Universidade de São Paulo – USP
Escola de Engenharia de São Carlos
Departamento de Engenharia Elétrica e Computação
Programa de Pós-Graduação Em Engenharia Elétrica

Antonio Henrique Figueira Louro

**A suavização Gaussiana como método de marcação
de características de fronteira entre regiões
homogêneas contrastantes**

São Carlos
2016

Antonio Henrique Figueira Louro

A suavização Gaussiana como método de marcação de características de fronteira entre regiões homogêneas contrastantes

Tese de doutorado apresentada à Escola de Engenharia Elétrica de São Carlos como parte dos requisitos para a obtenção de título de Doutor em Ciências, pelo Programa de Engenharia Elétrica.

Área de concentração: Processamento de Sinais e Instrumentação.

Orientador: Prof. Dr. Adilson Gonzaga

São Carlos
2016

Trata-se de uma versão corrigida da tese. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTA TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

F892a Figueira Louro, Antonio Henrique
A suavização Gaussiana como método de marcação de
características de fronteira entre regiões homogêneas
contrastantes / Antonio Henrique Figueira Louro;
orientador Adilson Gonzaga. São Carlos, 2016.

Tese (Doutorado) - Programa de Pós-Graduação em
Engenharia Elétrica e Área de Concentração em
Processamento de Sinais e Instrumentação -- Escola de
Engenharia de São Carlos da Universidade de São Paulo,
2016.

1. Difusão. 2. Curvatura. 3. Entropia. 4. Diferença
de Gaussianas. 5. Pontos Dominantes. 6. Campos
Receptivos. 7. Percepção de Formas. 8. Visão. I.
Título.

FOLHA DE JULGAMENTO

Candidato: Tecnólogo **ANTÔNIO HENRIQUE FIGUEIRA LOURO**.

Título da tese: "A suavização Gaussiana como método de marcação de características de fronteira entre regiões homogêneas contrastantes".

Data da defesa: 18/05/2016

Comissão Julgadora:

Resultado:

Prof. Associado **Adilson Gonzaga (Orientador)**
(Escola de Engenharia de São Carlos/EESC)

APROVADO

Prof. Associado **Evandro Luís Linhari Rodrigues**
(Escola de Engenharia de São Carlos/EESC)

APROVADO

Dr. **Paulo Estevão Cruvinel**
(EMBRAPA)

APROVADO

Prof. Dr. **Luiz Antônio Pereira Neves**
(Universidade Federal do Paraná/UFPR)

APROVADO

Prof. Dr. **Ricardo José Ferrari**
(Universidade Federal de São Carlos/UFSCar)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Associado **Luís Fernando Costa Alberto**

Presidente da Comissão de Pós-Graduação:
Prof. Associado **Paulo César Lima Segantine**

Este trabalho é humildemente dedicado à minha família.

Agradecimentos

Ao Prof. Adilson por me aceitar como seu aluno, por sua orientação, confiança, camaradagem e por todos os documentos que lhe fiz assinar. Foi um prazer estar ao seu comando. Sentirei saudades desses bons tempos.

Se entregar a um doutorado requer o abandono de parte de sua vida e poucos têm a sorte de ter uma esposa e sogra que cuidem de você e de toda essa parte, principalmente na educação e diversão dos filhos. Muito obrigado Priscila e Dona Lúcia. Também agradeço aos meus filhotes, que são muito compreensivos.

Aos meus pais Antonio e Angélica, e à minha tia Irene, por tudo.

Aos meus amigos Antonio Lemão e Alessandro Peixe, que sempre me apoiaram no que precisei aqui em São Carlos.

Ao Prof. Edson Rodrigues. Espero que este agradecimento chegue até aí, muito obrigado por me apresentar ao mundo acadêmico.

À Profa. Mabel Rodrigues.

Aos Profs. Marcelo Vieira, Homero Schiabel e Maria Stela.

Aos Profs. Evandro e Odemir pelas sugestões no trabalho de qualificação.

Aos Profs. Paulo Cruvinel e Luciano Fontoura por me apresentarem ao mundo da visão computacional.

Ao Prof. Gustavo Braga da UESC.

À PROP e à Reitoria da UESC

À Fundação de Amparo à Pesquisa do Estado da Bahia.

Resumo

Louro, A.H.F. A suavização Gaussiana como método de marcação de características de fronteira entre regiões homogêneas contrastantes. 210 p. Tese de Doutorado – Escola de Engenharia de São Carlos, Universidade de São Paulo.

Este trabalho mostra que a suavização Gaussiana pode exercer outra função além da filtração. Considerando-se imagens binárias, este processo pode funcionar como uma espécie de marcador, que modifica as feições das fronteiras entre duas regiões homogêneas contrastantes. Tais feições são pontos de concavidades, de convexidades ou de bordas em linha reta. Ou seja, toda a informação necessária para se caracterizar a forma bidimensional de uma região. A quantidade de suavização realizada em cada ponto depende da configuração preto/branco que compõe a vizinhança onde este se situa. Isto significa que cada ponto sofre uma quantidade particular de modificação, a qual reflete a interface local entre o objeto e o fundo. Então, para detectar tais feições, basta quantificar a suavização em cada ponto. No entanto, a discriminação pixel a pixel exige que a distribuição Gaussiana apresente boa localização, o que só acontece em escalas muito baixas ($\sigma \approx 0,5$). Assim, propõe-se uma distribuição construída a partir da soma de duas Gaussianas. Uma é bem estreita para garantir a boa localização e a outra possui abertura irrestrita para representar a escala desejada. Para confirmar a propriedade de marcação dessa distribuição, são propostos três detectores de *corners* de contorno, os quais são aplicados à detecção de pontos dominantes. O primeiro utiliza a entropia de Shannon para quantificar a suavização em cada ponto. O segundo utiliza as probabilidades de objeto e de fundo contidos na vizinhança observada. O terceiro utiliza a diferença entre Gaussianas (DoG) para determinar a quantidade suavizada, porém com a restrição de que uma das versões da imagem tenha suavização desprezível, para garantir a boa localização. Este trabalho se fundamenta na física da luz e na visão biológica. Os ótimos resultados apresentados sugerem que a detecção de curvaturas do sistema visual pode ocorrer na retina.

Palavras-chave: Difusão. Curvatura. Entropia. Diferença de Gaussianas. Pontos Dominantes. Campos Receptivos. Percepção de Formas. Visão.

Abstract

Louro, A.H.F. The Gaussian smoothing as a method for marking boundary features between contrasting homogeneous regions. 210 p. Ph.D. Thesis – São Carlos School of Engineering, University of São Paulo, 2016.

This work shows that the Gaussian smoothing can have additional function to filtration. Considering the binary images, this process can operate as a kind of marker that changes the features of the boundaries between two contrasting homogeneous regions. These features are points of concavities, convexities or straight edges, which are all the necessary information to characterize the two-dimensional shape of a region. The amount of smoothing performed at each point depends on the black/white configuration that composes the neighborhood where the point is located. This means that each point suffers a particular modification, which reflects the local interface between object and background. Thus, to detect such features, one must quantify the smoothing at each point. However, pixel-wise discrimination requires that the Gaussian distribution does not suffer flattening, which occurs in very low scales ($\sigma \approx 0.5$), only. Thus, it is proposed a distribution built from the sum of two Gaussians. One must be very narrow to ensure good localization, and the other is free to represent the desired scale. To confirm the property of marking, three boundary based *corner* detectors are proposed, which are applied to the detection of dominant points. The first uses the Shannon's entropy to quantify the smoothing at each point. The second uses the probabilities of object and background contained in the local neighborhood. The third uses the difference of Gaussians (DoG) to determine the amount of smoothing. This Work relies on the physics of light and biological vision. The presented results are good enough to suggest that the curvature detection, in visual system, occurs in the retina.

Keywords: Diffusion. Curvature. Entropy. Difference of Gaussians. Dominant Points. Receptive Fields. Shape Perception. Vision.

Lista de Figuras

Figura 1.1- Possível efeito da PSF sobre os detalhes do objeto	29
Figura 1.2- Relação ângulo-entropia	32
Figura 1.3- Ilusão face-vaso.	33
Figura 2.1 – O gato de Attneave.....	41
Figura 2.2 – Importância dos vértices para o reconhecimento de objetos.....	41
Figura 2.3 – Os caminhos corticais	43
Figura 2.4 – Quadrados aninhados	44
Figura 2.5 – Classes de métodos de detecção de <i>corners</i>	47
Figura 2.6 – Aproximação poligonal.....	52
Figura 2.7 – Conceitos de curvatura.....	53
Figura 2.8 – O <i>chain code</i>	55
Figura 2.9 – Região de suporte (RoS) do ângulo θ_{ik} :	56
Figura 3.1 – Relação entre os conceitos abordados nesta tese.	62
Figura 3.2 – Difusão isotrópica:	65
Figura 3.3 – Pintura com pontilhismo de George Seurat, "Un dimanche après-midi à l'Île de la Grande Jatte" (1884-1886).	68
Figura 3.4 – Entropia para duas possibilidades com probabilidades p e $(1-p)$	70
Figura 3.5 – Difração em uma abertura circular.....	73
Figura 3.6 – Simplificação da conexão entre fotorreceptores e uma célula bipolar.....	76
Figura 3.7 – Simplificação da conexão entre fotorreceptores e uma célula horizontal.....	76
Figura 3.8 – Campo receptivo <i>center-surround</i> para células bipolares:.....	77
Figura 3.9 – A diferença entre Gaussianas para uma imagem homogênea é zero.	78
Figura 3.10 – Detecção de bordas com DoG <i>off-center</i>	79
Figura 3.11 – Detecção de bordas com DoG on-center.....	79
Figura 3.12 – Ampliação de parte das bordas detectadas por DoG off-center e on-center.	80
Figura 3.13 – Contorno duplo na imagem entropia.....	81
Figura 4.1 – Relação ângulo-entropia: núcleo proposto versus núcleo Gaussiano	84
Figura 4.2 – Modificação de Gaussiana para obter a distribuição desejada.....	88
Figura 4.3 – Comparações entre os resultados das suavizações com SoG e com Gaussiana modificada.	88

Figura 4.4 – Ajuste de distribuição Gaussiana em janela quadrada.	90
Figura 4.5 – Suavizações de uma imagem binária com os núcleos da Tabela 4.1.	91
Figura 4.6 – Distribuição de fotorreceptores na retina.	92
Figura 4.7 – Núcleos de convolução gerados por SoG.	92
Figura 4.8 – Imagem original de uma cena com objetos distantes.	93
Figura 4.9 – Versão suavizada com o núcleo H143 ($\sigma=23,8333$).	93
Figura 4.10 – Versão suavizada com Gaussiana ($\sigma=2,8333$).	94
Figura 4.11 – Respostas dos campos receptivos <i>center-surround</i> de células ganglionares.	96
Figura 4.12 – Respostas das células ganglionares em relação ao tamanho da área excitada pelo estímulo luminoso.	97
Figura 4.13 – Relação complementar entre convexidades e concavidades.	98
Figura 4.14 – Comparação de detecção de vértices entre DoG e DoG_zero.	101
Figura 5.1 – Ângulos padrão em uma janela 3x3	108
Figura 5.2 – Influência da grade de amostragem.	108
Figura 5.4 – Imagem entropia gerada pelo EML na escala mais baixa de operação.	111
Figura 5.5 – Diferentes escalas de uma mesma configuração angular.	113
Figura 5.6 – O vértice de um retângulo observado através de seis janelas circulares.	115
Figura 5.7 – Variação de entropia do vértice observado na Figura 5.6.	115
Figura 5.8 – Evolução multiescala de convexidades: ponto dominante versus ruído.	117
Figura 5.9 – Evolução multiescala de ruídos convexos localizados em borda em linha reta.	118
Figura 5.10 – Evolução multiescala de concavidades: ponto dominante versus ruído.	119
Figura 5.11 – Interferência de regiões adjacentes no comportamento multiescala.	120
Figura 5.12 – Interferência de regiões adjacentes na comparação de candidatos.	122
Figura 5.13 – A sequência das comparações interfere no resultado.	126
Figura 5.14 – Saídas das fases 2 e 3 do EML.	127
Figura 5.15 – Fase-4: pontos dominantes eleitos.	127
Figura 6.1 – Comportamento multiescala dos ângulos padrão (sem correção).	131
Figura 6.2 – Comportamento multiescala dos ângulos padrão (linearizados).	131
Figura 6.3 – Relação ângulo – peso do objeto (preto).	132
Figura 6.4 – Evolução multiescala: entropia versus peso.	133
Figura 6.5 – Saída da quarta fase dos detectores EML e CLP.	136
Figura 7.1 – Resultado da diferença: “imagem suavizada ($\sigma=0,5$) – imagem original”	142
Figura 7.2 – Resultado da diferença: “imagem suavizada– imagem original” em seis escalas diferentes.	142

Figura 7.3 – Concavidade versus convexidade.	144
Figura 7.4 – Comportamento multiescala das respostas do DoG_zero para os ângulos padrão:	144
Figura 7.5 – Resposta do DoG_0 para a imagem da Figura 7.3.....	145
Figura 7.6 – Evolução multiescala: entropia versus DoG_zero.	148
Figura 7.7 – A propagação do primeiro extremo aplicada à Figura 7.6.b	150
Figura 7.8 – Resultados do algoritmo de avaliação multiescala.....	152
Figura 7.10 – Resultados do algoritmo de avaliação multiescala para as convexidades do fundo branco	155
Figura 7.11 – Pontos dominantes detectados pelo DoG_zero.	155
Figura 8.1 – Semelhança na detecção de vértices de convexidades e ruídos.	160
Figura 8.2 – O <i>ground-truth</i> das oito imagens.	166
Figura 8.3 – Os erros de detecção do EML.	166
Figura 8.4 – Os erros de detecção do CPL.	167
Figura 8.5 – Os erros de detecção do DoG_zero.	167
Figura 8.6 – O <i>ground-truth</i> da imagem F.	170
Figura 8.7 – Evolução multiescala do ponto X da Figura 8.6 sob o ponto de vista dos três detectores propostos.	171
Figura 8.8 – Ampliação de parte da imagem H.	171
Figura 8.9 – Evolução multiescala dos pontos A, B e Z da Figura 8.8.	172
Figura 8.10 – O <i>ground-truth</i> da imagem E.	172
Figura 9.1 – Contornos ilusórios (adaptado de Yang, Yue e Wu (2015)).	179
Figura 9.2 – Hipercolumnas do córtex visual V1 (adaptado de Ursino e Lara (2004)).	179
Figura 9.3 – A proximidade dos indutores fortalece a ilusão.	180
Figura 9.4 – A ativação dos ângulos no córtex visual V1:	181
Figura 9.5 – A ativação dos ângulos nas hipercolumnas:	181

Lista de Tabelas

Tabela 4.1 – Núcleos de convolução (SoG).	90
Tabela 5.1 – Ângulos padrão e suas entropias em 7 escalas diferentes.....	107
Tabela 5.2 – Fator de correção de entropia por ângulo e escala.....	114
Tabela 6.1 – Ângulos padrão e seus pesos em 7 escalas diferentes.	130
Tabela 6.2 – Fatores para a linearização de pesos.	132
Tabela 6.3 – Ângulos padrão ajustados pelo fator de linearização.	134
Tabela 7.1 – Ângulos padrão e suas DoG_zero para 7 escalas diferentes.....	147
Tabela 7.2 – Fator de correção de resposta DoG_zero por ângulo e escala.	148
Tabela 8.1 – Interseção das respostas dos três detectores	159
Tabela 8.2 – Interseção das respostas dos detectores EML e CLP.....	159
Tabela 8.3 – Interseção das respostas dos detectores EML e DoG_zero (continua).....	159
Tabela 8.4 – Interseção das respostas dos detectores CLP e DoG_zero	160
Tabela 8.5 – Resultados das comparações das respostas dos detectores propostos com o ground-truth.	162
Tabela 8.6 – Respostas do detector EML. (continua).....	162
Tabela 8.7 – Respostas do detector CLP. (continua).....	163
Tabela 8.8 – Respostas do detector DoG_zero. (continua)	164
Tabela 8.9 – Resultados das comparações das respostas de 13 detectores com o <i>ground-truth</i>	168

Lista de Siglas

2D	Bidimensional
CBIR	Recuperação de imagens baseada em conteúdo (Content-based Image Retrieval)
CLP	Computação Local de Pesos
CSS	Espaço-escala de curvaturas (Curvature Scale Space)
Diag	Diagonal
DoG	Diferença de Gaussianas
DoG_zero	Diferença de Gaussianas onde uma delas tem escala próximo à zero
EML	Entropia da Mistura Local
FN	Falso Negativo
FP	Falso Positivo
H/V	Horizontal/Vertical
LoG	Laplaciano da Gaussiana
PSF(FEP)	Função Espalhamento de Ponto
SoG	Soma de Gaussianas

Sumário

Capítulo 1 – Introdução	24
1.1 Considerações Iniciais	24
1.2 Objetivos	35
1.3 Contribuições	35
1.4 Organização do texto	36
Capítulo 2 – Revisão bibliográfica	39
2.1 Introdução	39
2.2 Aspectos psicofísicos	39
2.3 Aspectos biológicos	42
2.4 Detecção de características locais	45
2.4.1 Detecção de pontos dominantes	47
2.5 Conclusão do capítulo	57
Capítulo 3 – Fundamentação teórica	61
3.1 Introdução	61
3.2 Difusão	62
3.3 Entropia	66
3.4 Relações entre PSF e difusão	71
3.5 PSF neural, produção de entropia e diferença de Gaussianas	75
3.6 Conclusão do capítulo	82
Capítulo 4 – Núcleo de convolução proposto	83
4.1 Introdução	83
4.2 Desvantagens da suavização Gaussiana	83
4.3 O núcleo proposto para a computação de entropia	85
4.4 O núcleo proposto para a DoG	94
4.5 Conclusão do capítulo	102

Capítulo 5 – Proposta do detector EML	105
5.1 Introdução	105
5.2 Fase 1 – computação da entropia na menor escala	109
5.3 Fase 2 – computação multiescala dos candidatos	112
5.4 Fase 3 – análise da evolução dos candidatos através das escalas	112
5.5 Fase 4 – determinação dos pontos dominantes	124
5.6 Conclusão do capítulo	128
Capítulo 6 – Proposta do detector CLP	129
6.1 Introdução	129
6.2 Conclusão do capítulo	137
Capítulo 7 – Proposta do detector DoG_zero	141
7.1 Introdução	141
7.2 Funcionamento do DoG_zero	146
7.3 Conclusão do capítulo	156
Capítulo 8 – Resultados	157
8.1 Introdução	157
8.2 Comparando as saídas dos detectores propostos	157
8.3 Comparando as respostas dos detectores propostos com um <i>ground-truth</i>	161
8.4 Comparação entre os detectores propostos e outros disponíveis na literatura	165
8.5 Discussão	168
Capítulo 9 – Conclusões e trabalhos futuros	175
9.1 Conclusões	175
9.3 Trabalhos futuros	178
9.3 Publicações	182
Referências	185
ANEXO A – Imagens utilizadas	209

Capítulo 1 – Introdução

1.1 Considerações Iniciais

O tema desta tese está inserido no contexto da visão computacional. Uma disciplina que incorpora o conhecimento de diferentes áreas, com o intuito final de construir sistemas capazes de enxergar, seja no espectro visível ou em outras frequências.

Uma das grandes linhas de pesquisa em visão computacional é a análise de formas. Esta tese trata diretamente com formas bidimensionais, mais precisamente com as propriedades que podem auxiliar em sua percepção e distinção.

O conceito de forma tem sido de importância central na reflexão científica e filosófica desde o seu princípio (ALBERTAZZI, 1999). Em visão computacional e em visão natural há muita pesquisa dedicada à compreensão das formas. Em geral, o objetivo do primeiro está na detecção e reconhecimento de objetos, de seres vivos, de tumores, de lesões, de comportamento, de emoções, de gestos, de caracteres, ou de verificar se determinado produto possui a forma adequada para a sua comercialização.

Embora os objetos do mundo real sejam tridimensionais, suas projeções na retina (ou numa matriz de fotorreceptores artificiais) possuem, apenas, duas dimensões. Contudo, transportam informação suficiente para caracterizar a forma geral de um objeto.

Em visão computacional há um grande esforço dedicado à pesquisa em análise de formas bidimensionais. Um dos objetivos é buscar maneiras para simplificar a forma, sem distorcer o seu aspecto geral. A finalidade disto é descrever a forma usando-se, somente, os seus detalhes relevantes. As formas precisam ser representadas de maneira única e completa a fim de permitir a distinção entre objetos similares. Porém, a representação requer simplicidade, evitando detalhes desnecessários que possam dificultar os processos computacionais de comparação, armazenamento ou transmissão.

Não há definição precisa para o conceito genérico de forma. Num dicionário da língua portuguesa, esta palavra é definida como sendo os limites exteriores da matéria de que é constituído um corpo, e que conferem a este um feitio, uma configuração, um aspecto particular (AURÉLIO, 1999). Na literatura de análise de formas, o conceito de forma costuma ser definido como a informação que se mantém invariante às transformações geométricas de rotação, translação e variação de escala (DRYDEN; MARDIA, 1998). As imagens binárias,

ou silhuetas, se adaptam bem a essas definições e costumam ser a modalidade de imagens escolhida para o estudo da forma. Esta tese trabalha apenas com imagens binárias.

De acordo com a definição dada pelo dicionário, os limites exteriores se referem ao contorno do objeto, embutindo a ideia de quanto e em qual direção a matéria se expandiu, ou se contraiu para formar cada uma das saliências do objeto. Na segunda definição, a ideia de invariância à escala sugere que a forma geral do objeto é dada pelas grandes saliências, que se mantêm visíveis, mesmo que se aumente a distância em que o objeto é observado.

Uma saliência é uma dada região da imagem cujo aspecto a faz se sobressair em relação às regiões vizinhas. No contexto de formas bidimensionais, as saliências são protuberâncias e reentrâncias no contorno. Nos pontos do contorno onde se inicia uma saliência ocorre uma mudança de direção. Ao mudar a direção é criado um ângulo entre a direção atual e a direção anterior. Assim, a ideia de se procurar por mudanças de direção, ou ângulos ao longo do contorno é bastante útil para descrever a forma de um objeto. A variação na direção pode ser medida pela curvatura, quanto mais rápida é a variação na direção, maior é a magnitude da curvatura.

A maioria das técnicas envolvidas com a simplificação de formas lida com a busca por pontos de alta curvatura ao longo de contornos, por exemplo, em Mokhtarian e Mackworth (1986). Tais técnicas também levam em consideração o tamanho das saliências delimitadas pelos pontos de alta curvatura detectados.

As palavras convexidade e concavidade são mais adequadas para se referir às saliências geradas a partir de expansões ou contrações de matéria. O vértice de uma convexidade, ou de uma concavidade é um ponto do contorno onde ocorre uma mudança abrupta na direção. Tais pontos são conhecidos na literatura pertinente como *corners* de contorno. Caso a saliência observada tenha importância no aspecto global da forma, o *corner* de contorno pode receber a denominação de ponto dominante.

Pode-se dizer, então, que para se descrever a forma geral de um objeto, as técnicas de simplificação de formas buscam por vértices de convexidades e de concavidades relevantes do contorno.

Embora as ideias sugeridas pelas definições acima pareçam ser perfeitamente lógicas, não foram elas que inspiraram as técnicas de detecção de curvaturas em contornos. O marco no qual se baseia a maior parte dessas técnicas é o trabalho de Attneave (1954), onde são relatados experimentos psicofísicos, com a finalidade de comprovar certas leis da *Gestalt*¹ e

¹ *Gestalt* (“forma” em alemão) é uma linha da psicologia que se destina a estudar a percepção humana.

cujos resultados foram fundamentados pela teoria da informação. Esses experimentos mostraram que a incerteza sobre a direção do contorno é maior nos pontos onde a direção varia mais acentuadamente, pois causam uma ruptura de continuidade². Esses raros pontos de mudança abrupta na direção, sozinhos, são perfeitamente capazes de aproximar o aspecto geral da forma. Necessitando, apenas, que sejam unidos com segmentos de retas ou arcos.

A associação de incerteza com alta curvatura motivou a construção de muitos algoritmos de detecção de pontos dominantes fundamentados na análise de curvaturas. No entanto, o aspecto da incerteza recebeu muito menos importância na detecção de tais pontos. Os resultados apresentados nesta tese mostram que explorar o aspecto da incerteza é mais natural, mais simples e mais preciso do que as técnicas envolvidas com cálculos de curvatura.

Attneave utilizou a entropia de Shannon (SHANNON, 1948) para medir a incerteza na “adivinhação³” dos pontos do contorno (ATTNEAVE, 1954) e relacionou as mudanças de direção com valores altos de entropia. Isto significa que num segmento de alta curvatura é necessário mais informação para decidir a direção correta do contorno.

A entropia computada nas respostas desses experimentos está relacionada ao estado “consciente” do observador, pois decisões envolvem raciocínio. Contudo, antes da forma de um objeto se tornar consciente, a sua imagem sofre diferentes processamentos ao longo do caminho visual, incluindo a detecção de ângulos de contorno. “Poderia, então, o sistema visual utilizar algo semelhante à entropia para detectar curvaturas?”

As técnicas utilizadas para saber se uma saliência é importante, ou melhor, para verificar se a saliência sobrevive ao aumento de escala, são conhecidas como técnicas multiescala. Uma delas é o espaço-escala Gaussiano. Em Koenderink (1984) é feita uma equiparação entre o espaço-escala Gaussiano e a difusão do calor. Até então, o autor desta tese pensava em espaço-escala como uma iteração de suavizações, uma mera ferramenta de processamento de imagens. Isto desencadeou uma série de indagações, tais como: “é possível, na Natureza, ocorrer a difusão de uma imagem? O que é difundido nesse processo? Por que existe essa difusão? Quais as suas consequências na percepção das imagens?”

Percebeu-se que as questões sobre entropia e difusão mencionadas nos dois últimos parágrafos estão fortemente relacionadas. Uma das características de uma imagem binária é que as regiões de fundo e de objeto estão bem definidas, isto é, não há dúvidas sobre a qual

² Continuidade é uma das leis da Gestalt.

³ Nos experimentos de Attneave, apenas uma pequena região da imagem ficava visível ao observador. A tarefa deste era a de prever a cor, ou a direção do contorno das regiões adjacentes, a partir da porção visível da imagem.

região um determinado pixel possa pertencer. A suavização Gaussiana é uma instância de um processo de difusão. O efeito de uma suavização desse tipo, numa imagem binária, é o de transferir “partículas de cor” do objeto para o fundo e vice-versa. Esta transferência bidirecional, ou mistura de partículas, introduz incerteza nos pixels da fronteira entre o objeto e o fundo. Quanto mais acentuada for a suavização (aumentando-se a abertura da Gaussiana ou repetindo-se o processo de suavização), mais intenso se torna o grau de incerteza, bem como a sua abrangência, trazendo um número de pixels cada vez maior para o estado de incerteza. Este processo de suavização pode ser continuado até que todos os pixels da imagem apresentem o mesmo nível de cinza. A partir deste instante não há mais sentido em continuar a suavização, pois não ocorrerá qualquer modificação adicional. Diz-se que o sistema (imagem) entrou em seu estado de equilíbrio.

Portanto, onde há difusão, também pode haver o aumento de incerteza. Esta tem início a partir da fronteira, onde é mais intensa, e se propaga em duas frentes com sentidos opostos. Uma em direção ao interior do objeto e a outra em direção às partes do fundo mais afastadas da fronteira. A concentração da incerteza na fronteira está em harmonia com os resultados relatados em Attneave (1954), onde se diz que o contorno é a parte menos redundante da imagem, ou seja, onde se encontra mais variação, mais diferenças, mais contraste.

A difusão pode ser medida através do cálculo de entropia e, em alguns casos, pela entropia de configuração (*configurational entropy*) (LAMBERT, 2007), principalmente, quando se fala em aumento de desordem ou de baralhamento. A entropia de configuração não considera a dinâmica das partículas, apenas a distribuição espacial das mesmas e a sua equação é a mesma da entropia de Shannon.

Havendo relação entre difusão e incerteza, abre-se o caminho para verificar onde pode ocorrer difusão, ou suavização, num sistema de visão, bem como a possibilidade de inferir curvaturas ao longo de contornos, medindo-se a incerteza causada por tais processos.

A simples observação de uma cena através de um sistema óptico natural ou artificial introduz incerteza. A informação contida na cena e a informação obtida com o aparato óptico não são idênticas. A imagem sempre será uma suavização da cena real, independentemente da precisão do sistema utilizado. A causa dessa distorção se deve, principalmente, ao comportamento da luz ao interagir com a abertura do sistema (difração) e com o material das lentes (refração). O efeito de suavização resultante é a resposta impulsiva do sistema ou, equivalentemente, a função de espalhamento pontual (FEP ou PSF do inglês *point spread function*). Em geral, a abertura de um sistema óptico é circular, o que produz difração circular

e um efeito de suavização circularmente simétrica na imagem (COLE; JINADASA; BROWN, 2011; SMITH, 1997).

Um pixel de uma imagem representa uma região minúscula da cena em observação, que a partir de agora será chamada de região mínima. Hipoteticamente, esta região pode ser homogênea possuindo relevo e propriedades físicas idênticas em toda a sua extensão. Isto pode significar que a luz refletida por essa região mínima transporta informação de um único tipo. No caso da região mínima ser heterogênea, haverá informações distintas transportadas pela luz refletida.

Um pixel só pode ter um único valor, ele é o menor item de informação em uma imagem, não há como representar duas ou mais informações distintas simultaneamente num mesmo pixel. Para o caso de uma região mínima homogênea, o valor do pixel representa fielmente a informação transportada pelo reflexo desta região. No entanto, para que o valor de um pixel represente as informações refletidas por uma região mínima heterogênea, tais informações precisam ser combinadas. Considera-se, aqui nesta tese, que as regras de combinação (as ponderações) são ditadas pela PSF do sistema.

Aproximando-se a PSF com uma distribuição bidimensional circularmente simétrica e considerando-se uma cena hipotética composta por um objeto de cor preta num fundo branco, a aquisição da imagem poderá produzir os seguintes comportamentos:

- a) As regiões homogêneas da cena, isto é, o interior do objeto e o interior do fundo não sofrem qualquer modificação com a suavização da PSF;
- b) A suavização produzirá seus efeitos mais notáveis na fronteira entre o objeto e o fundo;
- c) Quando a PSF estiver centrada na fronteira, o efeito da suavização variará de acordo com as feições da fronteira, ou melhor, de acordo com as porções convexas e côncavas do objeto (ou do fundo).

Essas ideias ficam mais claras observando-se a Figura 1.1. As circunferências demarcam a janela de observação do sistema óptico, o qual é representado por um cilindro. Este possui a sua PSF, representada por uma Gaussiana. O sistema varre a cena, a qual é representada por uma estrela com superfície homogênea preta e um fundo homogêneo branco. Considera-se que a vizinhança demarcada por círculos (região mínima da cena) seja pequena o suficiente para ser representada por um único pixel de saída, o qual é um dos retângulos da grade de saída. Para cada configuração espacial de região mínima haverá um valor de saída próprio (pixel de saída). Se a região mínima for homogênea, isto é, composta apenas por

objeto preto, ou apenas por fundo branco, o valor de saída para essa região mínima será idêntico ao valor de entrada, não ocorrendo mistura alguma. Este é o caso dos círculos 1 e 2 desta figura. A combinação de informações (mistura) ocorre quando a região mínima da cena compreende, simultaneamente, informações de objeto e de fundo. A janela de observação centrada num ponto da fronteira faz com que o sistema produza um pixel cujo valor representa as proporções de objeto e de fundo da região mínima. Devido à distribuição de pesos da PSF ser bidimensional e circularmente simétrica, o valor do pixel de saída está fortemente atrelado à configuração espacial da região mínima. É muito importante observar que ao se deslocar a janela de observação ao longo do contorno, as proporções de objeto e fundo só podem variar de uma única maneira, que é através da variação de setores circulares.

Figura 1.1- Possível efeito da PSF sobre os detalhes do objeto. Os círculos numerados correspondem às configurações de entrada do sistema óptico, o qual produz um pixel com uma tonalidade correspondente ao detalhe observado. A forma de sino ao lado do cilindro é a representação da PSF do sistema.

Os ângulos dos setores circulares definem o aspecto da interface entre o objeto e o fundo, isto é, se a fronteira forma uma borda em linha reta, uma concavidade ou uma convexidade. Isto pode ser observado nos círculos 3, 4 e 5 da Figura 1.1. Em outras palavras, está sendo sugerido que a suavização circularmente simétrica pode fornecer informações que possibilitem inferir os ângulos do contorno de um objeto.

Anteriormente foi mencionado que a suavização causa incerteza, produzindo dúvidas se um pixel pertence ao objeto ou ao fundo. Se a região mínima for homogênea, tem-se a

certeza de que ela pertence exclusivamente ao objeto (ou exclusivamente ao fundo). Neste caso não há produção de incerteza em seus pixels de saída. O que era um “ponto” do objeto na cena continua sendo um “ponto” do objeto na imagem. Porém, nos “pontos” sobre a fronteira, a formação de suas imagens produz incerteza nos pixels de saída, a qual varia com o tipo de interface que o objeto forma com o fundo. A proporção de objeto capturada pela janela de observação é o fator que determina o grau de incerteza produzido.

A incerteza, então, pode ser usada para quantificar a suavização produzida, em imagens binárias e consequentemente caracterizar o ângulo de cada ponto do contorno. Porém, a incerteza não é a única maneira de se medir o efeito da suavização. Pode-se, por exemplo, usar somente o peso estatístico do objeto contido na região mínima e associá-lo ao ângulo do contorno. Quanto menor for o peso, mais agudo é o ângulo. Na verdade este procedimento é uma simplificação do cálculo da entropia de Shannon.

Outra maneira de quantificar a suavização produzida é através da diferença de Gaussianas (*Difference of Gaussians* - DoG), por exemplo, a diferença entre a versão suavizada e a versão original. Esta operação informa quanto a imagem foi modificada pela suavização em cada um de seus pixels, isto é, a subtração informa onde houve mistura e qual foi a sua intensidade.

Quando a abertura do núcleo de suavização é estreita, o resultado da subtração é zero em todos os pixels, exceto nos pixels da fronteira. Vale notar que as células ganglionares da retina são modeladas como DoG e, em geral, são chamadas de detectores de bordas.

Nesta tese foram construídos três detectores de *corners* de contorno, cada um usando uma das ideias mencionadas acima: entropia, probabilidades e DoG.

Para detectar *corners* de contorno válidos (pontos dominantes) é necessário verificar o tamanho da saliência do contorno e não apenas o seu ângulo. Para isso é necessário utilizar alguma técnica multiescala. Nesta tese optou-se em desenvolver um método multiescala próprio, cujas características principais são:

- a) Apenas os candidatos a pontos dominantes, detectados em baixa escala, são observados em escalas mais altas;
- b) As escalas são discretas e de quantidade limitada;

- c) O núcleo de convolução⁴ é uma modificação da distribuição Gaussiana para evitar o efeito de achatamento⁵.

A característica do item *c* é essencial para se obter uma relação ordenada entre a abertura do ângulo e a sua entropia (incerteza). Os itens *a* e *b* da Figura 1.2 demonstram essa relação ordenada. A entropia decresce monotonicamente com o aumento da abertura do setor. Vale salientar que o setor de interesse pode ser o de cor preta ou o de cor branca. Isto produz dois pontos de vista diferentes, porém complementares. Independentemente da cor escolhida, a diminuição da abertura do setor produzirá um pixel de saída com maior entropia. Isto ocorre devido à distribuição de pesos adotada, cujo peso central é muito alto em relação aos outros, e sempre pertencerá à região que está sendo observada (objeto ou fundo). Ou seja, quanto menor for a abertura do setor, mais próximo se estará da configuração de equiprobabilidade, a qual produz máxima entropia.

O achatamento da Gaussiana torna impossível criar uma relação confiável entre o ângulo do setor e a sua entropia. Especificamente, o achatamento permite que dois ou mais ângulos diferentes tenham o mesmo valor de entropia, invalidando a detecção angular. A solução para este problema foi o uso de mistura de Gaussianas para gerar a distribuição desejada, a qual possui o centro sempre íngreme, independentemente da escala utilizada. A mistura de Gaussianas foi manipulada para que a distribuição resultante possua uma relação equiprovável entre o peso do centro (0.5) e o peso total da periferia (0.5), o que confere entropia máxima (valor igual a 1) para as configurações mostradas no item *c* da Figura 1.2.

Num sistema de visão natural, o estágio seguinte ao da aquisição de imagens é formado por um banco de filtros, o qual é composto por células retinais de diferentes tipos. As células do tipo bipolar e ganglionar se destacam por possuírem campos receptivos compostos por regiões de sensibilidade antagonistas, dispostas de forma concêntrica. Este tipo de campo receptivo é chamado de *center-surround*. Um modelo para este tipo de campo utiliza a diferença entre Gaussianas (DoG) (RODIECK, 1965; ENROTH-CUGELL; ROBSON, 1966). Pode-se dizer que o campo receptivo de uma célula é a janela de observação da célula. No entanto, esta janela não recebe luz, mas sinais elétricos provenientes de outras células. De

⁴ Núcleo de convolução é a janela de observação que é deslocada ponto a ponto ao longo da cena. Matematicamente, é uma distribuição estatística cujos pesos são combinados aos pontos da cena através de produto interno.

⁵ A distribuição Gaussiana possui uma forma de sino. Aumentar a escala significa alargar a base desse sino, o que implica num ajuste interno dos pesos que diminui a altura do sino. O achatamento da altura deixa os pesos com valores muito próximos, que ao serem multiplicados pelos pontos da cena (ou imagem) deixam as regiões com valores muito parecidos, dificultando a sua distinção – perda de localização espacial.

qualquer maneira, permanece idêntica a ideia de se capturar um conjunto de valores de entrada e transformá-lo num único valor de saída. Neste caso, não se usa a PSF para ponderar a combinação dos sinais, mas a “sensibilidade” do campo receptivo, que também é distribuída de maneira circularmente simétrica (ou quase), por isso o uso de Gaussianas em sua modelagem.

As células ganglionares formam o último estágio de processamento dentro da retina. Seus axônios formam o nervo óptico, por onde a informação processada na retina é enviada ao cérebro. As ganglionares respondem aos estímulos com trens de pulsos, cuja frequência (ou outra característica) está associada ao estímulo de entrada. Isto é, existe um vocabulário, um código neural, que representa diferentes instâncias de alguma informação presente no estímulo. Acredita-se que uma dessas informações é o contraste, ou melhor, a borda entre duas regiões contrastantes, que é uma característica possível de ser extraída pela arquitetura *center-surround*.

Figura 1.2- Relação ângulo-entropia. a) a incerteza em relação à abertura do setor circular, a qual pode ser vista tanto pelo lado da porção preta quanto da porção branca; b) valores de entropia para os ângulos do item a; c) as configurações de entropia máxima (1) para o núcleo de convolução proposto; d) configurações de entropia mínima (0), não há incerteza sobre o ponto pertencer ao objeto ou ao fundo.

A forma é um conceito de alto nível, o qual, necessariamente, deriva do processamento de contraste que ocorre na retina. Nos estágios iniciais não existe a ideia de forma, apenas regiões homogêneas contrastantes. A forma a ser gerada em níveis mais altos do cérebro pode ser produzida pela região escura, ou pela região clara da cena. A escolha dependerá da estrutura que for mais semelhante aos itens armazenados na memória do

sistema. Há casos em que ambas as regiões podem ter significado para o observador, causando um chaveamento mental entre as duas formas, isto pode ser percebido na Figura 1.3.

Da retina ao cérebro existem caminhos específicos para conduzir o código neural que descreve as regiões clara e escura. Assim, sinais que representam ambas as regiões são enviados, em paralelo, ao córtex visual para decidir a classe de cada região. Os caminhos que levam essas informações são formados por duas variações de um mesmo tipo de célula ganglionar. Um dos caminhos é formado pelas células *on-center off-surround*, que são sensíveis a objetos claros em fundo escuro. O outro caminho é composto pelas células *off-center on-surround* sensíveis a objetos escuros em fundo claro (SODERQUIST, 2002).

Voltando à Figura 1.2, no item a, é possível encontrar essa ideia de processamento simultâneo para objetos pretos e brancos (a produção de incerteza pode ser medida separadamente para os dois tons). Uma semelhança ainda mais sugestiva está nos itens c e d da mesma figura. As configurações de entropia máxima e mínima coincidem com as respostas máximas e mínimas das células ganglionares. As células *on-center off-surround* apresentam sua resposta máxima quando se ilumina apenas o centro. As células *off-center on-surround* apresentam a sua resposta máxima quando apenas a periferia é iluminada. Os dois tipos de célula quando completamente iluminados, ou completamente obscurecidos, apresentam resposta mínima.

Essas semelhanças sugerem a possibilidade de que na retina exista o mecanismo necessário para inferir as curvaturas ao longo de contornos. A proposta de um detector de *corners* baseado em um modelo semelhante (DoG) ao das células ganglionares corrobora essa possibilidade.

Figura 1.3- Ilusão face-vaso. Adaptado de Vases and Faces (2016)

Supondo-se que existam células ganglionares que gerem código neural para representar “valores de curvatura”, o sistema visual precisará enfrentar um novo problema: a natureza isotrópica dos campos receptivos torna o valor detectado invariante à rotação. Isto é, têm-se as coordenadas espaciais do ponto, o “valor” do seu ângulo, porém não se sabe a orientação deste. O ângulo detectado pode estar em qualquer posição ao redor do círculo. Sem conhecer a orientação, não é possível reconstruir a forma do objeto observado.

Nesta tese, em trabalhos futuros, são apresentados os rudimentos de um modelo que sugere uma solução para a falta de uma convenção que informe onde se situa o ângulo 0° e qual o sentido para o seu crescimento. Para isso, é apresentada a hipótese de que a orientação de um ângulo pode ser obtida alinhando-o, simultaneamente, com um ângulo imediatamente anterior e com um ângulo imediatamente posterior, da mesma maneira como ocorre com os ângulos internos de figuras geométricas planas. Propõe-se que o mecanismo de alinhamento encontra-se, em sua maior parte, no córtex visual primário (V1) composto por estruturas conhecidas como hipercolunas. A criação da hipótese foi motivada pelo efeito conhecido por ilusão de Kanizsa (KANIZSA, 1955), na qual o alinhamento de ângulos faz surgir uma figura. Considera-se, aqui, que tal ilusão é o reflexo do funcionamento desse mecanismo de alinhamento angular.

A Figura 1.4 mostra o surgimento de uma figura a partir do alinhamento angular. No item a, apenas os vértices das convexidades do objeto são utilizados. No item b são considerados os vértices das convexidades e concavidades do objeto. No item c são utilizados os vértices das convexidades do objeto e das convexidades do fundo. Este se baseia na existência dos canais *on-center* e *off-center*, o que permite reduzir a representação angular para a faixa de 0° a 180° no código neural. Ao se observar as Figuras 1.1 e 1.2, percebe-se que um vértice de convexidade do objeto também pode ser o vértice de uma concavidade do fundo, e vice-versa. A restrição da faixa angular para 180° reduz a redundância e consequentemente reduz o consumo de energia metabólica, evitando que recursos do sistema sejam utilizados desnecessariamente. Além disso, a representação de um vocabulário menor torna o código menos complexo.

Figura 1.4 – Ilusões de Kanizsa. a) a ilusão é gerada apenas por vértices de convexidade do objeto; b) a ilusão é gerada por vértices de convexidade e concavidade do objeto e c) a ilusão é gerada por vértices de convexidades do objeto e do fundo.

Em resumo, o trabalho desenvolvido nesta tese lida com a detecção de *corners* de contorno, que é um processamento fundamental para fornecer subsídios para tarefas de visão computacional de nível mais alto, por exemplo, o reconhecimento de objetos. Os métodos de detecção desenvolvidos foram baseados no estudo da física da luz e nos resultados das pesquisas em visão biológica. Em essência, a principal hipótese é de que o efeito de suavização isotrópica embute a informação necessária para se inferir as curvaturas do contorno da imagem bidimensional de um objeto.

1.2 Objetivos

O objetivo principal desta pesquisa é comprovar que a suavização do tipo Gaussiana é essencial para distinguir os detalhes fronteiros entre duas regiões homogêneas contrastantes, funcionando como uma espécie de marcador, similar aos marcadores biológicos utilizados para realçar as características de alguma célula. Para comprovar essa hipótese são propostos três detectores de pontos dominantes fundamentados na mesma ideia: a suavização Gaussiana.

Como objetivo secundário é sugerido, com base nos ótimos resultados obtidos, que a detecção de curvaturas pode ser realizada na retina.

1.3 Contribuições

São propostos três métodos inovadores para a detecção de *corners* de contorno, principalmente por se basearem num conceito muito simples e de uso corriqueiro em visão

computacional e processamento de imagens, que é a suavização de imagens com núcleos de convolução circularmente simétricos.

A utilização da soma de Gaussianas, ou mistura de Gaussianas em escala, também é inovadora no sentido de utilizá-la para evitar o efeito de achatamento associado ao aumento de escala. Da mesma forma, é inovadora a utilização da produção de entropia máxima como um guia para obter a distribuição ideal a partir dessa soma de Gaussianas. A entropia máxima numa configuração circularmente simétrica é obtida quando os pesos central e periférico são equiprováveis.

A associação da produção de entropia com a diferença de Gaussianas é uma contribuição muito interessante, pois permitiu idealizar campos receptivos do tipo *center-surround* como mecanismos que obedecem à produção máxima de entropia. Em outras palavras, sugere-se que o peso probabilístico do centro da distribuição seja igual, ou maior que a totalização dos pesos de sua periferia. Esta ideia permitiu tornar o modelo DoG em um detector de curvaturas.

Assim, esta tese também contribui com novas ideias para a visão natural, indicando a possibilidade de ocorrer detecção de curvaturas na retina, o que poderia explicar, por exemplo, a ilusão de Kanizsa como o reflexo de um mecanismo de alinhamento angular situado no córtex, necessário para reconstruir a forma a partir dos ângulos detectados.

1.4 Organização do texto

O presente trabalho é composto por mais oito capítulos:

O capítulo 2 apresenta uma revisão bibliográfica contendo as principais ideias usadas na detecção de pontos dominantes nos últimos quarenta anos, incluindo aspectos psicofísicos e biológicos.

O capítulo 3 apresenta os fundamentos básicos a respeito de difusão, difração, entropia da termodinâmica e entropia da informação, com as possíveis relações entre elas, bem como aspectos básicos de visão biológica.

O capítulo 4 apresenta a proposta do núcleo de convolução, bem como todos os aspectos biológicos que motivaram a sua construção.

Os capítulos 5, 6 e 7 respectivamente descrevem os métodos propostos para a detecção de pontos dominantes baseado em entropia, computação local de pesos e diferença de Gaussianas.

O capítulo 8 apresenta os resultados com suas respectivas discussões.

O capítulo 9 apresenta as conclusões, trabalhos futuros e produção bibliográfica.

Capítulo 2 – Revisão bibliográfica

2.1 Introdução

O interesse em descobrir quais são os aspectos que mais influenciam na percepção de formas é bastante antigo, há quase mil anos Alhazen, o “pai da óptica moderna”, declarou que as concavidades e convexidades ao longo do contorno de um objeto desempenham um papel de grande importância na percepção de sua forma (NORMAN; PHILLIPS; ROSS, 2001).

Pode-se dizer que a necessidade em compreender o funcionamento da percepção de formas guiou a pesquisa científica através de três caminhos distintos: experimentos psicofísicos, investigações biológicas e visão computacional.

O objetivo deste capítulo é apresentar uma revisão bibliográfica a respeito de pontos dominantes, sua fundamentação psicofísica, os possíveis mecanismos biológicos mencionados na literatura e algumas das várias implementações em visão computacional, desenvolvidas a partir da década de 1970.

2.2 Aspectos psicofísicos

A Psicofísica é a primeira e mais antiga disciplina do campo da psicologia experimental, estuda a relação entre os estímulos físicos (entrada) e a experiência sensorial (saída). Por exemplo, uma determinada luminosidade produz uma determinada sensação de claridade (BERNARDINO, 2015). É uma abordagem não invasiva para estudar os processos sensoriais, comparando modelos teóricos com a experiência visual das pessoas (BRUCE; GREEN; GEORGESON, 2003).

Em Attneave (1954) foram conduzidos experimentos psicofísicos sobre a percepção visual, em especial sobre a capacidade humana em prever as direções em um contorno. Foi constatado que a informação a respeito de uma forma bidimensional se concentra em seu contorno, porém não de maneira uniforme. Os pontos do contorno que concentram a maior parte da informação são aqueles onde a sua direção muda abruptamente.

De acordo com a lei da Gestalt de “boa continuidade”, os elementos arranjados numa reta ou curva são percebidos como mais relacionados entre si do que aqueles que não estão nessas configurações, permitindo reduzir a incerteza e destacar grupos correlacionados em meio a uma grande quantidade de elementos aleatórios (WAGEMANS et al., 2012). Isto é,

permite a previsão das direções em um contorno. Aquilo que pode ser previsto é considerado redundante, com baixo teor informativo e consequentemente não necessita ser representado. O fator que limita tais previsões é a surpresa causada por uma mudança abrupta na direção, tornando as posições de incerteza (ou surpresa) como as mais informativas de um contorno. Visto que, matematicamente, uma mudança de direção corresponde a uma alta curvatura, tornou-se corriqueiro o fato de que os pontos de alta curvatura concentram a maior parte da informação a respeito de uma forma bidimensional.

O experimento relatado acima teve seus resultados analisados sob o ponto de vista da teoria da informação de Shannon (SHANNON, 1948). Aliás, esse experimento de Attneave é uma adaptação, para contornos, do experimento de Shannon sobre a capacidade humana em prever a próxima letra de uma palavra. Shannon escolhia um livro aleatoriamente em sua estante, o qual era aberto em uma página qualquer, da qual era escolhida aleatoriamente uma palavra. Esta era parcialmente coberta com o dedo e era indagado ao observador (a esposa de Shannon) qual seria(m) a(s) letra(s) oculta(s) (GLEICK, 2011).

A ideia de Attneave em fundamentar seus resultados com a teoria da informação produziu um aumento de confiabilidade na Gestalt que já estava a ponto de ser “sepultada”, pelo menos na América do Norte. Seu trabalho (ATTNEAVE, 1954) tornou-se altamente referenciado em visão computacional e percepção visual. A Figura 2.1 apresenta o gato de Attneave, o qual é uma aproximação poligonal a partir dos 38 pontos de maior curvatura do contorno da imagem de um gato.

Os vértices (*corners*) são características importantes para o sistema visual humano, remove-los das imagens reduz drasticamente o desempenho da tarefa de reconhecimento, o que não ocorre se for removida grande parte das informações de borda (BIEDERMANN, 1987). Na Figura 2.2 são confrontadas imagens de objetos que tiveram parte de seus contornos removidos. Na coluna central (b), as regiões removidas não incluem vértices, o que não modifica significativamente a capacidade humana em reconhecer os objetos. Na coluna da direita (c), as regiões removidas incluem os vértices dos contornos, causando um aumento de incerteza no reconhecimento dos objetos. Pode-se ter a falsa impressão de que o reconhecimento baseado em (c) é tão fácil quanto em (b). Mostrando apenas o item (c), peça para que alguém reconheça os objetos. Repita o teste mostrando o item (b).

Figura 2.1 – O gato de Attneave. Adaptada de Attneave (1954).

Figura 2.2 – Importância dos vértices para o reconhecimento de objetos. a) contorno intacto; b) remoção de partes do contorno não incluindo vértices; c) remoção de vértices. Figura adaptada de Biedermann (1987).

Em Norman, Phillips e Ross (2001) é reportado um experimento onde os participantes deveriam marcar em cada uma das doze imagens apresentadas (sombas de batatas doces) os dez pontos mais salientes. A análise dos resultados indicou que os participantes marcaram consistentemente as regiões de maior curvatura absoluta, confirmando os resultados em Attneave (1954).

Em De winter e Wagemans (2008) foi realizado um experimento com 161 participantes e 260 contornos de objetos do dia a dia. Os resultados indicaram a importância dos extremos de curvatura.

Em Wolfe, Yee e Friedman-Hill (1992) são apresentados vários experimentos associados à busca visual. Os resultados suportam a hipótese de que a curvatura se comporta como uma característica (feature) básica nesses processos de busca, sugerindo que em algum lugar no sistema visual deva existir um mecanismo específico para a detecção de curvaturas.

2.3 Aspectos biológicos

Não se sabe ao certo onde estão localizados os mecanismos de detecção de curvatura, nem se estes existem de forma explícita. Há uma tendência em acreditar que as informações de curvatura sejam extraídas a partir das informações de orientação. Uma parte da literatura tende a considerar as células *endstopped* (ou hipercomplexas) da área V2 do córtex visual como detectores de curvatura. Em Dobbins, Zucker, e Cynader (1987, 1989) foi proposto que as células *endstopped* serviriam como base para a estimação de curvaturas. Em Versavel, Orban e Lagae (1990) propuseram a existência de três tipos de células relacionadas com o sistema de detecção de curvatura derivada da orientação, sendo que o conjunto de células sensível a altas curvaturas seria formado por células *endstopped*, as quais são sintonizadas a orientação e ao tamanho do estímulo. Em Rodríguez-Sánchez e Tsotsos (2011, 2012) é apresentado um modelo computacional onde as células *endstopped* desempenham um papel crítico na seletividade de formas via computação de curvaturas.

Outra parte considerável da literatura relata a sensibilidade que os neurônios da área V4 têm em relação às curvaturas. Alguns dos trabalhos envolvidos nessas investigações são encontrados em: Gallant, Braun e Van Essen (1993); Gallant et al. (1996); Pasupathy e Connor (1999, 2001); Habak et al. (2004); Oliveira, Costa e Roque (2005); Connor, Brincat e Pasupathy (2007); Yue et al. (2014); Rodríguez-Sánchez, Neumann e Piater (2014).

A Figura 2.3 resume a localização das áreas do córtex visual envolvidas no reconhecimento de objetos, nas quais pode ocorrer a detecção de curvaturas. As investigações

fisiológicas sugerem a existência de uma bifurcação no processamento de objetos. O caminho dorsal é responsável em produzir sensações de orientação espacial, de percepção de profundidade, de localização, de direção, de movimento e de velocidade de um objeto no espaço (CORTICAL PATHWAYS, 2015). O caminho ventral é especializado em reconstruir e identificar os objetos a partir de suas características.

Figura 2.3 – Os caminhos corticais. Adaptado de Bachatene; Bharmauria e Molotchnikoff (2012).

Parece não haver muito interesse em investigar se ocorre detecção de curvaturas na retina, talvez exista a certeza de que ali a detecção não seja factível. Até onde se pesquisou, só foram encontradas quatro referências (três dos mesmos autores) em que se menciona a possibilidade de detecção de *corners*, de ângulos ou de curvaturas nas células da retina. No entanto, é mostrado nesta tese que a estrutura *center-surround* dos campos receptivos da retina fornece o mecanismo fundamental para se extrair informações angulares, sugerindo que as células retinais possam codificar valores angulares de convexidades e concavidades do contorno de objetos.

Em Lettvin et al.(1959) foram investigadas as informações no nervo óptico de um sapo, dentre elas encontrou-se informações sobre convexidades. Em Troncoso, Macknik e Martinez-conde (2005, 2009) e Troncoso et al. (2007) foram conduzidos experimentos psicofísicos com base na ilusão de Vasarely, mostrando que os *corners* são mais salientes que

bordas em linha reta e que quanto mais agudo é o ângulo de um *corner*, maior é a sua saliência. Os autores, então, sugerem a possibilidade de que os campos receptivos da retina (do tipo *center-surround*) tenham evoluído para extrair a informação angular a partir do contraste. A Figura 2.4 apresenta uma ilusão de quadrados aninhados “do tipo Vasarely”. O importante é perceber a formação (ilusão) de linhas diagonais com intensidade do brilho destacada. Este realce na intensidade é o indicativo de que os *corners* se sobressaem em relação às bordas em linha reta. Quanto mais agudo for o ângulo, mais destacada se torna a intensidade. Pode ser interessante observar que há certa semelhança entre este tipo de ilusão e o resultado de transformadas como a *grassfire transform* e a *medial axis transform*, que geram os esqueletos das imagens (*skeletonization*). Alguns detalhes a respeito dessas transformadas podem ser encontrados em Costa e Cesar (2001).

Figura 2.4 – Quadrados aninhados. Figura adaptada de Troncoso, Macknik e Martinez-conde (2005).

2.4 Detecção de características locais

Em aplicações comuns de visão computacional, um estágio fundamental, após o tratamento da imagem, é a extração de informações que a representem de uma maneira apropriada para a realização de uma tarefa específica. Em geral, essas informações ou padrões são chamados de características ou feições da imagem (*image features*). Alguns tipos de características são bordas, contornos, *corners*, junções e *blobs*. Ou seja, em vez de se usar a imagem como um todo se usa um vetor de características para representá-la.

A literatura sobre detecção de características locais é muito vasta. As investigações a esse respeito iniciaram-se a partir dos resultados dos experimentos relatados em Attneave (1954), os quais indicaram a importância dos pontos de alta curvatura. Assim, inicialmente, a ideia era buscar por características locais ao longo de contornos de imagens, as quais estavam associadas a mudanças abruptas na direção local de um contorno, isto é, onde se formavam esquinas ou cantos. Os algoritmos que realizam tal busca são chamados de detectores de *corners*. Devido a esse aspecto angular, a definição comum para *corner* é a interseção entre duas bordas (PATEL; PANCHAL, 2014). Nesta categoria de detecção surgiu o termo ponto dominante, que se refere ao *corner* mais importante de uma dada região do contorno.

Uma desvantagem da detecção de *corners* em contornos é a necessidade prévia de segmentação da imagem, ou de detecção de bordas. Para tornar essa etapa desnecessária, as investigações evoluíram para a busca de *corners* diretamente em imagens em tons de cinza. As características importantes são posições na imagem onde ocorrem mudanças significativas de sinal. Para detectá-las não é necessário localizar fronteiras entre regiões, como mencionado acima, a busca por altas curvaturas é feita diretamente no gradiente da imagem. A ideia básica é computar as diferenças de brilho em torno de uma pequena localidade da imagem. Se as diferenças forem irrelevantes em todas as direções, pode significar que a localidade é interna a uma região homogênea da imagem. Se as diferenças forem grandes numa única direção e irrelevantes nas outras, pode significar que a localidade se encontra numa borda. Se as diferenças forem grandes em quase todas as direções, pode significar que a localidade é um *corner*. No entanto, nem sempre as grandes variações ao redor de uma localidade caracterizam um *corner*. Tais variações podem estar indicando a existência de um *blob*, uma pequena região que se destaca do seu entorno, por exemplo, uma mancha clara em meio a uma região escura, ou vice-versa. Então, os resultados produzidos por esses detectores podem indicar a presença de *corners*, *blobs* e bordas. Devido a essa variedade, os algoritmos receberam o nome de detectores de pontos de interesse, ou de pontos interessantes. Porém, a

tradição fez com que continuassem a ser chamados de detectores de *corners*. Um exemplo típico é o detector Harris (HARRIS; STEPHENS, 1988).

Outra classe de detectores de características tem o interesse em detectar pequenas regiões distintas da imagem, isto é, a detecção de *blobs* com a subsequente análise de seu comportamento interno, por exemplo, a direção principal de sua textura. Uma ideia fundamental é obter características que se mantenham invariantes a diferentes tipos de modificações que venham a ser produzidas na imagem, por exemplo, mudanças de escala e deformações geométricas. Essas regiões são chamadas de pontos chave (*keypoints*). Em Tuytelaars e Mikolajczyk (2007) são avaliados diferentes algoritmos envolvidos nesse tipo de detecção. Um exemplo típico desta categoria de detectores é o SIFT, descrito em Lowe (2004).

Percebe-se, então, que há uma nomenclatura um tanto confusa. Por exemplo, um detector de *corners* pode ser usado para detectar *blobs*, os quais podem ser pontos interessantes, ou pontos chave. A confusão pode aumentar devido à proximidade dos significados linguísticos dos termos pontos interessantes, pontos chave e pontos dominantes. Todos transportam a ideia sobre a importância de um ponto. Embora a ideia de ponto só exista matematicamente, este se torna ainda mais inadequado para se referir aos *blobs*, principalmente quando funcionam como pontos chave, pois são tratados explicitamente como regiões.

Apesar da nomenclatura confusa, em geral, a detecção de cada tipo de característica local tem detectores e finalidades específicas. A detecção de *corners* em contornos, ou em silhuetas e a subsequente escolha dos pontos dominantes estão diretamente relacionadas com análise e percepção de formas, que compõe o assunto de interesse desta tese.

Revisões e avaliações sobre técnicas de detecção de *corners* em contornos (ou curvas planas), conhecidos por *boundary based corner detectors*, podem ser encontradas em: Teng, Najmus Sadat e Lu (2015); Awrangjeb (2008); Awrangjeb, Lu e Frase (2010, 2012); Awrangjeb e Lu (2013); Kahaki, Nordin e Ashtari (2014); Mokhtarian e Mohanna (2006); Tuytelaars e Mikolajczyk (2007);

Uma excelente revisão sobre detecção de características locais, incluindo *keypoints*, está em Tuytelaars e Mikolajczyk (2007). Revisões e avaliações sobre técnicas de detecção de pontos de interesse são encontradas em: Chen et al. (2009); Patel e Panchal (2014); Tissainayagam e Suter (2004); Rockett (2003); Dutta, Chatterjib e Kar (2008); Zukal e Cika (2012); Zheng, Wang e Khwang Teoh (1999); Schmid, Mohr e Bauckhage (1998, 2000);

Existem, também, revisões de detectores com o propósito específico em verificar qual é o mais apropriado para uma dada aplicação. Em Li, Liu e Gong (2013) são comparados alguns detectores de pontos interessantes e detectores de *corners* para aplicações de reconhecimento de linguagens de sinais. Em Lee, Lee e Xu (2014) é proposto um método geométrico para detecção de veias em imagens de retina, o qual é comparado com detectores de pontos de interesse e detectores baseados em *keypoints*. Em Cumani e Guiducci (2009) são avaliados alguns detectores baseados em *keypoints* para uso em hodometria visual, isto é, o acúmulo de movimentos relativos, estimados a partir de características da cena enquanto um veículo, provido de visão estéreo, se movimenta. Em Zukal et al. (2013) são avaliados detectores de pontos interessantes para uso em imagens médicas de ultrassom. Em Nowicki e Skrzypczynski (2014) são avaliados detectores baseados em *keypoints* para navegação visual usando *smartphones*.

Devido à existência de grande quantidade de métodos de detecção de *corners*, é necessário organizá-los em grupos para que se tenha uma noção geral dos trabalhos já desenvolvidos. A figura 2.5 mostra uma possível organização.

Figura 2.5 – Classes de métodos de detecção de *corners*. Adaptado de (AWRANGJEB, 2008).

2.4.1 Detecção de pontos dominantes

Nesta tese, o objetivo é determinar quais são os pontos mais salientes ao longo de contornos de formas bidimensionais (de silhuetas). A saliência dos pontos de um contorno foi

quantificada por seu conteúdo informativo, com base na teoria da informação, em Attneave (1954), Singh (2015), Feldman e Singh (2005). Outros trabalhos em psicofísica, que utilizaram mais participantes e mais imagens, reforçaram a ideia de que os pontos mais salientes ao longo de um contorno são os pontos de mais alta curvatura (DE WINTER; WAGEMANS, 2008; NORMAN; PHILLIPS; ROSS, 2001).

Uma revisão de detectores de *corners* (*corners* de contorno e pontos interessantes) em imagens digitais listou cento e catorze algoritmos diferentes, desenvolvidos no período de 1977 a 2006, dos quais cinquenta e quatro se dedicam à detecção de *corners* em contornos realizada no domínio espacial (DUTTA; KAR; CHATTERJI, 2008). Algumas publicações mais recentes, que incluem detecção de pontos dominantes, são encontradas em (AWRANGJEB; LU; FRASER, 2010; PARVEZ; MAHMOUD, 2010; ZHU; WANG; LIU, 2010; NGUYEN; RENNESSON, 2011; PEDROSA; BARCELOS; BATISTA, 2011; PAULA et al., 2011; PRASAD, 2012; TAHAEI et al., 2012; POTOČNIK, 2012; BACKES; BRUNO, 2013; SARFRAZ; SWATI, 2013; LIU; LI, 2013; LABOURE; AZEMA; FOURNEL, 2014).

A grande maioria dos algoritmos estima as curvaturas ao longo de contornos (ou curvas planas). O maior problema da computação da curvatura é a sua sensibilidade a ruídos e pequenas perturbações (KERAUTRET; LACHAUD; NAEGEL, 2008). Estimar curvaturas, bem como localizar precisamente a sua posição espacial são tarefas muito difíceis, mesmo em contornos perfeitamente digitalizados (UTCKE, 2003).

O termo ponto dominante foi utilizado pela primeira vez em Rosenberg (1972). Um ponto dominante é um marco onde a natureza da curva se modifica significativamente (ponto de alta curvatura). Pontos dominantes representam características importantes de um objeto e desempenham um papel importante na percepção de formas (ASADA; BRADY, 1986). Em Guru, Dinesh e Nagabhushan (2004) é declarado que a informação sobre a forma está concentrada nos pontos dominantes, os quais são primitivas descritivas para a representação de formas e interpretação de imagens. Os termos ponto dominante, *corner*, ponto crítico e *break point* são usados na literatura como tendo o mesmo significado (ZHANG; ZHAO, 1997). Um ponto dominante é um *corner* especial, é o vértice da estrutura mais importante de uma dada vizinhança. A importância ou saliência de tal estrutura é dada pela combinação de sua curvatura e seu tamanho. Este último é identificado pela quantidade de escalas em que a estrutura se mantém proeminente.

De acordo com Awrangjeb, Lu e Fraser (2010) os pontos dominantes apresentam algumas vantagens em relação aos pontos interessantes, são elas: a) são visualmente distinguíveis (intuitivos) e mais robustos que os pontos interessantes; b) a quantidade de

pontos dominantes é muito menor que a de pontos interessantes; c) os pontos dominantes podem ser ordenados de acordo com sua importância (saliência); d) a quantidade de pontos dominantes pode ser controlada por limiar. A principal desvantagem na detecção de *corners* em contornos é a necessidade de segmentação prévia, que aumenta o tempo de processamento e pode gerar formas distorcidas.

Ao se refletir sobre a questão apresentada em Mandelbrot (1967), percebe-se que o problema de se medir uma zona costeira é o mesmo de se medir o contorno de um objeto, ambos podem apresentar quantidades imensas de protuberâncias e reentrâncias, limitadas, apenas, pela capacidade tecnológica em observá-las. Isto pode conferir à forma uma complexidade infinita, tornando impossível (ou quase) encontrar dois objetos que sejam idênticos em seus mínimos detalhes. Em aplicações de reconhecimento ou classificação de formas, duas ou mais formas costumam ser comparadas. Se cada detalhe de contorno for utilizado na comparação, concluir-se-á que todas as formas são distintas. Isto faz da detecção de pontos dominantes um problema não trivial, pois os ângulos gerados por ruído ou por detalhes insignificantes têm que ser distinguidos daqueles que representam mudanças significativas na direção das curvas (LIU; SRINATH, 1990). A detecção de falsos positivos e falsos negativos, descrita em vários artigos, é um indicativo da dificuldade enfrentada pelos detectores ao realizar essa tarefa. A ideia, então, é a de se evitar a descrição desses pequenos detalhes, isto é, tornar a forma o mais suave possível sem distorcê-la significativamente. Para isso, tornou-se imprescindível a utilização de esquemas multiescala embutidos nos detectores.

Os *corners* são invariantes a translação, rotação e mudanças de escala, por isso são largamente usados em representação e análise de imagens. A detecção de pontos dominantes é uma operação de baixo nível, cuja saída pode alimentar aplicações de nível superior. Alguns exemplos são: casamento de imagens (SMITH et al., 1998; VINCENT; LAGANIRE, 2005), decomposição de curvas digitais (ABE et al., 1993; MARJI; KLETTE; SIY, 2004; SARFRAZ, 2008), aproximação poligonal ou segmentação linear por partes (SARFRAZ; ASIM; MASOOD, 2004; MARCONDES; COSTA, 1995), visão estéreo (LI, 2007; VINCENT; LAGANIRE, 2001), recuperação de imagens baseada em conteúdo (CBIR) (TORRES; FALCÃO, 2007) e rastreamento de movimento (DRESCHLER; NAGEL, 1982; WANG; BRADY, 1995). Um dos principais usos de pontos dominantes é a simplificação da análise de formas, reduzindo a quantidade de dados a serem processados e ao mesmo tempo preservando as informações importantes sobre o objeto (LIU; SRINATH, 1990). Outros indicativos da importância da detecção de *corners* em contornos são:

- Considera-se que a extração de contornos seja uma das primeiras tarefas realizadas pelo sistema visual de recém-nascidos. Em Salapatek e Kessen (1966) foi observado que os movimentos dos olhos de crianças com poucos dias de idade se concentram sobre contornos ou *corners*.
- Uma quantidade de teorias sobre reconhecimento de objetos sugere que os seres humanos codificam as formas em termos de suas partes constituintes, as quais, por sua vez, coincidem com os valores máximos das convexidades encontradas em contornos (HAUSHOFER et al., 2008).

Basicamente existem duas abordagens diferentes para a detecção de *corners* ao longo de contornos. A primeira é conhecida como aproximação poligonal, cujo propósito é ajustar segmentos de reta às diferentes direções de um contorno. Para isso é definido um limiar para a distância máxima entre uma região do contorno e o segmento de reta que o aproxima. Exemplos são encontrados em Ramer (1972); Douglas e Peucker (1973); Pavlidis e Horowitz (1974); Sklansky e Gonzalez (1980); Wall e Danielsson (1984); Rosin e West (1989); Ray e Ray (1991); Aoyama e Kawagoe (1991); West e Rosin (1991). Nesta abordagem, os *corners* (*breaking points*) são os vértices do polígono ajustado ao contorno.

Na segunda abordagem, os *corners* surgem em consequência da estimação de curvatura (ou o seu equivalente) ao longo do contorno, os quais, eventualmente, podem ser usados como *breaking points* em aplicações de aproximação poligonal. Dentre o elevado número de algoritmos propostos para detecção de pontos dominantes baseada em curvatura, pode-se citar Rosenfeld e Johnston (1973); Rosenfeld e Weska (1975); Freeman e Davis (1977); Asada e Brady (1986); Mokhtarian e Mackworth (1986); The e Chin (1989); Rattarangsi e Chin (1992); He e Yung (2004).

As duas abordagens mencionadas podem se tornar interdependentes em alguns métodos, obscurecendo a fronteira entre elas. Sendo assim, é válido incluir as seguintes observações:

- a) É possível detectar pontos dominantes através de aproximação poligonal. Em virtude da quantidade de ruído de quantização presente em contornos digitais, a aproximação poligonal pode ser desejável, pois é mais confiável detectar retas do que *corners*. Referências clássicas são encontradas em Ramer (1972); Pavlidis e Horowitz (1974).
- b) É possível realizar uma aproximação poligonal após o processo de detecção de *corners*. Isto é, os *corners* servirão como *breaking points* para gerar um contorno simplificado, o qual pode ser utilizado em tarefas de nível mais alto. A compreensão tácita nesta

abordagem é a seguinte: visto que os pontos dominantes concentram a maior parte da informação existente na curva, é de se esperar que a união destes por segmentos de reta represente a curva de maneira desejável. Um exemplo de aplicação é encontrado em Garrido, Perez e Garca-Silvente (1998), onde as aproximações poligonais foram usadas para iniciar modelos deformáveis.

- c) É possível realizar uma detecção de *corners* intermediários (via detecção de curvatura), chamados de *anchor points*, e então aplicar alguma técnica de aproximação poligonal para encontrar os pontos dominantes. As técnicas que usam essa abordagem são chamadas de híbridas (ANSARI; DELP, 1991; ANSARI; HUANG, 1991; WU; WANG, 1993);

A Figura 2.6 mostra a aproximação poligonal realizada pelos algoritmos desenvolvidos em Ramer (1972) e em Pavlidis e Horowitz (1974), as quais são comparadas a uma aproximação poligonal realizada após a detecção de *corners*. As três técnicas apresentam um problema em comum, que é a análise feita em uma única escala, tornando a detecção de *corners* dependente da escala de observação.

Ao contrário das outras duas, a aproximação baseada em curvatura não sofre modificações significativas com a rotação da imagem. Ocorre, apenas, o deslocamento de alguns pontos, porém a causa está no reajuste do contorno à grade de amostragem e não na detecção dos *corners*.

Os principais problemas com a detecção baseada em aproximação poligonal são: a falta de invariância à escala, a dependência da escolha dos *breaking points* de entrada e a sensibilidade à rotação da imagem. Revisões e comparações de técnicas de aproximação poligonal são encontradas em The e Chin (1989) e em Rosin (1997).

Podem existir outros algoritmos diferentes das abordagens descritas acima. Um exemplo é tratar o conjunto de pontos, que aproxima o contorno, como se fosse um grafo e aplicar as propriedades pertinentes. Em Backer e Bruno (2013) é empregada a propriedade de *vertex betweenness* para alcançar a melhor aproximação poligonal.

De acordo com Awrangjeb (2008), em geral, as técnicas de detecção de pontos dominante baseadas em curvaturas são compostas de cinco estágios:

- Detecção de bordas e extração do contorno a partir da imagem em tons de cinza;
- Suavização do contorno;
- Estimação das curvaturas;

- Determinação dos pontos dominantes e
- Rastreamento dos pontos dominantes da alta escala em direção à escala mais baixa, a fim de melhorar a localização dos mesmos (válido somente para multiescala).

Figura 2.6 – Aproximação poligonal. a) contorno original, b) contorno original com rotação, c) contorno original aumentado. Na segunda linha (d,e,f) a aproximação poligonal é realizada após a detecção de curvatura. A terceira linha (g,h,i) apresenta a aproximação poligonal descrita em Ramer (1972), e na quarta linha (j,k,l) a aproximação é feita com o algoritmo *split-and-merge* (PAVLIDIS; HOROWITZ, 1974). As setas apontam para as distorções nas aproximações.

Discutir o estágio de segmentação ou detecção de bordas está fora do escopo desta tese. Porém é digno de nota que muitos dos detectores de pontos dominantes utilizam o detector de bordas de Canny.

Além do problema da escala, que é inerente a qualquer processo de detecção de *corners*, as técnicas que se baseiam em análise de curvatura necessitam enfrentar a falta de

definição para curvatura de curvas digitais. Os processos de amostragem e quantização envolvidos na digitalização requerem que a imagem de um objeto seja encaixada numa grade não isotrópica. Isto faz com que as informações exatas sobre o objeto sejam perdidas. Portanto, a curvatura não pode ser calculada com exatidão, apenas estimada.

Em geometria diferencial existem três conceitos inter-relacionados para definir curvatura de curvas planas. Os conceitos são baseados no recíproco do raio de círculos osculadores, variação do ângulo que a tangente forma com o eixo x (Figura 2.7), e uma equação criada a partir da parametrização da curva em relação ao comprimento de arco (s), a qual considera a variação do vetor tangente unitário em cada ponto da curva (equação 2.1). Esta equação tem sido bastante usada nos algoritmos de detecção de curvaturas, como os apresentados em Mokhtarian e Mackworth (1986) e Marcondes e Costa (1995).

As técnicas de detecção de pontos dominantes que estimam diretamente a curvatura não apresentam bons resultados (WORRING; SMEULDERS, 1993). Exemplos são encontrados em Rosenfeld e Johnston (1973); Rosenfeld e Weska (1975); Freeman e Davis (1977); Asada e Brady (1986); Mokhtarian e Mackworth (1986); Ansari e Huang (1991); Chetverikov e Szabo (2003); Sarfraz, Rasheed e Muzaffar (2005). No entanto, a técnica desenvolvida em Mokhtarian e Mackworth (1986) e usada em Mokhtarian, Abbasi e Kittler (1996) foi considerada como a melhor escolha para o MPEG-7 (MARTINEZ; KOENEN; PEREIRA, 2002; MARTINEZ; 2002b). Porém, em Zhang e Lu (2003) são apresentadas algumas falhas nos testes comparativos.

Figura 2.7 – Conceitos de curvatura: a) círculos osculadores - a curvatura em $T1$ é $k1=1/R1$ em $T2$ é $k2=1/R2$, com sinal inverso ao de $k1$; b) variação angular da tangente.

$$K(s) = \frac{\dot{x}(s) \cdot \ddot{y}(s) - \ddot{x}(s) \cdot \dot{y}(s)}{(\dot{x}(s)^2 + \dot{y}(s)^2)^{3/2}} \quad (2.1)$$

O ambiente hostil criado pela grade de amostragem faz com que os pesquisadores busquem por soluções diferentes. O Algoritmo descrito em Marcondes e Costa (1995) aplica a transformada de Fourier no contorno parametrizado, e então computa a curvatura utilizando a equação 2.1 associada às propriedades de derivada desse domínio. Outros pesquisadores fazem uso de técnicas que computam indiretamente a curvatura. Deve-se ressaltar que algumas dessas técnicas se originaram na abordagem de aproximação poligonal, a qual, de certa forma, estimava indiretamente a curvatura. Em Awrangjeb (2008) as técnicas indiretas são agrupadas em:

- a) Técnicas que utilizam a vizinhança à esquerda e à direita de cada ponto – é escolhida uma quantidade k de pontos vizinhos de cada lado do ponto do qual se deseja estimar a curvatura. Diferentes maneiras foram propostas para analisar essa vizinhança. Por exemplo, em Wu (2003) foi analisada a energia de dobramento dos braços à esquerda e à direita de cada ponto. Chama-se de braço a reta que une o ponto em análise ao ponto extremo de cada vizinhança.
- b) Técnicas que utilizam distância arco-corda – são computadas as distâncias do arco (pontos do contorno) a uma corda (braço). Exemplos são encontrados em Han e Poston (2001); Marji, Klette e Siy (2004) e Awrangjeb e Lu (2008).
- c) Técnicas que utilizam ajuste de curvas - em vez de aproximação poligonal, o ajuste é feito de maneira mais suave através de *splines*. Exemplos de técnicas baseadas em *spline* estão em Medioni e Yasumoto (1987) e em Park e Lee (2007).

Muitos dos algoritmos de detecção de pontos dominantes têm como entrada um contorno codificado em *chain code*. Em termos gerais essa codificação é uma forma de representação de percurso, indicando o sentido que se deve seguir ao longo do contorno. A Figura 2.8 compara o *chain code* às direções de uma rosa dos ventos e mostra um exemplo de contorno codificado por essa técnica. Em Liu e Srinath (1990) são comparadas algumas técnicas que usam codificação *chain code* antes de realizarem a detecção de *corners*.

Figura 2.8 – O *chain code*: o seu equivalente em rosa-dos-ventos e um contorno codificado em *chain code*.

Os primeiros algoritmos de detecção de pontos dominantes (baseados em curvatura) tinham como preocupação principal a precisão na estimação da curvatura. No entanto, constatou-se que o tamanho do detalhe do contorno onde está localizada a curvatura detectada, também, é de grande importância. O tamanho desse detalhe é comparável ao tamanho da região de suporte, isto é, ao comprimento dos segmentos de reta utilizados para computar os ângulos ao longo do contorno. Esses segmentos de reta, chamados de braços, se interceptam no ponto onde se deseja computar um ângulo. A Figura 2.9 esclarece esta ideia.

Os braços podem ter comprimentos de mesmo tamanho (simétricos) ou serem assimétricos; podem ter comprimento fixo ou adaptativo. Se os braços forem de comprimento fixo, a escolha de comprimentos longos acarretará na detecção dos detalhes maiores, enquanto os menores passarão despercebidos, assim como os ruídos de digitalização. Se a escolha for por braços muito curtos, todos os detalhes serão detectados, inclusive os ruídos.

A ideia de braço de comprimento fixo é equivalente à ideia de se observar a cena em uma escala fixa. Métodos que utilizam comprimentos adaptativos, ou esquemas multiescala, têm como objetivo a obtenção de imunidade a ruídos e ao mesmo tempo serem capazes de detectar todos os detalhes importantes do contorno. O primeiro trabalho a identificar a necessidade de se detectar ângulos em diferentes escalas foi o de Larry Davis em Davis (1977), o qual é a provável semente do espaço escala de curvaturas (CSS).

Figura 2.9 – Região de suporte (RoS) do ângulo θ_{ik} : Neste exemplo, os segmentos de reta (braços) são simétricos com tamanho de k pontos.

Uma categoria especial de algoritmos de detecção de *corners* aplica o conceito de representação espaço-escala para lidar com ruídos e detalhes de diferentes tamanhos. O espaço-escala de curvaturas é construído através da convolução de um contorno (1D) com uma família de funções Gaussianas (1D), cujo desvio padrão (σ) representa a escala. Os ruídos e detalhes do contorno com tamanho inferior ao da escala (σ) de filtragem são totalmente dissolvidos ao longo dos *pixels* vizinhos, deixando o contorno mais suave. A curvatura é computada em cada nível do espaço-escala e através do exame dos cruzamentos por zero, as convexidade e concavidades do contorno são encontradas. Cada escala de suavização possui o seu próprio conjunto de *corners* detectados. O espaço escala usado desta maneira gerou uma classe de detectores conhecidos por *curvature scale space* (CSS). Um dos algoritmos que recebeu maior destaque na literatura foi o apresentado em Mokhtarian e Mackworth (1986). Exemplos de esquemas de espaço-escala de curvatura são encontrados em Asada e Brady (1986); Rattarangsi e Chin (1992); Pei e Lin (1992); He e Yung (2004); Garrido, Perez e Garca-Silvente (1998); Mokhtarian e Mackworth (1992); Mokhtarian (1995); Marcondes e Costa (1995); Mokhtarian e Abbasi (2001); Zhang et al. (2007); Zhong e Liao (2007); Awrangjeb, Lu e Murshed (2007).

A deformação no contorno causada pela suavização Gaussiana fez com que se buscasse por outros métodos multiescala. O espaço escala morfológico, cuja modificação de escala é causada pelo tamanho do elemento estruturante, tem a característica de manter o contorno com sua forma original através das escalas. Técnicas de detecção utilizando o espaço-escala morfológico são encontradas em Leymarie e Levine (1988); Jalba, Wilkinson e Roerdink (2006).

Outro esquema multiescala é a transformada *wavelet*. As técnicas de detecção de *corners* que o utilizam são encontradas em Lee, Sun e Chen (1995); Antoine et al. (1997); Quddus e Fahmy (1999); Quddus e Gabbouj (2002); Gao et al. (2007); Paula Jr et al. (2011).

Além dos detectores de *corners* pertencentes aos grupos já apresentados, existem vários outros que possuem características próprias e não possuem uma classe definida. Em Zhu e Chirlian (1995) são usados algoritmos não lineares; em Tsai, Hou e Su (1999) os autovalores de matrizes de covariância são associados à ideia de curvatura. Existem, também, os detectores que se baseiam em mecanismos dos sistemas naturais de visão, por exemplo, em Hansen e Neumann (2002); Xu et al. (2010).

2.5 Conclusão do capítulo

Neste capítulo foi abordado o tema central desta tese, que é a detecção de pontos dominantes. Foram apresentadas revisões bibliográficas sobre os aspectos psicofísicos, fisiológicos e sobre algoritmos computacionais de detecção de *corners*/pontos dominantes, salientando-se alguns dos problemas enfrentados pelos detectores, como a estimação da curvatura em grade retangular e a observação em escala fixa. Foram mencionadas as técnicas utilizadas para estimar a curvatura bem como os esquemas multiescala adotados.

A detecção de pontos dominantes é uma área ativa de pesquisa com dezenas de métodos publicados, porém não foi proposto, ainda, um detector capaz de funcionar bem em todas as condições de ruído e deformações geométricas, tampouco conciliar velocidade de processamento e complexidade dos esquemas multiescala. Dentre os vários métodos existentes, os que mais se destacam são os baseados em CSS (espaço-escala de curvaturas).

As estrelas e as galáxias se apagaram e morreram, o espaço tornou-se negro após dez trilhões de anos de atividade.

Um a um, o Homem fundiu-se ao AC, cada corpo físico perdendo a sua identidade mental, acontecimento que era de alguma forma, benéfico.

A última mente humana parou antes da fusão, olhando para o espaço vazio a não ser pelos restos de uma estrela negra e um punhado de matéria extremamente rarefeita, agitada aleatoriamente pelo calor que aos poucos se dissipava, em direção ao zero absoluto.

O Homem disse, “AC, este é o fim? Não há como reverter este caos? Não pode ser feito?”.

O AC disse, “ainda não há dados suficientes para uma resposta significativa”.

ISAAC ASIMOV

What we call the past is built on bits.

What we call reality, arises in the last analysis from the posing of yes-no questions.

All things physical are information-theoretic in origin, and this is a participatory universe. The whole universe is thus seen as a computer—a cosmic information-processing machine.

Tomorrow, we will have learned to understand and express all of physics in the language of information. Every it—every particle, every field of force, even the space-time continuum itself—derives its function, its meaning, its very existence ... from bits.

JOHN ARCHIBALD WHEELER

The more energy, the faster the bits flip. Earth, air, fire, and water in the end are all made of energy, but the different forms they take are determined by information. To do anything requires energy. To specify what is done requires information.

SETH LLOYD

Capítulo 3 – Fundamentação teórica

3.1 Introdução

A finalidade deste capítulo é introduzir alguns conceitos fundamentais tratados por esta tese, não de forma a reproduzir o conteúdo dos livros texto disponíveis, mas para associá-lo às ideias propostas neste trabalho.

Este capítulo está organizado de tal forma a representar a sequência de procedimentos usados para tentar responder, pelo menos, parcialmente às principais questões que guiaram esta pesquisa, as quais estão citadas abaixo:

- A ideia de Attneave que relaciona curvaturas e incerteza pode ser transportada para os estados iniciais da visão e usada na percepção de formas?
- É possível, na Natureza, ocorrer difusão de imagens?
 - O que é difundido nesse processo?
 - Por que existe essa difusão?
 - Quais as suas consequências na percepção de formas?

As principais publicações que fizeram surgir tais questões foram os trabalhos onde o espaço-escala Gaussiano é equiparado à difusão (KOENDERINK, 1984) e onde a incerteza é associada aos ângulos de contorno (ATTNEAVE, 1954). Esses artigos foram estudados com o objetivo de se entender os efeitos resultantes da aplicação do cálculo de entropia em imagens binárias.

A divulgação informal do conceito de entropia é, às vezes, rodeada de mistério e pode provocar sentimentos de desolação e terror, principalmente se o leitor for uma criança. Por exemplo, em (ASIMOV, 1975) a entropia é associada à extinção do universo. Que entidade seria essa, capaz de estar presente em um fenômeno tão assombroso e ao mesmo tempo em resultados de processamento de imagens?

Num dado experimento, observou-se que ao filtrar uma imagem binária com filtros Gaussianos, a entropia se manifesta mais intensamente ao longo dos contornos, apresentando valores extremos (máximos e mínimos) nos pontos onde o contorno muda de direção. Isto fez perceber a existência de três elementos que se relacionam: entropia, filtragem Gaussiana e detecção de *corners*.

A finalidade da Figura 3.1 é de mostrar quais os assuntos que são estudados neste capítulo, bem como indicar a entropia como o elo entre a difusão e a detecção de *corners* de contorno. Por um lado a entropia pode ser vista como uma medida de difusão e pelo outro pode ser usada para medir a incerteza a respeito da direção da fronteira num ponto, isto é, a curvatura de um segmento do contorno. Os itens destacados por cor são os tópicos abordados neste capítulo. No lado esquerdo da Figura 3.1 e no mesmo nível estão os itens que se relacionam com a difusão. A função de espalhamento de ponto (PSF) foi mencionada em Koenderink (1984) como um núcleo de convolução para obter difusão, o que serviu como indicação de onde poderia ocorrer, naturalmente, a difusão de uma imagem. Os campos receptivos da retina também são vistos como fonte natural de difusão de imagens (ROMENY; FLORAK, 2000). O estudo da entropia tem como finalidade entender os seus conceitos do ponto de vista da difusão e da teoria da informação, e confirmar a possibilidade de utilizar a equação da entropia de Shannon tanto para medir a incerteza, quanto para medir a difusão em imagens binárias.

Figura 3.1 – Relação entre os conceitos abordados nesta tese.

3.2 Difusão

A difusão é um processo espontâneo de mistura decorrente do movimento aleatório de moléculas. Um processo pelo qual a matéria é deslocada de uma posição espacial para outra. Quando o transporte é de energia, o processo é chamado de condução de calor ou difusão térmica. Existem outras formas de se misturar substâncias ou conduzir calor, mas quando se fala de difusão, o transporte de energia ou matéria é causado exclusivamente por movimento molecular.

Não existe uma direção preferida para o movimento das moléculas, tornando imprevisível o caminho que uma determinada partícula irá percorrer. Porém, é possível afirmar que na média, o fluxo ocorrerá de lugares com alta concentração de partículas em direção a lugares com menor concentração (CRANK, 1975). Se for considerado o calor, a sua condução ocorre de regiões mais quentes para regiões mais frias.

Esse movimento aleatório é chamado de movimento Browniano em homenagem a Robert Brown (1773-1858) por ter observado, com microscópio, o movimento aleatório de partículas de pólen sobre a água. Em 1905, Albert Einstein publicou um artigo dando um embasamento atômico para o movimento Browniano, reforçando a teoria da existência do átomo (MEHRER, 2007), a qual sofria grande opressão dos cientistas mais conservadores daquela época.

O deslocamento errático das moléculas é mais facilmente descrito por uma distribuição estatística de deslocamento. Esta serve para descrever a proporção de moléculas que se move numa direção específica e numa distância específica (HAGMANN et.al., 2006). Quanto maior for o intervalo de tempo dedicado à observação da difusão, maior será o deslocamento das moléculas. É comum, em muitos meios, não haver “preferências” pela direção de deslocamento, as moléculas se movem em todas as direções com igual probabilidade. É um deslocamento isotrópico, o qual costuma ser modelado por distribuição Gaussiana. A abertura da Gaussiana (variância) é proporcional ao intervalo de tempo em que a difusão foi considerada.

Então, pode-se dizer que para cada intervalo de tempo imaginável haverá uma distribuição circularmente simétrica, cujo raio de alcance (onde se encontram as moléculas que mais se afastaram) é proporcional ao intervalo decorrido. Em outras palavras, há um conjunto contínuo, em escala, de distribuições Gaussianas. Onde a área (espaço) afetada pela difusão está associada à escala (abertura) da Gaussiana. Esta ideia forma a base do conceito de espaço-escala Gaussiano. Quanto maior for a escala de observação, maior será a quantidade de informação espacial misturada pela difusão.

A difusão pode ser representada pela equação 3.1, onde u representa a matéria e D é o coeficiente de difusão (ou de condução). Este representa o efeito de um conjunto de propriedades da matéria, as quais facilitam ou dificultam a difusão e definem se esta ocorre igualmente em todas as direções (isotrópica), ou se é dependente de direção (anisotrópica).

Na difusão Gaussiana, o processo ocorre em todas as direções sem qualquer controle, suavizando objetos, deteriorando bordas e fundindo regiões até que toda a estrutura da

imagem seja dissolvida, e todos os seus pixels apresentem o mesmo nível de cinza (o mesmo macro estado).

Nesta difusão isotrópica, o coeficiente D se torna um valor escalar, que depende do tipo de matéria em difusão. No caso específico de imagens, o valor do coeficiente de difusão D pode ser considerado igual a 1. Ainda na equação 3.1, ∇u representa a concentração de matéria (gradiente). O divergente (div) do gradiente é equivalente ao Laplaciano do campo escalar sobre o qual o gradiente foi computado (KREYSZIG, 1993). O Laplaciano pode ser considerado como uma generalização da derivada segunda para dimensões mais altas, servindo como uma ferramenta para caracterizar a concavidade de uma função (TANG, 2007).

$$\partial_t u = \text{div}(D \cdot \nabla u) \quad (3.1)$$

Do ponto de vista do processamento de imagens, o Laplaciano realiza uma comparação entre o valor do pixel central com a média de sua vizinhança. Isto informa o sentido e a quantidade de “partículas de brilho” a serem deslocadas, a fim de equalizar suas concentrações.

Portanto, a equação 3.1 diz que a taxa de variação das “partículas de luminância” por unidade de tempo é proporcional ao Laplaciano da luminância. Se o Laplaciano for zero, nenhuma alteração irá ocorrer na vizinhança, indicando que esta tem luminância homogênea (único nível de cinza). Se o Laplaciano for negativo, ocorrerá um fluxo de partículas do centro em direção à vizinhança. Se o Laplaciano for positivo, indicará que a vizinhança tem maior concentração e por isso o fluxo será no sentido de aumentar a concentração do pixel central, isto é, da periferia em direção ao centro.

A luminância ou brilhância é uma medida fotométrica da intensidade luminosa, que descreve a quantidade de luz emitida ou refletida por uma determinada área. Isto é, diz quão brilhosa é uma superfície. Numa imagem binária há uma região com presença de intensidade luminosa e outra com ausência desta. Ao se permitir a difusão, cria-se um fluxo de luminância da região clara para a região escura. Ao passo que a região escura se torna mais clara por receber “partículas de brilho”, a região clara se torna mais escura por tê-las cedido.

Aparentemente, este fluxo cessará após não haver mais vestígios de concentração de brilho, o que determina o instante em que o sistema entra em equilíbrio e o núcleo difusor se torna incapaz de produzir qualquer modificação macroscopicamente perceptível. O

movimento das partículas não cessa (exceto se a difusão for desativada) e sua recombinação espacial continua indefinidamente. Porém, a grande quantidade de micro combinações ⁶, as quais produzem uma mesma percepção macroscópica, cria a sensação de que o sistema se tornou estático. Quando o sistema atinge esse ponto de equilíbrio, diz-se que ele atingiu o seu estado de entropia máxima. A figura 3.2 mostra uma imagem binária e seu estado de equilíbrio após 17.025 suavizações com um filtro Gaussiano ($\sigma=1,5$). Deste ponto em diante, quaisquer suavizações adicionais não produzirão alterações macroscópicas.

Figura 3.2 – Difusão isotrópica: a) imagem original; b) núcleo de convolução ($\sigma=1,5$), o qual foi aplicado iterativamente 17.025 vezes, até obter a imagem estabilizada; c) imagem estabilizada (único nível de cinza).

No presente trabalho não há interesse por essa difusão completa e contínua, mas por uma difusão controlada, que permita, em cada escala selecionada, observar e quantificar a dissolução dos detalhes ao longo do contorno de um objeto.

Inicialmente, quando a abertura do núcleo de convolução é pequena (baixa escala), a difusão se concentra na área ao longo da fronteira entre as regiões homogêneas. No entanto, a quantidade difundida não é igual em todos os seus pontos. Devido à variação de direções, cada vizinhança centrada ao longo da fronteira pode ser composta por quantidades diferentes de região homogênea, o que implica em quantidades difundidas distintas. A Figura 1.1 no capítulo 1 pode auxiliar na compreensão destas afirmações.

A ideia de difusão, em visão computacional, teve início com a introdução do espaço escala (no mundo ocidental) em Witkin (1983). A partir deste, muitas variações foram introduzidas e combinadas com outras ideias. Entre essas variações estão: difusões anisotrópicas (PERONA; MALIK, 1990; WEICKERT, 1995), difusões não lineares

⁶ A quantidade de micro combinações que geram um mesmo macroestado é chamada de multiplicidade, que é a quantidade medida pela entropia.

(ROMENY, 1994; WEICKERT, 1997; WEICKERT, 2000), equações diferenciais parciais (PDE) (KIMMEL; SOCHEN; WEICKERT, 2005) e métodos *level-set* (OSHER; SETHIAN, 1988).

Até onde se tem conhecimento, dentre todas as técnicas e aplicações baseadas em difusão, somente uma se dedicou à detecção de ângulos (CINQUE; LOMBARDI; ROSENFELD, 1995). No entanto, a abordagem utilizada no referido trabalho difere da abordagem adotada nesta tese. Algumas diferenças são:

- É utilizada a difusão Gaussiana cuja evolução é do tipo usado em equações diferenciais, onde o tempo (ou quantidade de iterações) é que determina a extensão da difusão. Além disso, o coeficiente de difusão D é variável;
- O objeto e o fundo são sistemas isolados um do outro, não ocorre difusão do objeto para o fundo ou vice-versa. O processo de difusão (condução do calor) ocorre do contorno do objeto em direção ao seu interior e as temperaturas são medidas durante essa evolução. O valor dos ângulos é dado em faixa de temperaturas;
- A imagem, mesmo sendo binária, necessita de um pré-processamento com filtros de Sobel para determinar o contorno, e dar aos seus pontos o valor inicial da temperatura;
- Não é utilizada a capacidade da difusão em determinar fronteiras.

A associação de difusão à forma (ou a curvaturas) parece estar latente no pensamento científico há anos. Em meio às justificativas para exaltar a importância de se estudar o calor, Fourier fez a seguinte declaração: “The forms of bodies are infinitely varied; the distribution of the heat which penetrates them seems to be arbitrary and confused; but all the inequalities are rapidly cancelled and disappear as time passes on” (FREEMAN, 1878, p. 8).

Então, parece que a difusão é uma maneira natural de se perceber a estrutura geométrica dos objetos. Se isto for verdade, a quantificação da difusão é necessária para distinguir as particularidades de cada um deles. O cálculo de entropia é uma das maneiras de se quantificar a difusão.

3.3 Entropia

Acima foi exposto que a suavização Gaussiana pode ser vista como uma instância de difusão, que o seu efeito em imagens é a mistura de regiões de cor diferente e que essa mistura é dependente de características espaciais da imagem (concavidades, convexidades e fronteiras em linha reta). Portanto, caso se deseje distinguir as características espaciais da

imagem é necessário encontrar maneiras de quantificar a mistura. A medida entropia é usada na difusão do calor (MAXWELL, 1902), bem como em avaliação de misturas de substâncias (CAMESASCA; KAUFMAN; MANAS-ZLOCZOWER, 2006). Assim, o objetivo desta seção é compreender a entropia e verificar se é uma medida válida para o problema de detecção de curvaturas.

Lembrando-se do que foi exposto no capítulo 1, a região mínima de uma cena é representada por um pixel da imagem. Devido à falta de resolução infinita dos sistemas de aquisição de imagens, os detalhes da região mínima não podem ser distinguidos, causando a mistura dessas informações num único pixel. Consequentemente, a localização espacial de tais detalhes é destruída. Então, pode-se dizer que o pixel percebido é gerado, apenas, pelas proporções de brilho da região mínima. A inexistência de informação de localização espacial implica na possibilidade de haver várias distribuições espaciais distintas capazes de produzir o mesmo brilho percebido. Em outras palavras, há incerteza a respeito de qual seria a configuração espacial exata da vizinhança “absorvida” pelo pixel.

Uma maneira simples de se passar as ideias de estados microscópicos e macroscópicos é através da Figura 3.3, a qual representa uma técnica de pintura iniciada no final do século XIX conhecida por pontilhismo, uma técnica semelhante ao “*dithering*” da computação gráfica. A distribuição espacial de pontos de cores diferentes produz a cor percebida de uma determinada região aparentemente homogênea. Para que surja a percepção dessa região, o que importa são as proporções de pontos coloridos e não a ordem em que estão distribuídos. Isto implica na existência de diferentes distribuições capazes de produzir o mesmo efeito percebido. As diferentes distribuições de pontos coloridos podem ser chamadas de microestados e a cor resultante é o seu macroestado correspondente.

Outro exemplo que pertence à mesma classe das situações expostas acima é o jogo de dados. Seja o caso onde se considera a soma dos resultados de dois dados. Os possíveis resultados são 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12. Cada um dos resultados pode ser chamado de macroestado, enquanto que o par de valores que o produziu pode ser chamado de microestado. Deve-se perceber que, excetuando-se os resultados extremos (2 e 12), todos os outros podem ser produzidos por dois ou mais pares diferentes. O resultado 7, por exemplo, pode ser produzido por qualquer um dos seguintes pares (1,6), (2,5), (3,4), (4,3), (5,2), (6,1). O macroestado 7 é o que possui a maior quantidade de microestados, o que lhe confere a maior probabilidade de ocorrer (16,67%).

Figura 3.3 – Pintura com pontilhismo de George Seurat, "Un dimanche après-midi à l'Île de la Grande Jatte" (1884-1886).

Na escala atômica, quantidades inimagináveis de partículas se combinam de diferentes maneiras para formar o universo macroscópico que pode ser medido ou sentido. A visão, o olfato, o tato, o paladar ou a audição não possuem resolução infinita, obrigando os seres vivos a experimentarem o universo como uma média ponderada de partículas e de seus movimentos. Seja, por exemplo, um contêiner contendo um gás ideal. Cada uma de suas moléculas, num dado instante, se movimenta com uma determinada velocidade e num determinado sentido. Não se tem resolução, muito menos memória e processamento suficientes para rastrear individualmente cada molécula. O que se pode fazer é medir as variáveis de estado macroscópicas, por exemplo, a temperatura e a pressão do gás como um todo. Os valores destas medidas estão associados à energia cinética das moléculas. Cada molécula tem a sua própria energia cinética, que varia a todo instante devido aos constantes choques com as outras partículas e com as paredes do contêiner.

Quando tal sistema, que está isolado do universo, entra em equilíbrio termodinâmico, as variáveis macroscópicas estacionam num valor constante. No entanto, a hostilidade de seu ambiente atômico-molecular não diminui, o movimento é incessante e os valores individuais de energia cinética continuam variando. Isto indica que as energias cinéticas de sextilhões de moléculas podem se combinar em quantidades extraordinárias e, mesmo assim, produzir o mesmo macroestado de equilíbrio. Este é o macroestado de maior probabilidade em que o sistema pode ser encontrado, pois possui a maior multiplicidade, isto é, a maior quantidade de microestados. A probabilidade de que o sistema saia desse estado é praticamente zero.

Um sistema chega ao estado de equilíbrio devido à difusão térmica, na qual parte da energia das moléculas mais agitadas é transferida às menos agitadas, isto é, a difusão gera a dissipação da energia. Num processo bastante semelhante, a Figura 3.2 (a) foi transformada na Figura 3.2(c).

Os exemplos citados acima: o brilho de um pixel, a cor aparente de uma região criada com pontilhismo, os resultados de um jogo de dados e a temperatura de um gás apresentam uma característica em comum: a incerteza sobre qual configuração microscópica exata é a responsável por gerar o estado macroscópico percebido. Quanto maior for a multiplicidade de um macroestado, maior será a incerteza de encontrar o microestado exato que o produziu. Esta incerteza pode ser medida através da entropia de Shannon (SHANNON, 1948) representada na equação 3.2, onde a_k pode ser um determinado símbolo, ou nível de cinza, e $p(a_k)$ é a sua frequência de ocorrência.

$$H = -\sum_{k=1}^N p(a_k) \cdot \log_2(p(a_k)) \quad (3.2)$$

Visto que nesta tese se trabalha, apenas, com os níveis branco e preto, cujas probabilidades são complementares, a equação 3.2 é transformada na equação 3.3.

$$H = -[p \cdot \log_2 p + (1 - p) \cdot \log_2(1 - p)] \quad (3.3)$$

A palavra entropia foi utilizada pela primeira vez por Rudolf Clausius em meados do século XIX para se referir a uma variável macroscópica de sistemas termodinâmicos, que indicava a quantidade de energia não disponível para o trabalho (perdas com a dissipação térmica) (CLAUSIUS, 1879; MAXWELL, 1902), por exemplo, para movimentar pistões de máquinas a vapor. Um pouco mais tarde Josiah Willard Gibbs e Ludwig Boltzmann chegaram a equações similares, que explicavam a entropia de Clausius como probabilidades associadas às energias cinéticas de moléculas (JAYNES, 1965). Tais equações são muito semelhantes à entropia de Shannon (equação 3.2). A diferença é que esta não utiliza grandeza energética. Há várias passagens históricas em que se interpretam as entropias de Gibbs e Boltzmann como medidas de incerteza, de maneira bastante semelhante aos exemplos apresentados nesta seção. Atualmente, parece haver uma forte tendência em admitir que a quantidade medida pela entropia seja, de fato, a incerteza associada ao macroestado, colocando a teoria da informação

em primeiro plano para explicar os fenômenos do universo (GLEICK, 2011; BEN-NAIM, 2008; BEN-NAIM, 2007; YU, 2000).

As palavras informação e incerteza estão fortemente associadas. De acordo com Shannon, sua entropia mede a quantidade de informação transmitida. Quanto maior é a multiplicidade de um dado macroestado, menor é a probabilidade de se encontrar o microestado correto (o espaço de busca é maior). Portanto, a quantidade de informação necessária para encontrá-lo é maior. Deste ponto de vista, a informação é inversamente proporcional à probabilidade de se encontrar o microestado correto, por isso o sinal negativo aparece no início da equação 3.2. A Figura 3.4 mostra o comportamento da entropia em relação à probabilidade de ocorrência de um evento com duas possibilidades (pixel de cor preta ou branca). Quando as possibilidades do evento possuírem a mesma probabilidade de ocorrer, a entropia, a incerteza ou a surpresa serão máximas, ou seja, é uma situação absolutamente imprevisível.

Figura 3.4 – Entropia para duas possibilidades com probabilidades p e $(1-p)$. Figura adaptada de Shannon (1948).

Imaginando-se as regiões de uma imagem binária (objeto e fundo) como contêineres isolados, o processo de suavização se torna o mecanismo que os conecta, permitindo a troca de partículas entre os mesmos. Este é um cenário semelhante àqueles comumente encontrados na literatura onde a entropia é aplicada. Cada vizinhança da fronteira, que se pretende medir, possui sua própria concentração de objeto e fundo. Em outras palavras, cada vizinhança contém probabilidades próprias de objeto e fundo, implicando em entropia própria para cada

uma delas. Isto torna a entropia uma medida viável para caracterizar o tipo de fronteira, se é em linha reta, ou se forma uma concavidade ou convexidade.

A afirmação acima não é uma verdade completa, pois podem existir várias configurações objeto-fundo capazes de produzir a mesma entropia. Há maneiras de se restringir essa quantidade de configurações. Nesta tese foram associadas três ideias: 1) definir as configurações espaciais que são possíveis de existir ao longo de fronteiras e excluindo-se aquelas que não se encaixam neste grupo; 2) utilizar vizinhanças pequenas fornece menos espaço para variações; 3) se a vizinhança for a menor possível, a utilização de uma distribuição de probabilidades circularmente simétrica garante que, para um dado valor de entropia, só exista uma configuração espacial. Os detalhes a respeito dessas ideias são apresentados ao longo do capítulo 5, dedicado ao detector de *corners* de contorno baseado em entropia.

Até o momento foram apresentados: o conceito de difusão, a noção de que em seus estágios iniciais a difusão isotrópica se concentra ao longo das fronteiras e que é possível utilizar a entropia, ou melhor, a incerteza para medi-la. Entre as questões que guiaram a pesquisa desta tese estão: “*onde uma imagem pode, naturalmente, sofrer difusão?*” e “*o que está sendo difundido?*”.

3.4 Relações entre PSF e difusão

De acordo com Thibos (1989), o processamento de imagens pelo olho é composto pela concatenação de filtros lineares passa-baixas. O primeiro filtro é óptico, caracterizado por uma PSF óptica e o segundo filtro é neural, caracterizado por uma PSF neural. Esta seção lida exclusivamente com a PSF óptica, deixando a PSF neural para a seção seguinte.

O funcionamento da difusão molecular faz surgir a ideia de que a difusão em imagens possa estar relacionada à natureza corpuscular da luz. Ao encontrar a entrada de um sistema óptico ou uma fenda num anteparo qualquer, os fótons da luz refletida por um objeto sofrem perturbações ao interagir com a matéria que compõe as paredes da fenda. A perturbação é tal, que um fóton capturado numa dada posição no plano da imagem pode não corresponder, exatamente, à mesma posição na superfície do objeto de onde o fóton foi emitido. Isto é, a trajetória dos fótons é modificada, resultando num espalhamento da luz (*scattering*), o qual é representado pela PSF do sistema óptico, ou da fenda (BERGÉ, 2013; WESTHEIMER, 2009; SUPPES; DE BARROS, 1994).

Embora exista uma semelhança em termos de distribuição de probabilidades, não se está afirmando que as modificações de trajetória dos fótons é uma difusão. Devido ao comportamento exótico da luz e à complexidade envolvida em seu estudo, será considerada, apenas, a aproximação da PSF com uma distribuição de probabilidades, que é capaz de produzir suavização circularmente simétrica.

A resposta impulsiva de um sistema óptico é conhecida por função de espalhamento de ponto (PSF). É a medida do poder de resolução de um sistema óptico, quanto mais estreita for a sua PSF, melhor será a sua resolução, isto é, menor será o detalhe da cena capaz de ser reproduzido na imagem. A PSF pode representar o efeito combinado de diferentes fenômenos, como, difração, refração ou aberrações de lentes. Simplificadamente, o seu efeito é transformar um ponto da cena em uma pequena mancha circular na imagem, como se o brilho do ponto fosse dissipado do centro em direção à periferia. Este efeito é consequência da filtragem espacial. Na propagação da luz entre uma cena e a sua imagem, as componentes de alta frequência são filtradas, impedindo que o ponto seja reconstruído com precisão total (NOVOTNY; HECHT, 2006). A convolução de cada ponto da cena com a PSF do sistema produz uma versão suavizada da cena, isto é, a sua imagem.

Em um sistema óptico ideal, o qual possui lentes perfeitas, ou não possui lentes (é apenas uma entrada ou fenda), o espalhamento da luz é um resultado exclusivo do fenômeno de difração. Neste caso, a PSF do sistema representa a difração, a qual costuma ser explicada através do modelo ondulatório da luz. A explicação baseada em ondas não impede que exista difração em partículas (POLLOCK, 1993; SUPPES; DE BARROS, 1994). O padrão de difração, ou figura de difração, representa a probabilidade de se detectar fótons no plano da imagem. Então, a PSF pode ser vista como uma função de distribuição espacial de probabilidades de se registrar fótons (LARSON, 2010). A Figura 3.5 mostra o padrão de difração gerado pela passagem da luz através de uma abertura circular. O disco central da difração (ou lobo central) concentra 84% da intensidade óptica total (ZAPPE, 2010; DRAGESCO, 1995). Este padrão foi descrito matematicamente, pela primeira vez, pelo astrônomo e matemático inglês George Biddell Airy (1801-1892). Por isso, o disco central é chamado de disco de Airy e os anéis concêntricos de padrão de Airy.

Figura 3.5 – Difração em uma abertura circular. O disco central, limitado pelo primeiro anel escuro, é conhecido por disco de Airy. O quadro a direita mostra a difração em 3D. O lobo central (disco de Airy) se assemelha a uma distribuição Gaussiana.

É comum aproximar a PSF com uma distribuição Gaussiana (BARTEN, 1999; HE, 2011; HUSSEIN, 2011; ZHANG et al., 2007; ELLIOTT; JENKINS, 1990; ZHANG; ZERUBIA; OLIVO-MARIN, 2006; ROTH, 2006). Aproximar a PSF com distribuições circularmente simétricas embute a ideia de que um sistema óptico não trata igualmente a informação transportada pela luz. A informação que flui no centro de um feixe de luz sofre menos espalhamento, o qual vai aumentando gradativamente em direção à periferia do feixe. Um peso maior para o centro pode significar que a informação ali presente representa com maior fidelidade o ponto da cena que está sendo observado.

Pode-se dizer, então, que a PSF de uma abertura circular e a difusão isotrópica possuem distribuições de probabilidades similares. Na difusão as partículas são “embaralhadas” por movimento térmico. Num sistema óptico, a luz que o penetra tem seus fótons embaralhados em virtude de sua interação com as paredes da abertura. Trocando o verbo embaralhar pelo sinônimo misturar, pode-se dizer que os fótons são misturados obedecendo às regras da PSF. Isto é, de acordo com suas probabilidades distribuídas de forma circularmente simétrica.

Supondo-se que as distribuições da difusão e da difração (PSF) sejam realmente semelhantes, acredita-se que a entropia usada para avaliar a difusão, também seja útil para avaliar a difração. A seguir são apresentadas algumas referências que reforçam essas ideias:

- Quando a imagem de um objeto, iluminado por luz incoerente, é capturada por um sistema óptico, cada ponto luminoso (x,y) na superfície do objeto gera uma

distribuição de intensidades (a imagem difração do ponto) na superfície da imagem (LINFOOT, 1955).

- A distribuição de intensidades através de uma imagem, ou a PSF, pode ser interpretada como funções de densidade de probabilidade, desde que sejam adequadamente normalizadas (TORRABA; RABAL; RUIZ, 1992).
- Na superfície da retina onde se encontram os cones, a PSF se torna a densidade de probabilidade que indica a possibilidade de onde um fóton surgirá (DEERING, 2005).
- De certa forma, é possível interpretar a expansão espacial do ponto, causada pela difração, como um aumento de entropia (BARAKAT, 1998).
- A entropia pode ser considerada como o estudo das distribuições de probabilidade. Neste sentido, a PSF devidamente normalizada apresenta as propriedades formais de uma função de densidade de probabilidade (PDF) e pode ser tratada com análise entrópica (BARAKAT, 1998).

Um aspecto bastante interessante dos sistemas visuais biológicos, pelo menos nos vertebrados, é que há reproduções aproximadas da PSF pelas redes neurais da retina, chamadas de PSFs neurais. Seria uma maneira de a Natureza biológica tentar entender a Natureza da luz? Ou seriam as distribuições circularmente simétricas uma ocorrência comum na Natureza, e as PSFs óptica e neural não têm qualquer relação uma com a outra? No caso de haver essa relação, como os sistemas biológicos tomariam consciência da existência de uma PSF óptica, visto que a realidade perceptiva está depois da abertura do sistema óptico, isto é, a cena suavizada é a fundação sobre a qual a realidade perceptiva é construída? Os sistemas visuais biológicos levam tempo para amadurecer. Seria esse amadurecimento um tipo de calibração entre os sistemas sensoriais, permitindo averiguar discrepâncias entre as realidades perceptivas e físicas?

Infelizmente, da mesma forma que a compreensão da Natureza da luz, a compreensão da existência da PSF neural está fora do escopo desta tese. Lembra-se que o objetivo deste capítulo é constatar que existe suavização da cena e que isto promove um aumento de incerteza. O que será visto na seção seguinte é que as ideias expostas anteriormente sobre difusão e entropia, também, são válidas para a PSF neural. A teoria de espaço-escala tem como base a PSF neural (ROMENY, 2003) e o entendimento sobre difusão e entropia poderia ter sido explorado diretamente sobre esta PSF.

3.5 PSF neural, produção de entropia e diferença de Gaussianas

Num sistema visual biológico, o plano da imagem é a retina, uma película composta por dois tipos de fotorreceptores e uma variedade de células neurais dispostas em camadas. Sobre essa película é projetada a versão suavizada da cena. Cada um dos fotorreceptores se encarrega da transdução da luz em sinal neural, ou seja, a imagem é quebrada em pequenos pedaços (amostras) e uma nova representação da informação é enviada às camadas seguintes. Deve-se considerar que um fotorreceptor também tem uma entrada, o que causa difração local, isto é, em cada fotorreceptor será promovida uma “mistura” de fótons.

Em geral os fotorreceptores se conectam às células neurais da camada seguinte numa relação de muitos para um. Assim os valores neurais provenientes de uma vizinhança de fotorreceptores são conduzidos para uma única célula. Isto permite que a célula possua uma “janela visual” maior, composta pelas janelas de vários fotorreceptores. Esta janela é conhecida por campo receptivo da célula.

Os sinais neurais provenientes de uma vizinhança de fotorreceptores são tratados de forma ponderada pela célula neural. Isto é, o fotorreceptor no centro da vizinhança recebe o maior peso, e os mais distantes do centro recebem pesos cada vez menores. A Figura 3.6 mostra a conexão parcial entre fotorreceptores e uma célula bipolar, bem como a distribuição dos pesos da conexão, a qual é modelada como uma distribuição Gaussiana.

Os fotorreceptores se conectam às células bipolares através de conexões diretas, como as da Figura 3.6, e através de conexões indiretas via células horizontais. As Figuras 3.6, 3.7 e 3.8 são construções didáticas para mostrar a existência desses dois tipos de conexões e como são tratadas pela célula. As células bipolares distinguem as conexões diretas das indiretas atribuindo sinais opostos⁷ às mesmas.

⁷ Biologicamente, a atribuição de sinal positivo ou negativo advém de processos eletroquímicos no interior da célula conhecidos respectivamente por despolarização e hiperpolarização.

Figura 3.6 – Simplificação da conexão entre fotorreceptores e uma célula bipolar.

Figura 3.7 – Simplificação da conexão entre fotorreceptores e uma célula horizontal.

O campo receptivo da célula bipolar, então, é o resultado da combinação das distribuições Gaussianas que modelam as conexões diretas e indiretas, cujos sinais opostos geram duas regiões concêntricas com comportamentos antagônicos. Esse campo receptivo é dito ser do tipo *center-surround* (HARTLINE, 1940; KUFFLER, 1952; WIESEL, 1960; BARLOW, 1953; HUBEL; WIESEL, 1960; BARLOW; HILL; LEVICK, 1964; GRAHAM; CHANDLER; FIELD, 2006), e o seu esquema simplificado é mostrado na Figura 3.8. Deve-se considerar que todos os fotorreceptores se conectam com a célula horizontal, permitindo a sua modelagem com uma Gaussiana, como é mostrado na Figura 3.7. O tamanho da vizinhança (ou do campo receptivo) é demarcado por esta célula

As células bipolares podem ser classificadas quanto à relação antagônica das regiões que compõem seus campos receptivos. O tipo *on-center* responde ativamente quando a região central é estimulada com luz, e a região periférica com ausência de luz. O tipo *off-center* funciona de maneira inversa, respondendo ativamente quando a região periférica é iluminada e o centro é mantido no escuro. Como foi dito acima, os campos *center-surround* são modelados através da combinação das Gaussianas que representam as distribuições de pesos das duas regiões, usando-se o modelo de diferença de Gaussianas (DoG) (RODIECK, 1965; ENROTH-CUGELL; ROBSON, 1966). O item b da Figura 3.8 mostra a DoG de um campo receptivo *off-center* e o item c mostra a DoG do *on-center*.

Figura 3.8 – Campo receptivo *center-surround* para células bipolares: a) conexões diretas (seta azul) e indiretas (seta vermelha) entre uma vizinhança de fotorreceptores e uma célula bipolar; b) resultado da diferença entre Gaussianas para uma célula bipolar *off-center* e c) idem para uma célula *on-center*.

A camada seguinte de células neurais da retina é composta por células ganglionares e células amácrinas. As ganglionares também têm a sua janela visual, a qual é fornecida por uma ou mais células bipolares, permitindo que a ganglionar tenha acesso a uma área maior da imagem. Um padrão de conexões semelhante ao anterior, também, é encontrado nesta camada. As células bipolares trazem a informação da imagem para as células ganglionares, através de conexões diretas, ou indiretas via células amácrinas, fazendo uso do mesmo tipo de

ponderação radial, o que gera o mesmo tipo de campo receptivo *center-surround*. As células ganglionares também funcionam nas modalidades *on-center* e *off-center*.

Deve-se perceber que sinais provenientes de uma vizinhança de entrada são combinados de forma radialmente ponderada para formar um único sinal de saída. Este, por sua vez, é combinado de forma radialmente ponderada com outros sinais de saída, com a finalidade de criar uma vizinhança de entrada para a camada seguinte. Um processo que lembra a auto similaridade fractal, que se repete desde os reflexos provenientes da cena até a última camada de células que desembocam no córtex visual. O que é sempre encontrado nessa repetição são as distribuições circularmente simétricas, as mesmas que se encontram em difusões isotrópicas.

Então, pode-se considerar que um campo receptivo realiza um processo de suavização devido a sua característica de combinar muitos valores de entrada em um único de saída. No entanto, os campos do tipo *center-surround* não são apenas suavizações. Estes são modelados com diferença entre Gaussianas, ou melhor, com diferença entre suavizações. O que esses campos querem obter com tal procedimento? A resposta fica muito mais clara através de figuras. A Figura 3.9 mostra a diferença entre duas versões suavizadas de uma imagem homogênea. A quantidade de suavização, ditada pelo desvio padrão (σ), é irrelevante para regiões homogêneas, pois a suavização não causa modificação alguma. Isto significa que campos receptivos *center-surround* não respondem a regiões homogêneas.

Figura 3.9 – A diferença entre Gaussianas para uma imagem homogênea é zero.

A Figura 3.10 mostra o resultado de diferença entre Gaussianas quando a imagem é composta por duas regiões homogêneas (objeto preto sobre fundo branco). Nesta figura, a Gaussiana mais fechada (menor desvio padrão) é a componente negativa da DoG, isto é, a região central é negativa e a periférica é positiva, simulando uma célula *off-center*. Pode ser percebido que um campo receptivo *center-surround* responde bem ao contraste entre regiões,

detectando a fronteira entre elas (foi necessário retocar o contorno da imagem devido ao baixo contraste).

Figura 3.10 – Detecção de bordas com DoG *off-center*.

A Figura 3.11 apresenta uma DoG modelando as respostas de células *on-center*, cuja componente negativa é a região periférica. Neste caso, também, ocorre detecção de bordas, porém aparecem deslocadas, as quais não são consideradas como parte do objeto, mas pertencentes ao fundo (foi necessário retocar o contorno da imagem devido ao baixo contraste).

Figura 3.11 – Detecção de bordas com DoG *on-center*.

As duas modalidades de célula (*on-center* e *off-center*) detectam bordas. As células *on-center* respondem ao contraste positivo e as *off-center* respondem ao contraste negativo. A detecção de cada tipo de contraste depende, apenas, do posicionamento das células. Se a

região branca estiver projetada sobre o centro da célula *on-center* e a região escura sobre a sua periferia, o contraste positivo é detectado. Se a região escura estiver projetada sobre o centro de uma célula *off-center* e a região clara cair sobre a sua periferia, o contraste negativo é detectado. Visto que todo o plano da imagem está coberto por campos receptivos das células *on-center* e *off-center*, certamente os dois contrastes serão detectados. Por isso existem duas bordas em posições diferentes, uma pertencente ao objeto e a outra pertencente ao fundo. No caso específico de objeto escuro em fundo claro, a borda deslocada é detectada pela célula *on-center*. Quando o objeto for claro e o fundo for escuro, a borda deslocada será detectada pelas células *off-center*. A Figura 3.12 mostra mais nitidamente essas duas bordas.

Figura 3.12 – Ampliação de parte das bordas detectadas por DoG off-center e on-center. O pixel do vértice do retângulo foi marcado com um círculo e as bordas reforçadas para melhor visualização e percepção da diferença de posicionamento das bordas detectadas por cada modalidade de DoG.

Então, o que a DoG (ou campos receptivos *center-surround*) realiza é a busca por modificações causadas pela suavização. Tais modificações só ocorrem se houver regiões em contraste. Se as suavizações forem fracas, ou melhor, de curto alcance (desvio padrão baixo), as modificações se concentrarão ao longo da fronteira entre as regiões. Pode-se dizer, que a DoG quantifica a suavização produzida, e a detecção de bordas é uma consequência disso.

É muito importante observar que a DoG e a entropia servem para a mesma finalidade, quantificar a difusão (ou suavização). Ao longo desta tese já foi mencionado que a difusão tem início a partir da fronteira entre duas regiões. Porém, uma fronteira, em geral, tem um comportamento variado. Em determinadas posições aparecem convexidades e em outras concavidades. Tais reentrâncias diferem entre si em ângulo e tamanho. Se a configuração espacial da fronteira varia, a sua difusão também irá variar. Nas Figuras 3.10 e 3.12 é possível perceber essa diferença no brilho da fronteira, que é mais intenso nos vértices e permanece constante nas direções horizontal e vertical. O quadro de valores da DoG *off-center*, apresentado na Figura 3.12, indica numericamente essa diferença. A produção de suavização no vértice é quase o dobro da produzida nas direções horizontal e vertical.

A Figura 3.13 mostra a imagem entropia gerada pelo detector de *corners* de contorno baseado em entropia, que está sendo proposto nesta tese. O detector foi aplicado na mesma imagem binária das Figuras 3.10 e 3.11. Pode-se observar na imagem entropia a existência de uma fronteira (contorno) com espessura dupla (dois pixels). Na realidade são duas fronteiras, uma externa e outra interna, que correspondem exatamente às fronteiras obtidas com DoG *on-center* e *off-center* (compare com as ampliações da Figura 3.12). A entropia registra simultaneamente os dois fluxos de partículas gerados pela suavização, um que vai do objeto para o fundo e outro que vai do fundo para o objeto. Embora não sejam os mesmos valores, a entropia tem comportamento idêntico ao da DoG. Nos vértices há difusão mais intensa do que nas direções horizontal e vertical.

Figura 3.13 – Contorno duplo na imagem entropia.

Nesta seção, as descrições biológicas foram mantidas ao mínimo, a fim de evidenciar, apenas, os processos de suavização existentes na retina. É importante acrescentar que existem diferentes tipos de células bipolares e uma quantidade ainda maior de tipos de células ganglionares (GOLLISCH; MEISTER, 2010), e que todas têm em comum a característica de possuir as modalidades *on-center* e *off-center*, formando dois caminhos diferentes até o córtex (SODERQUIST, 2002). A quantidade de células *off-center* é quase o dobro de *on-center*.

3.6 Conclusão do capítulo

Neste capítulo a principal preocupação foi a de fundamentar teoricamente as ideias a respeito de como a difusão poderia ocorrer num sistema visual, como seria o seu aspecto, como poderia ser medida e como poderia ser utilizada para derivar medidas a respeito das curvaturas ao longo de contornos.

Foi mostrado que medir a produção de suavização permite detectar o contraste entre regiões homogêneas, bem como de caracterizar os detalhes da fronteira entre elas. A diferença entre Gaussianas (DoG) e a entropia de Shannon são ferramentas viáveis para caracterizar os detalhes de contornos.

Capítulo 4 – Núcleo de convolução proposto

4.1 Introdução

Todos os detectores que estão sendo propostos nesta tese são baseados na quantificação de produção de suavização, seja através da entropia, ou da diferença entre suavizações. Isto indica que ocorre algum tipo de suavização durante o processo de detecção de *corners* de contorno. Visto que as funções de espalhamento de ponto (PSF) costumam ser aproximadas com distribuição Gaussiana, esta, inicialmente, foi admitida como núcleo de convolução. A distribuição Gaussiana se mostrou bastante adequada enquanto se mantinha baixo o seu desvio padrão, isto é, para escalas muito baixas. No entanto, os detalhes associados a baixas escalas não são representativos. É necessário observá-los em escalas mais altas, a fim de verificar se possuem relevância na descrição global da forma do objeto em estudo. Sendo assim, outra distribuição de probabilidades deveria ser escolhida. Este capítulo tem o objetivo de mostrar como se chegou à distribuição definitiva.

4.2 Desvantagens da suavização Gaussiana

Um dos principais usos da suavização Gaussiana é a eliminação de ruídos. Se o ruído estiver presente na forma de pequenas estruturas, este será atenuado, porém, a Gaussiana não difere ruídos de estruturas reais da imagem. Quanto mais abrangente for o núcleo Gaussiano, isto é, quanto maior for a sua escala σ , maior serão as estruturas atenuadas.

Sabe-se que a suavização Gaussiana e a precisão na localização de bordas são incompatíveis. Os detectores de borda baseados em derivadas (Laplaciano) são muito sensíveis a ruídos, por isso são combinados com uma suavização Gaussiana resultando no Laplaciano da Gaussiana (LoG). O deslocamento de bordas é comum na resposta de detectores que fazem uso dessa suavização (SHEN; CASTAN, 1992; LU; JAIN, 1992).

Outro efeito indesejável da suavização Gaussiana é a contração de contornos (*contour shrinking*), afetando diretamente na estimação de curvaturas (COSTA; CESAR Jr., 2001).

Visto que as distribuições Gaussianas, usadas como núcleo de convolução, são normalizadas, a amplitude está vinculada à abertura. Não há como aumentar a escala sem reduzir a amplitude. A suavização com uma Gaussiana de baixa amplitude tende a igualar os valores de uma vizinhança (é o efeito da difusão em destruir concentrações). Obviamente, tal

equalização reduz a distinção entre os pontos de uma vizinhança, causando a perda de localização, além da atenuação do detalhe ali existente.

Especificamente para as detecções de *corners* de contorno aqui realizadas, a suavização Gaussiana, a partir de uma escala relativamente baixa (em torno de $\sigma=0,8$), introduz ambiguidades nas relações ângulo-entropia. Ou seja, o achatamento da Gaussiana impede a associação correta entre a suavização produzida e os ângulos de setores circulares. A Figura 4.1 facilita a compreensão do que está sendo relatado. O gráfico da relação ângulo-entropia computada com o núcleo proposto (Figura 4.1.b) decresce monotonicamente, isto é, cada ângulo só possui um valor de entropia e a cada valor de entropia só é associado um ângulo. No entanto, quando esta mesma relação é computada usando-se um núcleo Gaussiano (Figura 4.1.c), um valor de entropia pode ser associado a dois ângulos distintos. Os setores circulares do item *a* desta figura são produzidos ao centrar o núcleo de convolução sobre a fronteira das regiões clara e escura.

Figura 4.1 – Relação ângulo-entropia: núcleo proposto versus núcleo Gaussiano. a) ângulos da; b) Relação ângulo-entropia gerada com o núcleo proposto; c) Relação ângulo-entropia gerada com um núcleo Gaussiano.

4.3 O núcleo proposto para a computação de entropia

Considerando-se que o problema na construção de uma relação ângulo-entropia válida reside no vínculo entre a escala (σ) da Gaussiana normalizada e a sua amplitude central, a solução é a busca por distribuições de probabilidades não suscetíveis à contração de amplitude. No entanto, a Gaussiana e suas derivadas apresentam uma quantidade de características muito especiais que permitem o modelamento das aberturas através das quais os seres vivos e os instrumentos ópticos observam o universo. Em Romeny e Florak (2000) destaca-se o seguinte parágrafo:

[...] All *partial derivatives* of the Gaussian kernel are solutions too of the diffusion equation. So the first important result is that we have found the Gaussian kernel and all of its partial derivatives as the *unique* kernel for a front-end visual system that satisfies the constraints "no preference for location, scale and orientation" and linearity. We have found a one-parameter *family* of kernels, where the scale σ is the free parameter. This is a general feature of the biological visual system: the exploitation of *ensembles* of aperture functions, which are mathematically modeled by families of kernels for a free parameter, e.g. for all scales, derivative order, orientation, stereo disparity, motion velocity etc. The Gaussian kernel is the unique kernel that generates no *spurious resolution* (e.g. the squares so familiar with zooming in on pixels). It is the physical *point operator*, the Gaussian derivatives are the physical *derivative operators*.

Diante da importância dos núcleos Gaussianos decidiu-se não buscar por novas distribuições, mas modificá-los de tal forma que mantivessem alta amplitude independente da escala adotada. Para isso é necessário que a amplitude e a abertura da Gaussiana sejam desvinculadas.

Uma maneira simples de resolver o problema é a utilização de soma de Gaussianas (SoG – *Sum of Gaussians*) para aproximar a curva desejada. Esse tipo de aproximação faz parte das técnicas estatísticas conhecidas como *finite mixture models* ou *Gaussian mixture models* (GMM) (GOSHTASBY; ONEILL, 1994; MCLACHLAN; PEEL, 2000). Para os casos desta tese, o parâmetro que varia é a escala (σ). A ideia, então, é somar duas Gaussianas de escalas distintas, uma larga com a escala suficiente para encobrir a vizinhança desejada e a outra estreita para garantir que a amplitude do núcleo resultante se mantenha íngreme. Desta forma a abertura do núcleo e a sua amplitude se tornam independentes, interrompendo os efeitos indesejáveis produzidos pela suavização Gaussiana. Ademais, as características desejáveis da Gaussiana são mantidas, visto que a soma de Gaussianas produz outra Gaussiana (WEISSTEIN, 2016).

Os detectores propostos nesta tese funcionam em multiescala, o que implica na utilização de vários núcleos de convolução, uma família de núcleos. Visto que cada um deles é construído com a soma de duas Gaussianas, optou-se por padronizar essa construção. Como foi mencionado acima, cada Gaussiana que compõe o núcleo possui uma área de atuação bem definida. A influência da Gaussiana mais larga se concentra na região periférica do núcleo, enquanto a Gaussiana estreita influencia na região central.

Essas duas regiões do núcleo gerado com a SoG faz lembrar os campos receptivos *center surround*, porém sem o uso de sinais opostos. Neste tipo de configuração, a entropia máxima (igual a 1) é obtida quando o somatório dos pesos⁸ da região periférica se iguala ao somatório dos pesos da região central. Assim, decidiu-se usar esta restrição de equiprobabilidade para guiar a construção dos núcleos de convolução. Então, independentemente do tamanho do núcleo, o seu ponto central possuirá valor igual a 0,5, assim como o somatório de todos os outros pontos do núcleo.

Para a construção do núcleo deve-se obedecer a seguinte lista de passos:

1. Construir a Gaussiana na escala (σ_{larga}) desejada.
2. Descobrir a escala (σ_{estreita}) da segunda Gaussiana, que somada à primeira produzirá uma terceira distribuição cujo peso central tem valor 1;
3. Normalizar a terceira distribuição, isto é, dividir por dois. O que produzirá o núcleo desejado com valor central igual a 0,5.

O passo número 2 descreve uma tarefa que consome tempo e paciência, pois o valor adequado para σ_{estreita} se obtém através de tentativa e erro. Por isso, é proposto, também, um algoritmo capaz de modificar uma distribuição Gaussiana. Este algoritmo produz um núcleo muito semelhante ao construído a partir de soma de Gaussianas e com todas as suas características desejáveis. A ideia se resume em “puxar para cima” o peso central da Gaussiana até que este atinja o valor 0,5, sem que a escala seja alterada. Puxar para cima significa aumentar o valor do peso central, o que torna o peso total da distribuição maior do que a unidade. Isto exige que a distribuição resultante sofra normalização. As equações 4.1, 4.2 e 4.3 descrevem como obter o valor a ser adicionado ao peso central da Gaussiana.

De acordo com a restrição adotada de entropia máxima, o peso central (C_f), da distribuição desejada, precisa apresentar valor 0,5, então é necessário somar um valor Δ ao

⁸ A palavra peso deve ser subentendida como peso probabilístico de uma distribuição de probabilidades. Por isso, logo após, se usou o termo equiprobabilidade.

peso central atual (C_a). Para calcular o valor de Δ é necessário considerar que o peso final ($C_f = 0,5$) só é obtido após a normalização da nova distribuição (dividir por N). O valor de N corresponde à totalização dos pesos da nova distribuição não normalizada, a qual corresponde à totalização da distribuição Gaussiana original, acrescida do valor Δ . A equação 4.3 surge da combinação das equações 4.1 e 4.2. Os pesos C_a e C_f estão ilustrados na Figura 4.2.

$$C_f = \frac{C_a + \Delta}{N} = 0,5 \quad (4.1)$$

$$N = 1 + \Delta \quad (4.2)$$

$$\Delta = 1 - 2 \times C_a \quad (4.3)$$

A sequência de passos para modificar uma Gaussiana e gerar a distribuição desejada está listada a seguir:

1. Construir a Gaussiana na escala desejada;
2. Verificar o valor do peso central atual;
3. Determinar o valor de Δ ;
4. Somar Δ ao valor central atual;
5. Normalizar a distribuição.

A aproximação realizada pelo algoritmo descrito acima gera um núcleo muito semelhante ao núcleo gerado pela soma de Gaussianas (SoG). A Figura 4.3 apresenta uma comparação visual entre as suavizações produzidas pelas duas versões do núcleo de convolução proposto. As imagens nos itens *b* e *c* desta figura são praticamente iguais. A maior diferença produzida pela subtração destas é 0,00028. Tal semelhança não é uma surpresa, visto que a componente mais estreita ($\sigma \leq 0,5$) utilizada nas SoGs se concentra quase que totalmente em um único pixel. É quase uma soma entre uma Gaussiana e um impulso localizado no pixel central. É exatamente a ideia de puxar o centro da Gaussiana para cima.

Figura 4.2 – Modificação de Gaussiana para obter a distribuição desejada. O peso central da Gaussiana (C_a) é incrementado por um valor Δ e após a normalização é obtida a distribuição desejada com o peso central $C_f=0,5$.

Figura 4.3 – Comparações entre os resultados das suavizações com SoG e com Gaussiana modificada. a) Imagem original; b) Suavização com SoG ($\sigma_1=23,833$ e $\sigma_2 = 0,1952$); c) Suavização com Gaussiana modificada ($\sigma=23,833$); d) Suavização com Gaussiana ($\sigma=23,833$); e) Ampliação da região do bico do avião, os valores são válidos para as suavizações em b e c.

Pode-se perceber através dos itens *b* e *c* da Figura 4.3 que o núcleo proposto não serve para filtrar ruídos ou dissolver detalhes do contorno do objeto, como ocorre no item *d*. Aqui, o objetivo da suavização é produzir um “fluxo de partículas de brilho” do fundo para o objeto, ou vice-versa, mantendo a localização original de cada pixel. Assim, utilizando-se a entropia é possível quantificar a mistura de brilho que ocorre individualmente em cada um dos pixels da imagem. O valor dessa quantificação depende da configuração espacial da vizinhança na qual o pixel está inserido. Por exemplo, se o pixel estiver no vértice de uma convexidade do objeto, a sua entropia será alta. Isto pode ser inferido a partir da Figura 4.1, observando-se os itens *a* e *b*.

Os esquemas multiescala, como o espaço-escala, levam em conta que numa cena coexistem elementos de várias escalas. Por exemplo, na cena de uma floresta é possível observar a floresta, uma de suas árvores, as folhas desta árvore, ou ainda, as nervuras de cada folha.

Para se enxergar tais nervuras, o sistema visual necessita estar bem próximo à folha. Neste caso é dito que o sistema está ajustado à baixa escala, na qual os campos receptivos da retina integram apenas as informações locais da folha. Ao afastar o sistema visual, a escala aumenta, e também aumenta a quantidade de informação integrada pelos campos receptivos. Isto faz com que os pequenos detalhes exerçam menos influência na integração de uma vizinhança grande e, assim, vão sendo atenuados com o aumento da escala.

Em termos de núcleo Gaussiano, quanto maior é a sua escala (σ), maior é a vizinhança de pixels abrangida, e maior é a suavização produzida. Os pequenos detalhes de uma vizinhança tendem a desaparecer com o aumento de escala. Assim, a observação multiescala permite descobrir quais são os detalhes relevantes da forma de um objeto.

Os detectores propostos funcionam em multiescala, mas os detalhes do contorno do objeto não desaparecem com o esquema utilizado. Nos itens *b* e *c* da Figura 4.3, os menores detalhes permanecem visíveis apesar da suavização com um núcleo bastante amplo. Porém, com a evolução multiescala, os seus valores de entropia indicam a perda gradual de proeminência. Este assunto será melhor discutido no capítulo 5, que apresenta o detector baseado em entropia.

Embora seja um assunto comum em processamento de imagens, é interessante mostrar como é feita a adaptação de uma distribuição Gaussiana a uma janela quadrada. Em probabilidade e estatística é dito que 99,7% dos pesos de uma distribuição Gaussiana se concentram na faixa de seis desvios-padrão (6σ), como pode ser visto na Figura 4.4. Isto significa que é “permitido” utilizar a extensão de 6σ como o diâmetro da distribuição. Para

descobrir o valor do desvio padrão (σ) para uma determinada janela quadrada, basta dividir a dimensão desta janela pelo diâmetro da distribuição. No entanto, nada impede que se ajuste a distribuição a uma janela maior que 6σ . Isto só aumentará a precisão do ajuste.

Figura 4.4 – Ajuste de distribuição Gaussiana em janela quadrada.

O esquema multiescala adotado utiliza sete escalas, as quais são produzidas pelos núcleos de convolução, do tipo SoG, listados na Tabela 4.1. Esta quantidade de escalas não é obrigatória, e tampouco fixa. Nada impede que sejam adotadas mais escalas, ou menos. Na realidade, é interessante que a escala máxima seja adaptativa à área do objeto. Por exemplo, poder-se-ia circunscrever um retângulo ao objeto, calcular a sua área em pixels, extrair a raiz quadrada e gerar a escala máxima a partir dessa informação.

Tabela 4.1 – Núcleos de convolução (SoG).

Núcleo (identificação)	Dimensões	$\sigma_{estreita}$	$\sigma_{periférico}$
H3	3x3	0,50000	0,65375
H5	5x5	0,43300	0,83333
H9	9x9	0,35450	1,50000
H17	17x17	0,30700	2,83333
H35	35x35	0,26900	5,83333
H71	71x71	0,23000	11,83333
H143	143x143	0,2286861	23,83333

A Figura 4.5 apresenta diferentes estágios de evolução na escala, isto é, suavizações de uma mesma imagem com os sete núcleos da Tabela 4.1. O objetivo desta figura é mostrar o fundo cedendo partículas de brilho ao objeto, sem que ocorra perda de localização ou destruição de detalhes da fronteira. A região do fundo que cede partículas se torna mais escura, e a região do objeto que as recebe se torna mais clara (a mesma ideia da difusão de calor). No canto superior esquerdo está a imagem original.

Figura 4.5 – Suavizações de uma imagem binária com os núcleos da Tabela 4.1.

Na Figura 4.6 é mostrada a distribuição de cones na retina, observe que a concentração na fóvea é muito maior do que na periferia, apesar da área periférica ser maior. Em outras palavras, o peso estatístico do centro é superior ao da periferia. Pode-se notar que essa descrição é bem semelhante ao núcleo proposto. Comparando as Figuras 4.6 e 4.7, não é difícil perceber a semelhança entre os núcleos de escala mais alta e a distribuição dos cones na retina. Provavelmente, um objeto observado a uma distância relativamente grande manterá sua forma reconhecível graças à alta concentração de cones na fóvea. Com esta distribuição é possível ter uma visão abrangente e detalhada simultaneamente. Por exemplo, a Figura 4.8 mostra objetos a uma distância razoavelmente grande, no entanto seus pequenos detalhes ainda se mantêm visíveis, como é o caso dos mastros das embarcações e os postes da ponte. A

Figura 4.9 é o resultado da suavização desta mesma cena com o núcleo H143 (Tabela 4.1). A Figura 4.10 mostra a suavização com uma Gaussiana na escala $\sigma = 2,8333$, que é bem menor que a escala de H143. Os mastros e os postes continuam visíveis na Figura 4.9, o que não ocorre na Figura 4.10. Isto está indicando que a restrição de pesos equiprováveis utilizada no núcleo proposto pode ter um fundamento biológico.

Figura 4.6 – Distribuição de fotorreceptores na retina. Adaptado de Osterberg (1935).

Figura 4.7 – Núcleos de convolução gerados por SoG.

Figura 4.8 – Imagem original de uma cena com objetos distantes.

Figura 4.9 – Versão suavizada com o núcleo H143 ($\sigma=23,8333$). Os objetos distantes ainda mantêm os seus detalhes visíveis.

Figura 4.10 – Versão suavizada com Gaussiana ($\sigma=2,8333$).

4.4 O núcleo proposto para a DoG

No capítulo 3 foi mencionado que a diferença entre Gaussianas mede a produção de suavização de uma escala para outra. Na secção acima foi apresentado um núcleo baseado em soma de Gaussianas (SoG), cuja característica é a de se manter íngreme em qualquer escala. Utilizando-se algum recurso capaz de medir o efeito de sua suavização, por exemplo, a entropia, obtém-se um método que permite estimar curvaturas. Percebe-se, então, que o uso da entropia e da operação de subtração, realizada na DoG, têm o mesmo objetivo de medir os efeitos provocados pela suavização. Além disso, os núcleos SoG e DoG são gerados pelos mesmos elementos, porém com sinais invertidos. Tais semelhanças levam a crer que a equiparação em magnitude do núcleo DoG com o SoG pode torna-lo imune à perda de localização e capacitá-lo a estimar curvaturas. Visto que os campos receptivos *center-surround* podem ser modelados com DoG, é possível que tenham a finalidade de detectar curvaturas ao longo de fronteiras entre regiões homogêneas.

O sucesso obtido com o núcleo SoG se deve à imposição de restrição de equiprobabilidade. A área de atuação da Gaussiana central, praticamente, se limita ao pixel do centro, isto é, todo o peso estatístico (0,5) está concentrado nesse pixel, garantindo a correta

localização de qualquer ponto da fronteira, independentemente do aumento da escala da Gaussiana periférica.

O sentido da palavra localização, mencionada acima, se relaciona ao grau de certeza de que o ponto observado pertence ao objeto. Porém, há outro sentido de maior importância, que é o da referência. Por exemplo, uma carta náutica mostra com certeza as regiões que pertencem à terra firme. No entanto, para o navegador determinar com boa precisão onde o seu navio se localiza naquele momento, ele não usa qualquer ponto de terra como referência. Os cálculos de navegação são baseados em pontos salientes de terra como pontas ilhas, ou de protuberâncias da costa. É necessário que seja uma característica que se destaque do resto, mesmo que seja algo artificial como um farol.

Há muitos exemplos que mostram a importância dos pontos de referência. Hipoteticamente, uma pessoa perdida no deserto do Saara, mesmo possuindo um meio de comunicação teria muita dificuldade em identificar sua posição a uma possível equipe de resgate, dada a homogeneidade do deserto. Caso essa pessoa conseguisse chegar ao litoral (fronteira entre duas regiões homogêneas), a busca se restringiria às regiões litorâneas do Atlântico, do Mediterrâneo ou do mar Vermelho, mesmo assim, muito tempo e energia seriam gastos devido à falta de um bom ponto de referência. Em um exemplo do mundo infantil, a característica principal de um mapa do tesouro é a identificação de uma grande quantidade de pontos de referência. Num exemplo menos fantasioso, uma empresa de entrega de mercadorias, mesmo de posse do endereço do destinatário, exige um ponto de referência.

Se a imagem da cena fosse conectada diretamente ao cérebro, toda a informação estaria ao seu dispor, isto é, ele estaria vendo a cena diretamente. No entanto, a arquitetura do sistema de visão é outra, na qual a retina é a responsável em fazer um resumo sobre as regiões e suas localizações na cena, e transmiti-lo ao cérebro. A retina possui bastante poder de fotorrecepção, porém sua capacidade de transmissão é cem vezes menor. Esta discrepância requer processos de compressão da informação. A detecção de contrastes, ou de fronteiras entre regiões homogêneas, efetuada pelos campos receptivos da retina, é uma maneira de reduzir a quantidade de informação. Porém, tal redução pode ser mais efetiva se for considerada a ideia de que a forma de um objeto pode ser aproximada através de pontos de alta curvatura (ATTNEAVE, 1954). Para isso, é necessário supor que os campos receptivos sejam especializados em detectá-los.

A retina não é capaz de identificar o que é o objeto, ou o que é o fundo. Sendo assim, neste nível não há como saber o que é uma convexidade ou concavidade. Porém, seus circuitos têm capacidade para detectar contrastes negativos e positivos. A Figura 4.11

apresenta os estímulos usados para produzir as respostas mínimas e máximas dos campos receptivos de células ganglionares *on-center* e *off-center* (HUBEL, 1995). Estas duas modalidades de célula ganglionar são dedicadas, respectivamente, a contrastes positivos (objeto claro em fundo escuro) e a contrastes negativos (objeto escuro em fundo claro). As respostas mínimas são obtidas quando as regiões central e periférica são excitadas com o mesmo estímulo (tudo escuro, ou tudo claro). A Figura 4.12 mostra o comportamento de uma célula ganglionar em relação à área ocupada pelo estímulo. Se, por exemplo, a célula for *on-center*, a resposta vai se tornando mais forte à medida que o estímulo ocupa uma fração maior da área central. Ao ocupar esta área por completo, a resposta é máxima. Se o estímulo continuar aumentando, invadindo a área periférica a resposta será atenuada, até atingir seu valor mínimo, que coincide com a iluminação total do campo receptivo (ou a escuridão total).

Figura 4.11 – Respostas dos campos receptivos *center-surround* de células ganglionares. As regiões com sinal positivo respondem a estímulos claros, e as com sinal negativo respondem à ausência de luz. As células do tipo *on-center* são representadas com sinal positivo na região central. As do tipo *off-center*, o sinal central é negativo.

Figura 4.12 – Respostas das células ganglionares em relação ao tamanho da área excitada pelo estímulo luminoso.

Em imagens binárias existe uma relação complementar entre as duas regiões que a formam. Com isto, uma convexidade do objeto pode ser interpretada como uma concavidade do fundo e vice-versa. A Figura 1.3 do capítulo 1 facilita entender essa dualidade. Considerando-se que ocorra sobreposição de campos receptivos na retina, esta característica complementar permite que se reduza ainda mais a quantidade de informação a ser transmitida ao cérebro, visto que apenas as convexidades necessitam ser transmitidas. A Figura 4.13 ilustra esta ideia. Os itens *a* e *b* apresentam a mesma vizinhança local de uma imagem, na qual não se sabe qual é a intensidade que representa o objeto, ou o fundo. O ponto vermelho situado, aproximadamente, no centro da vizinhança representa o centro de um campo receptivo. No item *a*, o campo receptivo é do tipo *on-center*. A resposta deste é composta pela ativação total do *center* somada à ativação quase total do *surround*. No item *b*, o campo é do tipo *off-center*, cuja resposta é a soma da ativação total do *center* com a atenuação quase total do *surround*. Os itens *c* e *d* desta mesma figura são explicados de forma semelhante. Isto sugere que a resposta mais forte é apresentada pelo campo receptivo cujo tipo (*on-center* ou *off-center*) coincide com o tipo de contraste (positivo ou negativo) da convexidade local. Deve-se observar que a transmissão da informação extraída das configurações representadas pelos itens *a* e *c* pode tornar desnecessária a transmissão da informação obtida em *b* e *d*.

É importante mencionar que os movimentos microsacádicos podem compor um mecanismo com o objetivo de ajustar a imagem aos campos receptivos, para obter a melhor resposta, conciliando o tipo do campo com o tipo de contraste. Afirma-se em Martinez-

Conde, Macknik e Hubel (2000) que campos receptivos do córtex estriado respondem ativamente após esses movimentos. É como se estivesse buscando pela melhor relação posição/resposta.

Figura 4.13 – Relação complementar entre convexidades e concavidades. Projeções de regiões contrastantes (convexas/côncavas) sobre campos receptivos *on-center* e *off-center* sobrepostos. Haverá redução de redundância se apenas os valores mais altos forem considerados. Do ponto de vista da convexidade (itens a e c), as saídas são mais altas.

A importância que os pontos de referência possuem em descrever a localização de algo; a capacidade que um ponto de alta curvatura tem em servir como referência de uma vizinhança local; a semelhança entre os modelos SoG e o DoG e a capacidade de compressão da informação que os pontos de alta curvatura oferecem, permitem sugerir que os campos receptivos *center-surround* tenham a competência de distinguir vértices de concavidades, de convexidades e pontos pertencentes a fronteiras em linha reta. Ou seja, o modelo DoG tem a capacidade latente de distinguir curvaturas nas duas modalidades de contraste (positivo e negativo).

Supõe-se, então, que em consequência⁹ da estreiteza do nervo óptico, a imagem apresentada aos circuitos da retina necessita ser varrida em busca de contrastes, os quais poderão servir como pontos de referência (pontos de alta curvatura) para capacitar o cérebro a construir a sua versão da cena observada.

⁹ Foi utilizada a palavra consequência, mas é possível que a detecção de contraste tenha ditado a largura do canal de comunicação e não o contrário.

Um ponto de referência necessita ser diferente do resto de sua vizinhança, por isso, o processo de suavização deve ser usado com cautela, visto que esta tende a igualar os vizinhos em consequência de seu efeito de achatamento.

Em visão computacional, os detectores de borda derivativos buscam por pontos distintos através do processo de diferenciação. Alguns embutem um processo de suavização como parte da detecção. O Laplaciano da Gaussiana (LoG), por exemplo, faz parte dessa classe de detectores, o qual utiliza a suavização Gaussiana para filtrar ruídos de alta frequência gerados pela diferenciação (MARR; HILDRETH, 1979).

A observação das Figuras 4.3, 4.5, 4.8 - 4.10 faz surgir a questão: as células da retina, especificamente as que compõem a região da fóvea, utilizam a suavização apenas para a filtragem de ruídos?

Talvez seja necessária uma nova interpretação para a fase de suavização do modelo DoG, a fim de harmonizá-la com a ideia de ponto de referência distinguível. Talvez a palavra marcar deva substituir a palavra suavizar. Os pesquisadores da área biológica e os profissionais da área médica costumam lidar com processos de marcação. O objetivo é detectar padrões específicos excitando-se alguma propriedade do tecido em estudo. Pode-se citar, por exemplo, aplicação de contraste para exames de ressonância magnética e marcadores biológicos para identificação de células ou características genéticas. Inspirando-se nesta ideia, sugere-se que a suavização Gaussiana possa agir como um processo de marcação, o qual é imprescindível para que ocorra a detecção de contraste.

Metaforicamente, a suavização Gaussiana funcionaria como um “pulverizador de brilho adaptativo”. Para cada vizinhança local da imagem seria construída uma intensidade de brilho própria, utilizando-se a média ponderada obtida pelo produto interno entre a janela Gaussiana e o brilho atual da vizinhança. Não há qualquer inovação neste procedimento: o pixel de saída é gerado a partir da média ponderada da vizinhança de entrada. No entanto, deve-se enfatizar que o valor do pixel de saída depende da configuração espacial da vizinhança de entrada (da intensidade de cada pixel que forma a vizinhança). Se esta for uma região homogênea, o pixel de saída terá o mesmo valor do pixel de entrada. Porém, se a vizinhança de entrada estiver centrada na fronteira entre duas regiões homogêneas, o valor do pixel de saída refletirá a configuração espacial dessa localidade da fronteira.

No capítulo 3 é feita a associação entre a difusão térmica e a suavização de imagens. O calor flui de uma região quente para uma região adjacente de menor temperatura. Esta se torna mais quente e a primeira mais fria. De maneira similar a “pulverização Gaussiana” ocorre da região clara para a região escura. Obedecendo a “conservação de brilho”, o pixel

que cede brilho se torna mais escuro, e o pixel que o recebe se torna mais claro. A quantidade de brilho recebido ou cedido reflete o tipo de vizinhança do pixel. Se for uma região homogênea, não haverá perda ou ganho. Se for região de fronteira, a perda, ou o ganho de brilho dependerá da configuração espacial da vizinhança.

Uma vez que cada pixel da imagem foi marcado de forma adaptativa, resta determinar a quantidade de marcação que cada um recebeu (do ponto de vista do contraste negativo), ou cedeu (do ponto de vista do contraste positivo). Então, o processo de diferenciação na DoG seria a quantificação de quanto brilho cada pixel cedeu ou recebeu. A Figura 4.11 mostra que um campo receptivo *center-surround* responde fortemente a regiões contrastantes e fracamente a regiões homogêneas, isto significa que esses campos “gostam” de diferenças, e estas precisam ser fortes o suficiente para gerar uma resposta do campo.

No modelo DoG tradicional não há qualquer restrição às escalas das Gaussianas que serão subtraídas. A subtração pode ser feita, por exemplo, entre duas versões fortemente suavizadas de uma imagem. Para a função de marcação aqui proposta, há dois problemas com esse modelo. O primeiro, e mais óbvio, é que ambas as versões já perderam a localização de seus pontos de fronteira, e a diferença entre elas continuará a apresentar essa perda. O segundo se relaciona à diminuição da diferença absoluta entre as versões suavizadas. Para que a marcação funcione da melhor maneira possível, o valor modificado deve ser comparado com o seu valor original. Isto é, o interessante é comparar a versão suavizada da imagem com a sua versão original.

Na fôvea, o centro do campo receptivo das células ganglionares do tipo *midget* é composto por um único cone (KAPLAN; LEE; SHAPLEY, 1990). Caso este cone concentre, pelo menos, a metade do peso sináptico do campo receptivo, então será possível inferir que as células ganglionares do tipo *midget* têm a configuração ideal para comparar as versões suavizada e original de uma imagem.

Devido ao modelo de diferença entre Gaussianas adotado nesta tese exigir que a escala de uma delas tenha valor próximo à zero (imagem original), este será referido por DoG_zero. Este compõe o núcleo central de um dos detectores de *corners* de contorno aqui propostos. A ideia básica de funcionamento do modelo DoG_zero é suavizar uma imagem binária com uma Gaussiana de qualquer escala e, então, efetuar a subtração entre as versões original e suavizada. A Figura 4.14 compara a detecção de um vértice entre a DoG e a DoG_zero. Para a versão DoG, a subtração foi realizada entre as imagens suavizadas com $\sigma = 2,833$ e com $\sigma = 1,5$. Para a versão DoG_zero, a subtração foi entre a imagem suavizada com $\sigma = 2,833$ e a

imagem original ($\sigma \cong 0$). Nesta figura, um dos vértices do retângulo tem as coordenadas espaciais 103,70. Deve-se perceber: 1) a perda de localização gerada pela suavização (item a) e a subsequente recuperação da localização promovida pela DoG_zero (item e); 2) a versão DoG_zero produz um vértice com valor alto e distinto (0,67) (item e). Na versão DoG o valor do vértice é muito baixo, menor que outros valores do objeto (item c).

Figura 4.14 – Comparação de detecção de vértices entre DoG e DoG_zero.

A versão DoG não é capaz de recuperar a localização, exceto se uma das imagens for suavizada com escala bastante reduzida. Como pode ser visto na Tabela 4.1, ao se aumentar a escala da Gaussiana periférica, é necessário reduzir a escala da Gaussiana central, a fim de que se mantenha a equiprobabilidade de pesos probabilísticos entre o centro e a periferia da SoG. Por exemplo, o núcleo H143 é composto por uma Gaussiana central cuja escala é $\sigma = 0,228$. Isto significa que a Gaussiana ocupa uma área menor do que 1,5 pixels. Este σ não causa mudanças na imagem. Sendo assim, não era necessário que se construíssem núcleos de convolução, bastaria somar a imagem original com uma versão suavizada qualquer (Gaussiana de qualquer escala), seguida por uma normalização (dividir cada intensidade por

dois), para que se obtivessem resultados semelhantes. No entanto, o núcleo SoG foi construído para ser usado num detector que utiliza os pesos do SoG para computar a entropia local.

A DoG_zero também poderia ser construída na forma de núcleos de convolução. Para isso, seriam utilizadas as Gaussianas listadas na tabela 4.1, porém em vez de somá-las, deve-se subtraí-las. No entanto, a informação que se deseja extrair é fornecida diretamente pela diferença, o que torna mais prático e mais fácil suavizar a imagem com uma Gaussiana, seguida da operação de subtração entre este resultado e a imagem original.

4.5 Conclusão do capítulo

Neste capítulo é mostrada a importância da suavização na detecção de curvaturas ao longo de fronteiras entre regiões homogêneas. A qual é imaginada como sendo um processo de marcação, a fim de facilitar a detecção e distinção dos valores de contraste. Porém, para que a suavização possua essa propriedade de marcação de pixels é necessário que a localização destes não seja perdida.

Foram identificadas duas maneiras de se quantificar a suavização efetuada. Uma é através do cálculo da entropia de Shannon e a outra através da diferença de Gaussianas (DoG).

Independente do método de quantificação utilizado é necessário manter a localização dos pixels intacta, não importando a escala considerada. Para o método de entropias foi proposto um núcleo construído a partir da soma de duas Gaussianas (SoG). Uma delas, a de maior abertura, é quem determina a escala e a outra, de menor abertura (quase sem efeito de suavização), é responsável em manter a localização dos pixels. Esta ideia é estendida para o modelo DoG, através da restrição de que uma das imagens não deve sofrer suavização, isto é, a diferença é realizada entre a versão suavizada da imagem e a imagem original.

Em meio à discussão do modelo DoG foram feitas alusões aos campos receptivos da retina, os quais poderiam ter capacidade de detectar curvaturas, o que auxiliaria na redução da quantidade de informação a ser transmitida ao córtex visual. Também foi dito que as duas vias de transmissão, baseadas em contraste positivo e negativo, podem estar relacionadas com a dualidade convexidade/concavidade existente em imagens binárias, e que isto reduziria ainda mais a quantidade de informação a ser transmitida, visto que os contrastes de convexidades produzem as respostas mais fortes, tornando desnecessária a transmissão de contrastes de concavidades. Para obter essas respostas mais fortes é necessário que os vértices das

convexidades coincidam com os centros dos campos receptivos apropriados para aquele contraste. Foi sugerido que a sobreposição dos campos receptivos e os movimentos micro sacádicos contribuem para que ocorra essa coincidência.

Capítulo 5 – Proposta do detector EML

5.1 Introdução

Como mencionado nos capítulos 3 e 4, o processo de difusão em imagens binárias (silhuetas), inicialmente, produz modificações (mistura das cores do fundo e do objeto) mais acentuadas ao longo da fronteira entre o objeto e o fundo. Também foi visto que tais modificações têm maior ou menor intensidade de acordo com a configuração espacial da fronteira. Os pixels do contorno de uma convexidade do objeto sofrem maior modificação que aqueles que estão numa fronteira em linha reta. Estes, por sua vez, sofrem maior modificação que os pixels de uma concavidade.

O detector de *corners* proposto neste capítulo se fundamenta na ideia de que a entropia pode ser usada para quantificar as modificações causadas por suavização. Este detector possui o seu próprio esquema multiescala, que o torna capaz de identificar os *corners* mais importantes ao longo do contorno. O núcleo multiescala utilizado pelo detector foi proposto no capítulo 4.

Vale mencionar que existem muitos trabalhos que utilizam o cálculo de entropia para analisar ou detectar determinadas características de uma imagem (CHAMOLI; KUKREJA; SEMWAL, 2014). Um trabalho que pode trazer dúvidas quanto à originalidade do detector aqui proposto é o relatado em Kadir e Brady (2001). O referido trabalho não relaciona valores angulares com valores de entropia, não há a preocupação com análise de formas e a entropia é calculada a partir da distribuição local dos níveis de cinza, ou seja, há uma distribuição diferente para cada vizinhança. No detector aqui proposto, a entropia é calculada sobre a distribuição de probabilidades imposta pelo núcleo de convolução. A cada pixel do contorno da imagem é associado um valor de entropia, o qual está diretamente relacionado com o ângulo local onde está situado o pixel.

Neste capítulo é apresentado um detector de pontos dominantes multiescala cujo funcionamento se baseia na quantificação da suavização através da entropia de Shannon. Para facilitar as referências, o detector é chamado de EML - Entropia da Mistura Local. O termo mistura vem da combinação das intensidades do objeto e do fundo em virtude da suavização.

Basicamente, o EML é composto de quatro fases. São elas:

1. Computação da entropia na escala mais baixa – O núcleo de convolução de escala mais baixa, ajustado a uma janela 3x3, é deslocado pixel a pixel pela imagem. A distribuição de probabilidades do núcleo é usada para determinar as probabilidades de fundo e de objeto contidos na vizinhança 3x3. Essas probabilidades são usadas para computar a entropia de cada pixel da imagem. Os pixels da imagem resultante com valores de entropia característicos de convexidades e concavidades serão considerados candidatos a pontos dominantes.
2. Computação multiescala dos candidatos – As coordenadas dos candidatos, obtidos na fase anterior, são usadas como ponteiros para que apenas estes sejam observados em escalas mais altas. Assim, para cada candidato é computada a entropia utilizando-se núcleos maiores. Então, para cada candidato é produzida uma lista contendo os seus valores de entropia computados em uma quantidade finita de escalas.
3. Análise da evolução do candidato através das escalas – A lista contendo as entropias do candidato é analisada com o objetivo de verificar possíveis decaimentos de saliência. Nesta fase são separados os verdadeiros candidatos daqueles que representam pequenos detalhes ou ruídos. A saída desta fase é uma lista dos possíveis pontos dominantes com seus respectivos valores de entropia e coordenadas espaciais.
4. Determinação dos pontos dominantes – O objetivo é determinar as regiões que possuem mais de um provável ponto dominante e compará-los para que reste apenas um em cada região.

É necessário considerar a restrição imposta pela fronteira entre duas regiões homogêneas às possíveis configurações espaciais locais. Isto é, nem todas as combinações de pixels brancos e pretos são válidas na região do contorno. Para determinar quais são as possíveis configurações, foi feita uma análise das possíveis combinações em uma matriz 3x3. Do total de 512 (2^9) combinações diferentes, a grande maioria não pôde ser considerada como uma fronteira entre objeto e fundo. As configurações possíveis estão representadas na Figura 5.1 e seus valores listados na Tabela 5.1. Deve-se notar que essas configurações podem sofrer rotação em torno do pixel central, mantendo seus valores de entropia inalterados.

A ideia original era a de promover o mesmo estudo em janelas maiores, porém mostrou-se impraticável devido a grande quantidade de configurações a serem analisadas. Optou-se, então, por avaliar em escalas mais altas as mesmas configurações mostradas na

Figura 5.1. Os valores resultantes estão listados na Tabela 5.1. Este conjunto de valores é utilizado como referência para auxiliar na terceira fase do EML, quando se avalia a evolução da entropia do candidato através das escalas. Este conjunto é denominado, aqui, por ângulos padrão.

Na Figura 5.1 percebe-se a existência de ângulos (90° e 180°) que possuem dois valores distintos de entropia. Isto ocorre devido à grade de amostragem não ser isotrópica, tornando a representação na direção diagonal diferente das representações nas direções horizontal e vertical. As versões na diagonal recebem o sufixo “Diag”, como pode ser visto na Tabela 5.1.

A Figura 5.2 representa mais claramente o efeito da grade de amostragem. O ajuste do núcleo de convolução a uma grade digital pode gerar comportamentos indesejáveis na relação ângulo-entropia. Quanto menor forem as dimensões da grade, maior será a diferença entre as entropias da borda horizontal e da borda diagonal. Isto acontece devido à baixa granularidade dos pesos na baixa escala. O aumento de escala promove o aumento da diversidade de pesos probabilísticos, o que permite uma melhor aproximação do núcleo Gaussiano ao mundo discreto.

A Figura 5.3 mostra a Tabela 5.1 na forma de gráfico. Nesta, Pode ser percebido que os ângulos formados por diagonais (90° Diag, 180° Diag e 270°) não têm um bom comportamento inicial. Por esse motivo, a avaliação da evolução nas escalas (fase 3) exclui a escala inicial. Também pode ser percebido que o aumento da escala atenua o problema da falta de isotropia da grade de amostragem.

Tabela 5.1 – Ângulos padrão e suas entropias em 7 escalas diferentes.

Ângulo	Entropia de Referência						
	3x3	5x5	9x9	17x17	35x35	71x71	143x143
(1) 45°	0,95443	0,95765	0,96815	0,97803	0,98326	0,98604	0,98738
(2) 90° Diag.	0,94276	0,92374	0,92792	0,93879	0,94595	0,95012	0,95226
(3) 90°	0,83405	0,87514	0,90695	0,93002	0,94184	0,94820	0,95134
(4) 135°	0,81127	0,81827	0,84194	0,86607	0,88006	0,88797	0,89199
(5) 180° Diag.	0,78679	0,74882	0,75701	0,77870	0,79336	0,80209	0,80664
(6) 180° H/V	0,58397	0,65780	0,71666	0,76115	0,78491	0,79806	0,80467
(7) 225°	0,54356	0,55595	0,59803	0,64134	0,66676	0,68129	0,68871
(8) 270°	0,50021	0,43271	0,44733	0,48584	0,51185	0,52730	0,53535

Figura 5.1 – Ângulos padrão em uma janela 3x3.

Figura 5.2 – Influência da grade de amostragem. a) aspectos da borda em diagonal e horizontal; b) núcleo de convolução: a soma dos pesos em amarelo (0.2351) corresponde ao fundo quando o núcleo está centrado sobre uma diagonal, e a soma dos pesos em rosa (0.1399) corresponde ao fundo quando o núcleo está centrado sobre uma borda horizontal (ou vertical); c) o motivo por haver pesos diferentes é que uma fração de cada um dos pixels mais externos na direção diagonal não está incluída na distribuição circular.

Figura 5.3 – Entropia dos ângulos padrão por escala.

5.2 Fase 1 – computação da entropia na menor escala

As imagens utilizadas são binárias, a região escura (nível 0) representa o objeto e a região clara (nível 255) representa o fundo. O tipo de dados usado para representar os valores das imagens é o tipo real (*double*). Antes de executar qualquer processamento na imagem é feito um *padding* para permitir que os pixels que estejam nos limites da imagem sejam trabalhados pelo núcleo de convolução. A construção do núcleo está descrita no capítulo 4.

O núcleo é deslocado pixel a pixel, varrendo toda a imagem. Em cada vizinhança delimitada pela janela 3x3, o operador determina quais são os pixels pertencentes ao objeto e quais os pertencentes ao fundo. No EML, isto foi implementado como uma operação lógica “E” entre a matriz da vizinhança (w) e o núcleo (H), produzindo a matriz lógica w_l .

$$w_l = w \wedge H$$

Se todos os valores de w_l forem de fundo (0), ou de objeto (1), será fornecido automaticamente o valor zero à entropia do pixel de saída. Para detectar essas situações, basta somar os valores de w_l , como é mostrado nas seguintes expressões:

$$\begin{aligned}
Total &= \sum_{i=1}^n \sum_{j=1}^n w_{l_{i,j}} \\
Max &= n^2 \\
se \ Total = 0 \ ou \ Total = Max \\
Entropia &\leftarrow 0
\end{aligned}$$

Quando a vizinhança envolvida pelo núcleo de convolução for mista, isto é, possuir pixels de fundo e de objeto, torna-se necessário contabilizar as concentrações de cada tipo para usá-las na equação de entropia de Shannon (3.3). As expressões a seguir descrevem como as probabilidades de objeto (p_{objeto}) e fundo (p_{fundo}) foram obtidas. O operador “ $\cdot*$ ” representa uma multiplicação matricial ponto a ponto, conhecida como produto de Hadamard. Visto que w_l é uma matriz de valores lógicos, a multiplicação ponto a ponto com o núcleo (H) irá evidenciar os pesos estatísticos referentes ao fundo. Para descobrir o peso total do fundo, basta somar os valores da matriz resultante w_p . Sabendo-se que o núcleo (H) é normalizado, o peso do objeto é o complemento do peso do fundo. As palavras peso, probabilidade e concentração estão sendo usadas, aqui, como sinônimos.

$$\begin{aligned}
w_p &= w_l \cdot* H \\
p_{fundo} &= \sum_{i=1}^n \sum_{j=1}^n w_{p_{i,j}} \\
p_{objeto} &= 1 - p_{fundo} \\
Entropia &= -[p_{fundo} \cdot \log_2 p_{fundo} + p_{objeto} \cdot \log_2 p_{objeto}]
\end{aligned}$$

A imagem de saída do processamento descrito acima será o contorno do objeto. Aliás, serão apresentados dois contornos paralelos. O mais externo representa a entropia a partir do ponto de vista do fundo. O mais interno representa a entropia a partir do ponto de vista do objeto. O contorno de interesse é este último. Para se livrar do contorno externo, basta subtrair a imagem original da imagem entropia. A Figura 5.4 mostra resultado deste processamento aplicado à imagem de um quadrado preto em fundo branco. Nesta são mostradas as imagens de saída com dois contornos e a imagem final contendo apenas o contorno interno, o qual é formado pelos pixels limítrofes do objeto. Vale mencionar que os números mostrados na ampliação são os valores de entropia, os quais correspondem aos ângulos 180° e 90° da Tabela 5.1.

Figura 5.4 – Imagem entropia gerada pelo EML na escala mais baixa de operação. A parte superior é a imagem resultante com contorno duplo, e sua região de vértice ampliada. A parte inferior contém apenas o contorno do objeto e a ampliação da região do vértice.

Após a detecção dos pixels de maior entropia, os quais compõem o contorno do objeto, resta selecionar aqueles que representam os vértices das concavidades e convexidades existentes nesta escala. Os dois intervalos de valores de entropia que representam essas saliências de contorno são respectivamente:

$$0,58 > H(x, y) > 0,50$$

$$1 > H(x, y) > 0,787$$

O produto final desta fase é uma lista contendo os valores compreendidos nesses intervalos, bem como suas coordenadas espaciais. Esses pontos são os candidatos a pontos dominantes.

5.3 Fase 2 – computação multiescala dos candidatos

O esquema multiescala adotado no EML aplica-se somente aos candidatos a ponto dominante. Cada candidato é observado em seis escalas adicionais, suas entropias são computadas e anexadas à lista de candidatos criada na primeira fase.

A janela quadrada de dimensões 3x3 é a menor janela que apresenta simetria em torno do pixel central (possui dimensões ímpares). Esta abriga o núcleo de menor escala ($\sigma = 0,5$) usado no EML. Adotou-se o critério de que a janela do núcleo seguinte deveria ser (aproximadamente) o dobro da anterior. Esta decisão não é baseada cientificamente, apenas define uma padronização. No entanto, não se pode negar que houve influência das pirâmides Gaussianas e Laplacianas descritas, por exemplo, em Burt e Adelson (1983) e Lowe (2004). Embora, ao contrário destas, a pirâmide aqui utilizada é criada a partir do ponto de vista do tamanho da abertura de observação, e não do tamanho da imagem que se reduz com a distância. Assim foram gerados núcleos para as seguintes janelas: 5x5, 9x9, 17x17, 35x35, 71x71 e 143x143, cujas escalas são, respectivamente, 0,8333; 1,5; 2,8333; 5,8333; 11,8333 e 23,8333. A relação janela-escala foi discutida no capítulo 4 e obedece à relação “*largura_da_janela = 6 σ* ”. As dimensões da janela de maior escala (143x143) correspondem, aproximadamente, à metade do tamanho médio das imagens que foram utilizadas nos experimentos (em torno de 300 x 300 pixels).

Nesta fase o funcionamento do EML se restringe a:

1. Utilizar as coordenadas dos candidatos para localizá-los na imagem original;
2. Gerar suas vizinhanças de acordo com as dimensões das janelas dos núcleos e
3. Computar suas entropias seguindo o método usado na primeira fase.

5.4 Fase 3 – análise da evolução dos candidatos através das escalas

Antes de iniciar a explicação do funcionamento da fase 3, é necessário que se faça algumas observações a respeito das distorções produzidas pela acomodação de objetos à grade digital.

Ao observar o comportamento do gráfico apresentado na Figura 5.3, percebe-se que a entropia de uma mesma configuração angular varia de uma escala para outra. A partir da equação da entropia de Shannon (SHANNON, 1948), a qual é baseada em probabilidades, deduz-se que tal comportamento não está correto. No EML, as probabilidades são traduzidas

como quantidades de preto e branco contidas em uma vizinhança circular. Quando esta vizinhança está centrada na fronteira, obrigatoriamente, haverá duas partes, dois setores. Um preto e outro branco. A expansão ou contração de tal vizinhança, mantendo a sua configuração angular constante, não pode alterar a razão entre os setores, ou melhor, o peso probabilístico de cada setor não pode ser alterado. Isto está ilustrado na Figura 5.5. Tal ideia parece não estar de acordo com os valores crescentes de entropia de uma escala para outra. Qual será a causa dessa discrepância?

Ao se observar o comportamento das entropias dos ângulos 90Diag e 90°, e dos ângulos 180Diag e 180°, percebe-se que quanto maior é a escala, mais próximas se tornam as entropias das versões diagonal e horizontal/vertical. Quanto maior é a escala, maior é o tamanho do núcleo e consequentemente maior é a quantidade de pixels para representá-lo. Como cada pixel representa uma subdivisão para abrigar um peso probabilístico, a precisão na representação do núcleo aumenta. Em outras palavras, há uma maior granularidade dos pesos, permitindo um melhor ajuste dos mesmos à forma circular. Então, o aumento de escala causa aumento de precisão na aproximação da distribuição circular. Este mesmo efeito é válido para a aproximação dos setores circulares.

Figura 5.5 – Diferentes escalas de uma mesma configuração angular. Não importa a escala, a área do setor preto equivale a $\frac{1}{4}$ da área do círculo.

Supõe-se, então, que essa variação nos valores de entropia é causada pela acomodação à grade digital. Assim, decidiu-se eliminar esse efeito. Para isso, assume-se que os valores de entropia para os ângulos padrão devam se manter constantes através das escalas. Isto pode ser feito, por exemplo, elegendo-se o valor da entropia da última escala como o “valor correto”, obrigando todos os valores obtidos nas escalas anteriores a serem “empurrados” para cima

através da multiplicação por um fator. Este é obtido pela razão entre a entropia da última escala (143×143) e cada uma das entropias das escalas anteriores (5×5 a 143×143). A razão é computada para cada ângulo padrão separadamente. Esta correção foi efetuada nos ângulos padrão que representam convexidades.

Para os ângulos padrão que representam concavidades fez-se o inverso. Os valores foram puxados para baixo, igualando-os ao valor obtido na escala 5×5 ($\sigma = 0,8333$). A Tabela 5.2 apresenta os fatores de correção para os ângulos padrão.

Os valores de entropia dos candidatos são corrigidos com base na classificação feita na escala 3×3 ($\sigma = 0,5$). Por exemplo, se um candidato detectado na escala 3×3 possui um valor de entropia igual ao do ângulo padrão 135° , então seus valores de entropia em escalas mais altas serão corrigidos com o fator 135 da Tabela 5.2.

Tabela 5.2 – Fator de correção de entropia por ângulo e escala.

Fator de correção	5x5	9x9	17x17	35x35	71x71	143x143
Fator45	1,031045	1,019863	1,00956	1,00419	1,001359	1
Fator90D	1,030874	1,026231	1,014348	1,006671	1,002252	1
Fator90	1,087072	1,048944	1,022924	1,010087	1,003312	1
Fator135	1,090093	1,059446	1,029928	1,013556	1,004527	1
Fat180D	1,077215	1,065561	1,03588	1,016739	1,005673	1
Fat180HV	1	0,917869	0,864219	0,838058	0,824249	0,817478
Fator225	1	0,929636	0,866857	0,833808	0,816025	0,807234
Fator270	1	0,967317	0,890643	0,845384	0,820614	0,808275

Aqui, a análise multiescala se destina a caracterizar a saliência de um ponto (candidato) baseando-se nas suas variações de entropia ao longo das escalas. O aumento de escala aumenta a vizinhança incluindo novos pixels, o que produz modificações em sua configuração angular.

Tais modificações ocorrem até mesmo em figuras geométricas bem definidas. Seja, por exemplo, o vértice do retângulo preto sobre fundo branco mostrado na Figura 5.6, onde os anéis concêntricos representam as escalas de observação. Percebe-se que do menor anel até aquele identificado pelo número 4, os setores circulares permanecem constantes em 90° (o setor de cor preta). Porém, o quinto e o sexto anéis apresentam um aumento considerável de área branca, alterando as concentrações (ou probabilidades) das duas cores e,

consequentemente, modificando o valor de entropia. O gráfico da Figura 5.7 mostra essa variação refletida nos valores de entropia (já corrigidos pela Tabela 5.2).

Figura 5.6 – O vértice de um retângulo observado através de seis janelas circulares. As janelas 5 e 6 apresentam variações na configuração angular em relação às quatro janelas menores, as quais se mantêm constantes em 90° .

Figura 5.7 – Variação de entropia do vértice observado na Figura 5.6.

É interessante notar que o aumento de espaço (fundo branco), que ocorre nos anéis 5 e 6 da Figura 5.6, se assemelha ao aumento de volume de um recipiente. É comum encontrar na literatura sobre entropia, por exemplo, em Ben-Naim (2007), experimentos indicando que o aumento de volume de um recipiente contendo gás ideal produz aumento de entropia. Tanto na imagem, quanto no gás, o aumento de entropia acontece devido ao aumento de microestados. O item 3.3 pode esclarecer tal ideia.

O objetivo desta terceira fase é analisar a evolução da entropia de cada candidato, a fim de obter dados válidos que permitam realizar um julgamento justo a respeito de sua saliência.

Com o aumento de escala, uma estrutura pode revelar-se como um pequeno detalhe, que rapidamente perde a sua saliência, ou como uma característica com importância mais global, a qual mantém a sua saliência por várias escalas.

Um ruído, ou pequeno detalhe, situado numa borda em linha reta apresenta forte saliência nas escalas mais baixas, porém sua saliência é drasticamente reduzida com o aumento da escala. Neste caso o aspecto global da região (reta) facilita a detecção do ruído.

No entanto, os “ruidos” podem aparecer em qualquer tipo de região. Quando o ruído e a região possuem comportamentos semelhantes, por exemplo, quando um ruído convexo está localizado numa convexidade, a sua saliência é intensificada com o aumento de escala. Em casos extremos, a saliência do ruído pode superar a do ponto dominante verdadeiro. O mesmo pode ocorrer com pequenas estruturas côncavas situadas em uma concavidade mais global.

Assim, por mais sutis que possam parecer, deve-se analisar todas as modificações causadas na saliência (variações de entropia) de uma estrutura. Isto significa que as informações pertencentes às diferentes escalas são importantes para escolher os bons candidatos a pontos dominantes. Isto tornou necessário analisar o comportamento de cada candidato e confrontá-lo com o comportamento de pontos dominantes reais obtidos via *ground-truth*¹⁰. A seguir são apresentados alguns exemplos de como a inclusão de informações de escalas mais altas podem interferir na saliência das estruturas.

A Figura 5.8 compara a evolução de dois pontos dominantes convexas e seus respectivos vizinhos, que são pequenas estruturas convexas situadas em regiões convexas. Em d , o eixo das escalas inicia na escala 5x5 representada por 1 e termina na escala 143x143 representada por 6. O ponto dominante1, na escala 3x3 (não representada) possui configuração angular idêntica ao ângulo padrão 90Diag (item 2 da Figura 5.1). A entropia desta configuração inicial evoluiu suavemente para um valor, que representa o aspecto global da região, um ângulo em torno de 73° (medido sem muita precisão). Na vizinhança do dominante1 encontra-se o ruído1, o qual, na escala 3x3, foi classificado como um ângulo de 90°. Esta pequena estrutura tende a perder a sua saliência, pois é apenas um detalhe em uma borda em linha reta. Sua entropia deveria decair até o valor que representa 180°, porém em meio ao decaimento, a entropia cresce até atingir um valor que representa um ângulo mais

¹⁰ *Ground-truth* ou verdade absoluta são pontos marcados por observadores humanos. Em geral, são eleitos os pontos que receberam o maior número de votos.

agudo que 90° . Isto ocorre devido à influência exercida pela configuração global da região. Em geral, um ruído tende a perder sua saliência, mesmo momentaneamente, antes de tirar proveito das características globais da região observada.

O par (determinante2, ruído2) possui comportamento semelhante ao do par anterior. O determinante2 foi classificado inicialmente como um ângulo de 135° , por isso tem um crescimento acentuado da 1ª à 3ª escala, e a partir desta começa a estabilizar no ângulo global da região (75°).

Figura 5.8 – Evolução multiescala de convexidades: ponto dominante versus ruído. a) Imagem original com marcações de duas regiões convexas; b) e c) Ampliações das regiões convexas indicadas em a; d) Evolução em escala dos pontos dominantes e seus respectivos vizinhos ruidosos.

A Figura 5.9 mostra a perda de saliência de pequenas estruturas convexas situadas em bordas retas. O gráfico mostra, até à quinta escala (71×71), um comportamento decrescente e idêntico para os dois pontos rotulados como ruído1 e ruído2. Desta escala em diante os valores voltam a crescer e com intensidades distintas. O crescimento ocorre devido à região em linha reta apresentar, superiormente, uma mudança em sua inclinação. Numa escala superior à mencionada, a vizinhança é composta, também, por parte desta região inclinada, o

que reduz a proporção de pixels pretos em relação aos pixels brancos. O resultado é uma entropia mais alta. Visto que o ruído1 está mais próximo dessa inclinação, os pesos ocupados pelos pixels brancos são ligeiramente mais significativos¹¹.

Figura 5.9 – Evolução multiescala de ruídos convexos localizados em borda em linha reta.

Na Figura 5.10 são comparados os comportamentos multiescala do ponto dominante de uma região côncava e de um ruído côncavo vizinho. Inicialmente, na escala 3x3 (não representada), o ponto dominante foi classificado como uma configuração angular de 270° (item 8 da Figura 5.1). Com o aumento da escala a configuração local foi se ajustando à configuração global da região côncava, a qual possui um ângulo aproximado de 290°.

Em relação à pequena estrutura côncava (ruído), esta é limitada pelos pixels rotulados como ruído1 e ruído2. Ambos podem servir como vértice da estrutura. Inicialmente, ambos foram classificados como vértices de uma configuração angular de 225° (item 7 da Figura 5.1). O aumento paulatino da escala incluiu mais pixels brancos do que pretos, fazendo com

¹¹ Isto ocorre devido à distribuição do núcleo ser circularmente simétrica, com os valores dos pesos diminuindo do centro para a periferia.

que a entropia aumentasse e se aproximasse cada vez mais da configuração de 180° (reta). Porém, em torno da quarta escala, a janela do núcleo de convolução iniciou a inclusão de pixels pretos pertencentes ao lado oposto da abertura, o que trouxe para baixo os valores de entropia.

No caso de concavidades, o aumento de entropia significa perda de saliência. A perda de saliência inicial pode indicar que o ponto observado se trata de um ruído.

Figura 5.10 – Evolução multiescala de concavidades: ponto dominante versus ruído.

As Figuras 5.6 a 5.10 demonstram o comportamento comum de estruturas submetidas ao processamento multiescala. Em geral, uma estrutura mantém a sua configuração angular, mais ou menos, inalterada enquanto as suas dimensões forem maiores, ou iguais às dimensões do núcleo de convolução. Quando o núcleo se torna maior que a estrutura, esta tende a se “ajustar” à configuração angular da sua região hierarquicamente superior. Isto é, o aspecto da região maior, onde se situa a estrutura, é quem ditará o comportamento nas próximas escalas de observação. Por exemplo, na Figura 5.9 as pequenas estruturas convexas só se mantiveram “ativas” na escala inicial. Nas escalas seguintes, o decaimento de saliência foi ditado pela região em linha reta onde se situavam tais estruturas.

Embora em muitos casos a estrutura seja maior, ou do mesmo tamanho que o núcleo de convolução, o percurso multiescala dos valores de entropia pode destoar do esperado. Percebeu-se que além da interferência hierárquica mencionada no parágrafo acima, existe a interferência entre estruturas adjacentes, que altera a quantidade de pixels de objeto envolvida pelo núcleo de convolução, e consequentemente altera os valores de entropia. Na imagem usada pelas Figuras 5.8-5.10 a interferência adjacente é mínima, o que é mostrado na Figura 5.11.

Figura 5.11 – Interferência de regiões adjacentes no comportamento multiescala.

Na Figura 5.11, a circunferência inscrita em um quadrado representa o núcleo de convolução ajustado à janela quadrada 143x143, que é a maior escala utilizada pelo EML. Esta janela está centrada no vértice de uma região convexa, envolvendo-a quase que totalmente, bem como uma parte da estrutura adjacente, a qual está realçada com uma textura diferente e indicada como área de interferência. Isto significa que ao computar a entropia do vértice nesta escala, parte dos pixels que formam a estrutura adjacente será incluída no cálculo. Embora a área realçada pareça ser significativa, nesta escala, os pesos que a compõem possuem valores extremamente baixos, que variam de 10^{-8} a 10^{-6} . Assim, o peso

total da área de interferência não alcança 0,1% do peso total. Então, pelo menos para o vértice observado, a interferência adjacente não causa problemas.

A Figura 5.12 mostra um exemplo mais representativo da ação das interferências. Nesta são apresentados os comportamentos dos pontos de maior saliência de uma região (turbina do avião): o ponto dominante de uma extremidade e “seu gêmeo”, o ponto dominante da outra extremidade e dois ruídos centrais.

Com as Figuras já apresentadas (5.6 a 5.10), percebe-se que para gerar um valor capaz de representar um candidato a ponto dominante é necessário que se combine informações de todas as escalas. O somatório das entropias de cada escala, a princípio, parece ser uma boa solução. Pelo menos, para os ângulos padrão, esta solução funciona bem, principalmente para a criação de limiares para aceitar ou rejeitar os ângulos dos candidatos. No entanto, a aplicação dessa totalização nos pontos da Figura 5.12 faz com que os ruídos sejam mais salientes em comparação a um dos pontos dominantes, devido aos seus totais serem mais altos.

Desta maneira, percebe-se que não se pode aplicar a totalização de entropias sem um tratamento prévio. As figuras anteriores também ajudaram a perceber que as estruturas pequenas tendem a apresentar decaimentos de saliência nas escalas iniciais. Assim, bastaria identificar se um candidato possui crescimento negativo nessas escalas ($escala_{n+1} - escala_n$). No entanto, isto não pode ser considerado uma regra. Para constatar, basta observar o comportamento do ponto determinante² na Figura 5.12. Visivelmente, este ponto é o centro de uma configuração de 90°, porém, com o crescimento inicial da escala, a região composta pelos dois ruídos é abrangida pelo núcleo de convolução, o que aumenta a quantidade de pixels pretos na configuração e, conseqüentemente, reduz o valor da entropia. Tal comportamento inicial é semelhante ao comportamento de ruídos. Sendo assim, a exclusão de candidatos que apresentem decaimento inicial, também é uma solução que não pode ser aplicada diretamente.

Figura 5.12 – Interferência de regiões adjacentes na comparação de candidatos.

A circunferência centrada no ponto dominante1, apresentada na Figura 5.12, está representando a quarta escala (35x35). Para os pontos dominante1 e gêmeo, que estão mais próximos à fuselagem do avião, a quarta escala é a primeira a sofrer interferências dessa região, como se pode constatar no gráfico da figura. Os outros três pontos só recebem interferências significativas a partir da quinta escala. Isto deixa o ponto dominante1 e seu gêmeo em desvantagens. Para piorar a situação, os ruídos têm decaimentos comparáveis ao decaimento do dominante2, impedindo que sejam excluídos. Para este caso, a única solução encontrada para evitar que ruídos sobrepujem pontos dominantes é compará-los usando apenas as escalas que não sofrem interferências de regiões adjacentes. Ou seja, a cada candidato se associa o número da última escala livre dessa interferência. Assim, quando chegar o momento de determinar a dominância de uma região (na fase-4 do EML), as entropias do grupo de candidatos em disputa serão totalizadas somente até a menor escala livre de interferências de regiões adjacentes encontrada no grupo.

As ações utilizadas no EML para a criação de um valor justo para a representação de candidatos são as seguintes:

1. Verificar se a queda de saliência nas três primeiras escalas é maior do que os limiares definidos empiricamente para cada uma delas (para a exclusão de ruídos);
2. Verificar se há queda significativa de saliência nas últimas escalas;
3. Informar a última escala livre de interferência de regiões adjacentes;
4. Totalizar as escalas úteis e verificar se esta representa um ângulo significativo.

O objetivo do item-1 é excluir os candidatos que representam pequenas estruturas (ruídos). Tais candidatos apresentam forte redução de saliência inicial (nas escalas 1, 2, e 3). A escolha dos limiares, que definem o que é uma “forte redução”, se baseou na observação do comportamento multiescala de diversos candidatos dominantes e não dominantes. Os valores escolhidos para os limiares não são ótimos, mas apresentam bom funcionamento. O valor para a “forte redução” é relativo à quantidade e à combinação de escalas envolvidas. Isto é, quanto menor o número de escalas que apresente decaimento de saliência, mais tolerante se torna o limiar.

Os itens 2, 3 e 4 trabalham em conjunto na tentativa de não incluir interferências externas à região em estudo. É como se isolasse tal região do restante da imagem para avaliar seus pontos de forma equânime.

Da mesma maneira como foi feito no item-1, no item-2 buscou-se a compreensão do que seria uma “forte redução”, a fim de se determinar limiares adequados para aceitar ou não a inclusão de uma, ou mais escalas finais. Isto é feito da terceira escala (17x17) em diante. Estruturas menores que 17 pixels são consideradas muito pequenas pelo EML.

Uma vez que as escalas úteis são identificadas, realiza-se o somatório de suas entropias e identifica-se o total obtido com o número da última escala útil incluída na operação. Esta identificação é necessária em duas operações posteriores. Numa delas a identificação é usada para determinar qual limiar será usado para aceitar ou rejeitar o candidato. A outra operação faz parte da fase-4 do EML, quando os candidatos a ponto dominante de uma região são comparados entre si.

Além de se excluir candidatos por serem identificados como ruídos, também se excluem aqueles que não apresentam um ângulo significativo. Para convexidades, o EML considera ângulos significativos aqueles com valor igual ou menor que 135° . Para concavidades, os ângulos precisam ser iguais ou maiores que 225° . A totalização dos valores de entropia desses ângulos é feita de acordo com a informação sobre a última escala útil do candidato em estudo. Por exemplo, se a última escala livre de interferências de um candidato

convexo for a escala-4, o limiar utilizado para averiguar a sua aceitação será construído somando-se os valores de entropia da primeira à quarta escala do ângulo padrão 135° .

Visto que os ângulos 135° e 225° podem ser muito restritivos e, também, pela existência de incerteza na precisão dos ângulos dos candidatos, foram criados parâmetros de tolerância para abrandar os limiares.

O algoritmo usado é baseado em instruções estáticas convencionais, resultando numa grande quantidade de instruções de desvio condicional (*IF/ELSEIF*), em especial na implementação do item-1. Deve-se reconhecer que a melhor estratégia seria o uso de técnicas de aprendizagem automática, o que será considerado em futuras versões deste detector.

Em suma, o primeiro procedimento é manter separados os candidatos convexos e os candidatos côncavos, pois o decaimento de saliência funciona de forma inversa para cada tipo. Para convexidades, o decaimento de entropia de uma escala para outra reflete a perda de saliência. Para concavidades a perda de saliência ocorre quando há aumento de entropia.

O núcleo desta fase é a análise de decaimentos de saliência. O comportamento dos decaimentos influencia diretamente na construção do valor final que representa o candidato. As funções do algoritmo desenvolvido para a fase 3 são:

1. Detectar decaimentos de saliência;
2. Verificar a quantidade de decaimentos;
3. Verificar como os decaimentos estão arranjados ao longo das escalas, isto é, se estão organizados de forma adjacente ou intercalados, se estão concentrados nas escalas iniciais ou finais;
4. Verificar a magnitude dos decaimentos;
5. Somar as entropias do candidato, levando-se em conta os aspectos observados nos itens anteriores;
6. Comparar a soma obtida no item 5 com um limiar;
7. Aceitar ou rejeitar o candidato baseado na comparação acima.

5.5 Fase 4 – determinação dos pontos dominantes

A entrada para esta fase é uma lista contendo os candidatos mais prováveis a pontos dominantes. Esta é composta pelas coordenadas de cada candidato, seis valores de entropia (um para cada escala), um número indicando a última escala útil, e a totalização das entropias até a escala indicada.

O procedimento básico é selecionar um candidato da lista, definir uma região de disputa em torno deste, verificar se há outros candidatos que possuem coordenadas na região e eleger o candidato mais saliente como o ponto dominante.

Em relação à definição da região de disputa, existem dois problemas a serem resolvidos. Um é como definir o tamanho da região. O outro é definir o tipo de região, isto é, se é bidimensional ou unidimensional.

A definição do tamanho da região foi solucionada fornecendo-se um parâmetro para o usuário interagir com o EML, indicando o tamanho desejado em pixels. Em relação ao tipo de região, optou-se pelo tipo unidimensional, visto que pode ocorrer disputa entre candidatos de regiões adjacentes ao se utilizar janelas 2D. Sendo assim, nesta fase ocorre extração de contornos.

Possuindo-se a lista de candidatos definitivos, o valor do parâmetro (j) que determina o tamanho da região e a lista de coordenadas que compõem o contorno da imagem, os passos seguintes são:

- Escolher um candidato da lista;
- Localizar o candidato no contorno;
- Verificar se na região de tamanho j (onde o candidato escolhido é o centro) há outros candidatos;
- Comparar os candidatos e escolher o dominante.

Devido às regiões de disputa (j) serem construídas a partir da localização dos candidatos, existe a possibilidade de que se sobreponham parcialmente umas às outras. Isto torna necessária a ordenação das listas de candidatos. A ordenação deve ser feita pela entropia total (nas seis escalas). No caso dos candidatos convexos a ordenação é crescente e no caso de candidatos côncavos a ordem é decrescente. Se a comparação de candidatos for feita de maneira aleatória, será maior a chance de que um candidato mais fraco seja eleito como dominante. A Figura 5.13 esboça essa ideia, onde está havendo a comparação entre três candidatos. O ponto central está localizado na interseção entre duas regiões. A ordem em que são feitas as comparações influencia no resultado final.

Iniciar as comparações pelos candidatos com menor saliência evita que se cometa injustiça na eleição do ponto dominante. Lembrando-se que no caso de convexidades, os candidatos de menor saliência apresentam menor entropia. Nas concavidades, menor saliência

significa maior entropia. Por isso a ordenação da lista de convexos deve ser inversa a dos côncavos.

Figura 5.13 – A sequência das comparações interfere no resultado. Se a sequência das comparações iniciar em A e seguir para B, o resultado será a eleição de dois pontos dominantes. Se a sequência for de B para A, só haverá um único dominante.

Considerando-se a lista de candidatos já ordenada, o processo de determinação dos pontos dominantes funciona da seguinte maneira:

1. Seleciona-se da lista de candidatos o menos saliente;
2. Localiza-se o candidato no contorno;
3. Define-se a região de disputa: $j/2$ pixels antes do candidato e $j/2$ pixels depois;
4. Cria-se um grupo com todos os candidatos que têm coordenadas na região de disputa;
5. Verifica-se qual é o candidato do grupo que possui a menor escala útil;
6. Totalizam-se as entropias, de cada candidato do grupo, até a menor escala útil;
7. Exclui-se da lista de candidatos, aquele que foi eleito o menos saliente;
8. Repete-se o procedimento (1-7) até que todos os candidatos da lista tenham sido examinados. Os que não forem excluídos são os pontos dominantes.

Para finalizar este capítulo, são apresentados os resultados de cada fase do EML. A Figura 5.14 apresenta o resultados dos candidatos computados no fim da fase 2 (item a) e da

fase 3 (item *b*). A fase 3 garante a redução da quantidade de candidatos, deixando apenas os mais prováveis para disputar a dominância das regiões. Deve-se perceber que no item *b*, todos os candidatos se concentram em áreas côncavas ou convexas.

Figura 5.14 – Saídas das fases 2 e 3 do EML: a) Saída da fase-2, todos os candidatos; b) Saída da fase-3, apenas os candidatos prováveis a pontos dominantes.

Figura 5.15 – Fase-4: pontos dominantes eleitos.

5.6 Conclusão do capítulo

Neste capítulo foi descrito o funcionamento do EML, um detector de *corners* de contornos, o qual se baseia na medição das modificações causadas pelo processo de suavização multiescala de imagens binárias. A medição utilizada pelo detector é a entropia de Shannon.

A forma circular dos núcleos de convolução, o tipo de imagem utilizada (binária) e o tipo de *corner* em estudo (situado em contornos) permitiram a construção de uma relação entre ângulos de setores circulares e a entropia produzida pela suavização dos mesmos.

Além do funcionamento do detector, foram expostos os diferentes problemas enfrentados em cada fase e como foram solucionados. Por exemplo, as distorções causadas pela grade de amostragem e as interferências entre regiões em virtude do aumento de escala. Este último se assemelha ao problema conhecido por escala natural de curvas, discutido em Rosin (1998).

Capítulo 6 – Proposta do detector CLP

6.1 Introdução

A criação do detector baseado na computação local de pesos, o CLP, deriva dos estudos realizados sobre difusão e entropia durante o desenvolvimento do detector EML (descrito no capítulo 5). O núcleo desses dois assuntos é a concentração de substâncias. O efeito da difusão é a dissolução das concentrações, e a entropia é usada para medir a difusão com base nas concentrações existentes. Assim, supõe-se que a concentração de uma substância é a principal informação e, por isso, seria possível descartar a computação do logaritmo que compõe a equação de entropia, reduzindo a complexidade computacional do algoritmo do detector. Ou seja, o CLP é, basicamente, o EML que não computa logaritmos.

No caso de imagens binárias, entende-se por concentração a quantidade de objeto presente numa dada vizinhança. As concentrações de objeto e fundo são complementares. A concentração de objeto depende da quantidade e da localização dos pesos ocupados pelos pixels do objeto dentro da vizinhança. Os pesos são distribuídos espacialmente, respeitando a distribuição circularmente simétrica do núcleo de convolução. Os termos concentração, peso e probabilidade são tratados, aqui, como sinônimos.

O CLP é composto pelas mesmas quatro fases que compõem o EML. São elas:

1. Computação do peso na escala mais baixa – O núcleo de convolução ajustado a uma janela 3x3 é deslocado pixel a pixel pela imagem. A distribuição de probabilidades do núcleo é usada para determinar o peso do objeto contido na vizinhança 3x3. O peso encontrado é usado para classificar o pixel como reto, candidato convexo ou côncavo.
2. Computação multiescala dos candidatos – As coordenadas dos candidatos, obtidos na fase anterior, são usadas como ponteiros para que apenas estes sejam observados em escalas mais altas. Assim, para cada candidato é computado o seu peso considerando-se vizinhanças maiores, com distribuição de pesos baseada em núcleos maiores. Então, para cada candidato é produzida uma lista contendo seus pesos computados em uma quantidade finita de escalas.
3. Análise da evolução do candidato através das escalas – A lista produzida na fase-2 é analisada com o objetivo de verificar possíveis decaimentos de saliência. Nesta fase são separados os verdadeiros candidatos daqueles que representam pequenos detalhes

ou ruídos. A saída desta fase é uma lista dos possíveis pontos dominantes com seus respectivos pesos e coordenadas espaciais.

4. Determinação dos pontos dominantes – O objetivo é determinar as regiões que possuem mais de um provável ponto dominante e compará-los para que reste apenas um em cada região.

O CLP utiliza os mesmos ângulos padrão que o EML, os quais são mostrados na Figura 5.1 (capítulo 5). Os pesos multiescala de tais ângulos são mostrados na Tabela 6.1. A Figura 6.1 mostra o comportamento multiescala dos pesos dos ângulos padrão. Tanto a Tabela 6.1, quanto a Figura 6.1 representam os valores originais, os quais necessitam ser linearizados. A Figura 6.2 apresenta a versão linearizada da evolução multiescala dos pesos. A Tabela 6.2 informa o valor do fator de linearização que deve ser aplicado a cada candidato em cada escala. Para escolher o fator, faz-se exatamente como no EML, verifica-se qual é o ângulo do candidato na escala 3x3.

Com o uso de pesos, em vez de entropia, a relação com os ângulos fica bem mais intuitiva. Por exemplo, quanto menor o peso, menor é o ângulo. A Figura 6.3 reflete esta ideia.

Tabela 6.1 – Ângulos padrão e seus pesos em 7 escalas diferentes.

Ângulo	Pesos de Referência						
	3x3	5x5	9x9	17x17	35x35	71x71	143x143
(1) 45°	0,62500	0,62050	0,60470	0,58700	0,57600	0,56940	0,56600
(2) 90° Diag.	0,63990	0,66110	0,65670	0,64460	0,63600	0,63070	0,62790
(3) 90°	0,73510	0,70500	0,67760	0,65450	0,64100	0,63320	0,62910
(4) 135°	0,75000	0,74550	0,72970	0,71200	0,70100	0,69440	0,69100
(5) 180° Diag.	0,76490	0,78610	0,78170	0,76960	0,76100	0,75570	0,75290
(6) 180° H/V	0,86010	0,83000	0,80260	0,77950	0,76600	0,75820	0,75410
(7) 225°	0,87500	0,87050	0,85470	0,83700	0,82600	0,81940	0,81600
(8) 270°	0,88990	0,91110	0,90670	0,89460	0,88600	0,88070	0,87790

Figura 6.1 – Comportamento multiescala dos ângulos padrão (sem correção).

Figura 6.2 – Comportamento multiescala dos ângulos padrão (linearizados).

Tabela 6.2 – Fatores para a linearização de pesos.

Classe do candidato Na escala 3x3	Valor do Fator por Escalas					
	5x5	9x9	17x17	35x35	71x71	143x143
45°	0,9121676	0,936001	0,964225	0,982639	0,994029	1
90° Diag.	0,9497807	0,956144	0,974092	0,987264	0,99556	1
90°	0,8923404	0,928424	0,961192	0,981435	0,993525	1
135°	0,9268947	0,946965	0,970506	0,985735	0,995104	1
180° Diag.	0,9577662	0,963157	0,9783	0,989356	0,996295	1
180° H/V	0,9085542	0,939571	0,967415	0,984465	0,994592	1
225°	0,9373923	0,954721	0,97491	0,987893	0,995851	1
270°	0,9635605	0,968236	0,981332	0,990858	0,996821	1

Figura 6.3 – Relação ângulo – peso do objeto (preto). Quanto maior é o ângulo, maior é o peso.

Pode-se perceber pelas Figuras 6.1, 6.2 e 6.3 que há uma inversão no comportamento dos valores em relação ao EML. Agora, as convexidades têm valor baixo (pois têm menos objeto) e as concavidades têm valores altos. Para converter o algoritmo do EML em CLP, além da modificação do cálculo de entropia por cálculo de pesos, foi necessário refazer os

limiares para se adequarem ao comportamento invertido. A Figura 6.4 faz uma comparação entre o gráfico da entropia e o gráfico de pesos. Ambos representam a evolução multiescala dos mesmos pontos mostrados na Figura 5.12. Basta rebater um dos gráficos em relação ao eixo horizontal para obter o outro gráfico.

Figura 6.4 – Evolução multiescala: entropia versus peso. a) Gráfico da entropia copiado da Figura 5.12; b) Gráfico da evolução dos pesos, relativos aos mesmos pontos da Figura 5.12.

A Tabela 6.3 apresenta os ângulos padrão com seus valores já linearizados.

Tabela 6.3 – Ângulos padrão ajustados pelo fator de linearização.						
Ângulo	Pesos de Referência					
	5x5	9x9	17x17	35x35	71x71	143x143
(1) 45°	0,566	0,566	0,566	0,566	0,566	0,566
(2) 90° Diag.	0,6279	0,6279	0,6279	0,6279	0,6279	0,6279
(3) 90°	0,6291	0,6291	0,6291	0,6291	0,6291	0,6291
(4) 135°	0,691	0,691	0,691	0,691	0,691	0,691
(5) 180° Diag.	0,7529	0,7529	0,7529	0,7529	0,7529	0,7529
(6) 180° H/V	0,7541	0,7541	0,7541	0,7541	0,7541	0,7541
(7) 225°	0,816	0,816	0,816	0,816	0,816	0,816
(8) 270°	0,8779	0,8779	0,8779	0,8779	0,8779	0,8779

Todas as considerações e suposições feitas na terceira fase do EML (item 5.4), também, são válidas para o CLP, incluindo interferências entre regiões adjacentes e busca pela escala adequada para comparar os candidatos de uma mesma região. Também é válido, aqui, o que foi feito na quarta fase do EML (item 5.5), a ordenação dos candidatos antes de compará-los. Por isso, tais descrições não serão repetidas aqui. No entanto, é válido que algumas ideias sejam reapresentadas. São elas:

- Decaimentos de saliência nas escalas iniciais podem indicar ruídos. Porém, isto não é uma regra, pois bons pontos podem ter decaimentos iniciais. A escolha dos limiares que excluirão tais estruturas precisa ser feita cuidadosamente. Tornando necessário o estudo de uma grande quantidade de pontos em diferentes imagens, a fim de entender as diferenças entre os bons e maus candidatos.
- Decaimentos nas escalas finais podem indicar que a estrutura é menor que essas escalas, ou está sofrendo interferências de estruturas adjacentes. Aqui também é necessário avaliar cuidadosamente a magnitude desses decaimentos. Pequenas flutuações não devem ser consideradas.
- Uma vez que esses decaimentos forem determinados, deve-se prover o candidato com a última escala usada na totalização de seus valores multiescala. Candidatos de uma mesma região estão sujeitos às mesmas influências, porém algum pode ser influenciado mais intensamente. Assim, na disputa pela dominância de uma região é

necessário totalizar até a escala do candidato mais influenciado, isto é, aquele que tem a sua existência limitada pela menor escala. Por exemplo, na Figura 6.4 alguns candidatos sobrevivem até a quinta escala. Outros sobrevivem até a quarta escala. Ao compará-los é necessário que, antes, todas as totalizações sejam feitas até a quarta escala.

Em termos de algoritmos computacionais, a terceira fase do CLP difere do EML. No detector baseado em entropia foi usada uma abordagem mais complexa, onde se tentou representar todas, ou quase todas as combinações de variações que os candidatos poderiam apresentar ao longo das escalas. Isto gerou uma grande quantidade de instruções de desvio aninhadas. Na terceira fase do CLP são realizadas as seguintes ações:

1. Linearização dos candidatos – aplicação da Tabela 6.2
2. É verificado se a saliência do candidato se mantém suficientemente forte em cada uma das seis escalas. Isto é, o peso em cada escala é comparado com um limiar obtido empiricamente, o qual embute uma tolerância de 2,5% em relação aos pesos dos ângulos de 135° e 225° . No EML há um parâmetro de tolerância disponível ao usuário. No CLP optou-se por fixá-la em 2,5% de acordo com os experimentos realizados.
 - a. Para convexidades, o peso em cada escala deve ser menor que 1,025 do peso do ângulo de 135° .
 - b. Para concavidades, o peso em cada escala deve ser maior que 0,975 do peso do ângulo de 225° .
3. Para os candidatos não reprovados é verificado se houve decaimento significativo de saliência nas escalas 5 e 6.
 - a. Se o decaimento for fraco, as seis escalas do candidato são totalizadas.
 - b. Se o decaimento for significativo, computa-se a média até a última escala sem decaimento (escala 4 ou 5), a qual é multiplicada por seis para simular a totalização.
 - c. Este passo (3) é executado para evitar que bons candidatos sejam excluídos no passo 4. Em geral, decaimentos abruptos nas escalas finais indicam interferências de estruturas adjacentes.
4. A totalização das escalas de cada candidato é comparada a um limiar obtido empiricamente.

- a. Para convexidades, a totalização das escalas deve ser menor que 1,023 da totalização do ângulo de 135° .
 - b. Para concavidades, a totalização das escalas deve ser maior que 0,99 da totalização do ângulo de 225° .
5. Para os candidatos que restarem, verifica-se o decaimento de saliência nas escalas 1 a 4. Tais decaimentos podem indicar que a estrutura em observação é um “ruído”. Os limiares usados foram obtidos experimentalmente.
 - a. Para convexidades a queda acumulada não pode ultrapassar o valor 0,31.
 - b. Para concavidades a queda acumulada não pode ser menor que o valor -0,17 e a média do candidato tem que ser superior ao peso do ângulo de 225° .

As respostas dos detectores EML (entropia) e CLP (peso) são quase idênticas. A Figura 6.5 mostra a comparação de suas respostas. Para a imagem da aeronave, todos os pontos coincidem.

Figura 6.5 – Saída da quarta fase dos detectores EML e CLP.

O CLP não utiliza parâmetros de tolerância para os limiares escolhidos, isto é, os limiares são fixos e idênticos para todas as imagens. Isto pode tornar o CLP mais permissivo, deixando que mais candidatos sejam aceitos para a fase de disputa de dominância, como é o caso da Figura 6.6. Nos itens *a* e *b* desta imagem, o parâmetro que define a região de disputa para candidatos convexos foi ajustado em 11 pixels. No EML, essa distância de 11 pixels foi o suficiente para escolher um único dominante na região. Uma consequência negativa da

fixação do parâmetro de tolerância é a possibilidade de que candidatos localizados em regiões não salientes sejam aceitos para a fase de disputa de dominância. A solução, então, é gerar grandes regiões de disputa, para que os bons candidatos possam eliminar os maus candidatos remanescentes. No item c da Figura 6.6, o tamanho da região de disputa foi ajustado em 50 pixels, o que permitiu a eliminação de quase todos os maus candidatos apontados no item b.

Figura 6.6 – “Divergências” entre detectores EML e CLP: a) e b) distância de disputa ajustada em 11 pixels; c) distância de disputa ajustada em 50 pixels.

6.2 Conclusão do capítulo

Neste capítulo é descrito o funcionamento do CLP, um detector de *corners* de contornos, o qual se baseia na computação do peso probabilístico da região do objeto abarcada pelo núcleo de convolução. O CLP deriva do EML (detector descrito no capítulo 5) e apresenta vantagens sobre este, são elas:

- Funcionamento mais intuitivo, pois o peso do objeto é diretamente proporcional ao espaço que ocupa dentro da vizinhança. Na entropia, essa relação é inversamente proporcional.
- Não há computação de logaritmos, reduzindo o esforço computacional.

As vantagens descritas acima não se relacionam com a essência do método, isto é, medir a modificação causada pela suavização. Neste sentido, os detectores têm o mesmo comportamento. Outro aspecto que os diferencia está na forma de analisar seu comportamento multiescala. A decisão de reduzir parâmetros de entrada (tolerâncias de decaimentos de saliência), fixando-os de acordo com limiares obtidos experimentalmente tem a vantagem de aumentar a automatização do funcionamento. Porém, maus candidatos, cujas localizações estão distantes das regiões salientes, podem ser aceitos para a fase de disputa de dominância. As consequências podem ser a necessidade de se aumentar exageradamente a região de disputa e a admissão de falsos positivos. Isto pode tornar o CLP menos preciso que o EML.

Capítulo 7 – Proposta do detector DoG_zero.

7.1 Introdução

Os primeiros testes realizados com o EML mostraram que a computação da entropia causava a extração de bordas, basta observar a Figura 5.4. Haja vista a existência de detectores de bordas derivativos, como o Laplaciano da Gaussiana, imaginou-se que poderia haver uma relação matemática entre a entropia e a derivação. Deve-se confessar que muito tempo infrutífero foi dedicado à busca dessa relação.

No entanto, ao se concentrar no significado físico das expressões “diferença de Gaussianas” e “entropia produzida”, concluiu-se que são maneiras distintas de se quantificar as modificações produzidas pela suavização Gaussiana.

Uma imagem binária sofre modificações quando suavizada por uma Gaussiana. De acordo com que já foi mencionado nos capítulos 3 e 4, tais modificações se concentram ao longo das fronteiras entre o fundo e o objeto. Nesses mesmos capítulos, foi visto que a entropia é uma ferramenta adequada para detectar modificações causadas por suavização.

Na mesma situação descrita acima, isto é, uma imagem binária e a sua versão suavizada, se for computada a subtração entre elas, o resultado será o mesmo que o da entropia. Devido às diferenças nos processos matemáticos, os valores computados pela entropia e pela subtração são distintos, porém identificam exatamente as mesmas regiões modificadas pela suavização.

Neste capítulo está sendo proposto um detector multiescala baseado em diferença de Gaussianas. Porém, uma das parcelas da subtração é a imagem original. Esta ideia já foi discutida no capítulo 4, em especial nos itens 4.3 e 4.4. A Figura 7.1 mostra o resultado da diferença entre uma imagem binária suavizada com Gaussiana ($\sigma=0,5$) e a sua versão original não suavizada. Percebe-se a intensidade mais forte nos pontos mais salientes, por exemplo, nos cantos das turbinas, no bico da aeronave e em todos os “ruídos convexos”.

Figura 7.1 – Resultado da diferença: “imagem suavizada ($\sigma=0,5$) – imagem original”.

A Figura 7.2 apresenta a resposta da DoG para seis escalas adicionais, as mesmas empregadas nos detectores propostos nos capítulos anteriores. Deve-se observar que a subtração da imagem original garante a exata localização dos pixels do contorno, independentemente da escala considerada.

Figura 7.2 – Resultado da diferença: “imagem suavizada– imagem original” em seis escalas diferentes. a) $\sigma=0,8333$; b) $\sigma=1,5$; c) $\sigma=2,8333$; d) $\sigma=5,8333$; e) $\sigma=11,8333$; f) $\sigma=23,8333$.

Para facilitar, este detector será chamado de “DoG_zero”. O seu funcionamento básico se assemelha aos outros apresentados nos capítulos anteriores. As saliências são detectadas na escala mais baixa, verifica-se a sua evolução multiescala e, então, compara-se os candidatos que ocupam uma mesma região a fim de se determinar os pontos dominantes.

O DoG_zero tem fortes influências biológicas, então, pretende-se manter uma certa fidelidade ao possível funcionamento dos campos receptivos das células ganglionares da retina. O básico desta fidelidade é não supor a existência de inteligência nesse nível, isto é, as células não são capazes de discernir, a partir dos sinais que chegam às suas entradas, o que é objeto, ou o que é fundo. Como mencionado no capítulo 4, em relação à capacidade de detecção de contraste, tais células são classificadas em dois tipos. O tipo *off-center* que detecta objeto escuro em fundo claro, e o tipo *on-center* que detecta objeto claro em fundo escuro.

Diretamente associado ao tipo de célula ganglionar, pode estar a noção de convexidade e concavidade. Numa imagem binária, designar o que é concavidade ou convexidade depende do referencial adotado. O vértice de uma concavidade escura pode ser também, o vértice de uma convexidade clara, ou vice-versa. Por exemplo, na Figura 7.3, quais são as concavidades e quais são as convexidades? Da mesma forma, qual região representa o objeto e qual região representa o fundo? É bem possível que este seja o tipo de problema enfrentado pela retina. A ideia, então, é detectar o que for mais relevante e deixar para os estágios superiores do processamento visual, a decisão sobre o que é objeto, ou o que é o fundo, num processo conhecido por *figure-ground organization* (WAGEMANS, 2015).

Neste estágio do sistema visual (retina), o mais relevante é a detecção de respostas fortes de contraste. As respostas de contraste dos campos receptivos costumam ser modeladas por diferença de Gaussianas (DoG). No caso do detector DoG_zero, as respostas fortes de contraste identificam ângulos estreitos, ou altas curvaturas (vértices de convexidades claras ou escuras), como pode ser visto na Figura 7.4.

O esquema *on-center* / *off-center* fornecido pelas células ganglionares resolve o problema da detecção de contraste negativo e positivo, e, também, a dualidade concavidade/convexidade. Em relação à eletrofisiologia das células ganglionares é possível que esse esquema facilite a criação de um código neural imune ao ambiente ruidoso do nervo óptico, porém não é relevante para esta tese.

O detector DoG_zero incorpora o esquema *on-center* / *off-center* através da simples comutação das parcelas de subtração. Para detectar as convexidades escuras (ou concavidades claras) faz-se a subtração “imagem suavizada - imagem original” e para detectar as

convexidades claras (ou as concavidades escuras) faz-se a subtração “imagem original - imagem suavizada”.

Figura 7.3 – Concavidade versus convexidade.

Figura 7.4 – Comportamento multiescala das respostas do DoG_zero para os ângulos padrão: a) valores não linearizados; b) valores linearizados.

A Figura 7.5 apresenta o resultado do DoG_zero para a imagem da Figura 7.3. Em 7.5.a é mostrada a resposta do esquema *off-center*, que marca mais intensamente os vértices das convexidades escuras, que coincidem com os vértices das concavidades claras. Em 7.5.b é mostrada a resposta do esquema *on-center*, que marca mais intensamente os vértices das convexidades claras, que coincidem com os vértices das concavidades escuras.

Outra indicação da dualidade convexidade/concavidade está na existência de dois contornos, ou duas fronteiras. Um contorno é interno, o qual é formado pelos pixels pertencentes ao objeto. O outro é externo, formado pelos pixels do fundo. Por exemplo, em 7.5.b, as circunferências vermelhas mostram as convexidades do contorno externo, as quais estão deslocadas um pixel para cima em relação aos “mesmos pontos” (concavidades da região preta) em 7.5.a.

Figura 7.5 – Resposta do DoG_0 para a imagem da Figura 7.3: a) convexidades da saída *off-center*; b) convexidades da saída *on-center*.

Ao admitir a hipótese de que a visão biológica possa seguir a ideia de detecção de convexidades exposta acima, o leitor poderá indagar-se a respeito da possibilidade da não detecção, caso a imagem da cena não seja projetada sobre os detectores corretos.

Visto que os campos receptivos estão fixos (*hardwired*) na retina, se a cena não for projetada exatamente sobre o tipo correto de detector de contraste, isto é, se o contraste negativo for projetado “erroneamente” sobre as células *on-center* e o contraste positivo sobre as células *off-center*, haverá a probabilidade de que os vértices das convexidades não sejam detectados.

Contudo, para reduzir tal preocupação, pode ser dito que nos últimos anos tem havido um interesse crescente em entender as vibrações que ocorrem no olho (MARTINEZ-CONDE; OTERO-MILLAN; MACKNIK, 2013; TIAN; CHEN, 2015; MEYBERG et al., 2015). Algumas pesquisas observaram que os movimentos microsacádicos podem estar relacionados com a tentativa de “encaixar” a cena nos circuitos corretos. Em Martinez-Conde, Macknik e Hubel (2000) observou-se que os sinais transmitidos pelas células ganglionares são mais representativos após os movimentos microsacádicos.

7.2 Funcionamento do DoG_zero

Exatamente como foi feito nos detectores propostos nos capítulos anteriores, detecta-se os pontos candidatos na escala mais baixa. Neste, uma Gaussiana com desvio padrão de meio pixel ($\sigma=0,5$) é adaptada a uma janela quadrada de dimensões 3x3 pixels. Com este núcleo, suaviza-se a imagem binária.

O próximo passo é obter os candidatos a vértices das convexidades da região preta e da região branca. Para o primeiro caso utiliza-se o método *off-center*, que é implementado através da subtração da “imagem suavizada - imagem original”. Para o segundo caso, utiliza-se o método *on-center*, efetuando-se a subtração “imagem original – imagem suavizada”. Para cada caso cria-se uma lista de candidatos aos vértices das convexidades. As listas serão compostas pelos pontos que apresentarem valores superiores ao valor que representa o ângulo de 180° (180Diag), o qual está listado na Tabela 7.1. Os valores desta tabela estão representados graficamente na Figura 7.4.a.

Tabela 7.1 – Ângulos padrão e suas DoG zero para 7 escalas diferentes.

Ângulo	Respostas DoG zero						
	3x3	5x5	9x9	17x17	35x35	71x71	143x143
(1) 45°	0,2855	0,5839	0,7345	0,8072	0,8439	0,8601	0,8677
(2) 90° Diag.	0,2741	0,5224	0,6378	0,695	0,7245	0,7377	0,744
(3) 90°	0,2017	0,4528	0,599	0,6744	0,7145	0,7328	0,7415
(4) 135°	0,1903	0,3913	0,5023	0,5622	0,5951	0,6104	0,6178
(5) 180° Diag.	0,179	0,3299	0,4055	0,45	0,4757	0,488	0,494
(6) 180° H/V	0,1065	0,2603	0,3667	0,4294	0,4657	0,4831	0,4916
(7) 225°	0,0952	0,1988	0,27	0,3172	0,3463	0,3607	0,3679
(8) 270°	0,0838	0,1373	0,1733	0,2049	0,2269	0,2383	0,2441

Uma vez que os candidatos estão identificados, o procedimento de subtração *on-center* e *off-center* é repetido para núcleos Gaussianos de escalas maiores ($\sigma = 0,8333; 1,5; 2,8333; 5,8333; 11,8333$ e $23,8333$) adaptados a janelas quadradas de dimensões maiores (5x5, 9x9, 17x17, 35x35, 71x71 e 143x143). Lembrando-se que a subtração é sempre feita entre a imagem suavizada e a imagem original. Os valores dos candidatos nas novas escalas são incluídos nas listas.

A próxima fase é a de analisar a evolução de cada candidato através da escala. Para amenizar a complexidade dessa análise, são usadas as mesmas facilidades discutidas nos dois detectores propostos nos capítulos 5 e 6. Uma é a “linearização” dos valores, isto é, transformar as curvas do gráfico apresentado na Figura 7.4.a em segmentos de retas que compõem o gráfico da Figura 7.4.b. A outra facilidade é deixar de incluir a escala inicial (3x3) na análise multiescala.

Para obter a linearização, o valor de cada escala de um candidato é multiplicado por um determinado fator de correção. O conjunto de fatores de correção usado para um dado candidato é decidido de acordo com a classificação angular que o candidato obteve na escala 3x3. Ou seja, o candidato quando é detectado na escala mais baixa, ele possui um dos quatro valores listados na primeira coluna da Tabela 7.1, que corresponde a um dos ângulos 45°, 90° Diag, 90° ou 135°. Nesta escala, uma convexidade só pode possuir um desses quatro valores. Tal restrição foi discutida no início do capítulo 5. Sabendo-se a classificação do candidato, escolhe-se o conjunto de linearização. A Tabela 7.2 apresenta esses fatores de correção. Esta tabela apresenta valores de correção para não convexidades, porém não são usados.

Tabela 7.2 – Fator de correção de resposta DoG zero por ângulo e escala.

Fator de correção	5x5	9x9	17x17	35x35	71x71	143x143
Fator45	1,486042	1,181348	1,07495	1,028202	1,008836	1
Fator90D	1,424196	1,16651	1,070504	1,026915	1,00854	1
Fator90	1,637588	1,237896	1,099496	1,037789	1,011872	1
Fator135	1,57884	1,229942	1,098897	1,038145	1,012123	1
Fat180D	1,497423	1,218249	1,097778	1,03847	1,012295	1
Fat180H	1,88859	1,340605	1,144853	1,055615	1,017595	1
Fator225	1,850604	1,362593	1,159836	1,062374	1,019961	1
Fator270	1,777859	1,40854	1,191313	1,075804	1,024339	1

Toda a complexidade relatada no item 5.4 é válida aqui também. A Figura 7.6 faz uma comparação entre o gráfico da entropia e o gráfico do DoG_zero. Ambos representam a evolução multiescala dos mesmos pontos mostrados na Figura 5.12. Uma parte da figura contendo os candidatos foi reproduzida para facilitar a observação.

Figura 7.6 – Evolução multiescala: entropia versus DoG_zero. a) Gráfico da entropia copiado da Figura 5.12; b) Gráfico da evolução DoG_zero, relativos aos mesmos pontos da Figura 5.12.

O algoritmo desenvolvido para avaliar a evolução multiescala dos candidatos no DoG_zero difere dos algoritmos usados no EML e no CLP. Consequentemente, podem surgir diferenças na detecção dos pontos dominantes, embora sejam mínimas. Porém, deve-se insistir que tais diferenças independem da teoria básica envolvida nos três detectores. A diferença está na forma como são interpretadas as variações de valores multiescala.

O algoritmo da análise multiescala do DoG_zero é composto de três partes. A primeira realiza a totalização dos valores multiescala dos candidatos, mas desconsiderando qualquer tipo de interferência que possa favorecer, ou desfavorecer o candidato. A segunda parte computa as perdas de saliência sofridas pelo candidato e a terceira verifica a velocidade da perda inicial de saliência. A seguir, essas três partes são discutidas mais detalhadamente.

Totalização dos valores multiescala

A direção de uma função pode ser determinada pelo sinal de sua derivada. A mudança de direção pode ser indicada pela troca do sinal da derivada, isto é, nos pontos de mínimos locais ou de máximos locais. Por exemplo, o gráfico da evolução do candidato *dominante2* na Figura 7.6.b pode ser encarado como uma função, a qual decai suavemente da escala 1 para a escala 3, onde atinge um mínimo local. A partir deste, a função cresce até um máximo na escala 5 e, então, decresce vertiginosamente para escala 6, onde atinge o seu mínimo global.

O alto valor inicial do ponto *dominante2* é consequência do ângulo de 90°. O decaimento suave que este sofre se deve a inclusão dos pixels *ruído_A* e *ruído_B* na janela do núcleo. O crescimento a partir da escala 3 se deve à maior quantidade de pixels de fundo¹², em relação à quantidade de objeto, envolvida pela janela do núcleo. O decaimento final é devido à janela de alta escala envolver boa parte da estrutura da asa e da fuselagem, aumentando consideravelmente a quantidade de objeto no interior da janela.

No relato acima está implícito que a interferência da vizinhança local (*ruído_A* e *ruído_B*) é válida, isto é, faz parte da estrutura observada (turbina). No entanto, as variações posteriores na direção da função indicam interferências de estruturas adjacentes (aumento da quantidade de objeto ou de fundo). Estas devem ser evitadas para que uma dada estrutura do objeto seja avaliada isoladamente.

Assim, computa-se a derivada de cada candidato (diferenças entre as escalas) e determina-se a primeira modificação do sinal (mínimo local ou máximo local). O valor do

¹² Deve-se perceber que o DoG_zero tem o mesmo comportamento do EML, quanto mais fundo (menos objeto) presente na vizinhança, mais convexa se torna a região do objeto.

mínimo local, ou do máximo local é, então, repetido nas escalas seguintes. Isto é, o valor do extremo é propagado até a última escala. O total do candidato é a soma dos valores anteriores ao extremo local com as repetições desse extremo. A Figura 7.7 mostra como ficou o gráfico da Figura 7.6.b após a repetição dos extremos. Observe como o comportamento final dos pontos dominantes ficou semelhante, assim como a dos candidatos considerados como ruídos.

Figura 7.7 – A propagação do primeiro extremo aplicada à Figura 7.6.b: a) evolução multiescala original; b) modificação usando a propagação do primeiro extremo detectado em cada candidato.

Deve-se notar que a propagação de extremos busca uma maneira de tornar mais justa a disputa pela dominância de uma região. A função da totalização é a de criar um valor único para que os candidatos possam ser confrontados entre si, bem como para criar um limiar capaz de excluir os candidatos menos salientes. Neste detector é disponibilizado um parâmetro para que o usuário ajuste o limiar (o parâmetro de tolerância usado no EML e abolido no CLP). Quando menor é o limiar, mais permissivo ele é, mais candidatos são admitidos para a disputa, inclusive candidatos que se encontram distantes das regiões convexas (ou côncavas). Tal distância pode comprometer a disputa, de forma a não incluir o candidato distante no confronto e, assim, aceita-lo como dominante de uma região não saliente. A solução (não desejável) seria estender o tamanho da região de disputa a fim de eliminá-lo.

Mesmos os ruídos podem ter valores representativos, por exemplo, os ruídos A e B da Figura 7.7.b são considerados bons candidatos, pois se mantêm tão salientes quanto um ângulo de 135° . Os ruídos que apresentam valores altos somente são eliminados durante a disputa pela dominância, isto se os mesmos estiverem próximos a uma saliência real, para serem enquadrados pela janela de disputa.

Não é raro encontrar ruídos que possuam boa totalização multiescala e se encontrem distantes de uma saliência real. Por isso, foi necessário buscar novas maneiras de eliminá-los. O DoG_zero inclui duas novas avaliações. Uma computa o acúmulo de perdas de saliência de cada candidato e a outra computa a velocidade desta perda entre a primeira e a segunda escala.

Acúmulo de perdas de saliência

Observou-se que as primeiras escalas definem a validade de um candidato, pois estas delimitam a vizinhança local, os limites da estrutura observada. Com base nisto, detecta-se o valor mínimo das quatro primeiras escalas e computa-se a diferença entre o valor da escala inicial e esse mínimo. Em média, um bom candidato tem perdas abaixo de 0,05, mas não é um valor ideal para ser transformado em um limiar para todas as imagens, assim, o acúmulo de perdas é comparado com um parâmetro ajustável, que na maioria das vezes terá o valor 0,05.

Velocidade de decaimento

Em geral, o decaimento de saliência de um ruído é mais acentuado do que o de um bom candidato, principalmente na evolução da primeira para a segunda escala. Então, computa-se a diferença entre o valor da segunda escala e o da primeira e compara-se a um limiar. Em geral, para os ruídos, a magnitude dessa diferença é maior que 0,05. Então, este é mais outro parâmetro ajustável utilizado pelo DoG_zero.

A Figura 7.8 apresenta o resultado do uso dos três parâmetros de eliminação de “maus” candidatos. O item *a* da figura é o resultado da totalização baseada em propagação de extremos utilizando o limiar 3,58. Todos os pontos dominantes marcados num *ground-truth* estão presentes nessa imagem, bem como outros não dominantes, incluindo ruídos afastados de saliências reais, os quais estão apontados por setas. Estes últimos não poderiam ser eliminados num confronto de disputa de dominância, exceto se fosse usada uma janela de disputa muito grande, o que poderia eliminar pontos dominantes de outras regiões. O item *b* da figura mostra o resultado da eliminação de candidatos que possuem acúmulo de decaimento superior a 0,05. Com este limiar, os ruídos distantes foram eliminados, porém pontos dominantes verdadeiros (do *ground-truth*) também foram excluídos (apontados pelas setas). A avaliação baseada em velocidade de decaimento não surtiu efeitos neste exemplo.

Figura 7.8 – Resultados do algoritmo de avaliação multiescala: a) contém os candidatos que possuem totalização superior a 3,58; b) contém os candidatos que possuem decaimento acumulado abaixo de 0,05.

Como pode ser percebido a partir da Figura 7.8, ainda não se obteve uma interpretação ideal a respeito da evolução multiescala. É comum encontrar pontos do *ground-truth* que possuem as mesmas propriedades de candidatos que não foram escolhidos como dominantes. Isto obriga a fazer escolhas: ou se considera o “ruído” como um ponto dominante, ou deixa-se de marcar o ponto dominante que possui as mesmas propriedades do ruído. A Figura 7.9 ilustra bem essa ideia. Deve-se notar que o comportamento nas quatro primeiras escalas é o mesmo e os ângulos dos pontos são muito próximos (138° e 134°), aliás, o ângulo do ruído é mais saliente. Os parâmetros utilizados pelo DoG_zero são incapazes de distinguir tais pontos.

A subjetividade do julgamento humano, a capacidade de decidir quais parâmetros são adequados numa dada posição da imagem, a liberdade de usar parâmetros diferentes em outra posição da imagem, incapacitam a construção de decisões artificiais que se comparem às decisões humanas, pelo menos, da maneira de como os estudos foram conduzidos para interpretar a evolução multiescala dos candidatos. É possível que o uso de alguma técnica de *machine learning* obtenha um melhor desempenho.

Figura 7.9 – Semelhanças entre pontos do *ground-truth* e ruídos.

Na construção dos detectores EML e CLP foram escritos algoritmos diferentes para a detecção de convexidades e para a detecção de concavidades. No DoG_zero é utilizado o mesmo algoritmo, pois só se detecta convexidades. Porém, se detecta as convexidades do objeto e as convexidades do fundo. Assim, tudo o que é feito para o objeto é feito para o fundo, no entanto os valores dos parâmetros não são idênticos. Por exemplo, para se detectar os bons candidatos de fundo da imagem apresentada na Figura 7.8, foi utilizado o limiar de totalização igual a 3,50 e limiar de decaimento = 0,05. A Figura 7.10, mostra os pontos detectados. Nesta imagem, o decaimento acumulado e a velocidade de decaimento não surtiram efeito na eliminação de candidatos. Todos os candidatos estão bem localizados e produzirão, após a disputa, pontos dominantes idênticos aos apontados pelo *ground-truth*.

Decisão sobre a dominância

Uma vez que todos (ou quase todos) os maus candidatos foram eliminados, resta, apenas, comparar os bons candidatos e decidir quem é o ponto dominante de uma dada região da imagem.

O algoritmo adotado aqui é o mesmo utilizado no EML e no CLP. A única diferença é a necessidade da extração de dois contornos. O contorno do objeto e o contorno do fundo. Os candidatos a vértices de convexidades do objeto são identificados no contorno do objeto. Então, estipula-se uma janela unidimensional com centro em cada candidato, por exemplo, seis pixels antes do candidato e seis pixels depois. Dentro deste intervalo averigua-se a existência de outros candidatos e suas totalizações são comparadas. O mais fraco é eliminado.

Como foi explicado no EML, no item 5.5, há a preocupação de que o posicionamento da janela obedeça a uma ordem crescente. Isto é, a janela é colocada, primeiro, sobre o candidato com a menor totalização, depois no candidato com a segunda menor totalização e assim sucessivamente, até que chegue a vez do candidato com a maior totalização. Este procedimento evita que eliminações fora de ordem excluam pontos dominantes reais.

Os mesmos procedimentos adotados para a determinação de pontos dominantes no contorno do objeto (convexidades do objeto) são utilizados na determinação de pontos dominantes no contorno do fundo (concavidades do objeto). A imagem 7.11 mostra o resultado da detecção de pontos dominantes usando o DoG_zero.

Figura 7.10 – Resultados do algoritmo de avaliação multiescala para as convexidades do fundo branco.

Figura 7.11 – Pontos dominantes detectados pelo DoG_zero. As setas marcam falsos positivos.

7.3 Conclusão do capítulo

Neste capítulo é proposto e descrito o funcionamento do detector DoG_zero, um detector de *corners* de contornos baseado no modelo de diferença de Gaussianas (DoG). A única diferença em relação ao modelo DoG é que uma das parcelas da diferença é a própria imagem binária (não suavizada).

A diferença entre uma imagem binária e sua versão suavizada revela as modificações causadas pela suavização da imagem, as quais são mais intensas nas protuberâncias convexas, que podem ser do fundo ou do objeto. Para detectar cada tipo, basta comutar as parcelas da diferença.

Indiretamente este capítulo sugere um modelo para as células ganglionares da retina, mostrando que estas teriam a capacidade de detectar convexidades ao longo de fronteiras entre regiões claras e escuras.

As respostas do detector DoG_zero e do detector EML não são idênticas devido às diferenças no algoritmo que interpreta o comportamento multiescala. Porém, no que se refere ao cerne dos detectores, isto é, todo o processo anterior às decisões multiescala, as respostas só diferem nos valores, porém o comportamento é idêntico. Isto indica que a suavização pode ser imprescindível (na visão biológica) para a detecção de *corners* ao longo da fronteira entre regiões contrastantes.

Capítulo 8 – Resultados

8.1 Introdução

Este capítulo apresenta os resultados obtidos com a aplicação dos detectores de *corners* (EML, CLP e DoG_Zero), propostos nesta tese, em um conjunto de imagens binárias. Este conjunto totaliza cinquenta e nove imagens diferentes e foi montado a partir da combinação de três bases de imagens distintas. Destas, oito imagens foram utilizadas em vários experimentos ao longo dos últimos quarenta anos e podem ser obtidas em Sarfraz (2008), juntamente com seu *ground truth* e os resultados de comparações entre alguns detectores que as utilizaram. Das cinquenta e uma imagens restantes, vinte foram extraídas da base de imagens em MPEG-7 (2013) e trinta e uma extraídas de Gestalt ReVision (2016). A combinação de diferentes bases de imagens gerou um conjunto bastante heterogêneo, cujos objetos contidos nas imagens possuem tamanhos variados. As imagens maiores, em geral possuem “ruídos” maiores, o que pode levar os detectores a não desprezá-los.

Neste capítulo são conduzidos três testes. O primeiro busca mostrar a equivalência nas respostas dos três detectores. O segundo compara as respostas dos três detectores com um *ground truth*, cujo objetivo é averiguar a semelhança entre suas respostas e as respostas de observadores humanos. No terceiro teste, outro *ground truth* é confrontado com as respostas dos detectores propostos e de outros detectores publicados na literatura de visão computacional.

8.2 Comparando as saídas dos detectores propostos

O objetivo principal da tese é mostrar que a suavização Gaussiana, em imagens binárias, fornece tudo o que é necessário para se detectar curvaturas ao longo das fronteiras entre fundo e objeto. O processo de suavização “marca” com intensidades distintas as características encontradas ao longo do contorno (concavidades, segmentos de reta e convexidades) essenciais para o entendimento de formas. Os detectores EML, CLP e DoG_zero são três maneiras distintas para ler tais marcações.

Esses detectores possuem o mesmo núcleo de convolução. No caso do DoG_zero, o núcleo embute o método de detecção, o qual é composto de Gaussianas cujas aberturas são as mesmas usadas na construção dos outros dois detectores. Nos capítulos 6 e 7, principalmente

em relação às Figuras 6.4 e 7.6, afirmou-se que as respostas dos três métodos seriam muito próximas. Assim, este teste tem a função de mostrar essa semelhança. Se dois detectores quaisquer realizam eficientemente suas funções, suas respostas devem ser semelhantes, porém imagina-se que os erros de detecção não sejam os mesmos, exceto se trabalharem sobre a mesma base teórica.

O funcionamento dos três detectores propostos se baseia na ideia de quantificação de suavização. Porém, a interpretação das respostas multiescala pode ser feita de muitas maneiras diferentes. Ao longo do desenvolvimento dos detectores, muitos algoritmos foram testados em busca de melhores resultados, porém, ainda não se chegou a um método ideal capaz de interpretar cada modificação produzida pela suavização multiescala. Cada detector proposto possui o seu algoritmo próprio para a análise multiescala. Isto dificulta a comparação com o intuito de mostrar a sua base teórica comum. A ideia, aqui, não é de obter a melhor detecção de pontos dominantes, mas a de mostrar que até os ruídos detectados são os mesmos para os três detectores.

As Tabelas 8.1 a 8.4 mostram as respostas coincidentes na detecção de pontos dominantes (e ruídos não eliminados) para os três detectores. Algumas das respostas são mostradas na Figura 8.1. Nesta, pode-se observar que a quantidade e a localização dos “erros” de detecção e dos vértices das regiões convexas são bem semelhantes. A Figura 8.1 está organizada em colunas, contendo as respostas dos detectores CLP, DoG_zero e EML, respectivamente.

Tabela 8.1 – Interseção das respostas dos três detectores

Imagem	EML	CLP	DoG_zero	Coincidências
002aviao	13	13	13	13
003jacare	17	15	14	14
007braço	8	8	8	8
008seta	5	5	5	5
012machado	6	6	6	6
028pardal	12	9	11	9
049gato	13	10	9	9
106luva	9	9	9	9
d.gif	19	17	17	17
device0-10	22	20	22	20
E.gif	10	11	9	9
f.gif	16	15	15	15
g.gif	21	20	21	20
h.gif	19	17	17	17
shape1.gif	17	17	15	15
Total de pontos detectados = 208 pontos, interseção = 186 pontos -> 89,42%				

Tabela 8.2 – Interseção das respostas dos detectores EML e CLP

Imagem	EML	CLP	Coincidências
002aviao	13	13	13
003jacare	17	15	15
007braço	8	8	8
008seta	5	5	5
012machado	6	6	6
028pardal	12	9	9
049gato	13	10	10
106luva	9	9	9
d.gif	19	17	17
device0-10	22	20	20
E.gif	10	11	10
f.gif	16	15	15
g.gif	21	20	20
h.gif	19	17	17
shape1.gif	17	17	17
Total de pontos detectados = 208, interseção = 191 pontos -> 91,83%			

Tabela 8.3 – Interseção das respostas dos detectores EML e DoG_zero (continua)

Imagem	EML	DoG_zero	Coincidências
002aviao	13	13	13
003jacare	17	14	14
007braço	8	8	8
008seta	5	5	5
012machado	6	6	6
028pardal	12	11	11
049gato	13	9	9
106luva	9	9	9
d.gif	19	17	17
device0-10	22	22	22
E.gif	10	9	8

Imagem	EML	DoG_zero	Coincidências
f.gif	16	15	15
g.gif	21	21	21
h.gif	19	17	17
shape1.gif	17	15	15

Total de pontos detectados = **207**, interseção = **190 pontos** -> **91,79%**

(conclusão)

Tabela 8.4 – Interseção das respostas dos detectores CLP e DoG_zero

Imagem	CLP	DoG_zero	Coincidências
002aviao	13	13	13
003jacare	15	14	14
007braço	8	8	8
008seta	5	5	5
012machado	6	6	6
028pardal	9	11	9
049gato	10	9	9
106luva	9	9	9
d.gif	17	17	17
device0-10	20	22	20
E.gif	11	9	9
f.gif	15	15	15
g.gif	20	21	20
h.gif	17	17	17
shape1.gif	17	15	15

Total de pontos detectados = **197**, interseção = **186 pontos** -> **94,42%**

Figura 8.1 – Semelhança na detecção de vértices de convexidades e ruídos.

8.3 Comparando as respostas dos detectores propostos com um *ground-truth*

A precisão de qualquer detector de pontos dominantes só pode ser julgada se as suas reais posições na imagem forem conhecidas. Um *ground-truth* foi criado para as cinquenta e uma imagens utilizadas neste teste. Seis observadores com idades variando de 12 a 60 anos, incluindo o próprio autor, marcaram os pontos do *ground-truth*. Os pontos mais votados foram os escolhidos.

As respostas dos três detectores propostos são confrontadas com esse *ground-truth*. Os resultados das comparações são avaliados com as medidas de efetividade precisão, revocação e acurácia, as quais são representadas pelas equações 8.1, 8.2 e 8.3 respectivamente. O parâmetro MP é a quantidade de pontos coincidentes entre o detector e o *ground-truth*, TD é a quantidade de pontos detectados e GT é a quantidade total de pontos do *ground-truth* (1143 pontos).

O valor perfeito para a precisão é 1.0, significando que cada ponto detectado é relevante. No entanto, este valor não possui qualquer significado ao se considerar a detecção de todos os pontos verdadeiros. Isto é, existe uma grande diferença entre afirmar que todos os pontos detectados são verdadeiros e afirmar que todos os pontos verdadeiros foram detectados.

Da mesma forma, o valor perfeito para a revocação é 1.0, significando que todos os pontos verdadeiros foram detectados. Porém nada é informado a respeito da detecção de pontos falsos. Assim, realizar comparações entre detectores utilizando, apenas, essas duas medidas não é ideal.

A medida acurácia (MOKHTARIAN; MOHANNA, 2006) informa o quanto as respostas do detector estão próximas às respostas do *ground-truth*. O Valor ideal para a acurácia é 100%. Este indica que todos os pontos verdadeiros e somente os pontos verdadeiros foram detectados.

$$precisão = \frac{MP}{TD} \quad (8.1)$$

$$revocação = \frac{MP}{GT} \quad (8.2)$$

$$acurácia = 100 \left(\frac{precisão + revocação}{2} \right) \quad (8.3)$$

A Tabela 8.5 apresenta os resultados dessas medidas para os três detectores propostos. As Tabelas 8.6 a 8.8 apresentam os pontos detectados por imagem e por detector, listando seus respectivos pontos coincidentes com o *ground-truth*, bem como os falsos positivos e falsos negativos. Tais valores são usados na construção da Tabela 8.5.

Tabela 8.5 – Resultados das comparações das respostas dos detectores propostos com o ground-truth.

Detector	Falsos Positivos	Falsos Negativos	Precisão	Revocação	Acurácia
DoG_zero	33	127	0,968541	0,888889	92,87%
EML	54	160	0,947927	0,860017	90,40%
CLP	45	208	0,954082	0,818023	88,61%

Tabela 8.6 – Respostas do detector EML. (continua)

Imagem	Pontos do Ground-Truth	Pontos Coincidentes	Falsos Positivos	Falsos Negativos
002aviao	25	19	0	6
003jacare	28	22	2	6
004ancor	23	14	4	9
007braço	13	11	0	2
008seta	7	7	0	0
012machado	8	8	0	0
013carrinho	23	20	0	3
024barata	28	26	0	2
028pardal	14	11	1	3
040Borbol	16	15	3	1
043camelo	25	21	5	4
048cenoura	27	23	3	4
049gato	16	13	0	3
088dedo	8	5	0	3
089peixe	20	16	0	4
105oculos	20	11	0	9
106luva	16	14	0	2
112revolver	17	15	1	2
114martelo	8	7	0	1
118chapeu	6	5	0	1
120helicopt	46	41	0	5
126canguru	26	22	0	4
134perna	9	6	0	3
169pinguim	15	12	1	3
174cachimbo	6	5	0	1
176alicate	8	8	0	0
177tomada	14	14	0	0
245guardchuva	12	11	1	1
256moinho	21	21	3	0
258taça	8	7	0	1
bat-4	28	25	0	3
bone-8	20	15	0	5
device0-7	10	10	0	0
device0-8	20	16	0	4
device0-9	20	13	0	7
device0-10	40	40	0	0
device0-11	80	74	0	6

Imagem	Pontos do Ground-Truth	Pontos Coincidentes	Falsos Positivos	Falsos Negativos
device0-12	93	81	12	12
device0-13	59	42	5	17
device0-16	20	16	0	4
device0-18	15	15	0	0
device1-4	79	78	5	1
device4-10	4	3	5	1
device5-3	16	16	0	0
elephant-2	33	26	0	7
fish-a026	19	18	0	1
fork-19	22	19	1	3
hammer-11	11	8	0	3
ja3-4	10	10	0	0
key-15	17	16	2	1
misk-11	14	12	0	2
Totais:	1143	983	54	160
(conclusão)				

Tabela 8.7 – Respostas do detector CLP. (continua)

Imagem	Pontos do Ground-Truth	Pontos Coincidentes	Falsos Positivos	Falsos Negativos
002aviao	25	19	0	6
003jacare	28	17	2	11
004ancor	23	14	4	9
007braço	13	11	0	2
008seta	7	7	0	0
012machado	8	8	0	0
013carrinho	23	19	0	4
024barata	28	25	1	3
028pardal	14	11	1	3
040Borbol	16	14	2	2
043camelo	25	18	1	7
048cenoura	27	19	2	8
049gato	16	13	2	3
088dedo	8	4	0	4
089peixe	20	15	0	5
105oculos	20	12	0	8
106luva	16	14	0	2
112revolver	17	15	2	2
114martelo	8	7	0	1
118chapeu	6	4	0	2
120helicopt	46	34	0	12
126canguru	26	22	1	4
134perna	9	6	0	3
169pinguim	15	12	1	3
174cachimbo	6	4	0	2
176alicate	8	6	0	2
177tomada	14	10	0	4
245guardchuva	12	10	1	2
256moinho	21	20	4	1
258taça	8	6	0	2
bat-4	28	24	0	4
bone-8	20	14	0	6
device0-7	10	10	0	0
device0-8	20	16	0	4
device0-9	20	10	0	10
device0-10	40	40	0	0

Imagem	Pontos do Ground-Truth	Pontos Coincidentes	Falsos Positivos	Falsos Negativos
device0-11	80	76	0	4
device0-12	93	77	7	16
device0-13	59	44	5	15
device0-16	20	18	0	2
device0-18	15	15	2	0
device1-4	79	76	1	3
device4-10	4	4	4	0
device5-3	16	16	0	0
elephant-2	33	27	2	6
fish-a026	19	14	0	5
fork-19	22	19	0	3
hammer-11	11	7	0	4
ja3-4	10	8	0	2
key-15	17	15	0	2
misk-11	14	9	0	5
Totais:	1143	935	45	208
(conclusão)				

Tabela 8.8 – Respostas do detector DoG zero. (continua)

Imagem	Pontos do Ground-Truth	Pontos Coincidentes	Falsos Positivos	Falsos Negativos
002aviao	25	25	0	0
003jacare	28	23	0	5
004ancor	23	16	3	7
007braço	13	12	0	1
008seta	7	7	0	0
012machado	8	8	0	0
013carrinho	23	20	4	3
024barata	28	23	0	5
028pardal	14	12	0	2
040Borbol	16	14	0	2
043camelo	25	22	1	3
048cenoura	27	22	1	5
049gato	16	14	0	2
088dedo	8	4	1	4
089peixe	20	15	0	5
105oculos	20	11	0	9
106luva	16	14	0	2
112revolver	17	15	1	2
114martelo	8	7	0	1
118chapeu	6	6	0	0
120helicopt	46	38	0	8
126canguru	26	23	0	3
134perna	9	9	0	0
169pinguim	15	13	2	2
174cachimbo	6	6	0	0
176alicate	8	8	0	0
177tomada	14	14	0	0
245guardchuva	12	12	1	0
256moinho	21	21	2	0
258taça	8	8	0	0
bat-4	28	25	1	3
bone-8	20	20	0	0
device0-7	10	10	0	0
device0-8	20	20	0	0

Imagem	Pontos do Ground-Truth	Pontos Coincidentes	Falsos Positivos	Falsos Negativos
device0-9	20	20	0	0
device0-10	40	40	0	0
device0-11	80	80	2	0
device0-12	93	77	8	16
device0-13	59	42	0	17
device0-16	20	20	0	0
device0-18	15	15	0	0
device1-4	79	71	1	8
device4-10	4	4	2	0
device5-3	16	16	0	0
elephant-2	33	27	2	6
fish-a026	19	19	1	0
fork-19	22	20	0	2
hammer-11	11	9	0	2
ja3-4	10	10	0	0
key-15	17	17	0	0
misk-11	14	12	0	2
Totais:	1143	1016	33	127
(conclusão)				

O *ground-truth* utilizado neste teste é apresentado no anexo A.

8.4 Comparação entre os detectores propostos e outros disponíveis na literatura

Neste teste é utilizado o *ground-truth* disponível em Sarfraz (2008), construído para um conjunto de oito imagens. Tanto os detectores propostos, quanto os detectores que tiveram seus resultados publicados neste mesmo conjunto de imagens são confrontados com o *ground-truth*. A Figura 8.2 apresenta o *ground-truth* utilizado (as marcações foram realçadas). As Figuras 8.3 a 8.5 mostram os resultados dos detectores EML, CLP e DoG_zero respectivamente, porém são mostradas apenas as imagens que não obtiveram resultado idêntico ao *ground-truth*.

Os detectores EML, CLP e DoG_zero têm seus resultados comparados com dez detectores diferentes cujos resultados estão publicados em Zhang et al (2010), Sarfraz (2008), Sobania e Evans (2005), Chetverikov e Szabo (1999), Smith e Brady (1997), Shi e Tomasi (1994), Beus e Tiu (1997), Freeman e Davis (1977), Rosenfeld e Weska (1975) e Rosenfeld e Johnston (1973). Os resultados das comparações estão registrados na Tabela 8.9. Para facilitar o entendimento, na Tabela 8.9 é registrado o nome do primeiro autor de cada trabalho.

Figura 8.2 – O *ground-truth* das oito imagens.

Figura 8.3 – Os erros de detecção do EML. As setas indicam os falsos positivos (FP) e os falsos negativos (FN).

Figura 8.4 – Os erros de detecção do CPL. As setas indicam os falsos positivos (FP) e os falsos negativos (FN).

Figura 8.5 – Os erros de detecção do DoG_zero. As setas indicam os falsos positivos (FP) e os falsos negativos (FN).

Tabela 8.9 – Resultados das comparações das respostas de 13 detectores com o *ground-truth*.

Detector	Falsos Positivos	Falsos Negativos	Precision	Recall	Accuracy
DoG_zero	4	3	0,960784	0,970297	96,55%
EML	6	3	0,942308	0,970297	95,63%
CLP	7	3	0,933333	0,970297	95,18%
Sarfraz	7	2	0,9346	0,980392	95,75%
Zhang	25	3	0,7984	0,970588	88,45%
Chetverikov	27	12	0,7692	0,882353	82,58%
Beus	10	29	0,8795	0,715686	79,76%
Freeman	15	28	0,8315	0,72549	77,85%
Rosenfeld 75	37	25	0,6754	0,754902	71,52%
Rosenfeld 73	37	30	0,6606	0,705882	68,32%
Sobania	17	10	0,6909	0,791667	74,13%
Shi	36	33	0,6571	0,676471	66,68%
Smith	464	5	0,1729	0,95098	56,19%

8.5 Discussão

O primeiro teste apresentado no item 8.2 mostra um forte relacionamento entre as respostas dos três detectores propostos: 89,42% do total de pontos detectados possui a mesma localização para os três detectores. A princípio, esta porcentagem pode parecer baixa, porém é necessário considerar a diferença nos intervalos numéricos utilizados por cada detector. A seleção dos pontos é baseada em limiares, o que define a quantidade de pontos aceitos como *corners*. Embora se tenha buscado por uma equivalência entre os limiares adotados nos detectores, não se obteve uma perfeita interpretação de como funciona a variação numérica dentro de cada intervalo. Por exemplo, todas as comparações onde o EML está presente, a quantidade de pontos coincidentes é menor (89,43%, 91,83% e 91,79% contra 94,42%). Conjectura-se que a função logarítmica utilizada na equação de entropia seja responsável por modificar a velocidade em que os números variam dentro do intervalo, tornando difícil obter um limiar proporcional aos usados nos outros detectores.

O segundo e o terceiro testes apresentados nos itens 8.3 e 8.4, mostraram bons resultados ao serem comparados com seus respectivos *ground-truth*. A maneira de como utilizar e combinar as informações multiescala interfere diretamente na escolha dos pontos dominantes. Por exemplo, na Figura 8.6 reproduz-se a imagem F do *ground-truth*, todos os três detectores deixaram de mostrar o ponto dominante indicado pela letra X. Este ponto é visível aos três detectores, porém possui uma vida útil curta, sua saliência decai quase totalmente a partir da terceira escala (17x17). Deve-se observar que o ponto em questão é o vértice de uma pequena convexidade, a qual está imersa em uma concavidade, significando que toda a estrutura adjacente a esta região interfere negativamente em sua evolução

multiescala. Uma das funções do algoritmo é permitir, ou não, que pontos desse tipo sejam aceitos. A implementação utilizada nestes testes não admite pontos com esse comportamento, o que não significa a incapacidade de detectá-lo. O funcionamento dos três métodos parte da escala mais baixa para a mais alta, isto significa que qualquer pequeno ruído é detectado. Os limiares adotados definem quais os pontos permanecerão “vivos” até o momento da disputa pela dominância. A Figura 8.7 mostra a evolução multiescala do ponto X nos três detectores.

Alguns fatores podem influenciar na construção do *ground-truth*, como a vivência de cada observador, as instruções dadas aos observadores de como devem ser escolhidos os pontos, a subjetividade e a liberdade de decisão humana.

Como um exemplo de vivência do observador é apresentado a Figura 8.8. Nesta é mostrada a ampliação de parte da imagem H do *ground-truth* (Figura 8.2). Um observador com boas noções de simetria, ou um aeromodelista, teria muita resistência em marcar o ponto A como o ponto dominante da região (o ponto A foi o escolhido para o *ground-truth*), o mais provável é que escolhesse o ponto B. Em versões anteriores do algoritmo para a análise multiescala, o ponto B era o escolhido. Para isso, bastava considerar o decaimento de saliência inicial que o ponto A apresenta. Para esse mesmo tipo de observador com experiência em desenho de aeronaves, a marcação do ponto C implicaria na marcação do ponto D. Para o *ground-truth* o ponto D é um falso positivo.

A Figura 8.9 mostra a evolução multiescala dos pontos A, B e Z nos três detectores propostos. O ponto Z é o simétrico do ponto B, o qual está marcado no *ground-truth*. Deve-se observar a semelhança nos comportamentos dos pontos B e Z ao longo de todas as escalas. O ponto A tem um comportamento inicial completamente destoante dos outros dois pontos. A perda de saliência inicial é comum em ruídos ou em estruturas bem pequenas. A saliência é perdida até que se chegue à “escala natural” da região. A partir da quarta escala todos os pontos iniciam perda de saliência. Isto ocorre devido à janela de convolução envolver uma concentração de fundo maior que a concentração de objeto, o que caracteriza uma convexidade. Então, o ponto que se inicia como um vértice de concavidade em baixa escala pode se transformar em um ponto qualquer de uma região convexa de alta escala.

Dependendo das instruções dadas aos observadores, é possível induzi-los a não marcar pontos em segmentos de arcos. Em relação à Figura 8.10, o que motivaria um observador a marcar o ponto R e não marcar o ponto S ou o ponto T? De forma bastante correlacionada, está a subjetividade e a liberdade das decisões humanas. Os limiares utilizados por um observador podem variar de região para região. Para tomar decisões semelhantes, um algoritmo teria que assumir limiares locais em vez de globais. Por exemplo, na Figura 8.11

uma análise local permitiria a marcação do M sem a obrigação de marcar o ponto N. Com o uso de limiar global, a não aceitação do ponto N acarreta a não aceitação do ponto M. Isto provocou um falso negativo nas respostas dos três detectores para o ponto M. Deve-se perceber que o ponto N é o vértice de um ângulo mais saliente do que o ponto M, embora diferença seja desprezível. O mesmo se aplica para a Figura 8.10, Não há como permitir a seleção do ponto R sem permitir a seleção dos pontos S e T.

Figura 8.6 – O *ground-truth* da imagem F.

Figura 8.7 – Evolução multiescala do ponto X da Figura 8.6 sob o ponto de vista dos três detectores propostos.

Figura 8.8 – Ampliação de parte da imagem H.

Figura 8.9 – Evolução multiescala dos pontos A, B e Z da Figura 8.8.

Figura 8.10 – O ground-truth da imagem E.

Figura 8.11 – A explicação de um falso negativo da imagem F.

Capítulo 9 – Conclusões e trabalhos futuros

9.1 Conclusões

Esta tese apresenta o desenvolvimento de três detectores multiescala de pontos dominantes, os quais têm seu funcionamento baseado na quantificação das modificações causadas pelo processo de suavização Gaussiana.

A detecção de pontos dominantes é um processamento primário, de baixo nível, cujos resultados podem ser aproveitados em diferentes tarefas de nível mais alto em análise de formas, por exemplo, a recuperação de imagens por conteúdo em uma base de dados de imagens (CBIR).

No desenvolvimento deste trabalho houve a preocupação em buscar embasamentos físicos, psicofísicos e biológicos. Pode-se dizer, então, que os detectores aqui propostos são inspirados e fundamentados em processos naturais.

Os resultados apresentados no capítulo 8 mostram que o contraste entre duas regiões homogêneas disponibiliza outras informações além da classificação de brilho de um pixel. Com a utilização de um núcleo de convolução adequado é possível extrair informações angulares, ou de curvatura, ao longo da fronteira entre as regiões contrastantes. Permitindo classificar cada um dos pixels fronteiros como vértices de concavidades, de convexidades, ou como pertencentes a um segmento de reta. Ou seja, toda a informação necessária para se descrever a forma bidimensional de um objeto.

A teoria que embasa todo o trabalho pode ser resumida na afirmação de que a suavização Gaussiana pode ter outra função além da tradicional filtração. A suavização pode ser utilizada como um processo de marcação, “pintando” cada pixel com a “tinta” obtida a partir da média ponderada de sua vizinhança local. Se o valor do pixel não sofre modificação é porque pertence a uma região homogênea. A quantidade de modificação que um pixel pode sofrer é gradual e proporcional à quantidade de pixels de mesmo valor existentes na vizinhança considerada. Se houver poucos pixels de um tipo e muitos de outro, a modificação será grande. Uma modificação grande significa que naquela vizinhança existe uma concavidade ou uma convexidade.

Uma vez que a marcação Gaussiana foi realizada, basta quantificá-la para identificar as características da fronteira. Os três detectores propostos representam os tipos de quantificação vislumbrados. O primeiro se baseia na computação da entropia de Shannon

(SHANNON,1948), a qual só foi utilizada após a comprovação de que poderia ser usada para quantificar misturas de brilho em imagens. No capítulo 3 é apresentado um estudo mostrando que o seu uso é apropriado.

Os outros dois detectores foram construídos a partir da compreensão do funcionamento do primeiro. No segundo (CLP), em vez de se computar a entropia, computa-se apenas as proporções de objeto e fundo presentes na vizinhança. O procedimento é idêntico ao do primeiro, só não se calcula os logaritmos presentes na equação de entropia, o que pode ser computacionalmente mais vantajoso.

O terceiro detector é baseado numa pequena modificação do modelo DoG, que obriga uma das Gaussianas envolvidas na subtração a possuir um desvio padrão bem baixo, próximo à zero. A suavização de imagens com Gaussianas muito estreitas (desvio padrão abaixo de 0,5) produz modificações desprezíveis. Por isso, a imagem suavizada com a Gaussiana mais estreita é substituída pela imagem binária original. Assim, a DoG se transforma na DoG_Zero, onde a subtração é feita entre uma versão suavizada da imagem e a própria imagem binária original.

Os resultados descritos no capítulo 8 indicam que o objetivo principal deste trabalho foi atingido. A suavização pode ser considerada como um meio necessário para se extrair as informações angulares de pixels de fronteira.

Os detectores propostos nesta tese possuem núcleos de convolução semelhantes. No EML e no CLP, o núcleo é baseado na soma de Gaussianas. No DoG_zero, o núcleo é baseado na diferença de Gaussianas. Porém é comum a todos, a restrição de que uma das Gaussianas deva ser bem estreita para promover a correta localização das características, independentemente da escala considerada.

Grande parte do trabalho, aqui apresentado, lida com a análise multiescala. Esta é necessária para identificar as estruturas da imagem que possuem real influência no aspecto global da forma. A técnica multiescala proposta para os detectores, bem como os núcleos de convolução são inovadores. A análise multiescala é realizada somente nos pontos que possuem a capacidade de se tornarem vértices de estruturas importantes da forma bidimensional. São usadas sete escalas que variam de 0,5 a 23,8333. Estas permitem avaliar adequadamente as estruturas contidas em imagens de até 512 x 512 pixels. Em imagens maiores, as estruturas pequenas ocupam muitos pixels. Para avalia-las mais precisamente seria necessário aumentar o desvio padrão para além de 23,83. É desejável a criação de um algoritmo que determine o tamanho do objeto contido na imagem e define as escalas a serem usadas.

A análise da evolução do candidato ao longo das escalas permite perceber quais os tipos de variação que podem ocorrer. Foram detectados dois tipos de variação. O primeiro ocorre devido à acomodação do candidato à configuração dominante da região em que este se situa. Esta configuração define o valor do candidato em sua escala mais alta. Por exemplo, um candidato situado sobre uma fronteira em linha reta apresenta decaimentos sucessivos até o seu valor se aproximar ao valor que representa o ângulo de 180° . O segundo tipo de variação é causada pela interferência de estruturas adjacentes. Esta ocorre quando a janela de suavização cresce além dos limites da estrutura observada, incluindo parte de outras estruturas adjacentes. A interpretação dada a esta interferência é a de que se está obtendo o limite da estrutura observada. Algo similar à escala natural da estrutura. Também foi identificada uma variação causada pela acomodação da distribuição dos pesos estatísticos do núcleo de convolução à grade digital. Esta foi neutralizada.

A ideia da suavização como um marcador de características de fronteira foi aplicada na detecção de pontos dominantes. Embora, os resultados apresentados coloquem os detectores propostos entre os melhores detectores de pontos dominantes disponíveis na literatura, ainda há muito que aperfeiçoar para se chegar aos resultados apresentados por observadores humanos. Foi percebido que para se detectar o ponto mais importante de uma região, os observadores humanos utilizam processos cognitivos, que ocorrem em estágios superiores do cérebro. Não é uma mera detecção, são comparações e decisões baseadas em experiências acumuladas. Pode-se conjecturar que a grande quantidade de detectores propostos na literatura e a não existência de um detector ótimo, sejam consequências da complexidade envolvida no processo de decisão humana.

A construção do núcleo se baseou na combinação de ideias extraídas da difração em aberturas circulares, do funcionamento dos campos receptivos *center-surround* da retina e da distribuição de cones na retina (os capítulos 3 e 4 apresentam os detalhes a esse respeito). Sendo assim, é possível que a retina possua a capacidade de extrair informações de curvatura a partir do contraste. Se essa hipótese for verdadeira, será necessário reinterpretar como o córtex visual trata as informações que chegam pelo nervo óptico. Este será o objetivo do próximo trabalho, elaborar um modelo capaz de receber informações angulares, alinhá-las e reconstruir a forma observada na cena. Na seção de trabalhos futuros são apresentados os rudimentos desse modelo.

9.3 Trabalhos futuros

Ao supor que a retina transmite informações angulares ao córtex, surge a primeira questão: O sistema visual possui alguma referência que indique onde inicia o 0° e qual é o sentido do crescimento do ângulo? A comunicação olho – córtex pode ser comparada com a comunicação via telefone entre duas pessoas. Imagine que uma pessoa necessite descrever para outra, via telefone, a forma bidimensional de um objeto. Imagine, também, que ambas possuam as coordenadas cartesianas dos pontos. Não se podem conectar os pontos com segmentos de retas, pois não se sabe a sequência das conexões. A cada coordenada cartesiana é associado um ângulo. Infelizmente, não há uma convenção preestabelecida entre os interlocutores a respeito de onde se inicia a contagem do ângulo e para qual direção ele cresce. Assim, o receptor tem as coordenadas, bem como o ângulo em cada uma delas. Este é um possível cenário para o córtex visual. Quais ações poderiam ser tomadas para solucionar o problema de reconstruir a forma a partir dessas informações?

Uma solução possível seria a de desenhar cada ângulo (como feixes de luz) em todas as possíveis direções e determinar aquela em que o ângulo se alinha aos ângulos adjacentes. A repetição deste processo para todos os ângulos alcançará o alinhamento do conjunto completo.

Aqui, levanta-se a hipótese de que algumas células do córtex visual implementam um sistema de alinhamento angular, e que a ilusão conhecida por contornos ilusórios, ou de Kanizsa seja o reflexo do funcionamento desse sistema. A Figura 9.1 apresenta alguns exemplos de contornos ilusórios. Deve-se observar que dependendo do alinhamento angular as bordas podem ser segmentos de retas ou arcos, podendo até representar circunferências. Vale mencionar que os detectores propostos são capazes de detectar *corners* em arcos e circunferências.

Imagens de ressonância magnética funcional (fMRI) incluindo mapeamento de populações de campos receptivos permitiram verificar, com clareza, que os contornos ilusórios ativam as células do córtex visual primário (V1) (KOK; DE LANGE, 2014).

O córtex visual primário é organizado em hipercolunas, onde várias células sensíveis a diferentes direções representam um único pixel do campo visual (URSINO; LARA, 2004). A Figura 9.2 mostra essa organização colunar. Deve-se notar que essa é uma organização apropriada para a implementação da ideia de se desenhar um mesmo ângulo em diferentes direções.

Figura 9.1 – Contornos ilusórios (adaptado de Yang, Yue e Wu (2015)).

Figura 9.2 – Hipercolunas do córtex visual V1 (adaptado de Ursino e Lara (2004)).

Na Figura 9.1, deve-se considerar que existem duas cores competindo por resposta, isto é, as duas estão ativando os campos receptivos da retina, a cor branca do papel e a cor preta dos indutores (*pacman*). Ambos são “enxergados”. Se os indutores não estiverem alinhados, serão enxergados como quatro objetos quaisquer num fundo branco. Porém, com o alinhamento, o fundo recebe uma limitação, tornando-o um candidato a objeto. Quanto mais

próximos estiverem os indutores alinhados, maior é a probabilidade de o espaço branco ser enxergado como um objeto. Caso os indutores alinhados se toquem mutuamente, o espaço branco estará totalmente limitado, isto é, se torna um objeto e deixa de ser ilusão. A Figura 9.3 mostra como a aproximação dos indutores torna a ilusão mais forte.

As Figuras 9.4 e 9.5 mostram uma possível arquitetura do córtex e a sua resposta às informações angulares que chegam da retina. Os itens *a* e *b* da Figura 9.4 podem ser uma possível representação da informação angular em quatro *corners* detectados na retina. Não se considerou as informações de borda entre os vértices para facilitar a representação de uma ideia ainda não muito clara. O item *c* desta mesma figura é uma visão de topo da ativação do córtex visual em quatro hipercolunas. A Figura 9.5 é uma representação de perfil do item *c* da Figura 9.4. A intenção é mostrar que o mesmo ângulo (90°) é ativado em níveis diferentes de cada hipercoluna, visto que cada nível representa a direção do ângulo na cena. Deve ficar claro que os prolongamentos que se interconectam nas Figuras 9.4 e 9.5 não representam dendritos de uma célula específica, mas um emaranhado de células simples interconectadas.

Está sendo considerado que o *corner* detectado é o vértice do ângulo. Os lados do ângulo são as bordas da convexidade, as quais podem ser representadas por uma sequência de informações angulares próximas a 180° .

Figura 9.3 – A proximidade dos indutores fortalece a ilusão.

Figura 9.4 – A ativação dos ângulos no córtex visual V1: a e b) possíveis estímulos angulares; c) resposta no córtex (vista de topo).

Figura 9.5 – A ativação dos ângulos nas hipercolumnas: para que os níveis ficassem visíveis, a conexão entre eles foi feita com um único “cabo” de maior diâmetro.

A proposta de um trabalho futuro dando continuidade ao que foi desenvolvido não implica em aceitar o trabalho realizado até o momento como imutável. Os resultados mostram exatamente o contrário, novas abordagens deverão ser utilizadas para melhorar os algoritmos de análise multiescala, bem como entender completamente as variações que ocorrem com a mudança de escala. Além disso, é necessário promover um estudo a respeito de como os sistemas visuais utilizam a cor para analisar cenas reais.

9.3 Publicações

LOURO, A.; MACHADO, W.; GONZAGA, A. Smoothing: A natural way to detect contour features. **Multimedia Tools and Applications**, v. 70, n. 3, p. 2111–2124, 2012.

LOURO, A.; GONZAGA, A. Multiscale Detection of Convexities and Concavities Based on Local Computation of Weights. X WVC - Workshop de Visão Computacional. **Anais...** , 2014.

LOURO, A.; GONZAGA, A. Detecção de *corners* baseada na produção de entropia. VIII WVC - Workshop de Visão Computacional. **Anais...** , 2012.

Referências

ABDEL-DAYEM, A.; EL-SAKKA, M. Fuzzy Entropy Based Detection of Suspicious Masses in Digital Mammogram Images. In: INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, p. 4017-4022, Shanghai, China, 2005.

ABE, K. et al. Comparison of methods for detecting corner points from digital curves-a preliminary report. In: PROCEEDINGS OF THE 2ND INTERNATIONAL CONFERENCE ON DOCUMENT ANALYSIS AND RECOGNITION, 1993, Tsukuba, Japan. p. 854-857, 1993.

ALBERTAZZI, L. (Ed.). **Shapes of forms: from Gestalt psychology and phenomenology to ontology and mathematics**. Netherlands: Springer, 1999. 378 p. (Synthese library: studies in epistemology, logic, methodology, and philosophy of science, 275). ISBN 978-90-481-5098-4

ANSARI, N.; DELP, E. On detecting dominant points. **Pattern Recognition**, v. 24, n. 5, p. 441-451, 1991.

ANSARI, N.; HUANG, K.-W. Non-parametric dominant point detection. **Pattern Recognition**, v. 24, n. 9, p. 849-862, 1991.

ANTOINE, J.-P. et al. Shape characterization with the wavelet transform. **Signal Processing**, v. 62, p. 265-290, 1997.

AOYAMA, H.; KAWAGOE, M. A piecewise linear approximation method preserving visual feature points of original figures. **CVGIP: Graphical Model Image Process**, v. 53, p. 435-446, 1991.

ASADA, H.; BRADY, M. The curvature primal sketch. **IEEE Trans. Pattern Analysis and Machine Intelligence**, v. 8, p. 2-4, 1986.

ASIMOV, I. **Nove amanhãs: a última pergunta**. Rio de Janeiro: Expressão e Cultura, 1975.

ATTNEAVE, F. Some informational aspects of visual perception. **Psychological Review**, v. 61, n. 3, p. 183-193, 1954.

AURÉLIO ELETRÔNICO SÉCULO XXI: Software versão 3.0 – 1999. MGB Informática Ltda.

AWRANGJEB, M.; LU, G.; MURSHED, M. An affine resilient curvature scale-space corner detector, in: PROC. OF THE IEEE INT. CONF. ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, VOL. 1, Hawaii, USA, 2007, p. 1233-1236.

AWRANGJEB, MOHAMMAD. **Contour-based Corner Detection and Robust geometric Point Matching Techniques**. 2008. 206. doutorado, tese - Monash University, Victoria, Austrália, 2008.

AWRANGJEB, M.; LU, G.. Robust image corner detection based on the chord-to-point distance accumulation technique. **IEEE Trans. Multimedia**. v. 10, n.6, p. 1059–1072, 2008.

AWRANGJEB, M.; LU, G.; FRASER, C. A comparative study on contour-based corner detectors. In: DIGITAL IMAGE COMPUTING: TECHNIQUES AND APPLICATIONS, 2010, Los Alamitos. IEEE Computer Society, p. 92-99.

AWRANGJEB, M.; LU, G.; FRASER, C. S. Performance comparisons of contour-based corner detectors. **IEEE transactions on image processing: a publication of the IEEE Signal Processing Society**, v. 21, n. 9, p. 4167–79, 2012. IEEE.

AWRANGJEB, M.; LU, G. A Performance Review of Recent Corner Detectors. 2013 International Conference on Digital Image Computing: Techniques and Applications (DICTA). **Anais...** . p.1–8, 2013. IEEE.

BACHATENE, L.; BHARMAURIA, V.; MOLOTCHNIKOFF, S. (Ed.). **Visual Cortex - Current Status and Perspectives: Adaptation and Neuronal Network in Visual Cortex**. InTech, 2012. DOI: 10.5772/46011. Disponível em: <http://www.intechopen.com/books/visual-cortex-current-status-and-perspectives/adaptation-and-neuronal-network-in-visual-cortex>

BACKES, A.R.; BRUNO, O. M. Polygonal approximation of digital planar curves through vertex betweenness. **Information Sciences**, v. 222, p. 795-804, 2013.

BALASUBRAMANIAN, V.; BERRY, M.J. Metabolically efficient codes in the retina. arXiv:cond-mat/0105128v1 [cond-mat.soft] 6 May 2001.

BALASUBRAMANIAN, V.; STERLING, P. Receptive fields and functional architecture in the retina. **J Physiol**, v. 587, n.12, p. 2753–2767, 2009.

BARAKAT, R. Some entropic aspects of optical diffraction imagery. **Optics Communications**, v. 156, p.235–239, 1998.

BARLOW, H.B. Summation and inhibition in the frog's retina. **Journal of Physiology-London**, v. 119, p. 69–88, 1953.

BARLOW, H.B.; LEVICK W.R. the mechanism of directionally selective units in rabbit's retina. **J. Physiol.**, v. 178, p. 477-504, 1965.

BARTEN, P. **Contrast Sensitivity of the Human Eye and Its Effects on Image Quality**, SPIE Press Book, 1999.

BATTIATO, S.; GALLO, G. An Information-Theoretical Approach to Saliency Maps Construction. IN PROCEEDINGS OF EUFIT98, p. 1375-1380, Aachen, 1998.

BEN-NAIM, A. **Entropy demystified : the second law reduced to plain common sense**. Singapore: World Scientific Publishing, 2007. ISBN-13 978-981-270-052-0

BEN-NAIM, A. **A farewell to entropy: statistical thermodynamics based on information**. Singapore: World Scientific Publishing, 2008. ISBN-13 978-981-270-706-2

BERGÉ, J. et al. An ultra-fast image generator (UFig) for wide-field astronomy. **Astronomy and Computing**, v. 1, p. 23-32, 2013.

BERNARDINO, L. Percepto: O que é Psicofísica? Disponível em: <<http://www.blogpercepto.com/2010/11/o-que-e-psicofisica.html>>. Acesso em: 24/9/2015.

BIEDERMAN, I. Recognition by components: A theory of human image understanding. **Psychological Review**, v. 94, n. 2, p. 115-147, 1987.

BLAKEMORE, C.; OVER, R. Curvature detectors in human vision? **Perception**, v. 3, n. 1, p. 3–7, 1974. SAGE Publications.

BRUCE, V.; GREEN, P. R.; GEORGESON, M. A. **Visual Perception: Physiology, Psychology, & Ecology**. New York: Psychology Press, 2003. ISBN 184169-237-9
BURT, P.; ADELSON, E. The Laplacian Pyramid as a Compact Image Code. **IEEE Transactions on Communications**, v. 31, n. 4, p. 532–540, 1983. IEEE.

CAMESASCA, M.; KAUFMAN, M.; MANAS-ZLOCZOWER, I. Quantifying Fluid Mixing with the Shannon Entropy. **Macromolecular Theory and Simulations**, v.15, n.8, p. 595–607, 2006.

CHAMOLI, N.; KUKREJA, S.; SEMWAL, M. Survey and Comparative Analysis on Entropy Usage for Several Applications in Computer Vision. **International Journal of Computer Applications**, v. 97, n. 16, p. 1–5, 2014.

CHANG, C.-I.; DU, Y.; WANG, J.; GUO, S.-M.; THOUIN, P.D. Survey and comparative analysis of entropy and relative entropy thresholding techniques. **Vision, Image and Signal Processing**, v. 153, n. 6, p. 837 – 850, 2006.

CHEN, J.; ZOU, L.; ZHANG, J.; DOU, L. The Comparison and Application of Corner Detection Algorithms. **Journal of Multimedia**, v. 4, n. 6, p. 435–441, 2009.

CHETVERIKOV, D.; SZABO, Z. A simple and efficient algorithm for detection of high curvature points in planar curves. In: PROC. OF 23RD WORKSHOP OF AUSTRALIAN PATTERN RECOGNITION GROUP, 1999, Steyr, p. 175–184, 1999.

CINQUE, L.; LOMBARDI, L.; ROSENFELD, A. Evaluating digital angles by a parallel diffusion process. **Pattern Recognition Letters**, v. 16, p. 1097–1104, 1995.

CLAUSIUS, R. **The mechanical theory of heat**. London: Macmillan and CO., 1879.

COLE, R. W.; JINADASA, T.; BROWN, C. M. Measuring and interpreting point spread functions to determine confocal microscope resolution and ensure quality control. **Nature protocols**, v. 6, n. 12, p. 1929–41, 2011.

CONNOR, C. E.; BRINCAT, S. L.; PASUPATHY, A. Transformation of shape information in the ventral pathway. **Current Opinion in Neurobiology**, v. 17, n. 2, p. 140–147, 2007.

CORTICAL PATHWAYS. In: Neuroscience online an electronic textbook for the neurosciences. Disponível em: < <http://neuroscience.uth.tmc.edu/s2/chapter15.html> >. Acesso em 04 out. 2015.

COSTA, L.F.; CESAR Jr., R.M. **Shape analysis and classification: theory and practice**. Boca Raton: CRC Press, 2001. (Image Processing Series). ISBN 0-8493-3493-4.

CRANK, J. **The mathematics of diffusion**. Oxford: Clarendon Press, 1975.

CUMANI, A.; GUIDUCCI, A. Selecting feature detectors for accurate visual odometry. **WSEAS Transactions on Circuits and Systems**, v. 8, n. 10, p. 822–831, 2009. World Scientific and Engineering Academy and Society (WSEAS).

DAVIS, L. S., Understanding shape: angles and sides. **IEEE Transactions on Computers**. v. C-26, n. 3, 1977.

DE WINTER, J.; WAGEMANS, J. Perceptual saliency of points along the contour of everyday objects: A large-scale study. **Perception & Psychophysics**, v. 70, n. 1, p. 50–64, 2008.

DEERING, M. F. A photon accurate model of the human eye. In: PROCEEDINGS OF ACM SIGGRAPH 2005, SESSION: PERCEPTION, 2005, p. 649-658.

DOBBINS, A.; ZUCKER, S. W.; CYNADER, M. S. Endstopped neurons in the visual cortex as a substrate for calculating curvature. **Nature**, v. 329, n. 6138, p. 438–441, 1987.

DOBBINS, A.; ZUCKER, S. W.; CYNADER, M. S. Endstopping and curvature. **Vision Research**, v. 29, n. 10, p. 1371–1387, 1989.

DOUGLAS, D.; PEUCKER, T. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. **Cartographica: The International Journal for Geographic Information and Geovisualization**, v. 10, n. 2, 1973.

DRAGESCO, J. **High resolution astro photography**. New York: Cambridge University Press, 1995. ISBN 521415888.

DRESCHLER, L.; NAGEL, H. On the selection of critical points and local curvature extrema of region boundaries for interframe matching. In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, ICPR, 1982, p. 542-544, 1982.

DRYDEN, I.L.; MARDIA, K.V. **Statistical Shape Analysis**. Chichester: John Wiley and Sons, 1998. ISBN 0-471-95816-6.

DUTTA A.; KAR, A.; CHATTERJI, B.N. Corner detection algorithms for digital images in last three decades. **IETE Tech Rev**, v. 25, n. 3, p. 123–132, 2008.

DUTTA, A.; CHATTERJIB, B. N.; KAR, A. Comparing and Evaluating Intensity Based Spatial Domain Corner Detectors. **International Journal of Information Processing**, v. 2, n. 4, p. 48–55, 2008.

ELLIOTT, D.F.; JENKINS, L. Convolutional relationships in EO sensors. In: CONF. RECORD TWENTY-FOURTH ASILOMAR COG. SIGNALS, SYSTEMS & COMPUTERS, 1990, p. 88-92.

ELLIOTT, D.F. et al. Accuracy of Gaussian approximation for simulating EO sensor response. In: CONFERENCE RECORD OF THE THIRTIETH ASILOMAR ON SIGNALS, SYSTEMS AND COMPUTERS, 1996, v. 2, n. 3-6, p. 868 – 872.

ENROTH-CUGELL, C.; ROBSON, J.G. the contrast sensitivity of retinal ganglion cells of the cat. **J. Physiol.**, v. 187, p. 517-552, 1966.

FERRARO, M; BOCCIGNONE, G.; CAELLI, T. On the representation of image structures via scale space entropy conditions. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, v. 21, n. 11, p. 1199 – 1203, 1999.

FERRARO, M; BOCCIGNONE, G. Image contrast enhancement via entropy production. **Real-Time Imaging**, v. 10, n. 4, p. 229-238, 2004.

FIELD, D.J.. Relations between the statistics of natural images and the response profiles of cortical cells. **Journal of Optical Society America**, v. 4, p. 2379–2394, 1987.

FIELD, D. J.; HAYES, A.; HESS, R. F. Contour integration by the human visual system: evidence for a local “association field”. **Vision research**, v. 33, n. 2, p. 173–93, 1993.

FREEMAN, A. **Fourier: The analytical theory of heat – translated with notes**. Edited for the syndics of the university press. Cambridge: UNIVERSITY PRESS. 1878.

FREEMAN, H.; DAVIS, L.S. A corner-finding algorithm for chain-coded curves. **IEEE Transactions on Computers**, v. 26, p. 297–303, 1977.

FRIEDEN, B. R. How well can a lens system transmit entropy? **Journal of the Optical Society of America**, v. 58, n.8, p. 1105-1112, 1968.

GALLANT, J.; BRAUN, J.; VAN ESSEN, D. Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. **Science**, v. 259, n. 5091, p. 100–103, 1993.

GALLANT, J. L.; CONNOR, C. E.; RAKSHIT, S.; LEWIS, J. W.; VAN ESSEN, D. C. Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. **Journal of neurophysiology**, v. 76, n. 4, p. 2718–39, 1996.

GAO, X. et. al. Multiscale contour corner detection based on local natural scale and wavelet transform. **Image and Vision Computing**, v. 25, p. 890–898, 2007.

GARRIDO, A.; PEREZ, N.; GARCA-SILVENTE, M. Boundary simplification using a multiscale dominant-point detection algorithm. **Pattern Recognition**, v. 31, p. 791–804, 1998.

GESTALT_PSYCHOLOGY. In: WIKIMEDIA FOUNDATION. Disponível em: <https://en.wikipedia.org/w/index.php?title=Gestalt_psychology&oldid=680321607>. Acesso em: 25/9/2015.

GESTALT REVISION - De Winter & Wagemans, BRMIC, 2004. Stimuli-sets & Datasets. Disponível em: <<http://www.gestaltrevision.be/en/resources/supplementary-material/129-de-winter-a-wagemans-brmic-2004-stimuli-sets-a-datasets>>. Acesso em: 11/2/2016.

GLEICK, J.; **The information: a history, a theory, a flood**. New York: Pantheon Books, 2011. eISBN 978-0-307-37957-3.

GOLLISCH, T.; MEISTER, M. Eye smarter than scientists believed: Neural computations in circuits of the retina. **Neuron**, v. 65, n. 2, p.150-164, 2010.

GOSHTASBY, A.; ONEILL, W.D. Curve Fitting by a Sum of Gaussians, **CVGIP: Graphical Models and Image Processing**, v. 56, n. 4, p. 281-288, 1994.

GRAHAM, D.J.;CHANDLER, D.M.; FIELD, D.J. Can the theory of “whitening” explain the center-surround properties of retinal ganglion cell receptive fields? **Vision Res.**, v.46, n.18, p. 2901–2913, 2006.

GRAHS, T.H.; SONAR,T.H. Discrete nonlinear filters for the numerical treatment of conservation laws. **PAMM**, v. 1, n. 1, p. 426–427, 2002.

GURU, D. S.; DINESH, R.; NAGABHUSHAN, P. Boundary-based corner detection and localization using new “cornerity” index: a robust approach. In: PROCEEDINGS OF 1ST CANADIAN CONFERENCE ON COMPUTER AND ROBOTIC VISION, CRV04, p. 417-423, 2004.

HABAK, C.; WILKINSON, F.; ZAKHER, B.; WILSON, H. R. Curvature population coding for complex shapes in human vision. **Vision research**, v. 44, n. 24, p. 2815–23, 2004

HAGMANN, P. et al. Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. **RadioGraphics**, v.26, p. s205-s223, 2006.

HAN, J.H.; POSTON, T.. Chord-to-point distance accumulation and planar curvature: a new approach to discrete curvature. **Pattern Recognition Letters**. v. 22, p. 1133–1144, 2001.

HANSEN, T.; NEUMANN, H. A biologically motivated scheme for the robust detection of junctions. In H. H. BÜLTHOFF, S.-W. LEE, T. A. POGGIO, & C. WALLRAVEN (EDS.), **BIOLOGICALLY MOTIVATED COMPUTER VISION (BMCV 2002)**, p. 16–26. LNCS 2525. Berlin: Springer-Verlag, 2002.

HARRIS, C.; STEPHENS, M. A combined corner and edge detector. In Proc. of Fourth Alvey Vision Conference. **Anais...** . p.147–151, 1988.

HARTLINE, H.K. The nerve messages of the fibres of the visual pathway. **J. Opt. Soc. Am.**, v. 30, p.239–247, 1940.

HAUSHOFER J.; BAKER C.I.; LIVINGSTONE, M.S; KANWISHER, N. Privileged Coding of Convex Shapes in Human Object-Selective Cortex. **J. Neurophysiol.** v. 100, n. 2, p. 753–762, 2008.

HE, B. B. Two-dimensional X-Ray Diffraction. New Jersey: John Wiley & Sons, 2011, p.426. ISBN: 978-0-470-22722-0

HE, N.H.C.; YUNG, X.C. Curvature scale space corner detector with adaptive threshold and dynamic region of support. In: IEEE PROCEEDINGS OF THE 17TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, ICPR 04, Cambridge, 2004, v. 2, p. 791–794. 2004.

HEAT EQUATION. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2014. Disponível em:< http://en.wikipedia.org/wiki/Heat_equation>. Acesso em: 16 jun. 2015.

HUBEL, D.H.; WIESEL,T,N. Receptive fields of optic nerve fibres in the spider monkey. **J.Phy8iol.**, v.154, p.572-580, 1960.

HUBEL, D.H. **Eye, brain and vision**. W. H. Freeman; 2nd edition (May 15, 1995). Scientific American Library Series (Book 22). P.256. ISBN 978-0716760092.

HUSSEIN, E.M.A.**Computed radiation imaging: physics and mathematics of forward and inverse problems**. Waltham: Elsevier inc, 2011. ISBN 9780123877772.

ITO, M.; GODA, N. Mechanisms underlying the representation of angles embedded within contour stimuli in area V2 of macaque monkeys. **The European journal of neuroscience**, v. 33, n. 1, p. 130–42, 2011.

ITO, M.; KOMATSU, H. Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. **The Journal of neuroscience : the official journal of the Society for Neuroscience**, v. 24, n. 13, p. 3313–24, 2004.

JALBA, C., WILKINSON, M. H. F.; ROERDINK J. B. T. M. Shape representation and recognition through morphological curvature scale-spaces. **IEEE Transactions on Image Processing**, v. 15, n. 2, 2006.

JAYNES, E.T. Gibbs vs Boltzmann Entropies. **American Journal of Physics**, v. 33, n. 5, p. 391-398, 1965.

KADIR, T.; BRADY, M. Saliency, Scale and Image Description. **International Journal of Computer Vision**, v. 45, n. 2, p. 83–105, 2001.

KAHAKI, S. M. M.; NORDIN, M. J.; ASHTARI, A. H. Contour-based corner detection and classification by using mean projection transform. **Sensors (Basel, Switzerland)**, v. 14, n. 3, p. 4126–43, 2014.

KAPLAN, E.; LEE, B. B.; SHAPLEY, R. M. New views of primate retinal function. **Progress in Retinal Research**, v. 9, p. 273-336, 1990.

KERAUTRET, B.; LACHAUD, J.-O.; NAEGEL, B. CURVATURE BASED CORNER DETECTOR FOR DISCRETE, NOISY AND MULTI-SCALE CONTOURS. **International Journal of Shape Modeling**, v. 14, n. 02, p. 127–145, 2008.

KIMMEL, R.; SOCHEN, N.; WEICKERT, J. (ED). Scale-space and PDE methods in computer vision. In: PROCEEDINGS OF THE 5TH INTERNATIONAL CONFERENCE, SCALE-SPACE 2005. Springer, 2005. ISBN: 978-3-540-25547-5. 634p.

KOENDERINK, J.J. The structure of images. **Biological Cybernetics**, v. 50, p. 363–370, 1984.

KOK, P.; DE LANGE, F. P. Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. **Current biology : CB**, v. 24, n. 13, p. 1531–5, 2014.

KREYSZIG, E. **Advanced engineering mathematics**. USA: John Wiley & Sons, 1993.

KUFFLER, S.W. Discharge patterns and functional organization of the mammalian retina. *Journal of Neurophysiology*, v. 16, p. 37–68, 1953.

LABOURE, M. J.; AZEMA, J.; FOURNEL, T. Detection of dominant points on a digital closed curve. *Acta Stereologica*, 2. dez. 2014

LAMBERT, F.L. Configurational entropy revisited. *Journal of Chemical Education*, v. 84, n. 9, p. 1548-1550, 2007.

LARSON, D.R. The economy of photons, *Nature Methods*, v. 7, p. 357-359, 2010.

LEE, J.S.; SUN, Y.N.; CHEN, C.H. Multiscale corner detection by using wavelet transform. *IEEE Trans. Image Process.*, v. 4, n. 1, p. 100–104, 1995.

LEE, J. A.; LEE, B. H.; XU, G.; et al. Geometric corner extraction in retinal fundus images. **Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference**, v. 2014, p. 158–61, 2014. IEEE.

LETTVIN, J.; MATURANA, H.; MCCULLOCH, W.; PITTS, W. What the Frog's Eye Tells the Frog's Brain. *Proceedings of the IRE*, v. 47, n. 11, p. 1940–1951, 1959.

LEYMARIE, F.; LEVINE, M. D. **Curvature morphology**. Technical Report, TR-CIM-88-26, Center for Intelligent Machines, McGill University, Montreal, Dec. 1988.

LI, H. Feature matching based on corner and edge constraints. *Electronic Imaging & Signal Processing*, SPIE Newsroom. DOI: 10.1117/2.1200705.0541, 2007.

LI, L.; LIU, Y.; GONG, H. Comparison and Evaluation on the Methods of Corner Detectors Used in Sign Language Recognition - ProQuest. *International Journal of Digital Content Technology and its Applications*, v. 7, n. 5, p. 943–951, 2013.

LINFOOT, E. H. Information Theory and Optical Images. *JOSA*, v. 45, n. 10, p. 808-818, 1955.

- LIU, G. S.; LI, B. L. Extraction of Optimal Contour Dominant Points Based on ICT Images in Reverse Engineering. **Applied Mechanics and Materials**, v. 423-426, p. 2570–2575, 2013.
- LIU, H. C.; SRINATH, M. D. Corner detection from chain-code. *Pattern Recognition*, v. 23, n. 1-2, p. 51-68, 1990.
- LONCARIC, S. A survey of shape analysis techniques. **Pattern Recognition**, v. 31, n. 8, p. 983–1001, 1998.
- LOWE, D. G. Distinctive Image Features from Scale-Invariant Keypoints. **International Journal of Computer Vision**, v. 60, n. 2, p. 91–110, 2004.
- Lu, Y.; Jain, R., Reasoning about edges in scale space. **Pattern Analysis and Machine Intelligence, IEEE Transactions on** , v.14, n.4, p.450-468, 1992.
- MANDELBROT, B. How long is the coast of Britain? Statistical self-similarity and fractional dimension. **Science (New York, N.Y.)**, v. 156, n. 3775, p. 636–638, 1967.
- MARCONDES R.; COSTA, L. Piecewise linear segmentation of digital contours in $O(N \cdot \log(N))$ through a technique based on effective digital curvature estimation. *Real-Time Imaging*, v. 1, n. 6, p. 409-417, 1995.
- MARJI, M.; KLETTE, R.; SIY, P. Corner detection and curve partitioning using arc-chord distance. In: INTERNATIONAL CONFERENCE ON COMBINATORIAL IMAGE ANALYSIS 10, 2004, Auckland. v. 3322, p. 512-521, 2004.
- MARR, D.; HILDREDTH, E. Theory of edge detection. Technical report A.I. memo no. 518. MIT, 1979. 64 p.
- MARTÍNEZ, J.M.; KOENEN, R.; PEREIRA, F. Overview of MPEG-7: the generic multimedia content description standard. **IEEE Multimedia**, v. 9, n. 2, pp. 78-87, 2002.
- MARTÍNEZ, J.M. Mpeg-7: Overview of mpeg-7 description tools, part 2. **IEEE multimedia**, v. 9, n. 2, p. 83-93, 2002B.
- MARTINEZ-CONDE, S.; MACKNIK, S. L.; HUBEL, D. H. Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. **Nature Neuroscience**, v. 3, n. 3, pp. 251-258, 2000.

MARTINEZ-CONDE, S.; OTERO-MILLAN, J.; MACKNIK, S. L. The impact of microsaccades on vision: towards a unified theory of saccadic function. **Nature reviews. Neuroscience**, v. 14, n. 2, p. 83–96, 2013.

MAXWELL, J. C. **Theory of heat**. London: Longmans, Green, and CO., 1902.

MCARTHUR, J.A.; MOULDEN, B. A two-dimensional model of brightness perception based on spatial filtering consistent with retinal processing. **Vision Research**, v. 39, p. 1199–1219, 1999.

MCLACHLAN, G. J.; PEEL, D. **Finite mixture models**. New York: Willey and Sons inc., 2000. ISBN 047100626-2.

MEDIONI, G.; YASUMOTO, Y. Corner Detection and Curve Representation Using Cubic B-Splines. **Computer Vision, Graphics and Image Processing**, v. 39, n. 3, p. 267-278, 1987.

MEHRER, H. **Diffusion in solids: fundamentals, methods, materials, diffusion-controlled processes**. Heidelberg: Springer, 2007. Springer Series in solid-state sciences 155. ISBN 978-3-540-71486-6.

MEYBERG, S.; WERKLE-BERGNER, M.; SOMMER, W.; DIMIGEN, O. Microsaccade-related brain potentials signal the focus of visuospatial attention. **NeuroImage**, v. 104, p. 79–88, 2015.

MIAO, L. ; QI, H ; SZU, H. A Maximum Entropy Approach to Unsupervised Mixed-Pixel Decomposition. **Image Processing, IEEE Transactions on**, v. 16, n. 4, p. 1008 – 1021, 2007

MOKHTARIAN, F.; MACKWORTH A. Scale-based description and recognition of planar curves and two-dimensional shapes. **IEEE Trans. Pattern Analysis and Machine Intelligence**, v. 8, n. 1, p. 34-43, 1986

MOKHTARIAN, F.; MACKWORTH, A.K. A theory of multiscale-based shape representation for planar curves. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 14, p. 789–805, 1992.

MOKHTARIAN, F. Silhouette-based isolated object recognition through curvature scale space. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 17, p. 539–544, 1995.

MOKHTARIAN, F.; ABBASI, S.; KITTLER, J. Efficient and robust retrieval by shape content through curvature scale space. In: PROCEEDINGS OF THE INTERNATIONAL WORKSHOP ON IMAGE DATABASES AND MULTIMEDIA SEARCH, 1996, Amsterdam, p. 35–42, 1996.

MOKHTARIAN, F.; ABBASI, S. Affine Curvature Scale Space with Affine Length Parametrisation. **Pattern Anal. Appl.**, v. 4, n. 1. p. 1-8, 2001.

MOKHTARIAN, F.; MOHANNA, F. Performance evaluation of corner detectors using consistency and accuracy measures. **Computer Vision and Image Understanding**, v. 102, n. 1, p. 81–94, 2006.

MPEG-7 Shape Part B database, http://www.imageprocessingplace.com/root_files_V3/image_databases.htm. Accessed on Mar 2013.

NGUYEN, T.P.; DEBLED-RENNESON, I. A discrete geometry approach for dominant point detection. **Pattern Recognition**, v. 44, n. 1, p. 32-44, 2011

NORMAN, J. F.; PHILLIPS, F.; ROSS, H. E. Information concentration along the boundary contours of naturally shaped solid objects. **Perception**, v. 30, n. 11, p. 1285–94, 2001.

NOWICKI, M.; SKRZYPCZYNSKI, P. Performance comparison of point feature detectors and descriptors for visual navigation on Android platform. 2014 International Wireless Communications and Mobile Computing Conference (IWCMC). **Anais...** . p.116–121, 2014. IEEE.

NOVOTNY, L.; HECHT, B. **Principles of nano-optics**. New York: Cambridge University Press, 2006. ISBN 978-0-521-83224-3.

OLIVEIRA, R. F.; COSTA, L. DA F.; ROQUE, A. C. A possible mechanism of curvature coding in early vision. **Neurocomputing**, v. 65-66, p. 117–124, 2005.

OLSHAUSEN, B.A.; FIELD, D.J. Vision and the Coding of Natural Images. **American Scientist**, v. 88, n. 3, p. 238-245, 2000.

OSHER, S.J., SETHIAN, J.A. Fronts propagating with curvature dependent speed: algorithms based on Hamilton–Jacobi formulations. **J. Comput. Phys.**, v. 79, p. 12–49, 1988.

OSTERBERG, G. Topography of the layer of rods and cones in the human retina. **Acta ophthalm. suppl.**, v.6, p. 11-97, 1935.

PARK, H; LEE, J-H. B-spline curve fitting based on adaptive curve refinement using dominant points. **Computer-Aided Design**, v. 39, n. 6, p. 439-451, 2007.

PARVEZ, M.T.; MAHMOUD S.A. Polygonal approximation of digital planar curves through adaptive optimizations. **Pattern Recognition Letters**, v. 31, p. 1997–2005, 2010.

PASUPATHY, A.; CONNOR, C. E. Responses to contour features in macaque area V4. **Journal of neurophysiology**, v. 82, n. 5, p. 2490–502, 1999.

PASUPATHY, A.; CONNOR, C. E. Shape representation in area V4: position-specific tuning for boundary conformation. **Journal of neurophysiology**, v. 86, n. 5, p. 2505–19, 2001.

PATEL, T. P.; PANCHAL, S. R. Corner Detection Techniques: An Introductory Survey. **IJEDR** v. 2, n. 4, p. 3680-3686, 2014.

PAULA, I.; MEDEIROS, F.N.S.; BEZERRA, F.N.; USHIZIMA, D.M. Corner detection within a multiscale framework. In: PROCEEDINGS OF SIBGRAPI 2011 XXIV CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES, Maceió, Brasil, 2011.

PAVLIDIS, T.; HOROWITZ, S. L. Segmentation of plane curves. **IEEE Transactions on Computers**, v. C23, n. 8, p. 860-870, 1974.

PEDROSA, G.V.; BARCELOS, C.A.Z.; BATISTA, M.A. An image retrieval system using shape salience points. IN: PROCEEDINGS OF 2011 IEEE INTERNATIONAL SYMPOSIUM ON CIRCUITS AND SYSTEM (ISCAS 2011), Rio de Janeiro, Brazil, pp. 2797–2800, 2011.

PEI, S.-C.; LIN, C.-N. The detection of dominant points on digital curves by scale-space filtering. **Pattern Recognition**, v. 25, n. 11, p. 1307-1314, 1992.

PRASAD, D. Assessing Error Bound For Dominant Point Detection. **International Journal of Image Processing**, V. 6, n. 5, 2012.

PERONA, P.; MALIK J. Scale-space and edge detection using anisotropic diffusion. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, V. 12 (7), p. 629-639. Jul. 1990.

POTOČNIK, B. Automated landmark points detection by using a mixture of approaches: the vole-teeth case. **Signal, Image and Video Processing**, v. 9, n. 1, p. 93–104, 2012.

QUDDUS, A.; FAHMY, M. Binary text image compression using overlapping rectangular partitioning. **Pattern Recognition Letters**, v. 20, n. 1, p. 81-88, 1999.

QUDDUS, A.; GABBOUJ, M. Wavelet-based corner detection technique using optimal scale. **Pattern Recognition Letters**, v. 23, n. 1-3, p. 215-220, 2002.

RAMER, U. An interactive procedure for the polygonal approximation of plane curves. **Computer Graphics and Image Processing**, v. 1, p. 244-256, 1972.

RATLIFF, F. **Mach Bands: quantitative studies on neural network in the retina**. San Francisco, CA: Holden-Day, 1965

RATTARANGSI, A.; CHIN, R.T. Scale-based detection of corners of planar curves. **IEEE Trans. Pattern Anal. Machine Intelligence**, v. 14, p. 430–449, 1992.

RAY, B. K.; RAY, K. S. A new approach to polygonal approximation. **Pattern Recognition Letters**, v. 12, n. 4, p. 229-234, 1991.

ROCKETT, P. I. Performance assessment of feature detection algorithms: a methodology and case study on corner detectors. **IEEE transactions on image processing : a publication of the IEEE Signal Processing Society**, v. 12, n. 12, p. 1668–76, 2003. IEEE.

RODIECK, R.W. Quantitative analysis of cat retinal ganglion cell response to visual stimuli. **Vision Research**, v. 5, n. 12, p. 583–601, 1965.

RODRÍGUEZ-SÁNCHEZ, A. J.; TSOTSOS, J. K. The roles of endstopped and curvature tuned computations in a hierarchical representation of 2D shape. **PloS one**, v. 7, n. 8, p. e42058, 2012.

RODRIGUEZ-SANCHEZ, A. J.; TSOTSOS, J. K. The importance of intermediate representations for the modeling of 2D shape detection: Endstopping and curvature tuned computations. CVPR 2011. **Anais...** p.4321–4326, 2011. IEEE.

RODRÍGUEZ-SÁNCHEZ, A.; NEUMANN, H.; PIATER, J. Beyond Simple and Complex Neurons: Towards Intermediate-level Representations of Shapes and Objects. **KI - Künstliche Intelligenz**, v. 29, n. 1, p. 19–29, 2014.

ROMENY, B. et al (Ed). Scale-space theory in computer vision: Proceedings of the first international conference, Scale-Space'97. Springer Berlin Heidelberg, 1997. ISBN 978-3-540-63167-5.

ROMENY, B. M. H. (Ed). **Geometry-driven diffusion in computer vision**. Dordrecht: Kluwer Academic Publishers, 1994. (Computational Image and Vision, 1). ISBN 0-7923-3087-0.

ROMENY, B. M. H. **Front-end vision and multi-scale image analysis: Multi-scale computer vision theory and applications, written in mathematica**. Netherlands: Springer Science + Business Media B.V., 2003. ISBN 978-1-4020-1507-6.

ROMENY, B. M. H.; FLORAK, L. M. J. Front-end vision, a multiscale geometry engine (lecture notes in computer science). In First IEEE International Workshop on Biologically Motivated Computer Vision, p. 1-35, 2000.

ROSENFELD, A.; JOHNSTON, E. Angle detection on digital curves. **IEEE Transactions on Computers**, v. C-22, p. 875–878, 1973.

ROSENFELD, A.; WESKA, J.S. An improved method of angle detection on digital curves. **IEEE Transactions on Computers**, v. C-24, p. 940–941, 1975.

ROSIN, P.L. Techniques for assessing polygonal approximations of curves. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, v. 19, n. 6, p. 659-666, 1997.

ROSIN, P. L. Determining local natural scales of curves. **Pattern Recognition Letters**, v. 19, n. 1, p. 63–75, 1998.

ROSIN, P. L.; West, G. A. W. Segmentation of edges into lines and arcs. **Image Vision Computing**, v. 7, n. 2, p. 109-114, 1989.

ROTH, M. PSF-fitting techniques for crowded field 3D spectroscopy. **New Astronomy Reviews**, v. 49, n. 10-12, p. 573–581, 2006.

RUDERMAN, D.L.; BIALEK, W. Statistics of natural images: scaling in the woods. **Physical Review Letters**, v. 73, n. 6, p. 814–817, 1994.

SALAPATEK,CP.; KESSEN, W. Visual scanning of triangles by the human newborn. **Journal of Experimental Child Psychology**, v. 3, p. 113-122, 1966.

SARFRAZ, M. **Interactive curve modeling and image processing with applications to computer graphics, vision and image processing**. London: Springer-Verlag, 2008. ISBN 978-1-84628-870-8.

SARFRAZ, M.; ASIM, M.R.; MASOOD, A. Piecewise polygonal approximation of digital curves. In: PROCEEDINGS OF THE 8TH INTERNATIONAL CONFERENCE ON INFORMATION VISUALIZATION, IV 04, 2004, p. 991-996, 2004.

SARFRAZ, M.; RASHEED, A.; MUZAFFAR, Z. A Novel Linear Time Corner Detection Algorithm. Computer Graphics, Imaging and Visualization – New Trends, Sarfraz, M., Wang, Y., and Banissi, E. (Eds.), ISBN: 3-7695-2392-7, IEEE Computer Society, USA, 2005, pp. 191-196.

SARFRAZ, M.; SWATI, Z. N. K. Mining Corner Points on the Generic Shapes. **Open Journal of Applied Sciences**, v. 03, n. 01, p. 10–15, 2013.

SCHMID, C.; MOHR, R.; BAUCKHAGE, C. Comparing and evaluating interest points. Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271). **Anais...** . p.230–235. Narosa Publishing House. 1998.

SCHMID, C.; MOHR, R.; BAUCKHAGE, C. Evaluation of Interest Point Detectors. **International Journal of Computer Vision**, v. 37, n. 2, p. 151–172. 2000.

SCHROEDER, D.V. **Introduction to thermal physics**. USA: Addison Wesley Longman, 1999. 435 p. ISBN: 0-201-38027-7

SHANNON, C. E. A Mathematical Theory of Communication. **Bell System Technical Journal**, v. 27, n. 3, p. 379–423, 1948.

SHEN, J; CASTAN, S. An optimal linear operator for step edge detection. **CVGIP: Graphical Models and Image Processing**, v. 54, n. 2, p. 112-133, 1992.

SKLANSKY, J.; GONZALEZ, V. Fast polygonal approximation of digitized curves. Pattern Recognition, v. 12, p. 327-331, 1980.

SIMONCELLI, E.P.; OLSHAUSEN, B.A. Natural image statistics and neural representation. **Annual Review of Neuroscience**, v. 24, p. 1193-1216, 2001.

SMITH, P. et al. Effective corner matching. In: PROCEEDINGS OF 9TH BRITISH MACHINE VISION CONFERENCE, 1998, v. 2, p. 545-556. Edited by Lewis, P.H. and Nixon, M.S.

SODERQUIST, D.R. **Sensory processes**. California: SAGE Publications Inc, 2002. 400 p. ISBN 0-7619-2333-0.

SRIKANTH, M. Image Indexing and Retrieval using the Cross-Entropy Measures. In: PROCEEDINGS OF THE HKK CONFERENCE, Waterloo, Ontario, Canada, 1999.

SRINIVASAN, M.V.; LAUGHLIN, S.B.; DUBS, A. Predictive Coding: A Fresh View of Inhibition in the retina. **Proc. R. Soc. Lond. B**, v. 216, p. 427-459, 1982.

SUPPES, P.; DE BARROS, J. A. Diffraction with well-defined photon trajectories: a foundational analysis. **Foundations of Physics Letters**, v. 7, n. 6, 1994.

TAHAEI, M. S.; HASHEMI, S. N.; MOHADES, A.; GHEIBI, A. Geometric algorithm for dominant point extraction from shape contour. **Pattern Analysis and Applications**, v. 17, n. 3, p. 481–496, 2012.

TANG, K.T. **Mathematical methods for engineers and scientists 3: fourier analysis, partial differential equations and variational methods**. New York: Springer-Verlag Berlin Heidelberg, 2007.

TENG, S. W.; NAJMUS SADAT, R. M.; LU, G. Effective and efficient contour-based corner detectors. **Pattern Recognition**, v. 48, n. 7, p. 2185–2197, 2015.

THE, C.H.; CHIN, R.T. On the detection of dominant points on digital curves. **IEEE Trans. Pattern Anal. Mach. Intell.**, v. 11, p. 859–872, 1989.

THIBOS, L. N. Image Processing by the Human Eye. Proc. SPIE 1199, Visual Communications and Image Processing IV, 1148 (November 1, 1989); doi:10.1117/12.970124.

TIAN, X.; CHEN, C.-Y. Probing perceptual performance after microsaccades. **The Journal of neuroscience : the official journal of the Society for Neuroscience**, v. 35, n. 7, p. 2842–4, 2015.

TISSAINAYAGAM, P.; SUTER, D. Assessing the performance of corner detectors for point feature tracking applications. **Image and Vision Computing**, v. 22, n. 8, p. 663–679, 2004.

TORROBA, R.; RABAL, H; RUIZ, B. An Entropy Approach to Light Propagation. **Journal of Modern Optics**, v. 39, n. 9, p. 1939-1946, 1992.

TORRES, R.S.; FALCAO, A.X. Contour salience descriptors for effective image retrieval and analysis. **Image and Vision Computing**, v. 25, n. 1, p. 3-13, 2007.

TRONCOSO, X. G.; MACKNIK, S. L.; MARTINEZ-CONDE, S. Novel visual illusions related to Vasarely's "nested squares" show that corner salience varies with corner angle. **Perception**, v. 34, n. 4, p. 409–20, 2005.

TRONCOSO, X. G.; TSE, P. U.; MACKNIK, S. L.; et al. BOLD activation varies parametrically with corner angle throughout human retinotopic cortex. **Perception**, v. 36, n. 6, p. 808–20, 2007.

TRONCOSO, X.; MACKNIK, S.; MARTINEZ-CONDE, S. Corner salience varies linearly with corner angle during flicker-augmented contrast: a general principle of corner perception based on Vasarely's artworks. **Spatial Vision**, v. 22, n. 3, p. 211–224, 2009.

TSAI, D-M.; HOU, H-T.; SU, H-J. Boundary-based corner detection using eigenvalues of covariance matrices. **Pattern Recognition Letters**, v. 20, p. 31-40 1999.

TUYTELAARS, T.; MIKOLAJCZYK, K. Local Invariant Feature Detectors: A Survey. **Foundations and Trends® in Computer Graphics and Vision**, v. 3, n. 3, p. 177–280, 2007. Now Publishers Inc.

URSINO, M.; LA CARA, G. E. A model of contextual interactions and contour detection in primary visual cortex. **Neural networks : the official journal of the International Neural Network Society**, v. 17, n. 5-6, p. 719–35, 2004.

UTCKE, S. Error-Bounds on Curvature Estimation. In: L. D. Griffin; M. Lillholm (Orgs.); **Scale Space Methods in Computer Vision**, Lecture Notes in Computer Science. v. 2695, p.657–666, 2003. Berlin, Heidelberg: Springer Berlin Heidelberg.

VASES and FACES. Disponível em: <<https://agilewarrior.wordpress.com/2014/03/16/vases-and-faces-an-exercise-for-the-double-brain/>>. Acesso em: 19/5/2016.

VINCENT, E.; LAGANIRE R. Matching featuring points in stereo pairs: a comparative study of some matching strategies. **Machine Graphics and Vision Journal**, v. 10, p. 237-259, 2001.

VINCENT, E.; LAGANIRE, R. Matching featuring points. *Journ. Visual Commun. Image Representation*, v. 16, n. 1, p. 38-54, 2005.

XU ET AL. A biologically motivated corner detection method based on the oriented receptive fields of simple cortical cells. In: PROCEEDINGS OF IEEE INT. CONF. ON BIOMEDICAL ENGINEERING AND COMPUTER SCIENCE (ICBECS 2010), 2010.

XUAN,Y.; WANHAI,Y. Adaptive multiscale edge detection using neighborhood entropy. in: SIGNAL PROCESSING PROCEEDINGS, 2000. WCCC-ICSP 2000. 5TH INTERNATIONAL CONFERENCE ON, v. 3, p. 1440 – 1443.

YANG, J.; YUE, Z.; WU, X. Independence of the completion effect from the noncompletion effect in illusory contour perception. **Journal of vision**, v. 15, n. 14, p. 6, 2015. The Association for Research in Vision and Ophthalmology.

YU, F.T.S. **Entropy and Information Optics**. Marcel-Dekker, New-York, 2000.

YUE, X.; POURLADIAN, I. S.; TOOTELL, R. B. H.; UNGERLEIDER, L. G. Curvature-processing network in macaque visual cortex. **Proceedings of the National Academy of Sciences of the United States of America**, v. 111, n. 33, p. E3467–75, 2014.

WAGEMANS, J.; ELDER, J. H.; KUBOVY, M.; et al. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. **Psychological bulletin**, v. 138, n. 6, p. 1172–217, 2012.

WAGEMANS, J. (ORG.). **The Oxford Handbook of Perceptual Organization**. Oxford University Press, 2015. ISBN 9780199686858.

WEISSTEIN, E. W. Normal Sum Distribution. . Wolfram Research, Inc. Disponível em: <<http://mathworld.wolfram.com/NormalSumDistribution.html>>. Acesso em: 23/5/2016.

WALL, K.; DANIELSSON, P. E. A fast sequential method for polygon approximation of digitized curves. **Computer Vision, Graphics and Image Processing**, v. 28, p. 220-227, 1984.

WANG, F. ET AL. Cumulative Residual Entropy, A New Measure of Information & its Application to Image Alignment. In: NINTH INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV2003), Nice, France, p. 548-553, 2003.

WANG, H.; BRADY, M. Real-time corner detection algorithm for motion estimation, **Image Vision Computing**, v. 13, n. 9, p. 695-703, 1995.

WEICKERT, J. A review of nonlinear diffusion filtering. Berlin: Springer, 1997, p. 3-28, 1997. (LNCS 1252-Scale-Space Theory in Computer Vision). Edited by B. ter Haar Romeny, L. Florack, Koenderink, J. and Viergever, M.

WEICKERT, J. Applications of nonlinear diffusion in image processing and computer vision. In: PROCEEDINGS OF ALGORITMY 2000, Acta Math. Univ. Comenianae, v. 70, n. 1, p. 33-50, 2000.

WEICKERT, J.; ISHIKAWA, S.; IMIYA, A. Linear scale-space has first been proposed in Japan. **J. Math. Imag. Vision**, v. 10, p. 237-252, 1999.

WEICKERT, J.; ISHIKAWA, S.; IMIYA, A. On the history of Gaussian scale-space axiomatics. Dordrecht: Kluwer, p. 45-59, 1997. (Gaussian scale-space theory).

WEST, G. A. W.; ROSIN, P. L.. Techniques for segmenting image curves into meaningful descriptions. **Pattern Recognition**, 24(7): 643-652, 1991.

WESTHEIMER, G. Visual acuity: Information theory, retinal image structure and resolution thresholds. **Progress in Retinal and Eye Research**, v. 28, n. 3, p. 178-186, 2009.

WIESEL, T. N. Receptive fields of ganglion cells in the cat's retina. **J Physiol.**, v. 153, n. 3, p. 583-594, 1960.

WITKIN, A.P. Scale-space filtering. In: PROCEEDINGS OF THE 8TH INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 1983, Karlsruhe, p. 1019-1021.

WOLFE, J. M.; YEE, A.; FRIEDMAN-HILL, S. R. Curvature is a basic feature for visual search tasks. **Perception**, v. 21, n. 4, p. 465-480, 1992.

WORRING, M.; SMEULDERS, A. W. M. Digital curvature estimation. **CVGIP: Image Understanding**, v. 58, p. 366-382, 1993.

WU, W-Y.; WANG, M-J. Detecting the dominant points by curvature-based polygonal approximation. **Graphical Models and Image Processing**, v. 55, n. 2, p. 79-88, 1993.

WU, W.Y.. Dominant point detection using adaptive bending value. **Image and Vision Computing**, v. 21, n. 6, p. 517-525, 2003.

ZAPPE, H. **Fundamentals of micro-optics**. New York: Cambridge University Press, 2010. ISBN 978-0-521-89542-2.

ZARSLAN, E.; VEMURI, B.C.; MARECI, T.H. Generalized scalar measures for diffusion MRI using trace, variance, and entropy. **Magnetic Resonance in Medicine**, v. 53, p. 866–876, 2005.

ZHANG, X.; ZHAO, D. A parallel algorithm for detecting dominant points on multiple digital curves. **Pattern Recognition**, v. 30, n.2, p. 239-244, 1997.

ZHANG, D.; LU, G. Evaluation of MPEG-7 shape descriptors against other shape descriptors. **Multimedia Systems**, v. 9, p. 15–30, 2003.

ZHANG, D; LU, G. Review of shape representation and description techniques. **Pattern Recognition** , v. 37, p. 1 – 19, 2004.

ZHANG, X. ET AL. Multi-scale curvature product for robust image corner detection in curvature scale space. **Pattern Recognition Letters**, v. 28, n. 5, p. 545-554, 2007.

ZHANG, W. ET AL. Point spread function characteristics analysis of the wavefront coding system. **Optics Express**, v. 15, n. 4, p. 1543-1552, 2007.

ZHANG, B; ZERUBIA, J.; OLIVO-MARIN, J.C. A study of Gaussian approximations of fluorescence microscopy PSF models, Three-Dimensional and Multidimensional Microscopy. In: IMAGE ACQUISITION AND PROCESSING XIII, PROCEEDINGS OF THE SPIE, v. 6090, 2006, p. 104-114, 2006.

ZHENG, Z.; WANG, H.; KHWANG TEOH, E. Analysis of gray level corner detection. **Pattern Recognition Letters**, v. 20, n. 2, p. 149–162, 1999.

ZHONG, B.; LIAO, W. Direct curvature scale space: theory and corner detection. **IEEE Trans. Pattern Anal. Mach. Intell**, v. 29, n. 3, p. 508–512, 2007.

ZHU, P.; CHIRLIAN, P.M. On critical point detection of digital shapes. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 17, n. 8, p. 737-748, 1995.

ZHU, Q; WANG, Y; LIU, H. Auto-Corner Detection Based on the Eigenvalues Product of Covariance Matrices over Multi-Regions of Support. **Journal of Software**, v. 5, n. 8, p. 907-914, 2010.

ZHU, S.C.; WU, Y.N.; MUMFORD, D.B. Minimax Entropy Principle and Its Applications to Texture Modeling. **Neural Computation**, v. 9, n. 8, p. 1627-1660, 1997.

ZUKAL, M.; CIKA, P. Corner detectors: Evaluation of information content. 2012 35th International Conference on Telecommunications and Signal Processing (TSP). **Anais...** . p.763–767, 2012. IEEE.

ZUKAL, M.; BENES, R.; CIKA, P.; QIU, X. Robustness evaluation of corner detectors for use in ultrasound image processing. 2013 36th International Conference on Telecommunications and Signal Processing (TSP). **Anais...** . p.763–767, 2013. IEEE.

ANEXO A – Imagens utilizadas

002.BMP

003.BMP

004.BMP

007.BMP

008.BMP

012.BMP

013.BMP

024.BMP

028.BMP

040.BMP

043.BMP

048.BMP

049.BMP

088.BMP

089.BMP

105.BMP

106.BMP

112.BMP

114.BMP

118.BMP

120.BMP

126.BMP

134.BMP

169.BMP

174.BMP

176.BMP

177.BMP

245.BMP

256.BMP

258.BMP

Bat-4.gif

Bone-8.gif

device0-7.gif

device0-8.gif

device0-9.gif

device0-10.gif

device0-11.gif

device0-12.gif

device0-13.gif

device0-16.gif

device0-18.gif

device1-4.gif

device4-10.gif

device5-3.gif

elephant-2.gif

Fish-a026.gif

Fork-19.gif

hammer-11.gif

jar-4.gif

key-15.gif

Misk-11.gif