *IEEE Transactions on Robotics and Automation*, 7(1):31–47, 1991.
- [72] A. S. Felinto. *Visão estéreo utilizando a curvatura como medida de similaridade*. PhD thesis, Instituto de Física de São Carlos, 2000.
- [73] M.M. Fleck. *Local rotational symmetries*. PhD thesis, Massachusetts Institute of Technology, 1985.
- [74] M.M. Fleck. Multiple widths yield reliable finite differences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(4):412–429, 1992.
- [75] L.M. Forster. Vision and prey-catching strategies in jumping spiders. *American Scientist*, 70(2), 70(2):165–175, 1982.
- [76] W. Forstner. Diagnostics and performance evaluation in computer vision. In *Performance versus Methodology in Computer Vision Workshop*, pages 11–25, Seattle, 1994.

- [77] W. Forstner. 10 pros and cons against performance characterization of vision algorithms. In *Proceedings of the ECCV Workshop on Performance Characteristics of Vision Algorithms*, Cambridge-UK, apr 1996.
- [78] H. Freeman. Computer processing of line drawing images. *ACM Computing Surveys*, 6(1):57–98, 1974.
- [79] B. Gatos, N. Papamarkos, and C. Chamzas. Using curvature features in a multiclassifier ocr system. *Engineering Applications of Artificial Intelligence*, 10(2):213–224, 1997.
- [80] P.J. Giblin and J.W. Bruce. Growth, motion and 1-parameter families of symmetry sets. *Proceedings of the Royal Society of Edinburgh Section A-Mathematics*, 104:179–204, 1986.
- [81] A. Glennerster. Computational theories of vision. *Current Biology*, 12:682–685, 2002.
- [82] C. M. Gold and J. Snoeyink. A one-step crust and skeleton extraction algorithm. *Algorithmica*, 30:144–163, 2001.
- [83] P. Golland and W.E.L. Grimson. Fixed topology skeletons. *Computer Vision and Pattern Recognition*, 1, 2000.
- [84] R.C. Gonzalez and R.E. Woods. *Digital Image Processing*. Addison-wesley Pub Co, apr 1992.
- [85] I.E. Gordon. *Theories of Visual Perception*. John Wiley & Sons, 1989.
- [86] K. Grammer and R. Thornhill. Human facial attractiveness and sexual selection: The roles of averageness and symmetry. *Journal of Comparative Psychology*, 108(3):233–242, 1994.
- [87] J.H. Han and T. Poston. Chord-to-point distance accumulation and planar curvature: A new approach to discrete curvature. *Pattern Recognition Letters*, 22(10):1133–1144, aug 2001.
- [88] R.M. Haralick. Measure for circularity of digital figures. *IEEE Transactions on Systems Man and Cybernetics*, SMC4(4):394–396, 1974.

- [89] I. Hargittai and M. Hargittai. *Symmetry: A Unifying Concept*. Shelter Pubns, 1994.
- [90] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [91] M.D. Heath, S. Sarkar, T. Sanocki, and K.W. Bowyer. Robust visual method for assessing the relative performance of edge-detection algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(12):1338–1359, 1997.
- [92] F. Heylighen. *Representation and Change - A Metarepresentational Framework for the Foundations of Physical and Cognitive Science*. Communication & Cognition, 1990.
- [93] P.A. Holtom. *Affine-Invariant Symmetry Sets*. PhD thesis, University of Liverpool - Available on: <http://www.liv.ac.uk/~pjgiblin/papers/phd.pdf>, sep 2000.
- [94] D.H. Hubel and T.N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160:106–154, 1962.
- [95] D.H. Hubel and T.N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195:215–243, 1968.
- [96] D.P. Illing and P.T. Fairney. Determining perceptually significant points on noisy boundary curves. *Pattern Recognition Letters*, 12(9):557–564, sep 1991.
- [97] Khorol Research Inc. The khoros page. <http://www.khorol.com/khoros/>.
- [98] Kitware Inc. The vtk page. <http://www.kitware.com>.
- [99] R.C. Jain and T.O. Binford. Ignorance, myopia, and naivet in computer vision systems. *Computer Vision, Graphics and Image Processing: Image Understanding*, 53(1):112–117, 1991.
- [100] R.C. Jain and T.O. Binford. Ignorance, myopia, and naivete' in computer vision systems. *CVGIP: Image Understanding*, 53(1):112–117, 1991.

- [101] M-S Jeong and B.F. Buxton. The affine skeleton of cross-sectional, 3-d human body scan data. Preprint, University College London, 2001.
- [102] C. Jordan. *Calculus of Finite Differences*. Chelsea, 3rd edition, 1965.
- [103] R.M. Cesar Jr. *Análise multi-escala de formas bidimensionais*. PhD thesis, Instituto de Física de São Carlos - USP, 1997.
- [104] R.M. Cesar Jr and L.F. Costa. Piecewise linear segmentation of digital contours in $O(n \log(n))$ through a technique based on effective digital curvature estimation. *Real-Time Imaging*, 1(6):409–417, 1995.
- [105] R.M. Cesar Jr and L.F. Costa. Towards effective planar shape representation with multiscale digital curvature analysis based on signal processing techniques. *Pattern Recognition*, 29(9):1559–1569, 1996.
- [106] R.M. Cesar Jr and L.F. Costa. The application and assessment of multiscale bending energy for morphometric characterization of neural cells. *Rev. Sci. Instr.*, 68(5):2177–2186, 1997.
- [107] A. Khotanzad and R.L. Kashyap. Feature-selection for texture recognition based on image synthesis. *IEEE Transactions on Systems Man and Cybernetics*, 17(6):1087–1095, 1987.
- [108] D.G. Kirkpatrick. Efficient computation of continuous skeletons. In *Proceedings of the 20th Annu. Symp. Found. Computer Sci.*, 1979.
- [109] L. Kitchen and A. Rosenfeld. Gray-level corner detection. *Pattern Recognition Letters*, 1(2):95–102, 1982.
- [110] J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.
- [111] D. Koller and M. Sahami. Toward optimal feature selection. In *International Conference on Machine Learning*, pages 284–292, 1996.
- [112] V. Kovalevsky. Curvature in digital 2d images. *International Journal of Pattern Recognition and Artificial Intelligence*, 15(7):1183–1200, 2001.

- [113] R. Kowner. Facial asymmetry and attractiveness judgment in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, 22:662–675, 1996.
- [114] M. Kunt. Comments on “dialogue,” a series of articles generated by the paper entitled “ignorance, myopia, and naivete in computer vision”. *CVGIP - Image Understanding*, 54(3):428–429, 1991.
- [115] L. Lam, S.W. Lee, and C.Y. Suen. Thinning methodologies-a comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(9), 1992.
- [116] D.T. Lee. Medial axis transformation of a planar shape. *IEEE Trans. PAMI*, 4(4):363–369, 1982.
- [117] E. Lendvai. *Symmetries of Music: An Introduction to Semantics of Music*. Kodály Institute, Kecskemét, 1993.
- [118] Z. Les. Shape understanding system: understanding the thin object. *Computers & Graphics-UK*, 26(6):951–970, dec 2002.
- [119] G. Levi and U. Montanari. A grey-weighted skeleton. *Information Control*, 17:62–91, 1970.
- [120] M.S. Lew, T.S. Huang, and K. Wong. Learning and feature-selection in stereo matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(9):869–881, 1994.
- [121] F. Leymarie and M. D. Levine. Curvature morphology. In *Proc. of the SPIE Conf. on Intelligent Robots & Computer Vision VIII - Algorithms & Techniques*, 1989.
- [122] F. Leymarie and M. D. Levine. Skeletons from snakes. In V. Cantoni et al., editor, *Proc. of the 5th International Conference on Image Analysis and Processing*, pages 186–193, Positano-Italy, 1989.
- [123] M. Leyton. Symmetry-curvature duality. *CVGIP*, 38:327–341, 1987.

- [124] M. Leyton. A process grammar for shape. *Artificial Intelligence Journal*, 34(2):213–247, 1988.
- [125] M. Leyton. *Symmetry, Causality, Mind*. MIT Press, 1992.
- [126] T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, 1994.
- [127] M. Livingstone and D. Hubel. Segregation of form, color, movement, and depth - anatomy, physiology, and perception. *Science*, 240(4853):740–749, 1988.
- [128] S. Loncaric. A survey of shape analysis techniques. *Pattern Recognition*, 31(8):983–1001, 1998.
- [129] D. G. Lowe. Organization of smooth image curves at multiple scales. In *Second International Conference on Computer Vision, Tampa*, pages 558–567, 1988.
- [130] M. Luppe, L.F. Costa, and V.O. Roda. Parallel implementation of exact dilations and multi-scale skeletonization. *Real-Time Imaging*, in Press, 2003.
- [131] S. A. Mahmoud. Arabic character recognition using fourier descriptors and character contour encoding. *Pattern Recognition*, 27(6):815–824, 1994.
- [132] R. Malladi and J. A. Sethian. Flows under min/max curvature and mean curvature. *CVGIP: Graphical Model and Image Processing*, 58(2):127–141, 1996.
- [133] D. Marr. The computation of lightness by the primate retina. *Vision Research*, 14:1377–1388, 1974.
- [134] D. Marr. Early processing of visual information. *Phil. Trans. R. Soc. London*, B275:483–524, 1976.
- [135] D. Marr. *Vision: A computational investigation into the human representation and processing of visual information*. W. H. Freeman, 1982.
- [136] D. Marr and L. Vaina (Editor). *From the Retina to the Neocortex: Selected Papers of David Marr*. Springer Verlag, 1991.

- [137] D. Marr and E. Hildreth. Theory of edge detection. In *Proceedings of the Royal Society, London, B207*, pages 187–217, 1980.
- [138] D. Marr and H.K. Nishihara. Representation and recognition of the partial organisation of three-dimensional shapes. In *Proceedings of the Royal society London*, pages 269–294, 1978.
- [139] S. Marshall. Review of shape coding techniques. *Image and Vision Computing*, 7(4):281–294, 1989.
- [140] G. Matheron. *Examples of topological properties of skeletons*, chapter 11. Image Analysis and Mathematical Morphology, Volume 2: Theoretical Advances. Academic Press, London, 1988.
- [141] R. McClamrock. Marr’s three levels: a re-evaluation. *Minds and Machines*, 1:185–196, 1991.
- [142] G. Medioni and Y. Yasumoto. Corner detection and curve representation using cubic b-splines. *CVGIP*, 39(3):267–278, 1987.
- [143] S.N. Meshkat and C.M. Sakkas. Voronoi diagram for multiply-connected polygonal domains, ii. implementation and application. *IBM J. Res. Devel.*, 31(3):373–381, 1987.
- [144] E. Milios. Shape matching using curvature processes. *Computer Vision, Graphics and Image Processing*, 47:203–226, 1989.
- [145] R. Mohr and B. Triggs. Projective geometry for image analysis. A Tutorial given at International Symposium of Photogrammetry and Remote Sensing, Vienna, Jul 1996.
- [146] L. Moisan. Affine plane curve evolution: A fully consistent scheme. *IEEE Transactions on Image Processing*, 7(3):411–420, 1998.
- [147] F. Mokhtarian and A. Mackworth. A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 14:789–805, 1992.

- [148] F. Mokhtarian and S. Abbasi. Affine curvature scale space with affine length parametrisation. *Pattern Analysis & Applications*, 4:1–8, 2001.
- [149] F. Mokhtarian, S. Abbasi, and J. Kittler. Indexing an image database by shape content using curvature scale space. In *Digest of IEE Colloquium on Intelligent Image DataBases*, pages 4/1–4/6, 1996.
- [150] F. Mokhtarian, N. Khalili, and P. Yuen. Curvature computation on free-form 3-d meshes at multiple scales. *Computer Vision and Image Understanding*, 83:118–139, 2001.
- [151] F. Mokhtarian and A. Mackworth. Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):34–43, 1986.
- [152] F. Mokhtarian and R. Suomela. Robust image corner detection through curvature scale space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1376–1381, 1998.
- [153] U. Montanari. A method for obtaining skeletons using a quasi-euclidean distance. *Journal of the Association for Computing Machinery*, 15(4):600–624, Oct 1968.
- [154] J. Moreira. *Uma proposta de estruturação e integração de processamento de cores em sistemas artificiais de visão*. PhD thesis, Instituto de Física de São Carlos, 1999.
- [155] M.E. Mortenson. *Geometric Modeling*. John Wiley & Sons, 1985.
- [156] R. Nevatia and T.O. Binford. Description and recognition of curved objects. *Artificial Intelligence*, 8(1):77–98, 1977.
- [157] R.L. Ogniewicz. Skeleton-space: a multiscale shape description combining region and boundary information. In *Proceedings of Computer Vision and Pattern Recognition*, pages 746–751, 1994.
- [158] R.L. Ogniewicz and O. Kübler. Hierarchic voronoi skeletons. *Pattern Recognition*, 28(3):343–359, 1995.

- [159] L. O’Gorman. An analysis of feature detectability from curvature estimation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 235–240, 1988.
- [160] A. Okabe, B. Boots, K. Sugihara, and S.N. Chiu. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. John Wiley, Chichester, 2nd edition, 2000.
- [161] D.W. Paglieroni. Distance transforms - properties and machine vision applications. *CVGIP-Graphical Models and Image Processing*, 54(1):56–74, 1992.
- [162] S. K. Pal. Fuzzy skeletonization of an image. *Pattern Recognition Letters*, 10(1):17–23, 1989.
- [163] S.K. Pal and A. Rosenfeld. A fuzzy medial axis transformation based on fuzzy disks. *Pattern Recognition Letters*, 12:585–590, 1991.
- [164] S.K. Pal and L. Wang. Fuzzy medial axis transformation (fmat) - image representation, skeleton extraction and uncertainty management in computer vision. *Sadhana-Academy Proceedings in Engineering Sciences*, 18:251–277, 1993.
- [165] J.S. Park and J.H. Han. Contour motion estimation from image sequences using curvature information. *Pattern Recognition*, 31:31–39, 1998.
- [166] F. Pasan and C. Vuerli. Core-line tracing for fuzzy image subsets. *Pattern Recognition Letters*, 4:93–98, 1986.
- [167] T. Pavlidis. *Structural Pattern Recognition*. Springer-Verlag, 1977.
- [168] T. Pavlidis. Algorithms for shape analysis of contours and waveforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(4):301–312, 1980.
- [169] T. Pavlidis. *Algorithms for Graphics and Image Processing*. Computer Science Press, 1982.
- [170] S.C Pei and C.N Lin. The detection of dominant points on digital curves by scale-space filtering. *Pattern Recognition*, 25(11):1307–1314, nov 1992.

- [171] J. Perloff. *Tornado in a Junkyard: The Relentless Myth of Darwinism*. Refuge Books, 1999.
- [172] E. Persoon and K-S. Fu. Shape discrimination using fourier descriptors. *IEEE Transactions on Systems, Man, and Cybernetics*, 7(3):170–179, 1977.
- [173] A. Pikaz and I.H. Dinstein. Using simple decomposition for smoothing and feature point detection of noisy digital curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8):808–813, aug 1994.
- [174] C.H. Portezani. Investigação e modelagem da percepção de saliências em formas bidimensionais. Master's thesis, Instituto de Física de São Carlos, 1999.
- [175] F.P. Preparata and M.I. Shamos. *Computational Geometry: An Introduction*. Springer-Verlag, 2nd edition, 1990.
- [176] A. Rosenfeld. Compact figures in digital pictures. *IEEE Transactions on Systems Man and Cybernetics*, MC 4(2):221–223, 1974.
- [177] A. Rosenfeld and E. Johnston. Angle detection on digital curves. *IEEE Transactions on Computers*, 22:875–878, 1973.
- [178] A. Rosenfeld and A. C. Kak. *Digital Picture Processing*. Academic Press, 2nd edition, 1982.
- [179] A. Rosenfeld and J.L. Pfaltz. Sequential operations in digital picture processing. *Journal of the ACM*, 13(4), 1966.
- [180] A. Rosenfeld and J. S. Weszka. An improved method of angle detection on digital curves. *IEEE Transactions on Computers*, 24:940–941, 1975.
- [181] G. Sapiro and P.J. Giblin. Affine-invariant distances, envelopes and symmetry sets. *Geometriae Dedicata*, 71(3):237–261, jul 1998.
- [182] R.J. Schalkoff. *Digital Image Processing and Computer Vision*. John Wiley & Sons, Inc. Publishers, 1989.

- [183] K.E. Selkirk. *Pattern and place. An introduction to the mathematics of geography*. Cambridge University Press, 1982.
- [184] J. A. Sethian. *Level set methods: Evolving interfaces in geometry, fluid mechanics, computer vision, and materials science*. Cambridge University Press, 1996.
- [185] C. E. Shannon. The mathematical theory of communication. *Bell System Tech. Journal*, 27, 1948.
- [186] L. Shao and H. Zhou. Curve fitting with bézier cubics. *Graphical Models and Image Processing*, 58(3):223–232, 1996.
- [187] E.C. Sherbrooke, N.M. Patrikalakis, and F-E. Wolter. Differential and topological properties of medial axis transforms. *Graphical Models and Image Processing*, 58(6):574–592, 1996.
- [188] K. Siddiqi, A. Shokoufandeh, S. J. Dickinson, and S. W. Zucker. Shock graphs and shape matching. *International Journal of Computer Vision*, 35(1):13–32, 1999.
- [189] M. Sonka, V. Hlavac, , and R. Boyle. *Image Processing, Analysis and Machine Vision*. PWS Publishing, 1998.
- [190] V. Srinivasan and L.R. Nackman. Voronoi diagram for multiply-connected polygonal domains, i. algorithmi. *IBM J. Res. Devel.*, 31(3):361–372, 1987.
- [191] B.M. ter H. Romeny, L. Florack, J.J. Koenderink, and M.A. Viergever, editors. *Scale-Space Theory in Computer Vision, First International Conference, Scale-Space'97, Utrecht, The Netherlands, July 2-4, 1997, Proceedings*, volume 1252 of *Lecture Notes in Computer Science*. Springer, 1997.
- [192] S. M. Thomas and Y. T. Chan. A simple approach for the estimation of circular arc center and its radius. *Computer Vision Graphics and Image Processing*, 45, 1989.
- [193] O.D. Trier and A.K. Jain. Goal-directed evaluation of binarization methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(12):1191–1201, dec 1995.

- [194] D. M. Ushizima and L. F. Costa. Jumping spider (salticidae) ethology with computer aid. In *II Encontro de Aracnólogos do Cone Sul*, volume 2, Porto Alegre-RS-Brazil, 1999.
- [195] F. Vesely. *Computational Physics - An Introduction*. Kluwer, 2nd edition, 2001.
- [196] D.K. Washburn and D.W. Crowe. *Symmetries of Culture: Theory and Practice of Plane Pattern Analysis*. University of Washington Press, Seattle, 1988.
- [197] A. P. Witkin. Scale-space filtering. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1019–1022, 1983.
- [198] F-E. Wolter. Cut locus and medial axis in global shape interrogation and representation. Technical report, MIT Ocean Engineering Design Laboratory - revised version, 1993.
- [199] M. Worring and A.W.M. Smeulders. Digital curvature estimation. *CVGIP-Image Understanding*, 58(3):366–382, nov 1993.
- [200] M.W. Wright. *The extended Euclidean distance transform*. PhD thesis, University of Cambridge, 1995.
- [201] I.T. Young, J.E. Walker, and J.E. Bowie. An analysis technique for biological shape i. *Information and Control*, 25(4):357–370, 1974.
- [202] C.T. Zahn and R.Z. Roskies. Fourier descriptors for plane closed curves. *IEEE Transactions on Computers*, C-21(3):269–281, 1972.