

**UMA PROPOSTA DE ESTRUTURAÇÃO E INTEGRAÇÃO DE  
PROCESSAMENTO DE CORES EM SISTEMAS ARTIFICIAIS DE VISÃO**

*Jander Moreira*

Tese de Doutorado apresentada ao Instituto de Física de São Carlos, da Universidade de São Paulo, para a obtenção do título de Doutor em Ciências: Física Aplicada (Física Computacional).

Orientador: *Prof. Dr. Luciano da Fontoura Costa*

*Dez Ok*

USP/IFSC/SBI



8-2-001276

**São Carlos  
1999**

Moreira, J.

Uma proposta de estruturação e integração de processamento de cores em sistemas artificiais de visão. / Jander Moreira — São Carlos: IFSC-USP, 1999.

285p.

Tese (Doutorado) — Instituto de Física de São Carlos, 1999.

Orientador: Prof. Dr. Luciano da Fontoura Costa.

1. Processamento de Imagens – Técnicas digitais.
2. Segmentação. 3. Imagens coloridas. I. Autor. II. Título.



**MEMBROS DA COMISSÃO JULGADORA DA TESE DE DOUTORADO DE JANDER MOREIRA APRESENTADA AO INSTITUTO DE FÍSICA DE SÃO CARLOS, DA UNIVERSIDADE DE SÃO PAULO, EM 05 DE JULHO DE 1999.**

COMISSÃO JULGADORA:

Prof. Dr. Luciano da Fontoura Costa/IFSC-USP

Prof. Dr. Roland Koberle/IFSC-USP

Prof. Dr. Arnaldo de Albuquerque Araújo/UFGM

Prof. Dr. Adilson Gonzaga/EEESC-SEL-USP

Prof. Dr. Roberto de Alencar Lotufo/UNICAMP

*A Sandra*  
*Aos meus pais, Jurandir e Zilda*

## AGRADECIMENTOS

*A* DEUS, pela iluminação e onipresença.

*A* Sandra, que sempre me incentivou, consolou e, principalmente, acreditou em mim.

*A* meus pais, pelo apoio, solidariedade e carinho.

*A*o Prof. Dr. Luciano da Fontoura Costa, não somente pela dedicação e orientação, mas também pela paciência e grande consideração.

*A* Alan Salvany Felinto, pela cooperação no trabalho de integração sensorial, essencial a esta tese, pelo conhecimento compartilhado e pela convivência.

*A* Odemir Martinez Bruno, pelo fornecimento do componente que implementa a comunicação entre programas, sem o qual a integração sensorial teria sido muito dificultada, e pelas discussões sobre os aspectos da comunicação, além da convivência pessoal.

*A* Luís Augusto Consularo, pelos inestimáveis auxílios em programação e pela amizade.

*A*o pessoal do laboratório, pelos anos de convívio amistoso e alegre e pelas discussões tão produtivas.

*A* Marcos Roberto Gonçalves, pela ajuda em muitas e muitas situações.

*A* Mara e todo o pessoal da Biblioteca do Instituto de Física de São Carlos, pela competência e presteza na aquisição de material bibliográfico.

*A* Wladerez, pelo apoio nas questões burocráticas.

*A* todos os amigos, pelo incentivo.

*A*o Departamento de Computação da Universidade Federal de São Carlos, pela oportunidade deste trabalho.

*A*o Instituto de Física de São Carlos, pelas estruturas e instalações.

## SUMÁRIO

LISTA DE FIGURAS .....	iv
LISTA DE TABELAS .....	xiii
LISTA DE ALGORITMOS .....	xiv
RESUMO .....	xv
ABSTRACT .....	xvii
<b>1 INTRODUÇÃO E OBJETIVOS .....</b>	<b>1</b>
1.1 O SISTEMA DE VISÃO ARTIFICIAL <i>CYVIS-1</i> .....	2
1.2 OBJETIVOS .....	5
1.2.1 <i>Segmentação de imagens coloridas</i> .....	5
1.2.2 <i>Comparação de mapas de bordas</i> .....	5
1.2.3 <i>Integração ao Cyvis-1</i> .....	5
1.3 ORGANIZAÇÃO DO TEXTO .....	5
<b>CONCEITOS BÁSICOS SOBRE CORES .....</b>	<b>7</b>
2.1 ONDAS ELETROMAGNÉTICAS .....	8
2.2 VISÃO COLORIDA .....	9
2.2.1 <i>Cor através dos três tipos de cones</i> .....	11
2.2.2 <i>Oposição cromática</i> .....	14
2.2.3 <i>Comentários sobre as abordagens tricromática e de oposição cromática</i> .....	16
2.3 REPRESENTAÇÃO DAS CORES .....	17
2.3.1 <i>Nomenclatura básica</i> .....	17
2.3.2 <i>Colorimetria e emparelhamento de estímulos coloridos</i> .....	17
2.3.3 <i>Espaços de cores</i> .....	21
2.4 COMENTÁRIOS FINAIS .....	26
<b>3 SEGMENTAÇÃO DE IMAGENS COLORIDAS: UMA REVISÃO .....</b>	<b>27</b>
3.1 TÉCNICAS BASEADAS EM BORDAS .....	28
3.1.1 <i>Detecção de bordas por métrica ou redução no espaço de cor</i> .....	28
3.1.2 <i>Detecção de bordas através do tratamento isolado das componentes cromáticas</i> .....	30
3.1.3 <i>Detecção de bordas no espaço vetorial das cores</i> .....	33
3.2 TÉCNICAS BASEADAS EM REGIÕES .....	40
3.2.1 <i>Segmentação por classificação e agrupamento</i> .....	40
3.2.2 <i>Segmentação por técnicas de divisão/fusão e crescimento de regiões</i> .....	52
3.3 SEGMENTAÇÃO POR APRENDIZADO COMPETITIVO E REDES NEURAIS .....	54
3.4 ABORDAGENS BASEADAS EM PROPRIEDADES FÍSICAS DA CENA .....	57
3.5 OS PINGÜINS E OS URSOS POLARES .....	59
<b>4 SEGMENTAÇÃO DE IMAGENS POR MAPAS AUTO-ORGANIZÁVEIS .....</b>	<b>63</b>
4.1 CARACTERIZAÇÃO DA METODOLOGIA .....	64

4.1.1	<i>Mapas auto-organizáveis</i> .....	64
4.1.2	<i>Técnica de agrupamento</i> .....	74
4.1.3	<i>Segmentação em regiões e extração de bordas</i> .....	78
4.2	RESULTADOS DE DESEMPENHO E DISCUSSÃO DAS CARACTERÍSTICAS DO MÉTODO PROPOSTO.....	80
4.2.1	<i>Avaliação da classificação</i> .....	80
4.2.2	<i>Tamanho e dimensão do mapa auto-organizável</i> .....	83
4.2.3	<i>Taxas de decaimento</i> .....	89
4.2.4	<i>Uso da cromaticidade</i> .....	93
4.3	APLICAÇÃO.....	97
4.4	COMPARAÇÃO DA CLASSIFICAÇÃO COM MAPA AUTO-ORGANIZÁVEL COM AS CLASSIFICAÇÕES BAYESIANA E POR K VIZINHOS MAIS PRÓXIMOS.....	98
4.5	COMENTÁRIOS FINAIS.....	104
<b>5</b>	<b>ESTUDO COMPARATIVO DE TÉCNICAS DE SEGMENTAÇÃO.....</b>	<b>107</b>
5.1	MODELO COMPARATIVO.....	108
5.1.1	<i>Atributos locais</i> .....	109
5.1.2	<i>Obtenção das bordas a partir de regiões</i> .....	112
5.1.3	<i>Outras técnicas comparadas</i> .....	113
5.1.4	<i>Descrição do modelo de comparação</i> .....	117
5.1.5	<i>Classes de imagens</i> .....	120
5.2	RESULTADOS DA COMPARAÇÃO E DISCUSSÃO.....	122
5.2.1	<i>Taxas de coincidência e excesso</i> .....	124
5.2.2	<i>Coefficiente acumulativo</i> .....	143
5.2.3	<i>Coefficiente proporcional</i> .....	149
5.3	CONSIDERAÇÕES GERAIS SOBRE OS COEFICIENTES ACUMULATIVO E PROPORCIONAL.....	153
5.4	COMBINAÇÃO DE BORDAS BASEADA NO CONTEXTO LOCAL.....	156
5.5	RESULTADOS DA COMBINAÇÃO DE BORDAS E DISCUSSÃO.....	157
5.6	COMENTÁRIOS FINAIS.....	163
5.6.1	<i>Comentários sobre a comparação das técnicas de segmentação</i> .....	164
5.6.2	<i>Comentários sobre a composição seletiva de bordas</i> .....	165
<b>6</b>	<b>INTEGRAÇÃO ENTRE MÓDULOS: COR E ESTÉREO.....</b>	<b>167</b>
6.1	INTEGRAÇÃO DE MODALIDADES.....	168
6.1.1	<i>Caracterização e estruturação básica da comunicação</i> .....	168
6.1.2	<i>Estruturação geral do sistema</i> .....	169
6.2	DESCRIÇÃO DO ERRO DO EMPARELHAMENTO.....	174
6.3	RESULTADOS E DISCUSSÃO.....	177
6.4	COMENTÁRIOS FINAIS.....	183
6.4.1	<i>Detecção de segmentos de retas</i> .....	183
6.4.2	<i>Mapas auto-organizáveis</i> .....	184
6.4.3	<i>Sobel</i> .....	184
6.4.4	<i>Considerações gerais sobre os resultados</i> .....	185
6.4.5	<i>Comentários gerais</i> .....	186
<b>7</b>	<b>CONCLUSÕES.....</b>	<b>187</b>
7.1	CONCLUSÕES DO TRABALHO.....	187
7.1.1	<i>Método de classificação com mapa auto-organizável</i> .....	188

7.1.2	<i>Comparação de desempenho de técnicas de detecção de bordas e combinação baseada em contexto local</i> .....	190
7.1.3	<i>Integração sensorial entre cor e estéreo</i> .....	191
7.2	CONCLUSÕES GERAIS.....	192
7.3	CONTINUIDADE.....	194
<b>8</b>	<b>BIBLIOGRAFIA</b> .....	<b>197</b>
<b>A.</b>	<b>GLOSSÁRIO</b> .....	<b>203</b>
<b>B.</b>	<b>CONVENÇÃO DE NOTAÇÃO</b> .....	<b>207</b>
<b>C.</b>	<b>A TRANSFORMADA DE HOUGH</b> .....	<b>209</b>
I.	RASTREAMENTO ( <i>BACKMAPPING</i> ).....	211
II.	GENERALIZAÇÕES DA TRANSFORMADA DE HOUGH .....	211
<b>D.</b>	<b>RESULTADOS PARA AS DIFERENTES TÉCNICAS DE SEGMENTAÇÃO</b> .....	<b>213</b>
<b>E.</b>	<b>RESULTADOS COMPARATIVOS DO COEFICIENTE ACUMULATIVO PARA AS CLASSES DE IMAGENS</b> .....	<b>251</b>
<b>F.</b>	<b>RESULTADOS COMPARATIVOS DO COEFICIENTE PROPORCIONAL PARA AS CLASSES DE IMAGENS</b> .....	<b>269</b>

## LISTA DE FIGURAS

Figura 1.1. Esquema da estrutura geral do sistema de visão <i>Cyvis-1</i> . O símbolo  indica a existência de uma base de dados associada. (Adaptado de Costa <i>et al.</i> [Costa <i>et al.</i> (1994)].).....	4
Figura 2.1 Distribuição da quantidade de luz relativa para uma lâmpada incandescente e a luz branca (solar) [Goldstein (1989)]. .....	9
Figura 2.2. Experimento de emparelhamento de cores, no qual a soma dos comprimentos de onda $\lambda_1$ , $\lambda_2$ , e $\lambda_3$ à esquerda fornecem a mesma cor que o comprimento de onda monocromático $\lambda$ projetado na direita. Por exemplo, a luz com comprimento de onda 500 nm pode ter sua equivalente perceptiva obtida a partir da mistura aditiva dos comprimentos 420, 560 e 640 nm [Goldstein (1989)]......	10
Figura 2.3. Imagens para (a) pós-imagem e (b) contraste simultâneo. (Adaptado de Goldstein [Goldstein (1989)].).....	11
Figura 2.4. Representação das células receptoras de energia luminosa da retina: cone (acima) e bastonete.....	12
Figura 2.5. Curvas de refletância para alguns vegetais. [Goldstein (1989)].....	12
Figura 2.6. Espectro de absorção (normalizado) dos quatro pigmentos dos cones e bastonetes. ....	13
Figura 2.7. Dois estímulos distintos em termos de composição espectral podem provocar uma mesma resposta nos diferentes tipos de cones, não permitindo a distinção entre os estímulos.....	13
Figura 2.8. Esquema de resposta de neurônio com oposição cromática. ....	15
Figura 2.9. Campo receptivo com oposição cromática dupla. O centro responde positivamente à luz vermelha enquanto é inibido pela luz verde. A circunvizinhança, por sua vez, é excitada por verde e inibida por vermelho. ....	16
Figura 2.10. Esquema de ligação do mecanismo tricromático para estabelecer um dos tipos de oposição cromática. Símbolos “+” indicam conexão sináptica excitatória, enquanto “-” representam conexões inibitórias.....	17
Figura 2.11. Distribuição hipotética das intensidades (potência da radiação) em função do comprimento de onda, para o intervalo visível do espectro [ $\lambda_a$ , $\lambda_b$ ]. .....	19
Figura 2.12. Distribuição de energia constante $E(\lambda)$ , na qual $E(\lambda) = 1$ para cada $\lambda$ monocromático. ....	21
Figura 2.13. Valores dos estímulos tricromáticos para o estímulo monocromático $e_\lambda$ , de intensidade unitária com respeito aos estímulos de cores primárias $r$ , $g$ e $b$ de comprimentos 700, 546,1 e 435,8 nm, respectivamente. [Wyszecki & Stiles (1982)].....	22
Figura 3.1. Resultados da detecção de bordas pelo operador de Prewitt. (a) Imagem original. (b) Aplicação à versão monocromática da imagem de testes; (c) Aplicação ao módulo do vetor (R,G,B) calculado ponto a ponto sobre a versão original. Notam-se diferenças mínimas, sendo os resultados praticamente os mesmos.....	29
Figura 3.2. Borda ideal de Hueckel. (a) Posição e orientação da borda; (b) Perfil da luminância. ....	32
Figura 3.3. Representação das matrizes e submatrizes dos operadores de Scharcanski e Venetsanopoulos [Scharcanski & Venetsanopoulos (1997)]. $H_{\pm} = \eta_{\pm}(i, j)$ . (a) Matriz genérica para a direção horizontal; (b) Exemplo para uma matriz $5 \times 5$ ( $w = 2$ ), direções horizontal e vertical, respectivamente.....	35
Figura 3.4. Ilustração do novo espaço de representação de cores HV $\gamma$ . A área hachurada representa a porção do espaço no qual o grau de cromaticidade é mais baixo.....	39

Figura 3.5. Estabelecimentos dos agrupamentos no espaço $L^*a^*b^*$ . (a) Os modos mais proeminentes ( $w_1$ e $w_2$ ) são detectados em um dos histogramas unidimensionais — por exemplo $L^*$ . Para cada modo é identificado o melhor pico nos outros dois histogramas — componentes $a^*$ e $b^*$ , neste exemplo. (b) Os intervalos definidos pelos três picos nos histogramas delimitam um volume circular cilíndrico no espaço de cor. $V_1$ é definido em relação a $w_1$ e $V_2$ em relação a $w_2$ . .....	45
Figura 3.6. Para uma dada resolução do histograma, são estabelecidas proporções entre a altura do vale e picos no histograma, além de ser considerada uma distância mínima entre estes picos. ....	46
Figura 3.7. Representação da pirâmide de resolução. Cada elemento de imagem de um nível superior da pirâmide compreende a média de quatro elementos do nível imediatamente inferior. Ou seja, a cada nível a imagem tem seu tamanho reduzido pela metade tanto em altura quanto em largura. ....	48
Figura 3.8. As linhas indicam os cruzamentos em zero da derivada do histograma, filtrado com variâncias distintas. O retângulo indica uma das regiões definidas por pares de cruzamentos em zero considerados significativos. ....	49
Figura 3.9. Espaço definido pelas componentes cromáticas $e_r$ e $e_g$ , além da componente $h$ , a qual corresponde ao número de ocorrências das outras duas componentes. ....	51
Figura 3.10. Estrutura de árvore estabelecida segundo vários planos paralelos ao plano $e_r e_g$ . A ilustração mostra uma visão na direção do eixo $e_g$ . ....	51
Figura 3.11. Redução da árvore. (a) Árvore inicial. (b) Após a remoção das folhas conectadas às ramificações; (c) “Redução” do nível dos nós de ramificação; (d) Redução das ramificações (árvore final). ....	52
Figura 3.12. Esquema de rede neural para segmentação de imagens coloridas (adaptado de Zheng [Zheng (1995)]). ....	56
Figura 3.13. Segmentação de imagem colorida. (a) Versão monocromática da imagem colorida. (b) Resultado da segmentação considerando cor e textura. [Zheng (1995)] .....	57
Figura 3.14. As duas componentes de reflexão do modelo de reflexão dicromática: especular e do corpo. ....	58
Figura 3.15. O urso polar e o pingüim. ....	59
Figura 4.1. Esquema de processamento. ....	64
Figura 4.2. Esquema de interconexão de entradas com os neurônios. A interconexão é completa. (a) Unidimensional; (b) bidimensional (com as conexões parcialmente representadas). ....	65
Figura 4.3. Vizinhança em mapas auto-organizáveis, indicada pela região escura envolvendo os nós. O nó hachurado representa o nó vencedor; os nós em cinza claro indicam os nós pertencentes à vizinhança; os nós em branco não são afetados pela ativação do nó vencedor. (a) Estruturação unidimensional do mapa; (b) Estruturação bidimensional com grade regular e vizinhança circular. ....	66
Figura 4.4. Esquemas de influência conforme a distância do nó vencedor. Decaimentos segundo uma função (a) linear e (b) gaussiana. ....	67
Figura 4.5. Vizinhança circular (em branco): (a) influência truncada além das bordas do mapa; (b) estruturada na forma de um hipermapa. ....	68
Figura 4.6. Estruturação bidimensional regular da rede que representa o mapa auto-organizável. ....	70
Figura 4.7. (a) Imagem original. (b) Representação da imagem com cores normalizadas (cromaticidade). ....	70
Figura 4.8. Ilustração do decaimento percentual considerando-se uma taxa de redução de 1%, ao longo de 5000 unidades de tempo. ....	71

Figura 4.9. Evolução da configuração de um mapa auto-organizável de tamanho 10×10. Cada unidade da rede armazena como pesos os valores RGB, representada graficamente pela cor correspondente. (a) Imagem 180×240 pontos. (b)-(g) Evolução dos valores do mapa para tempos 100, 500, 1000, 4000, 5000 e 10000, respectivamente. (h) Mapa após a convergência (24460 passos).....	72
Figura 4.10. Evolução da configuração de um mapa auto-organizável de tamanho 5×5, com representação similar à Figura 4.9. (a) Imagem utilizada com tamanho de 180×240. (b)-(g) Evolução dos valores do mapa para tempos 100, 500, 1000, 4000, 10000 e 28042, respectivamente. (h) Estado final para outra execução do procedimento (29632 passos até a convergência).....	73
Figura 4.11. Resultados finais para mapas lineares, com 16 unidades, considerando as componentes RGB como dados para os pesos internos dos nós. ....	74
Figura 4.12. Ilustração de um comportamento hipotético para parâmetro $\beta$ , segundo a variação do número de agrupamentos em que os dados são divididos. ....	75
Figura 4.13. Representação do eixo principal de uma distribuição de dados (duas dimensões), com indicação da média e primeiros centróides, diametralmente opostos ao longo do eixo e equidistantes em relação à média. Áreas mais escuras representam maior densidade de dados. ....	77
Figura 4.14. Exemplo da classificação através dos centróides detectados pela metodologia proposta. (a) Imagem original; (b) imagem resultante da classificação dos pontos; (c) representação do mapa auto-organizável após a convergência; (d) detalhe da borda entre duas regiões; (e) mesmo detalhe após a aplicação do filtro da mediana; (f) e (g) Detalhes de outra parte da imagem, respectivamente antes e após a aplicação do filtro da mediana. ....	79
Figura 4.15. Máscara usada para determinar pontos de borda. P: Centro da máscara, sobre o ponto sendo analisado. N, O, S e L: vizinhos nas direções cardeais. Áreas hachuradas não são utilizadas. ....	80
Figura 4.16. Esquema da fronteira entre duas regiões, indicadas por tons de cinza diferentes. Contornados pela linha mais espessa se encontram os pontos considerados como bordas entre elas. ....	80
Figura 4.17. Exemplo de rotulação para a imagem House06. Níveis de cinza indicam classes distintas. ....	81
Figura 4.18. Imagens para os exemplos da segmentação para diferentes tamanhos de mapas. (a) Imagem original e (b) imagem de referência para imagem House03. (c) Imagem original e (d) imagem de referência para imagem Cube014. (e) Imagem original e (f) imagem de referência para imagem Pyram01. (g) Imagem original e (h) imagem de referência para imagem Sign04.....	82
Figura 4.19. Taxa de acerto na classificação (em porcentagem) conforme variação do tamanho usado para mapa auto-organizável: (a) linear e (b) quadrado. Imagem House03. ....	84
Figura 4.20. Taxa de acerto na classificação (em porcentagem) conforme variação do tamanho usado para mapa auto-organizável: (a) linear e (b) quadrado. Imagem Cube011. ....	85
Figura 4.21. Taxa de acerto na classificação (em porcentagem) conforme variação do tamanho usado para mapa auto-organizável: (a) linear e (b) quadrado. Imagem Pyram01. ....	86
Figura 4.22. Taxa de acerto na classificação (em porcentagem), conforme variação do tamanho usado para mapa auto-organizável: (a) linear e (b) quadrado. Imagem Sign04.....	87
Figura 4.23. Representação de um mapa auto-organizável 32×32, processado com taxa de decaimento muito alta para o raio da vizinhança. ....	88

Figura 4.24. Figuras usadas nos experimentos dos gráficos das Figuras 4.26 e 4.27. (a) e (b): Imagem Cube104 e sua imagem de regiões (manual), respectivamente. (c) e (d): Idem para a imagem Frogs01.....	89
Figura 4.25. Taxa de acerto na classificação (em porcentagem) conforme variação da taxa de decaimento utilizada (a) para o raio da vizinhança e (b) para a taxa de aprendizado. Decaimento expresso em porcentagem. Imagem House03.....	90
Figura 4.26. Taxa de acerto na classificação (em porcentagem) conforme variação da taxa de decaimento utilizada (a) para o raio da vizinhança e (b) para a taxa de aprendizado. Decaimento expresso em porcentagem. Imagem Cube104.....	91
Figura 4.27. Taxa de acerto na classificação (em porcentagem) conforme variação da taxa de decaimento utilizada (a) para o raio da vizinhança e (b) para a taxa de aprendizado. Decaimento expresso em porcentagem. Imagem Frogs01.....	92
Figura 4.28. Resultados para a segmentação através de mapas auto-organizáveis com incorporação da intensidade luminosa. Imagens originais: (a): Aerial01; (d): Cube009; (g): Cube010; (j): Cube014; (m): Cube102; (p): Cube104; (s): Cube105; (v): Cube203; (y): House02; (ab): House03; (ae): House05; (ah): Pyram01; (ak): Pyram04; (an): Pyram06; (aq): Donna; (at): Flowers. (b), (e), (h), (k), (n), (q), (t), (w), (z), (ac), (af), (ai), (al), (ao), (ar), (au): Imagens segmentadas utilizando intensidade e cromaticidade, relativas às imagens da primeira coluna, respectivamente. (c), (f), (i), (l), (o), (r), (u), (x), (aa), (ad), (ag), (aj), (am), (ap), (as), (av): Imagens segmentadas utilizando intensidade, relativas às imagens da primeira coluna, respectivamente.....	93
Figura 4.29. Imagem House06.....	96
Figura 4.30. Comparação das as taxas de acerto para as classes das figuras do conjunto de testes, apresentadas por imagem. Em branco, a taxa obtida usando a segmentação proposta; em cinza médio, usando a classificação bayesiana; e em cinza escuro as taxas produzidas pela classificação com $k$ vizinhos mais próximos. (a) Aerial01. (b) Aerial02. (c) Balloons. (d) Cube009. (e) Cube010. (f) Cube011. (g) Cube014. (h) Cube015. (i) Cube101. (j) Cube102. (k) Cube104. (l) Cube105. (m) Cube202. (n) Cube203. (o) Cube204. (p) Flowers. (q) Flower02. (r) Frogs02. (s) House01. (t) House02. (u) House03. (v) House05. (w) House06. (x) House07. (y) Pyram01. (z) Pyram03. (aa) Pyram04. (ab) Pyram06. (ac) Sign01. (ad) Sign02. (ae) Sign03. (af) Sign04. (ag) Toy. (ah) Wall01. (ai) Wall02.....	98
Figura 5.1. Círculos discretos de raios 1, 2 e 3.....	110
Figura 5.2. Exemplo das variâncias locais. (a) Imagem original Aerial01. (b) Versão em níveis de cinza. (c) Mapa da variância dos níveis de cinza. (d) Mapa da variância das cores RGB. (e) Mapa de variância da oposição $R$ e $G$ . (f) Mapa de variância da oposição $Y$ e $B$ . Valores mais altos são representados por tons mais escuros, após a aplicação de um logaritmo e a escalação para o intervalo $[0, 255]$ . O raio utilizado para o cálculo foi de 3 pontos.....	112
Figura 5.3. Histogramas dos valores da magnitude do gradiente calculados para as imagens (a) Cube011 e (b) Flower02. Os valores da magnitude foram escalados para o intervalo $[0, 255]$ .....	115
Figura 5.4. (a) Primeira derivada para a curva do histograma da Figura 5.3(a). (b) Segunda derivada da mesma curva. O ponto indicado (18) representa o ponto de inflexão no histograma. Ambas as curvas foram escaladas para o intervalo $[-1, 1]$ .....	116
Figura 5.5. Imagens utilizadas nos exemplos das Figuras 5.6 a 5.8. Imagens originais: (a) Aerial01; (c) Pyram04 e (e) Sign03. (b), (d) e (f): Respetivos mapas de bordas de referência.....	121

Figura 5.6. Histogramas das variâncias nos pontos de borda definidos manualmente para a imagem <i>Aerial01</i> : (a) <i>g</i> ; (b) <i>c</i> ; (c) <i>rg</i> ; e (d) <i>yb</i> .....	122
Figura 5.7 Histogramas das variâncias nos pontos de borda definidos manualmente para a imagem <i>Pyram04</i> : (a) <i>g</i> ; (b) <i>c</i> ; (c) <i>rg</i> ; e (d) <i>yb</i> .....	123
Figura 5.8 Histogramas das variâncias nos pontos de borda definidos manualmente para a imagem <i>Sign03</i> : (a) <i>g</i> ; (b) <i>c</i> ; (c) <i>rg</i> ; e (d) <i>yb</i> .....	124
Figura 5.9. Histogramas das taxas de coincidência (à esquerda) e taxas de excesso (à direita) para os mapas de bordas gerados por várias técnicas. A imagem é <i>Aerial01</i> e o atributo considerado é a variância dos níveis de cinza. (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por <i>k</i> vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global. ....	126
Figura 5.10. Histogramas das taxas de coincidência (à esquerda) e taxas de excesso (à direita) para os mapas de bordas gerados por várias técnicas. A imagem é <i>Aerial01</i> e o atributo considerado é a variância dos vetores de cor (atributo <i>c</i> ). (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por <i>k</i> vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global. ....	129
Figura 5.11. Histogramas das taxas de coincidência (à esquerda) e taxas de excesso (à direita) para os mapas de bordas gerados por várias técnicas. A imagem é <i>Aerial01</i> e o atributo considerado é a variância da oposição cromática <i>RG</i> . (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por <i>k</i> vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global. ....	132
Figura 5.12. Histogramas das taxas de coincidência (à esquerda) e taxas de excesso (à direita) para os mapas de bordas gerados por várias técnicas. A imagem é <i>Aerial01</i> e o atributo considerado é a variância da oposição cromática <i>YB</i> . (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por <i>k</i> vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global. ....	134
Figura 5.13. Desempenho dos métodos de segmentação para a imagem <i>Aerial01</i> , considerando o atributo <i>g</i> . (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por <i>k</i> vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.....	137
Figura 5.14. Desempenho dos métodos de segmentação para a imagem <i>Aerial01</i> , considerando o atributo <i>c</i> . (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por <i>k</i> vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.....	139
Figura 5.15. Desempenho dos métodos de segmentação para a imagem <i>Aerial01</i> , considerando o atributo <i>rg</i> . (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por <i>k</i> vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.....	140

Figura 5.16. Desempenho dos métodos de segmentação para a imagem <i>Aerial01</i> , considerando o atributo <i>yb</i> . (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por <i>k</i> vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.....	142
Figura 5.17. Legenda das cores utilizadas na comparação das técnicas.....	144
Figura 5.18. Porcentagem do número médio de vezes em que uma técnica apresentou o melhor desempenho relativo à taxa de coincidência para as variâncias (a) <i>g</i> , (b) <i>c</i> , (c) <i>rg</i> e (d) <i>yb</i> . ...	145
Figura 5.19. Porcentagem do número médio de vezes em que uma técnica apresentou o melhor desempenho relativo à taxa de excesso (menor taxa) para as variâncias (a) <i>g</i> , (b) <i>c</i> , (c) <i>rg</i> e (d) <i>yb</i> . ....	146
Figura 5.20. Combinação do número relativo de melhores taxas de coincidência e menores taxas de excesso, utilizando o atributo <i>c</i> . (a) $p = 0,2$ , (b) $p = 0,4$ , (c) $p = 0,6$ e (d) $p = 0,8$ . ...	147
Figura 5.21. Combinação do número relativo de melhores taxas de coincidência e menores taxas de excesso, utilizando o atributo <i>yb</i> . (a) $p = 0,2$ , (b) $p = 0,4$ , (c) $p = 0,6$ e (d) $p = 0,8$ ...	148
Figura 5.22. Variação dos coeficientes proporcionais segundo diferentes proporções de combinação. As distribuições referem-se ao atributo <i>g</i> . (a)-(f) $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5, e 1, respectivamente. ....	150
Figura 5.23. Variação dos coeficientes proporcionais segundo diferentes proporções de combinação. As distribuições referem-se ao atributo <i>c</i> . (a)-(f) $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5, e 1, respectivamente. ....	152
Figura 5.24. Variação dos coeficientes proporcionais segundo diferentes proporções de combinação. As distribuições referem-se ao atributo <i>rg</i> . (a)-(f) $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5, e 1, respectivamente. ....	153
Figura 5.25. Variação dos coeficientes proporcionais segundo diferentes proporções de combinação. As distribuições referem-se ao atributo <i>yb</i> . (a)-(f) $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5, e 1, respectivamente. ....	154
Figura 5.26. Imagens utilizadas para os resultados apresentados nas Figuras 5.27 a 5.34. (a) <i>Cube011</i> ; (b) <i>Flower02</i> ; (c) <i>House03</i> ; (d) <i>Pyram04</i> e (e) <i>Sign04</i> .....	156
Figura 5.27. Representação da combinação de bordas para a imagem <i>Sign04</i> , utilizando o coeficiente acumulativo. (a) Atributo <i>g</i> , com $p = 0,2$ ; (b) Atributo <i>g</i> , com $p = 0,5$ ; (a) Atributo <i>rg</i> , com $p = 0,1$ ; e (a) Atributo <i>rg</i> , com $p = 0,5$ . O código de cores é o definido na Figura 5.17.....	157
Figura 5.28. Combinações de bordas para a imagem <i>House03</i> . (a)-(c): atributo <i>g</i> ; (d)-(f): atributo <i>c</i> ; (g)-(i): atributo <i>rg</i> ; e (j)-(l): atributo <i>yb</i> . As proporções de combinação para cada triade de imagens é, respectivamente, 0, 0,2 e 1. ....	158
Figura 5.29. Taxas de coincidência (à direita) e de excesso para as combinações de bordas da Figura 5.28(a) a (c), respectivamente. (a) $p = 0$ ; (b) $p = 0,2$ ; e (c) $p = 1$ .....	159
Figura 5.30. Combinações de bordas para a imagem <i>Pyram04</i> . Proporções: (a) 0; (b) 0,3 e (c) 1.....	160
Figura 5.31. Representação da combinação de bordas para a imagem <i>Sign04</i> , utilizando o coeficiente proporcional calculado sobre o atributo <i>rg</i> : (a) $p = 0$ ; (b) $p = 0,3$ . O código de cores é o definido na Figura 5.17. ....	160
Figura 5.32. Representação da combinação de bordas para a imagem <i>Sign04</i> , utilizando o coeficiente proporcional calculado sobre o atributo <i>rg</i> : (a) $p = 0$ ; (b) $p = 0,2$ . O código de cores é o definido na Figura 5.17. ....	161

Figura 5.33. Combinações de bordas utilizando o coeficiente proporcional. (a)-(c): Combinação baseada na classe <i>geral</i> ; (d)-(f): Combinação baseada na classe <i>laboratório</i> ; (g)-(i): Combinação baseada na classe <i>cabos</i> . Para cada tríade foram utilizadas as proporções $p$ iguais a 0, 0,2 e 1. Todas as bordas consideram o atributo $c$ .....	162
Figura 5.34. Combinações de bordas utilizando o coeficiente proporcional. (a)-(c): Combinação baseada na classe <i>geral</i> ; (d)-(f): Combinação baseada na classe <i>externas</i> . Para cada tríade foram utilizadas as proporções $p$ iguais a 0, 0,2 e 0,4. Todas as bordas consideram o atributo $y_b$ .....	163
Figura 6.1. Esquema geral do sistema de visão <i>Cyvis-1</i> . (Adaptado de Costa <i>et al.</i> [Costa <i>et al.</i> (1994)].) .....	170
Figura 6.2. Esquema da estruturação entre os módulos para a integração entre estéreo e cor. Setas sólidas indicam troca de dados, enquanto as tracejadas indicam trocas de sinais de controle.....	170
Figura 6.3. Representação da simplificação de segmentos de reta. (a) Dois pares de segmentos consecutivos e (b) sua simplificação para dois segmentos. (c) Dois pares de segmentos coincidentes e (d) sua simplificação. As retas sólidas indicam a posição de segmentos reais e as tracejadas a posição original dos segmentos que foram desconsiderados ou modificados. ....	171
Figura 6.4. Segmentos que descrevem os modelos da base de dados. ....	173
Figura 6.5. Três segmentos, em destaque, emparelhados com as arestas dos modelos. ....	175
Figura 6.6. Exemplo de um prisma. (a) e (b): Par de imagens em estéreo; (c) Imagem em (a) segmentada por mapas auto-organizáveis utilizando cromaticidade; (d) imagem em (a) segmentada por Sobel. ....	176
Figura 6.7. Imagens dos objetos e suas segmentações por mapas auto-organizáveis e por Sobel. (a), (d), (g), (j) e (m): Imagens direitas do par estéreo. (b), (e), (h), (k) e (n): Mapas de bordas por mapas auto-organizáveis. (c), (f), (i), (l) e (o): Mapas de bordas por Sobel. ....	178
Figura 6.8. Segmentos de reta detectados nos mapas de bordas, apresentados. São considerados os seguintes mapas: (a) Figura 6.7(e); (b) Figura 6.7(f); (c) Figura 6.7(k); e (d) Figura 6.7(l).....	179
Figura 6.9. Esquema de uma borda com certa espessura em um vértice da imagem. Destaca-se, tracejada, a reta desejada, que corresponde à melhor orientação para determinar a aresta. A reta sólida indica a orientação da reta obtida por Hough. ....	180
Figura 6.10. Segmentos detectados na imagem da Figura 6.7(j), utilizando a segmentação por mapas auto-organizáveis da Figura 6.8(c), sob três pontos de vista diferentes.....	181
Figura 6.11. Exemplos de reconhecimento de objetos com superposição dos segmentos detectados ao modelo reconhecido. (a), (d) e (g): Imagens direitas do par estéreo; (b), (e) e (h): superposição para mapas auto-organizáveis e (c), (f) e (i): superposição para Sobel. Para os segmentos do objeto, a cor magenta indica os que foram emparelhados com o modelo e a cor preta os restantes. Para o modelo, a cor azul representa as arestas emparelhadas e a cor ocre aquelas que não foram utilizadas.....	182
Figura 6.12. Ilustração da extensão de um segmento de reta (em preto) para o caso de serem toleradas descontinuidades de até três pontos. Os pontos hachurados representam as posições nas quais existem bordas. ....	185
Figura C.1. (a) Conjunto de cinco pontos no espaço $xy$ . (b) Curvas no espaço de parâmetros correspondentes a estes pontos. ....	210
Figura E.1. Coeficiente acumulativo do atributo $g$ para a classe <i>casas</i> . (a)-(f): $p$ com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.....	252

Figura E.2. Coeficiente acumulativo do atributo <i>c</i> para a classe <i>casas</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	253
Figura E.3. Coeficiente acumulativo do atributo <i>rg</i> para a classe <i>casas</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	254
Figura E.4. Coeficiente acumulativo do atributo <i>yb</i> para a classe <i>casas</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	255
Figura E.5. Coeficiente acumulativo do atributo <i>g</i> para a classe <i>cubos</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	256
Figura E.6. Coeficiente acumulativo do atributo <i>c</i> para a classe <i>cubos</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	257
Figura E.7. Coeficiente acumulativo do atributo <i>rg</i> para a classe <i>cubos</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	258
Figura E.8. Coeficiente acumulativo do atributo <i>yb</i> para a classe <i>cubos</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	259
Figura E.9. Coeficiente acumulativo do atributo <i>g</i> para a classe <i>externas</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	260
Figura E.10. Coeficiente acumulativo do atributo <i>c</i> para a classe <i>externas</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	261
Figura E.11. Coeficiente acumulativo do atributo <i>rg</i> para a classe <i>externas</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	262
Figura E.12. Coeficiente acumulativo do atributo <i>yb</i> para a classe <i>externas</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	263
Figura E.13. Coeficiente acumulativo do atributo <i>g</i> para a classe <i>laboratório</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	264
Figura E.14. Coeficiente acumulativo do atributo <i>c</i> para a classe <i>laboratório</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	265
Figura E.15. Coeficiente acumulativo do atributo <i>rg</i> para a classe <i>laboratório</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	266
Figura E.16. Coeficiente acumulativo do atributo <i>yb</i> para a classe <i>laboratório</i> . (a)-(f): <i>p</i> com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente. ....	267
Figura F.1. Coeficiente proporcional do atributo <i>g</i> para a classe <i>casas</i> . (a)-(f): <i>p</i> com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	269
Figura F.2. Coeficiente proporcional do atributo <i>c</i> para a classe <i>casas</i> . (a)-(f): <i>p</i> com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	270
Figura F.3. Coeficiente proporcional do atributo <i>rg</i> para a classe <i>casas</i> . (a)-(f): <i>p</i> com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	271
Figura F.4. Coeficiente proporcional do atributo <i>yb</i> para a classe <i>casas</i> . (a)-(f): <i>p</i> com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	272
Figura F.5. Coeficiente proporcional do atributo <i>g</i> para a classe <i>cubos</i> . (a)-(f): <i>p</i> com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	273
Figura F.6. Coeficiente proporcional do atributo <i>c</i> para a classe <i>cubos</i> . (a)-(f): <i>p</i> com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	274
Figura F.7. Coeficiente proporcional do atributo <i>rg</i> para a classe <i>cubos</i> . (a)-(f): <i>p</i> com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	275
Figura F.8. Coeficiente proporcional do atributo <i>yb</i> para a classe <i>cubos</i> . (a)-(f): <i>p</i> com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	276

Figura F.9. Coeficiente proporcional do atributo $g$ para a classe <i>externas</i> . (a)-(f): $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	277
Figura F.10. Coeficiente proporcional do atributo $c$ para a classe <i>externas</i> . (a)-(f): $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	278
Figura F.11. Coeficiente proporcional do atributo $rg$ para a classe <i>externas</i> . (a)-(f): $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	279
Figura F.12. Coeficiente proporcional do atributo $yb$ para a classe <i>externas</i> . (a)-(f): $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	280
Figura F.13. Coeficiente proporcional do atributo $g$ para a classe <i>laboratório</i> . (a)-(f): $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	281
Figura F.14. Coeficiente proporcional do atributo $c$ para a classe <i>laboratório</i> . (a)-(f): $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	282
Figura F.15. Coeficiente proporcional do atributo $rg$ para a classe <i>laboratório</i> . (a)-(f): $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	283
Figura F.16. Coeficiente proporcional do atributo $yb$ para a classe <i>laboratório</i> . (a)-(f): $p$ com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente. ....	284

## LISTA DE TABELAS

Tabela 2-I. Comparação básica dos cones e bastonetes. (Adaptado de Goldstein [Goldstein (1989)].).....	14
Tabela 2-II. Relação entre comprimento de onda e cor percebida. (Adaptado de Goldstein [Goldstein (1989)].).....	14
Tabela 2-III. Principais termos utilizados em cor e iluminação, com suas definições. ....	18
Tabela 5-I. Síntese das medidas de desempenho propostas. ....	120
Tabela 5-II. Classes de imagens. ....	120
Tabela 5-III. Variações máximas dos atributos locais para a base de imagens.....	145
Tabela 5-IV. Sumário do desempenho relativo entre as técnicas de segmentação utilizadas. ....	155
Tabela 7-I. Síntese das contribuições deste trabalho. ....	193
Tabela D-I. Descrição das legendas para a figuras do apêndice. ....	213

## LISTA DE ALGORITMOS

Algoritmo 3-1. Extração de bordas por gradiente cromático [Lee & Cok (1991)].....	38
Algoritmo 3-2. Procedimento básico do algoritmo $k$ médias. (Adaptado de Duda e Hart [Duda & Hart (1973)]). .....	42
Algoritmo 3-3. Segmentação por agrupamento de Coleman e Andrews [Coleman & Andrews (1979)]. .....	43
Algoritmo 3-4. Segmentação por agrupamento no espaço $L^*a^*b^*$ . [Celenk (1990)].....	44
Algoritmo 3-5. Procedimento básico para o agrupamento $c$ médias difuso. ....	48
Algoritmo 3-6. Algoritmo de classificação difusa $c$ médias com seleção automática do número de classes [Kamei-Harada <i>et al.</i> (1994)].....	50
Algoritmo 3-7. Simplificação da árvore de distribuição cromática [Uchimura (1994)]. .....	52
Algoritmo 3-8. Segmentação por divisão de regiões através da análise de histogramas [Ohlander <i>et al.</i> (1978)]. .....	53
Algoritmo 3-9. Segmentação por aprendizado competitivo [Uchiyama & Arbib (1994)]. .....	55
Algoritmo 4-1. Fase de aprendizado para a formação do mapa auto-organizável. (Adaptado de Fausett [Fausett (1994)]). .....	68
Algoritmo 4-2. Segmentação por agrupamento de Coleman e Andrews [Coleman & Andrews (1979)]. .....	76

## RESUMO

Esta tese descreve uma abordagem para a utilização da informação de cores no sistema de visão artificial com inspiração biológica denominado *Cyvis-1*. Considerando-se que grande parte da literatura sobre segmentação de imagens se refere a imagens em níveis de cinza, informações cromáticas na segmentação permanecem uma área que ainda deve ser melhor explorada e para a qual se direcionou o interesse da presente pesquisa. Neste trabalho, o subsistema de cor do *Cyvis-1* é definido, mantendo-se o vínculo com os princípios que inspiraram o sistema de visão como um todo: hierarquia, modularidade, especialização do processamento, integração em vários níveis, representação efetiva da informação visual e integração com conhecimento de nível alto. O subsistema de cor se insere neste escopo, propondo uma técnica para segmentação de imagens coloridas baseada em mapas auto-organizáveis para a classificação dos pontos da imagem. A segmentação incorpora a determinação do número de classes sem supervisão, tornando o processo mais independente de intervenção humana. Por este processo de segmentação, são produzidos mapas das regiões encontradas e um mapa de bordas, derivado das regiões. Uma segunda proposta do trabalho é um estudo comparativo do desempenho de técnicas de segmentação por bordas. A comparação é feita em relação a um mapa de bordas de referência e o comportamento de várias técnicas é analisado segundo um conjunto de atributos locais baseados em contrastes de intensidade e cor. Derivada desta comparação, propõe-se também uma combinação para a geração de um mapa de bordas a partir da seleção das técnicas segundo seus desempenhos locais. Finalmente, integrando os aspectos anteriores, é proposta uma estruturação do módulo de cor, adicionalmente com a aquisição de imagens, a análise de formas e o reconhecimento de objetos poliédricos. Há, neste contexto, a integração ao módulo de estéreo, que proporciona o cálculo de dados tridimensionais, essenciais para o reconhecimento dos objetos. Para cada parte deste trabalho são propostas formas de avaliação para a validação dos resultados, demonstrando e caracterizando a eficiência e as limitações de cada uma.

## ABSTRACT

This thesis describes an approach to color information processing in the biologically-inspired artificial vision system named *Cyvis-1*. Considering that most of the current literature in image segmentation deals with gray level images, color information remains an incipient area, which has motivated this research. This work defines the color subsystem within the *Cyvis-1* underlying philosophy, whose main principles include hierarchy, modularity, processing specialization, multilevel integration, effective representation of visual information, and high-level knowledge integration. The color subsystem is then introduced according to this framework, with a proposal of a segmentation technique based on self-organizing maps. The number of regions in the image is achieved through a unsupervised clustering approach, so no human interaction is needed. Such segmentation technique produces region oriented representation of the classes, which are used to derive an edge map. Another main topic in this work is a comparative study of the edge maps produced by several edge-oriented segmentation techniques. A reference edge map is used as standard segmentation, to which other edge maps are compared. Such analysis is carried out by means of local attributes (local gray level and "color" contrasts). As a consequence of the comparison, a combination edge map is also proposed, based on the conditional selection of techniques considering the local attributes. Finally, the integration of two above topics is proposed, which is characterized by the design of the color subsystem of *Cyvis-1*, altogether with the modules for image acquisition, shape analysis and polyhedral object recognition. In such a context, the integration with the stereo subsystem is accomplished, allowing the evaluation of the three-dimensional data needed for object recognition. Assessment and validation of the three proposals were carried out, providing the means for analyzing their efficiency and limitations.

*"The soul never thinks without a picture."*  
*Aristóteles*

# **Introdução e revisão bibliográfica**

“‘Begin at the beginning,’ the King said, gravely, ‘and go on  
till you come to the end; then stop’.”  
Lewis Carrol — *Alice in the Wonderland*

## 1 Introdução e objetivos

Muitos são os animais dotados de visão, uns com sistemas mais simples e outros mais sofisticados, uns mais específicos, outros mais gerais. A visão é usada para inúmeras tarefas, muitas delas relacionadas à sobrevivência e perpetuação da espécie. Sistemas de visão de mamíferos superiores, como os primatas, apresentam uma grande complexidade e mecanismos sofisticados de reconhecimento de padrões [Costa *et al.* (1994)]. Para estes animais, a qualidade da visão é cada vez mais exigida, incluindo uma componente diferencial que é a visão colorida. Embora animais como abelhas e camarões reconheçam as diferenças nos comprimentos de onda, a visão nos primatas assume um estágio de maior aprimoramento [Levine (1985); Goldstein (1989); Tovée (1996)].

As cores são elementos importantes no cotidiano, sendo usadas para distinguir objetos, qualificá-los ou avaliar suas propriedades. A incorporação desta classe de informações em sistemas artificiais de visão cria uma expectativa de abertura de novos caminhos para o estudo não somente da visão artificial, mas para a compreensão dos mecanismos biológicos envolvidos. É viável, assim, potencializar as soluções que a evolução introduziu para sua incorporação aos sistemas artificiais.

Este contexto das cores na visão levou à proposição deste trabalho, que abrange, em primeira instância, uma revisão dos mecanismos de visão em cores e das técnicas de processamento e análise de imagens que utilizam tal informação. Em uma segunda instância, a incorporação de uma abordagem específica da cores ao sistema de visão *Cyvis-1* é considerada, buscando mecanismos de segmentação para a extração de informações básicas das imagens. Além disto, a avaliação da qualidade das informações da segmentação é incluída, envolvendo regiões e bordas, juntamente com a proposta de uma organização para integração das cores a outras modalidades (atributos) da visão. Em particular, o interesse foi direcionado à integração com aspectos tridimensionais, que estão relacionados à visão estereoscópica. A integração compreende uma perspectiva particularmente interessante do trabalho, proporcionando novos

elementos para a busca de soluções, além de envolver aspectos de paralelismo, hierarquia e comunicação, entre outros.

O restante deste capítulo descreve o sistema de visão ao qual este trabalho se vincula (seção 1.1), caracterizando-o de forma geral. A seção 1.2, então, introduz os objetivos do trabalho e a organização do texto como um todo é apresentada na seção 1.3.

### 1.1 O sistema de visão artificial *Cyvis-1*

Um sistema de visão com inspiração em características de processamento biológico foi proposto por Costa, Roda e Köberle [Costa *et al.* (1994)], incorporando uma série de particularidades que o tornam um sistema de alta integração e grande versatilidade.

Em contínuo desenvolvimento, o sistema *Cyvis-1* incorpora vários princípios inspirados na organização e processamento biológicos. Uma relação sinóptica destes princípios que norteiam o desenvolvimento do sistema é apresentada na seqüência.

- a) *Resolução alta para representação de imagens.* Embora sistemas computacionais raramente utilizem imagens maiores que  $512 \times 512$ , o olho humano se utiliza de 126 milhões de receptores para amostrar seu campo visual. A utilização de uma “sobreamostragem” induz à expectativa de melhoramentos no reconhecimento de padrões.
- b) *Hierarquia.* Sistemas naturais de visão processam a informação de forma hierárquica, sendo que em cada estágio da hierarquia a quantidade de dados é reduzida e, simultaneamente, ocorre o aumento de conhecimento sobre os objetos da imagem.
- c) *Retroalimentação.* O refluxo de informações entre camadas hierárquicas, a exemplo de sistemas biológicos, permite não somente que o processamento ocorra de baixo para cima, mas também de cima para baixo na hierarquia. Etapas posteriores do sistema podem interagir com camadas anteriores, modificando seu modo de processamento e contribuindo para um reconhecimento mais direcionado, permitindo, por exemplo, que conhecimento (informação de nível alto) influencie o processamento de níveis mais baixos (veja-se o item j).
- d) *Modularidade.* Sistemas naturais de visão apresentam um alto grau de modularidade de processamento. Subsistemas podem ser organizados de forma a operar em características específicas da informação visual, produzindo resultados parciais que poderão ser integrados ao resultado de outros subsistemas. Adicionalmente, subsistemas podem, de modo cooperativo, compartilhar informações que tornem mais efetivas suas atuações individuais.
- e) *Representação efetiva da informação visual.* O sistema visual de primatas captura informações com algo em torno dos já mencionados 126 milhões de fotorreceptores, mas o número de fibras que deixam o globo ocular já reduzem os dados para aproximadamente 800.000 axônios, havendo uma grande codificação da informa-

- ção. A forma com que a informação é representada em cada nível da hierarquia constitui uma condição de grande importância para o sistema como um todo.
- f) *Especialização de processamento.* A especialização de atributos a serem processados formam uma das características mais interessantes do sistema de visão de primatas [Zeki & Shipp (1988); Livingstone & Hubel (1988)]. Textura, bordas, cor e profundidade, por exemplo, caracterizam-se como atributos processados de forma especializada. Uma grande vantagem desta segregação de processamento é a possibilidade que vários subsistemas atuem de forma paralela.
  - g) *Integração multiescala.* Existe uma extensa interconexão entre os diversos elementos de processamento do cérebro de primatas, complementando a modularidade, a hierarquia, a retroalimentação e o processamento especializado. Através desta interconexão, módulos especializados podem fazer intercâmbio de informações e cooperar no processo global de reconhecimento da imagem.
  - h) *Visão foveal*<sup>\*</sup>. A resolução dos fotorreceptores para a varredura do campo visual não é uniforme, mas apresenta, próximo ao eixo de visão, uma quantidade diferenciada de receptores na região denominada fóvea. Nesta área da fóvea, a acuidade é bastante superior às regiões mais periféricas da retina, tanto em termos de resolução quanto de qualidade (cor, por exemplo). Mover os olhos para que a luz do objeto de interesse recaia sobre a fóvea é um mecanismo bastante interessante para o processamento da informação visual.
  - i) *Atenção seletiva e visão ativa.* Diretamente relacionada à visão foveal, a atenção seletiva refere-se à interpretação da imagem pela análise parte a parte, movendo a atenção para regiões específicas. Embora a atenção selecione regiões seqüencialmente, em níveis mais altos da hierarquia cada informação local é combinada para uma compreensão mais geral da cena.
  - j) *Integração com conhecimento de nível mais alto.* Um princípio biológico importante que pode ser utilizado é a incorporação, no processamento da informação visual, do conhecimento de nível alto, o que inclui o contexto no qual a imagem está inserida, sua história e também memória, que permite a comparação de um objeto identificado com outro similar já interpretado anteriormente.
  - k) *Processo evolucionário.* O processo de evolução envolve pequenas mudanças no sistema com o decorrer do tempo, cada uma delas avaliada severamente quanto a sua utilidade, funcionalidade e desempenho. Também sistemas artificiais de visão devem, sistematicamente, avaliar cada uma de suas alterações quanto à correta operação e funcionalidade. Sob este aspecto, é importante não somente a validação dos novos processos, mas também uma avaliação comparativa com seu antecessor ou técnicas alternativas.
  - l) *Soluções não convencionais.* Embora a visão humana possa ser considerada como o sistema de reconhecimento de padrões mais versátil e eficiente entre os animais,

---

<sup>\*</sup> Define-se o termo *foveal* como o que é relativo à fóvea da retina, ou apresenta suas características.

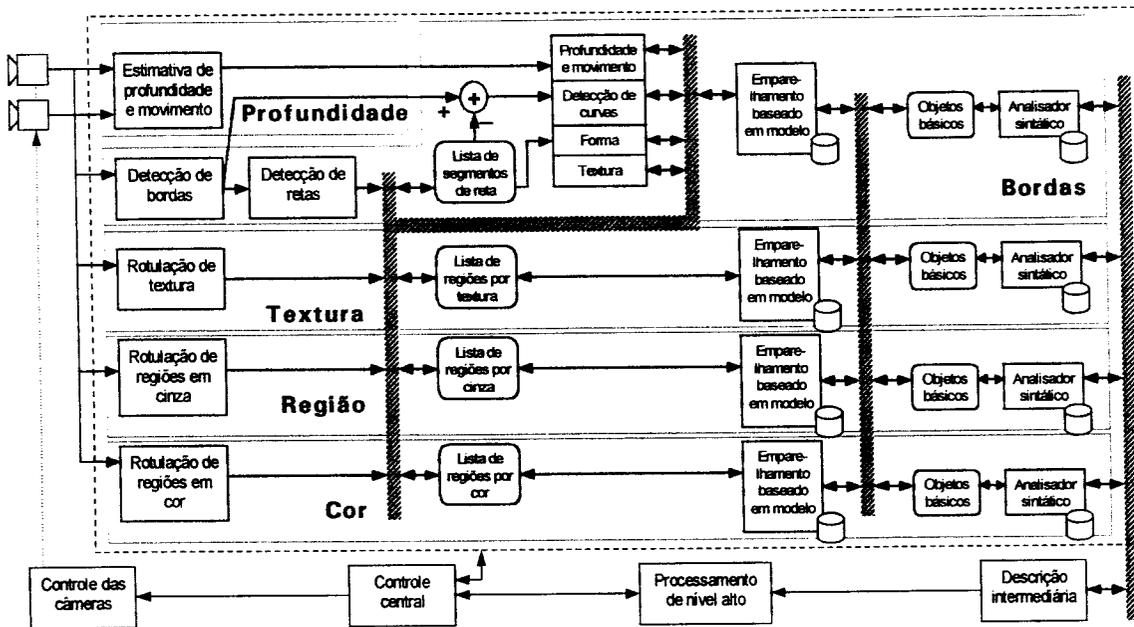


Figura 1.1. Esquema da estrutura geral do sistema de visão *Cyvis-1*. O símbolo  indica a existência de uma base de dados associada. (Adaptado de Costa *et al.* [Costa *et al.* (1994)].)

algumas soluções mais específicas também representam alternativas interessantes. Moscas, por exemplo, são capazes de captar energia eletromagnética na faixa do ultravioleta. Estas soluções não convencionais representam características importantes que podem ser incorporadas ao sistema de visão.

A Figura 1.1 mostra a estrutura geral do *Cyvis-1*, o qual foi projetado levando em conta critérios de desempenho, velocidade, custo e flexibilidade.

A entrada de informação no sistema é feita através de um par de câmeras de vídeo, que alimentam cinco subsistemas principais representando as especializações: profundidade, bordas, textura, regiões (em níveis de cinza) e cor. A estruturação é essencialmente hierárquica e mantém interconexões tanto de retroalimentação quanto de convergência de informações. A modularidade é empregada em todo o sistema, promovendo ao máximo a implementação de paralelismo.

Informações de nível baixo são obtidas dos módulos que processam diretamente a imagem, obtendo as bordas ou regiões (textura, níveis de cinza e cor). As informações extraídas são então comparadas a uma base de dados para identificação dos principais objetos da cena. A base de dados em questão tem nível mais específico, relacionando-se a um objeto específico (ou parte dele). Neste estágio, um analisador sintático interpreta os objetos no contexto completo da imagem, gerando uma descrição intermediária com o auxílio de uma outra base de dados de nível mais alto, a qual é repassada aos níveis posteriores da hierarquia para a tomada de decisões e controle do processo.

## 1.2 Objetivos

O presente trabalho tem como um dos seus principais objetivos a segmentação de imagens coloridas para a obtenção de um mapa das bordas presentes, além de uma análise comparativa entre alguns métodos de segmentação (também em relação a bordas). Adicionalmente incluem-se nos objetivos as investigações da integração entre módulos distintos do *Cyvis-1*, especialmente com o módulo de visão estereoscópica.

### 1.2.1 Segmentação de imagens coloridas

Para a segmentação de imagens considerando atributos de cor, o objetivo é a proposição, caracterização e análise de uma metodologia de segmentação baseada em mapas auto-organizáveis, envolvendo não somente a obtenção da segmentação por regiões, inerente ao método, mas permitindo também a derivação de um mapa de bordas a partir dos limites entre as regiões. Entre os objetivos secundários, relativos a esta metodologia de segmentação, incluem-se a avaliação das regiões obtidas, comparando-as a uma segmentação “ideal”, feita manualmente, e comparando-a com a obtida por outros métodos. Outros parâmetros como inclusão ou não do atributo iluminação no processo de segmentação também estão envolvidos, além da comparação dos resultados com outras técnicas de segmentação.

### 1.2.2 Comparação de mapas de bordas

Um segundo objetivo do trabalho é a comparação dos mapas de bordas produzidos por um conjunto de técnicas de segmentação. Assim, a proposição de um modelo de avaliação é considerada, visando também examinar como bordas derivadas da segmentação proposta se relacionam a outras técnicas existentes. A proposta envolve a comparação em relação a um conjunto de atributos locais, permitindo considerar relativamente os desempenhos de uma ou outra forma de segmentação. Incluem-se, portanto, as definições de medidas de erro que viabilizem esta confrontação das técnicas.

### 1.2.3 Integração ao *Cyvis-1*

Enquanto o *Cyvis-1* é um sistema bastante abrangente, a cor representa um de seus principais atributos. Considerando-se o sistema como um todo, além do processamento das informações coloridas existe o seu relacionamento com outros atributos. Objetiva-se, desta forma, uma estruturação para a integração de dois atributos visuais, especificamente *cor* e *estéreo*. Além da integração, voltada ao reconhecimento de objetos poliédricos, inclui-se como objetivo a definição da avaliação destes reconhecimentos, visando validar os resultados produzidos.

## 1.3 Organização do texto

O texto como um todo está organizado em duas partes principais: revisão e trabalho desenvolvido. No Capítulo 2, parte da revisão, é apresentada uma explicação sobre os conceitos básicos sobre as cores, sua percepção e cognição no sistema

visual de primatas, bem como uma modelagem numérica sob o prisma da colorimetria. Também está incluída uma síntese sobre os principais sistemas de representação de espaços de cores. Uma revisão das principais metodologias e técnicas para o processamento de imagens coloridas é feita no Capítulo 3, com direcionamento a segmentação de imagens. Uma subdivisão das abordagens de segmentação de imagens coloridas é feita incluindo-se ênfases em bordas (seção 3.1) e regiões (seção 3.2), além de redes neurais e modelagem física da cena (seções 3.3 e 3.4, respectivamente).

A segunda parte do texto inclui o trabalho desenvolvido, caracterizado em três capítulos. No Capítulo 4 é feita a descrição e análise da metodologia para a segmentação de imagens coloridas desenvolvida. Um estudo comparativo entre algumas técnicas de segmentação é exposto no Capítulo 5, comparando desempenhos relativos mapas de bordas gerados por técnicas distintas, dentro de um contexto de atributos locais das imagens. O Capítulo 6 apresenta, dentro do contexto do *Cyvis-1*, o esquema de cooperação entre os subsistemas de cor e visão estereoscópica, direcionado para o reconhecimento de objetos poliédricos.

As conclusões são apresentadas no Capítulo 7, o qual também discute as contribuições deste trabalho e apresenta alternativas de continuidade.

Os apêndices incluem tópicos adicionais ao texto. Um glossário é apresentado no Apêndice A, relacionando os principais termos utilizados no texto, caracterizando-os segundo a conotação adotada. Em virtude da necessidade de representação simbólica de vários conceitos utilizados, o Apêndice B traz uma descrição da notação utilizada tanto nas fórmulas quanto no texto, situando-a dentro dos respectivos temas. A transformada de Hough, utilizada no sistema de integração é apresentada de forma resumida no Apêndice C. Dados adicionais relativos aos resultados obtidos estão relacionados nos Apêndices D, E e F, sendo que o primeiro é relacionado às segmentações por mapas auto-organizáveis e os dois últimos são relativos ao estudo comparativo das bordas.

Escrito sob o enfoque de um trabalho científico, o texto permite leitura essencialmente linear. Entretanto, várias referências são feitas a informações contidas em outros capítulos ou seções, no intuito de prover os elementos necessários para a compreensão adequada de um assunto particular. As referências bibliográficas, indicadas entre colchetes, quando inseridas no meio das sentenças, referem-se somente ao assunto imediatamente exposto; quando posicionadas após o final do parágrafo, indicam que a referência aplica-se ao parágrafo como um todo.

*“Faith is a light of such supreme brilliance that it dazzles the mind and darkens all its visions of other realities, but in the end, when we become used to the new light, we gain a new view of all reality transfigured and elevated in the light itself.”*  
Thomas Merton

## 2 Conceitos básicos sobre cores

 estudo das cores atraiu um grande número de pesquisadores de renome, incluindo Aristóteles, Grimaldi, Newton, Goethe, Hegel, Schopenhauer, Young, Maxwell, Helmholtz, Hering e Schrödinger. Uma visão geral de suas participações no entendimento da cor é apresentada, baseada em boa parte na introdução de D. B. Judd à reedição do livro *Theory of colours* de Goethe [Goethe (1970)].

Aristóteles entendia a cor como um fenômeno da transição do claro para o escuro, dado que a intensidade da energia luminosa se reduz quando atravessa um material ou é refletida por um objeto. A cor seria, segundo sua visão, uma mistura, fusão ou justaposição de preto e branco, ou a degeneração da cor branca por certa quantidade de escuridão.

Com os experimentos de Newton, no século XVII, sobre a decomposição da luz branca (solar) em outras através de um prisma, a luz passou a ser vista como incolor e a cor, em si, passou a ser atribuída a alguma qualidade do material translúcido que degradava esta luz “pura”. A cor, assim, passou a ser entendida como uma propriedade da própria luz. Esta concepção de que a cor era exclusiva propriedade da luz difundiu-se entre os seguidores de Newton, embora ele próprio não concordasse com esta abordagem extrema\*. Esta opinião era compartilhada com Grimaldi que, também em 1665 (mesmo ano da publicação de Newton), expressou uma visão similar. Nesta época a teoria da luz como ondulações se propagando no espaço não havia sido ainda estabelecida, de forma que Newton enxergava a luz como corpúsculos ou partículas

---

\* “Os raios, para dizer apropriadamente, não são coloridos. Neles não há nada além de uma certa potência e disposição para provocar a sensação desta ou daquela cor. (...) então, cores em objetos não são nada mais que a disposição de refletir este ou aquele tipo de raio mais intensamente que os outros.” (Newton, *Optika*)

provindas de fontes de luz. A consolidação da teoria de ondas ocorreu somente em 1727, realizada por Young. A desaprovação da visão de Newton foi levantada também por Fresnel, que, paralelamente a Young, explicava também a natureza da luz como ondulações. Novamente passou a ser, em geral, difícil considerar a cor como uma sensação, mas simplesmente postulando que diferentes comprimentos de onda “possuem” determinadas cores.

Baseado somente em sua experiência e intensa observação pessoal, Goethe propôs sua *Theory of colours* (*Materialien zur Geschichte der Farbenlehre, Zur Farbenlehre* [Goethe (1810)]), livro que cobre grande parte dos fenômenos de cor conhecidos na época, ignorando completamente as descobertas de Newton. A visão filosófica empreendida no livro foi acolhida por Hegel, que compartilhava a concepção da lógica filosófica, mas estas concepções foram derrubadas frente às descobertas atuais.

Em 1802, Young, apoiado em experimentos psicofísicos de emparelhamentos de cores, propôs a teoria tricromática da visão. Considerando-se que, pela combinação de três cores diferentes em uma mesma área é possível obter a mesma aparência de uma quarta cor\*, Young postulou a percepção de cor baseada na existência de três mecanismos receptores com sensibilidades diferentes aos comprimentos de onda. Esta teoria foi compartilhada por Helmholtz em 1802, ficando conhecida também como a teoria de Young-Helmholtz. Hering, em 1878, observando uma relação de oposição entre cores como entre vermelho e verde e entre azul e amarelo, propôs a teoria de oposição cromática da visão. [Goldstein (1989)]

Estas teorias, que se antagonizaram, são ambas reforçadas por evidências fisiológicas, dados os três tipos de cones, responsáveis pela seleção espectral, e a existência de células que apresentam oposição cromática, presentes no núcleo geniculado lateral e no córtex visual.

Este capítulo cobre as principais informações sobre cor e visão colorida, principalmente quando relacionada aos primatas e, particularmente, aos seres humanos. São estabelecidos, inicialmente, os conceitos da luz como energia eletromagnética e sua caracterização por comprimentos de onda, indicando como o sistema visual responde a ela e a interpretação proporcionada. Além disto, os aspectos de representação das informações coloridas (espaços de cores) são expostos.

## 2.1 Ondas eletromagnéticas

O processo de visão tem origem na existência de energia eletromagnética em forma de ondas, com comprimentos de onda em faixas específicas. Estas ondas, com origem em fontes diversas (como Sol, lâmpadas, fogo etc.), são absorvidas e refletidas pelos diversos objetos. Devido às diferenças intrínsecas das propriedades dos corpos, como textura e opacidade, entre outras, cada um proporciona uma diferente gama de

---

\* Discutido na seção seguinte. Veja-se também a Figura 2.2.

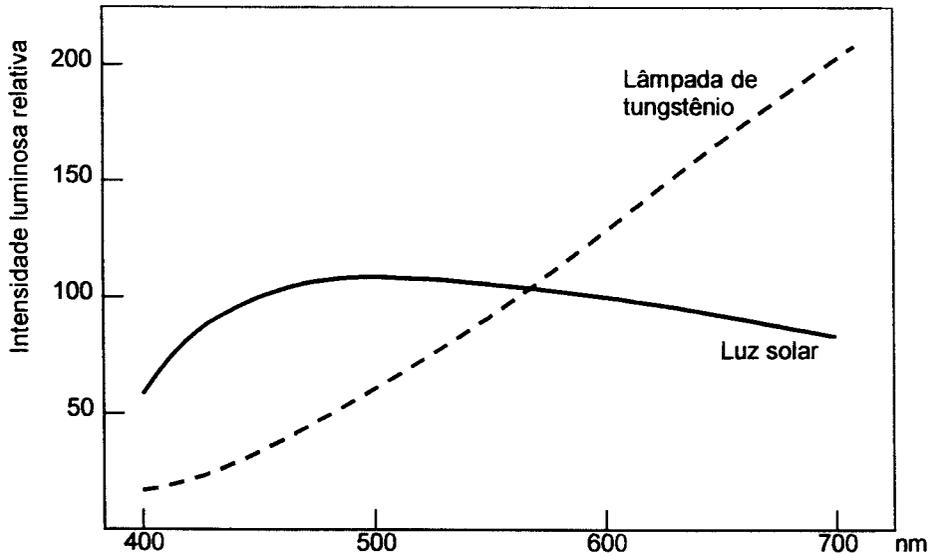


Figura 2.1 Distribuição da quantidade de luz relativa para uma lâmpada incandescente e a luz branca (solar) [Goldstein (1989)].

comprimentos de onda refletidos, em intensidades também distintas. É este universo de ondas que atinge os olhos e permite que o mundo seja percebido. Ainda assim, dentre a ampla variação de comprimentos das ondas eletromagnéticas, a visão humana é capaz de captar somente um intervalo ínfimo, com comprimentos de onda  $\lambda$  variando de 380 a 780 nm [Wyszecki & Stiles (1982)].

Cada fonte de luz apresenta características próprias, principalmente em relação à energia nos diversos comprimentos de onda visíveis. A Figura 2.1 mostra a diferença entre a luz branca (luz solar) e a produzida por uma lâmpada incandescente de tungstênio. Enquanto a distribuição de energia é mais homogênea para todos os comprimentos de onda que compõem a luz branca, para a luz da lâmpada a quantidade de energia para comprimentos de onda curtos é bem menor que para os longos. Desta forma, a energia nos diversos comprimentos de onda que são captadas pelos olhos varia não somente de acordo com as reflexões, absorções e refrações nos diversos objetos, mas também com a fonte de iluminação utilizada.

## 2.2 Visão colorida

O processo de visão compreende, em termos práticos, desde a aquisição da imagem a ser analisada até a cognição dos objetos e formas presentes. Neste processo existem várias etapas intermediárias, não sendo todas ainda bem conhecidas ou bem determinadas [Wyszecki & Stiles (1982); Zeki & Shipp (1988)]. Há, por exemplo, a detecção de bordas, separação do objeto do fundo, análise de cores, texturas e outras propriedades, tais como continuidade e suavidade de objetos, além da noção de profundidade proporcionada pela visão estereoscópica (binocular). Mecanismos de orientação também estão presentes, permitindo associações de ângulos na composição dos elementos presentes na cena. Em etapas posteriores do processo de visão, há o

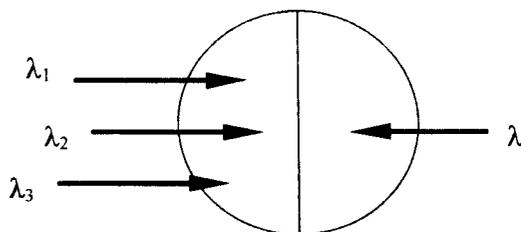


Figura 2.2. Experimento de emparelhamento de cores, no qual a soma dos comprimentos de onda  $\lambda_1$ ,  $\lambda_2$ , e  $\lambda_3$  à esquerda fornecem a mesma cor que o comprimento de onda monocromático  $\lambda$  projetado na direita. Por exemplo, a luz com comprimento de onda 500 nm pode ter sua equivalente perceptiva obtida a partir da mistura aditiva dos comprimentos 420, 560 e 640 nm [Goldstein (1989)].

reconhecimento dos objetos presentes, aos quais são atribuídos os caracteres de cor, textura etc. Esta última fase envolve, também, uma “memória” (ou banco de dados) que permite a associação de um objeto, ou partes dele, a elementos já conhecidos.

Os sistemas biológicos de visão presentes na grande maioria dos animais, incluindo o ser humano, apresentam características de particular eficiência, considerando seus objetivos funcionais, como é o caso da garantia da sobrevivência.

Não sendo diferente para o caso do primatas, a visão desempenha um papel fundamental na interação com o ambiente. Porém, parafraseando Mollon [Mollon (1982)], “todos são deficientes em visão colorida”, referindo-se ao fato de que os seres humanos não são capazes de distinguir entre luzes que lhes parecem visualmente iguais, apesar de terem composição distinta em termos de comprimentos de onda. Deficientes porque os humanos atribuem as mesmas características a fenômenos fisicamente diferenciados. Esta situação difere somente em grau de deficiências como, por exemplo, o daltonismo, que é uma deficiência visual que reduz o número de cores percebidas, com conseqüente dificuldade na distinção de cores que pessoas com visão normal percebem naturalmente.

Em experimentos de emparelhamento de cores, nos quais indivíduos são orientados a combinar as intensidades de três comprimentos de onda de forma que a cor produzida seja, em termos perceptivos, igual a um quarto (Figura 2.2), quase todas as cores podem ter sua equivalente a partir da combinação de outras três. É necessário, entretanto, que nenhuma delas seja obtida pela combinação das outras duas. No caso de se permitir que uma delas seja adicionada ao outro lado do campo de testes (duas de cada lado), então todas as cores podem ser emparelhadas e as duas partes do campo de testes parecerão iguais [Mollon (1982)]. Esta constatação levou à criação da teoria tricromática de cores (teoria de Young-Helmholtz), a qual postula que a luz estimula três mecanismos em graus diferentes e que a proporção destes estímulos fornece a percepção de uma cor\*.

Um fenômeno visual conhecido como pós-imagem levou à elaboração de uma outra teoria para a visão colorida. A pós-imagem é formada quando um indivíduo é orientado a observar uma figura de forma estável durante um período de 30 a 60 se-

\* Veja-se a Figura 2.6, na página 13.

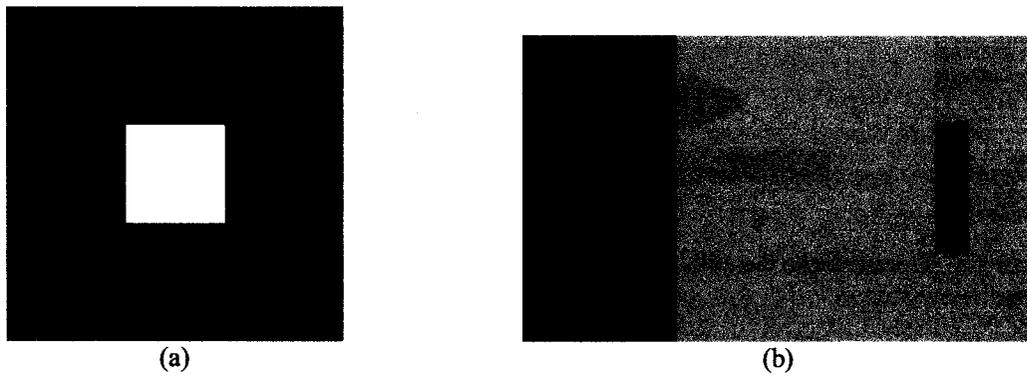


Figura 2.3. Imagens para (a) pós-imagem e (b) contraste simultâneo. (Adaptado de Goldstein [Goldstein (1989)].)

gundos e depois olhar para uma área branca. O efeito é a visualização da imagem, na qual partes claras ficam escuras, partes verdes ficam vermelhas e azuis ficam amarelas (e vice-versa). O efeito pode ser notado ao se realizar o experimento com a imagem da Figura 2.3(a). Ao se observar a pós-imagem fornecida, pode-se relacionar o vermelho e o verde de duas maneiras: através da figura que tem a coloração vermelha e, além disto, o quadrado branco em seu interior ganha uma coloração esverdeada. Este último efeito é denominado contraste simultâneo e reflete como a percepção de uma cor pode ser modificada conforme a cor se sua circunvizinhança. Na Figura 2.3(b) é possível notar o contraste simultâneo, sendo os dois retângulos pequenos de mesma cor, embora pareçam ter tons diferentes em virtude das cores que os cercam.

Esta teoria de oposição cromática propõe a existência de dois canais de visão colorida, um incorporando verde e vermelho e o outro amarelo e azul. Goldstein [Goldstein (1989)] propõe uma demonstração para este antagonismo através da visualização das cores. Não é difícil começar imaginando cores como vermelho e azul, que são primárias. Um segundo passo é a visualização de mistura de cores como vermelho amarelado ou verde azulado. Entretanto cores como vermelho esverdeado e azul amarelado são muito difíceis de visualizar, talvez impossíveis.

Ambas as teorias tricromática e de oposição cromática permaneceram separadas e conflitantes, até que descobertas permitiram comprovar que ambas estão corretas sob o ponto de vista fisiológico. A retina contém receptores cujos pigmentos absorvem a luz preferencialmente a comprimentos de onda curtos, médios e longos e, principalmente no LGN e no córtex, há células que respondem de forma antagônica à cor, criando as oposições vermelho/verde e azul/amarelo.

### 2.2.1 Cor através dos três tipos de cones

Na retina há a presença de três tipos distintos de receptores que respondem diferentemente a comprimentos de onda distintos. Estes receptores, denominados cones em função de sua morfologia, formam três grupos, sensíveis a comprimentos de onda curtos, médios e longos. Um quarto tipo de receptor, o bastonete, responde a intensidades luminosas, ou seja, à quantidade de energia luminosa que atinge a retina,

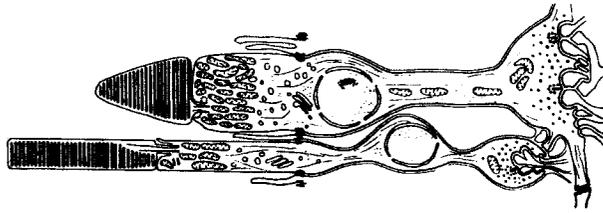


Figura 2.4. Representação das células receptoras de energia luminosa da retina: cone (acima) e bastonete.

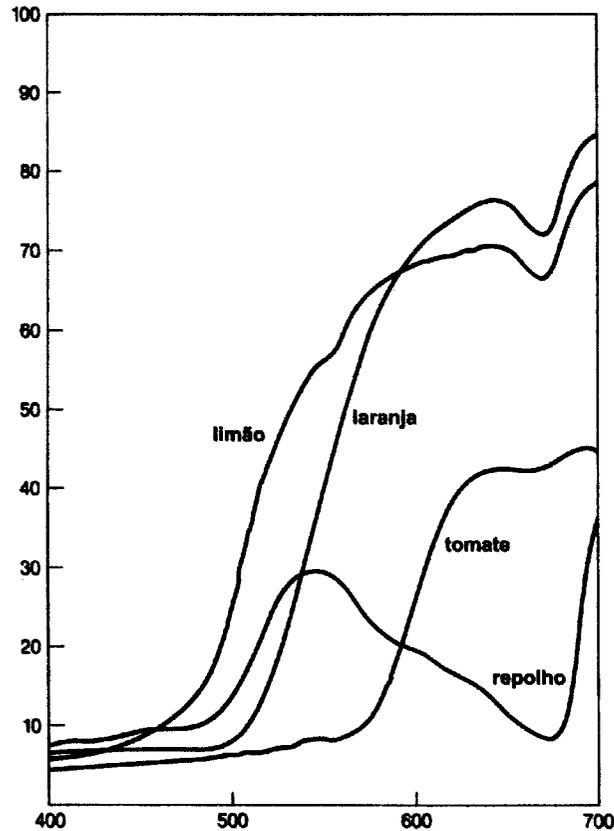


Figura 2.5. Curvas de refletância para alguns vegetais. [Goldstein (1989)]

não fazendo distinção dos comprimentos de onda que compõem a luz. A Figura 2.4 apresenta um esquema dos cones e bastonetes e a Tabela 2-I especifica suas diferenças comparativas básicas.

De acordo com o comprimento da luz que chega à retina uma certa cor é percebida, como ilustra a Tabela 2-II. A quantidade de luz refletida para vários comprimentos de onda permite, desta forma, associar cores aos pigmentos dos objetos. Na Figura 2.5 são apresentadas as curvas de refletância para alguns vegetais, que, se comparadas à Tabela 2-II, permitem verificar a relação entre a cor percebida e a distribuição dos comprimentos de onda. Isto, entretanto, é uma regra muito geral, uma vez que a percepção não depende somente do comprimento de onda, mas também de outros fatores, como, por exemplo, as cores da vizinhança.

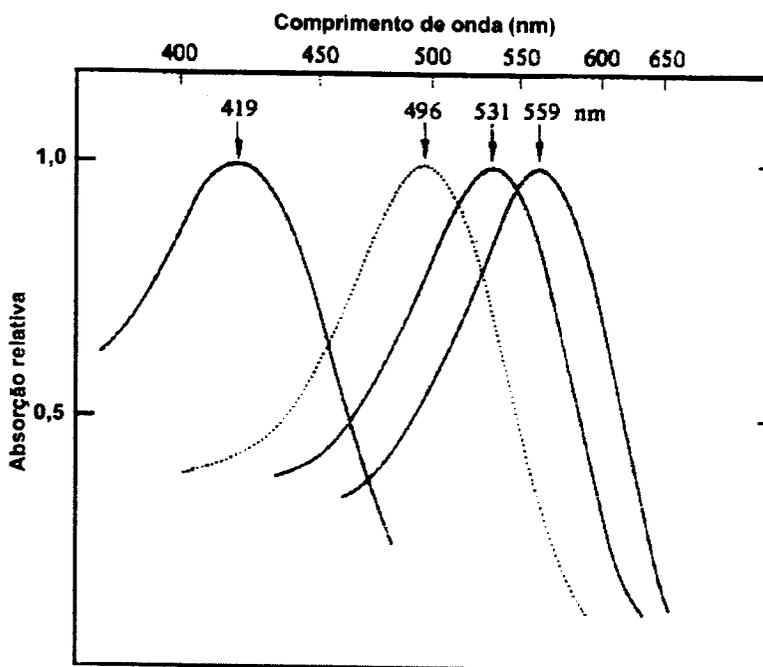


Figura 2.6. Espectro de absorção (normalizado) dos quatro pigmentos dos cones e bastonetes.

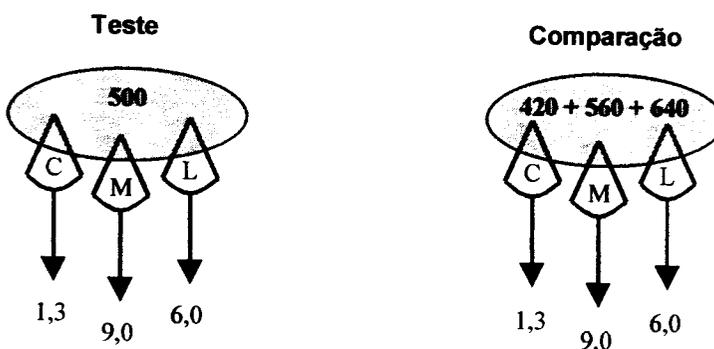


Figura 2.7. Dois estímulos distintos em termos de composição espectral podem provocar uma mesma resposta nos diferentes tipos de cones, não permitindo a distinção entre os estímulos.

Os três tipos de cones, por sua vez, têm espectros de absorção distintos, sendo que os sensíveis a comprimentos de onda curto apresentam pico de absorção em 420 nm, os sensíveis a comprimentos médios em 530 nm e os sensíveis a comprimentos longos em 560 nm [Mollon (1982)]. A Figura 2.6 mostra a absorção dos pigmentos que diferenciam cada tipo de cone e também os bastonetes. As linhas sólidas são relativas aos cones e a pontilhada aos bastonetes. A ordenada indica a absorção, ou seja, o logaritmo da razão da intensidade de luz incidente pela intensidade de luz transmitida, com valores normalizados. Há, ainda, diferenças em seu número e sua distribuição na retina. Cada retina tem, aproximadamente, 54% de cones sensíveis a ondas médias, 38% a longas e 13% a curtas, além do fato de que, na fóvea, a quantidade de cones sensíveis a ondas curtas está em torno de 3 a 4% do total. [Goldstein (1989)]

Tabela 2-I. Comparação básica dos cones e bastonetes. (Adaptado de Goldstein [Goldstein (1989)].)

<b>Propriedade</b>	<b>Cones</b>	<b>Bastonetes</b>
Formato do segmento externo	Similar a um cone	Similar a um bastão
Número	6 milhões em cada retina	120 milhões em cada retina
Distribuição	Fóvea e periferia	Periferia
Adaptação ao escuro	Rápida	Lenta
Sensibilidade espectral	Sensibilidade máxima em torno de 560 nm	Sensibilidade máxima em torno de 500 nm
Sensibilidade quando adaptado ao escuro	Mais baixa	Mais alta
Acurácia	Alta	Baixa

Tabela 2-II. Relação entre comprimento de onda e cor percebida. (Adaptado de Goldstein [Goldstein (1989)].)

<b>Comprimentos de onda refletidos</b>	<b>Cor percebida</b>
Curtos	Azul
Médios	Verde
Longos	Vermelho
longos e médios	Amarelo
longos e poucos médios	Laranja
longos e curtos	Roxo
longos, médios e curtos	Branco

A amostragem dos comprimentos de onda visíveis refletidos pelos objetos é feita, portanto, por três filtros caracterizados pelas curvas de absorção dos três tipos de cones. Conforme a proporção em que cada tipo de cone é estimulado, uma percepção de cor é estabelecida no cérebro. Goldstein [Goldstein (1989)], por exemplo, diz que, caso os cones sejam estimulados por um comprimento de onda de 500 nm, o grau de excitação relativo para os cones de comprimento de onda curtos pode ser 1,3, enquanto os graus para os de comprimento de onda médios e longos podem ser 9 e 6, respectivamente. Se o estímulo for substituído por uma mistura aditiva de cores de 420, 560 e 640 nm, os graus de excitação dos cones serão os mesmos, provocando a mesma sensação de cor que o estímulo de 500 nm isoladamente (Figura 2.7).

### 2.2.2 Oposição cromática

Seguindo o caminho visual, que combina através de sinapses os receptores da retina, já nas fibras do nervo óptico (formado pelos axônios das células ganglionares, que partem da retina) são encontradas células que respondem no esquema de oposição cromática a estímulos de luz de comprimentos de onda distintos. Nestas células a atividade aumenta para estímulos em uma extremidade do espectro e é inibida (abaixo da atividade espontânea) para estímulos na extremidade oposta. Uma função destes neurônios que apresentam oposição cromática, também encontrados no LGN e no córtex, é a de simplificar o sinal para a distinção de comprimentos de onda diferentes.

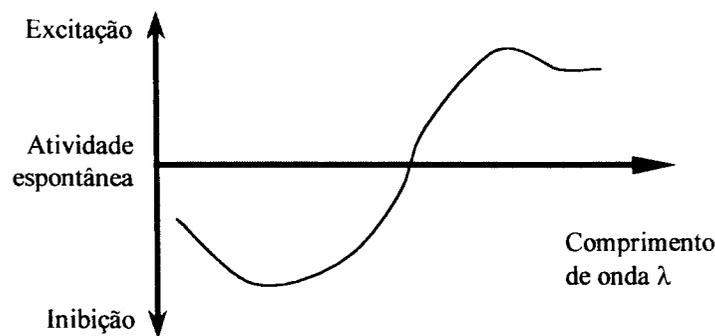


Figura 2.8. Esquema de resposta de neurônio com oposição cromática.

A Figura 2.8 mostra um esquema ilustrativo da resposta de uma célula que apresenta oposição cromática. No exemplo há inibição da atividade espontânea para comprimentos de onda curtos e excitação para comprimentos de onda longos.

Muitas células do córtex apresentam oposição cromática dupla, com campo receptivo formado por uma região central e uma região periférica. O centro é excitado por uma faixa do espectro e inibido por outra, enquanto a periferia mostra comportamento oposto. Assim, por exemplo, o centro é excitado por luz vermelha e inibido por luz verde (R+G-) e a periferia o inverso, ou seja, R-G+ (Figura 2.9). Estas células respondem a contrastes e são diretamente vinculadas à percepção de contraste simultâneo. [Goldstein (1989)]

Então, poder-se-ia postular que a percepção das cores resulta da composição espectral, em termos da energia nos diferentes comprimentos de onda, que atinge a retina. A percepção humana da cor, entretanto, não se dá somente pela análise das proporções das intensidades das respostas dos cones responsáveis por comprimentos de onda curtos, médios e longos. Uma folha de papel azul é percebida com o mesmo azul tanto sob uma iluminação de luz branca (com energia aproximadamente igual em todos os comprimentos de onda) quanto sob iluminação de uma lâmpada de tungstênio (com maior energia em comprimentos de onda longos). A informação de energia nos diferentes comprimentos de onda para as duas situações é diferente e a consequente composição espectral que atinge a retina. Ainda assim, o mesmo azul é percebido\*. Este fenômeno é conhecido como constância cromática.

Goldstein [Goldstein (1989)] observa que um fator importante para a obtenção de constância cromática é a adaptação cromática. Quando o campo de visão é coberto em grande parte por uma mesma cor, ela parece menos saturada. Quando um indivíduo sai de um ambiente de luz branca e move-se para outro com iluminação de lâmpada de tungstênio, a retina se adapta à nova composição espectral, tornando-se menos sensível aos comprimentos de onda mais longos, o que poderia compensar a em parte a nova composição espectral, mantendo a percepção. Outros fatores postos

\* Na realidade, o azul percebido não é exatamente igual, mas a variação é desprezível em termos de percepção.

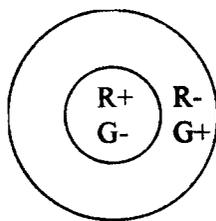


Figura 2.9. Campo receptivo com oposição cromática dupla. O centro responde positivamente à luz vermelha enquanto é inibido pela luz verde. A circunvizinhança, por sua vez, é excitada por verde e inibida por vermelho.

em consideração, também, são as cores na circunvizinhança e a memória de cor associada a um objeto. Neste segundo caso, mesmo sendo montada a partir de um mesmo papel vermelho, a silhueta de uma maçã (associada diretamente à cor vermelha) é percebida levemente mais vermelha que a silhueta de um sino, ao qual o vermelho não é diretamente atribuído\*.

### 2.2.3 Comentários sobre as abordagens tricromática e de oposição cromática

Evidências fisiológicas mostram que tanto o mecanismo tricromático quanto o de oposição cromática coexistem no sistema visual humano [Goldstein (1989)]. Em termos de cores, a captação da luz externa é essencialmente realizada por três filtros cromáticos estabelecidos pelos três tipos de cones citados, promovendo uma primeira codificação ou simplificação do sinal de entrada para o sistema visual. Então, em vez de uma complexa composição espectral (veja-se a Figura 2.5), apenas proporções por três canais distintos são interpretadas. Introduce-se, evidentemente, uma perda no sinal real, caracterizada principalmente pela incapacidade das pessoas em distinguir cores provenientes de um único comprimento de onda de composições de três comprimentos de onda independentes, desde que excitam o mecanismo dos cones da mesma maneira.

Após esta captura de informações externas, as conexões entre os neurônios também introduzem uma codificação das informações no mecanismo de oposição cromática. A Figura 2.10 apresenta um mecanismo esquemático de conexões sinápticas que produziriam células que responderiam de forma oponente aos estímulos luminosos. Y indica o canal amarelo, produzido pela composição aditiva das saídas dos canais vermelho e verde. São produzidos, então, mecanismos que respondem, ao mesmo tempo, positivamente ao azul das ondas de comprimento médio e negativamente a comprimentos médios e longos. Outras formas de oposição podem ser esquematizadas modificando-se os tipos de conexão.

\* Esta percepção não é facilmente distinguida em situações normais. Uma exposição do experimento psicofísico para memória de cor pode ser obtida no livro de Goldstein [Goldstein (1989)], pág. 138.

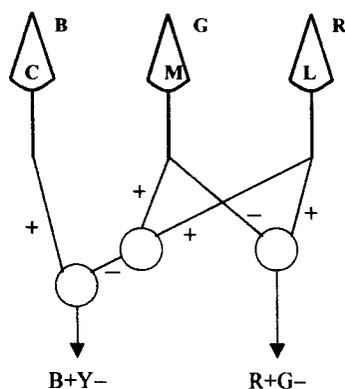


Figura 2.10. Esquema de ligação do mecanismo tricromático para estabelecer um dos tipos de oposição cromática. Símbolos “+” indicam conexão sináptica excitatória, enquanto “-” representam conexões inibitórias.

### 2.3 Representação das cores

O universo de cores que podem ser percebidas é, do ponto de vista do sistema visual dos primatas, inicialmente filtrado por três canais, cujas absorções máximas situam-se nas faixas do azul, do verde e do vermelho. A representação da informação cromática recebida é importante para se estabelecer um modelo da percepção da cor.

Esta seção apresenta, inicialmente, as definições de alguns termos importantes necessários à seção 2.3.2, que trata sobre colorimetria. A seção seqüente descreve sucintamente algumas formas mais comuns de representação de cores.

#### 2.3.1 Nomenclatura básica

Alguns termos são importantes e devem ter uma definição que os ponha no contexto desta seção. Embora estas definições não almejem ser absolutas do ponto de vista científico, um grau de precisão bastante criterioso foi utilizado. Estabelecem-se, assim, as definições da Tabela 2-III para termos sobre cor, iluminação, fotometria e cromaticidade\*.

#### 2.3.2 Colorimetria e emparelhamento de estímulos coloridos

Colorimetria, segundo Wyszecki e Stiles [Wyszecki & Stiles (1982)], é o ramo da ciência que estuda a cor em termos numéricos, permitindo definir um estímulo visual de maneira que: (a) o estímulo tenha a mesma percepção de cor para indivíduos com visão colorida normal, sob as mesmas condições de iluminação (emparelhamento de cor), (b) estímulos que sejam semelhantes tenham as mesmas especificações e (c) os valores que compõem a especificação sejam funções contínuas dos parâmetros físicos que definem a distribuição espectral de um estímulo. Assim, permite-se que sejam estabelecidos valores numéricos que representem os estímulos visuais.

\* O glossário do Apêndice A (pág. 203) apresenta os termos mais importantes desta seção e se mostra mais adequado para consultas rápidas.

Tabela 2-III. Principais termos utilizados em cor e iluminação, com suas definições.

Termo	Definição
Radiância	Medida da potência de fontes de luz relativa à potência emitida a partir de uma fonte (ou superfície). [Fairchild (1998)]
Irradiância	Medida da potência de fontes de luz relativa à potência incidente sobre uma superfície. [Fairchild (1998)]
Luz	Aspecto da energia radiante da qual um observador humano é ciente através das sensações visuais proporcionadas pela estimulação da retina do olho. [Goldstein (1989)]
Cor	Característica da luz visível na qual um observador pode distinguir diferenças conforme varia sua composição espectral [Goldstein (1989)]. Pode ser descrita por nomes cromáticos (amarelo, azul, marrom, rosa, roxo etc.) ou acromáticos (preto, cinza, branco) e qualificados por adjetivos como claro, escuro, brilhante etc. [Fairchild (1998)].
Estímulo de cor (cromático)	Potência radiante de certa magnitude e certa composição espectral que penetra no olho produzindo a sensação de cor. [Goldstein (1989)]
Estímulo monocromático	Estímulo de cor de certa magnitude e certo comprimento de onda que penetra no olho produzindo a sensação de luz ou cor. [Goldstein (1989)]
Estímulo acromático	Estímulo de cor que, nas condições de observação, proporcionam uma sensação desprovida de matiz [Goldstein (1989)]. Refere-se aos estímulos de diferentes intensidades de branco (branco, cinzas e preto).
Matiz	Atributo da sensação visual pela qual uma área se parece similar a uma das cores: azul, amarelo, verde e vermelho, ou a combinações delas. [Fairchild (1998)]
Saturação	Atributo da sensação visual que permite julgar o quanto um estímulo cromático se diferencia de um estímulo acromático [Goldstein (1989)]. Em outras palavras, equivale à quantidade de “branco” que um estímulo cromático possui: o rosa é o vermelho com menor saturação.
Iluminação	Atributo da sensação visual pelo qual um estímulo aparenta maior ou menor intensidade, isto é, parece emitir mais ou menos luz [Goldstein (1989)]. A especificação deste atributo, que independe do matiz e da saturação, varia de “claro” a “escuro”.

A *generalização tricromática* [Wyszecki & Stiles (1982)] postula que, sob condições variadas de observação, vários estímulos coloridos podem ser emparelhados com a mistura aditiva de três estímulos primários prefixados, cujas intensidades possam ser ajustadas apropriadamente (Figura 2.2). Assumindo-se leis de linearidade para as misturas, obtém-se a forma “forte” da generalização tricromática [Wyszecki & Stiles (1982)]. Estas leis incluem:

- a) *Simetria*. Se o estímulo colorido  $a$  se emparelha com o estímulo  $b$ , então  $b$  se emparelha com  $a$ .
- b) *Transitividade*. Se o estímulo  $a$  se emparelha com  $b$  e  $b$  se emparelha com  $c$ , então  $a$  se emparelha com  $c$ .
- c) *Proporcionalidade*. Se o estímulo  $a$  se emparelha com  $b$ , então  $va$  se emparelha com  $vb$ , sendo que  $v$  é um valor positivo que especifica o quanto a intensidade de um estímulo é aumentada ou reduzida.

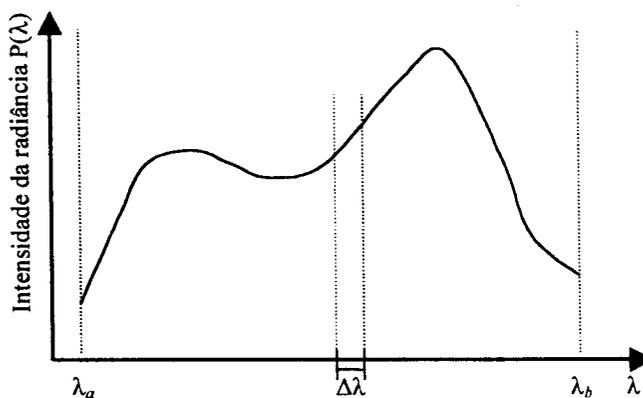


Figura 2.11. Distribuição hipotética das intensidades (potência da radiação) em função do comprimento de onda, para o intervalo visível do espectro  $[\lambda_a, \lambda_b]$ .

- d) *Adição.* Se, para estímulos  $a, b, c$  e  $d$ , os emparelhamentos de  $a$  com  $b, c$  com  $d$  e  $(a + c)$  com  $(b + d)$  valem, então  $(a + d)$  se emparelha com  $(b + c)$ .

As notações  $a, b, c$  e  $d$  são usadas para especificar estímulos coloridos arbitrários e  $(a + b)$  denota a soma aditiva de dois estímulos. A notação  $r, g$  e  $b$  é reservada para estímulos primários específicos utilizados em experimentos de emparelhamento de cores. Cada estímulo é composto por uma distribuição de intensidades de comprimentos de onda, denotada aqui por  $P(\lambda)$ . Nesta notação, um estímulo  $a$  qualquer tem distribuição  $P_a(\lambda)$  (Figura 2.11). Também os estímulos primários  $r, g$  e  $b$  apresentam distribuições  $P_r(\lambda), P_g(\lambda)$  e  $P_b(\lambda)$ , respectivamente. Dentro destas considerações, um estímulo  $q$  pode ser expresso (usando notação vetorial) em função dos estímulos primários pela Equação 2.1, na qual os valores escalares  $R_q, G_q$  e  $B_q$  indicam as intensidades dos respectivos estímulos primários.

$$q = R_q r + G_q g + B_q b \tag{2.1}$$

Em geral, assume-se que  $q$  possa ser obtido usando valores positivos para  $R_q, G_q$  e  $B_q$ . Entretanto, certos emparelhamentos somente podem ser obtidos adicionando-se um dos estímulos primários ao estímulo  $q$  (Equação 2.2), o que pode ser indicado de forma alternativa na Equação 2.3, na qual  $R_q \equiv -R_q^*$ . Para um subconjunto dos estímulos pode ser necessário migrar não somente um, mas dois dos estímulos primários para junto de  $q$ .

$$q + R_q^* r = G_q g + B_q b \tag{2.2}$$

$$\begin{aligned} q &= -R_q^* r + G_q g + B_q b \\ &= R_q r + G_q g + B_q b \end{aligned} \tag{2.3}$$

Um estímulo colorido  $\mathbf{q}$ , com distribuição espectral  $P_q(\lambda)$ , pode ser considerado como a soma aditiva de um conjunto de distribuições em  $n$  pequenas faixas  $\Delta\lambda$ . A Figura 2.11 exemplifica um dos possíveis intervalos  $\Delta\lambda$ . Assim,  $\mathbf{q}$  pode ser escrito como na Equação 2.4, na qual  $\mathbf{q}_i$  (Equação 2.5) representa a distribuição espectral para o intervalo  $\Delta\lambda_i$ .

$$\mathbf{q} = \sum_{i=1}^n \mathbf{q}_i = \left( \sum_{i=1}^n R_{q_i} \right) \mathbf{r} + \left( \sum_{i=1}^n G_{q_i} \right) \mathbf{g} + \left( \sum_{i=1}^n B_{q_i} \right) \mathbf{b} \quad (2.4)$$

$$\mathbf{q}_i = R_{q_i} \mathbf{r} + G_{q_i} \mathbf{g} + B_{q_i} \mathbf{b} \quad (2.5)$$

$$R_q = \sum_{i=1}^n R_{q_i}; \quad G_q = \sum_{i=1}^n G_{q_i}; \quad B_q = \sum_{i=1}^n B_{q_i} \quad (2.6)$$

Utiliza-se a notação  $P(\lambda)\Delta\lambda$  para indicar o estímulo no intervalo  $\Delta\lambda$ , centrado em  $\lambda$ , com intensidade média  $P(\lambda)$ . Quando o intervalo, suposto agora igual para todos os intervalos, tende a zero,  $P(\lambda)d\lambda$  passa a representar o estímulo monocromático de comprimento de onda  $\lambda$ , denotado por  $\mathbf{q}_\lambda$ , ou seja, o valor da radiância para um comprimento de onda  $\lambda$  específico.

Considerando-se um estímulo particular, no qual cada  $P(\lambda)$  seja igual a 1 (Figura 2.12), chamado de estímulo de energia constante  $E(\lambda)$ , pode-se descrever cada estímulo monocromático de comprimento de onda  $\lambda$  na forma da Equação 2.7. As funções  $\bar{r}(\lambda)$ ,  $\bar{g}(\lambda)$  e  $\bar{b}(\lambda)$  indicam os estímulos tricromáticos para  $\mathbf{e}_\lambda$ . A Figura 2.13 mostra os valores destas funções para o espectro visível  $[\lambda_a, \lambda_b]$ , considerando-se um observador com visão colorida normal. Valores negativos indicam que, para se obter o emparelhamento das cores, é necessário adicionar uma das componentes primárias juntamente com o estímulo  $\mathbf{e}_\lambda$ , de forma similar à discussão sobre as Equações 2.2 e 2.3.

$$\mathbf{e}_\lambda = \bar{r}(\lambda)\mathbf{r} + \bar{g}(\lambda)\mathbf{g} + \bar{b}(\lambda)\mathbf{b} \quad (2.7)$$

A Equação 2.8 representa a equação que envolve o emparelhamento de cores para um estímulo monocromático  $\mathbf{q}_\lambda$  com intensidade  $P(\lambda)d\lambda$ , usando a Equação 2.7 para especificá-lo em função dos estímulos primários.

$$\mathbf{q}_\lambda = (P(\lambda)d\lambda)\mathbf{e}_\lambda = (P(\lambda)d\lambda)\bar{r}(\lambda)\mathbf{r} + (P(\lambda)d\lambda)\bar{g}(\lambda)\mathbf{g} + (P(\lambda)d\lambda)\bar{b}(\lambda)\mathbf{b} \quad (2.8)$$

Assumindo-se que  $P(\lambda)$  seja uma função contínua no espectro visível, a Equação 2.8 é integrável, de forma que um estímulo complexo  $\mathbf{q}$  (conforme a Equação 2.1) pode ser escrito na forma da Equação 2.9.

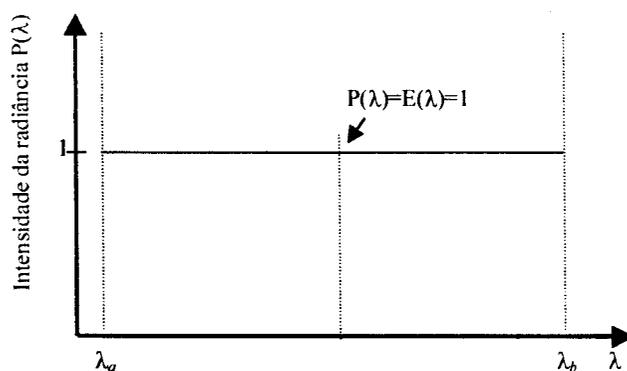


Figura 2.12. Distribuição de energia constante  $E(\lambda)$ , na qual  $E(\lambda) = 1$  para cada  $\lambda$  monocromático.

$$\mathbf{q} = \left( \int_{\lambda_a}^{\lambda_b} P(\lambda) \bar{r}(\lambda) d\lambda \right) \mathbf{r} + \left( \int_{\lambda_a}^{\lambda_b} P(\lambda) \bar{g}(\lambda) d\lambda \right) \mathbf{g} + \left( \int_{\lambda_a}^{\lambda_b} P(\lambda) \bar{b}(\lambda) d\lambda \right) \mathbf{b} \quad (2.9)$$

Desta forma, dados dois estímulos coloridos  $\mathbf{a}$  e  $\mathbf{b}$  definidos pelas distribuições de potência radiante  $P_a(\lambda)$  e  $P_b(\lambda)$ , respectivamente, pode-se afirmar que estarão emparelhados quando as expressões da Equação 2.10 forem satisfeitas simultaneamente.

$$\begin{cases} \int_{\lambda_a}^{\lambda_b} P_a(\lambda) \bar{r}(\lambda) d\lambda = \int_{\lambda_a}^{\lambda_b} P_b(\lambda) \bar{r}(\lambda) d\lambda \\ \int_{\lambda_a}^{\lambda_b} P_a(\lambda) \bar{g}(\lambda) d\lambda = \int_{\lambda_a}^{\lambda_b} P_b(\lambda) \bar{g}(\lambda) d\lambda \\ \int_{\lambda_a}^{\lambda_b} P_a(\lambda) \bar{b}(\lambda) d\lambda = \int_{\lambda_a}^{\lambda_b} P_b(\lambda) \bar{b}(\lambda) d\lambda \end{cases} \quad (2.10)$$

Considera-se que dois estímulos estão emparelhados em termos de cor, portanto, quando todas suas componentes em relação a três estímulos primários forem as mesmas.

### 2.3.3 Espaços de cores

Em termos computacionais, as representações numéricas das cores acabam sendo, ao invés de funções contínuas, funções discretas que quantizam os valores das componentes primárias, usualmente em 256 valores distintos.

A seguir, estão representados, sinteticamente, os principais espaços de cores que quantificam os estímulos cromáticos, permitindo verificar a vasta gama de representações, as quais podem ser convertidas entre si. As descrições dos espaços foram compiladas a partir do material apresentado por Ohta [Ohta *et al.* (1980)], Wyszecki e Stiles [Wyszecki & Stiles (1982)], Pratt [Pratt (1991)], Gonzales e Woods [Gonzales & Woods (1992)] e Castleman [Castleman (1996)].

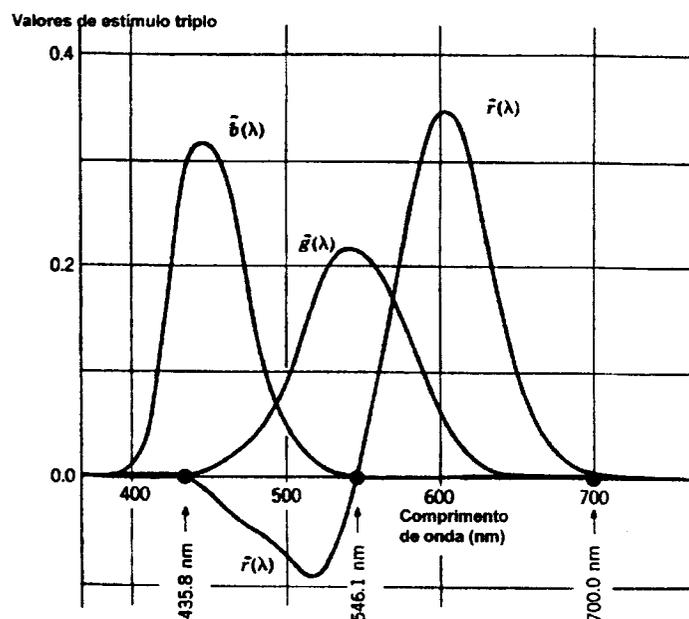


Figura 2.13. Valores dos estímulos tricromáticos para o estímulo monocromático  $e_\lambda$ , de intensidade unitária com respeito aos estímulos de cores primárias  $r$ ,  $g$  e  $b$  de comprimentos de onda de 700, 546,1 e 435,8 nm, respectivamente. [Wyszecki & Stiles (1982)]

### 2.3.3.1 RGB

O sistema RGB é o mais direto quando tomado sob o ponto de vista da visão tricromática. Consiste na combinação de cores primárias com comprimentos de onda de 700 (vermelho), 546,1 (verde) e 435,8 nm (azul). Os valores correspondem aos valores médios obtidos indiretamente por experimentos usando emparelhamento de cores em um grande número de observadores, estando sujeitos às condições estabelecidas pelo experimento. Esta resposta geral obtida corresponde ao que se convencionou chamar de *observador padrão* C.I.E.\* Este sistema, baseado nos estímulos  $r$ ,  $g$  e  $b$ , é referenciado como  $R_C G_C B_C$ .

Atualmente a nomenclatura RGB refere-se ao padrão utilizado para monitores de vídeo de raios catódicos, entre outros, nos quais as cores primárias utilizadas são escaladas para assumirem valores nos mesmos intervalos†. A obtenção dos valores a partir do  $R_C G_C B_C$  original pode ser feita através da Equação 2.11.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 0,842 & 0,156 & 0,091 \\ -0,129 & 1,320 & -0,203 \\ 0,008 & -0,069 & 0,897 \end{bmatrix} \begin{bmatrix} R_C \\ G_C \\ B_C \end{bmatrix} \quad (2.11)$$

\* C.I.E.: *Commission Internationale de l'Eclairage*.

† Geralmente são utilizados o intervalo contínuo  $[0, 1]$  ou o discreto  $[0, 255]$ .

### 2.3.3.2 YIQ

O sistema YIQ é um sistema projetado para transmissões de televisão nos Estados Unidos, coordenado pelo N.T.S.C.\* A componente  $Y$  representa diretamente a iluminação, enquanto as outras componentes,  $I$  e  $Q$ , representam juntas o matiz e a saturação. A possibilidade de obter diretamente das câmeras o sinal codificado em  $Y$ ,  $I$  e  $Q$  é bastante interessante, visto que  $Y$  individualmente pode ser utilizado em recepções em preto e branco. Além disto, é possível restringir bastante as bandas dos sinais  $I$  e  $Q$ , o que provoca degradação quase imperceptível na imagem. A correspondência entre o sistema RGB e o YIQ é estabelecida pela Equação 2.12.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ 0,596 & -0,274 & -0,322 \\ 0,211 & -0,523 & 0,312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.12)$$

### 2.3.3.3 XYZ

As representações dos sistemas C.I.E. e N.T.S.C. com três estímulos primários requerem, às vezes, que um dos valores seja negativo para que se obtenha um emparelhamento adequado. As primárias  $X$ ,  $Y$  e  $Z$  não são reais, isto é, não correspondem a estímulos que possam ser concebidos na prática, mas o sistema artificial define que os valores necessários para o emparelhamento sejam todos positivos. Todos os estímulos reais, entretanto, podem ser representados no XYZ.

O sistema é estabelecido de forma que a componente  $Y$  represente a intensidade da cor a ser emparelhada. As componentes  $X$  e  $Z$  representam a informação das cores. Os sistemas XYZ e RGB são relacionados pela transformação da Equação 2.13.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,607 & 0,174 & 0,200 \\ 0,299 & 0,587 & 0,114 \\ 0,000 & 0,066 & 1,116 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.13)$$

### 2.3.3.4 YUV

O espaço YUV foi a proposta inicial do N.T.S.C. para uso em transmissão de sinais de televisão, abandonado posteriormente em favor do YIQ. A componente  $Y$  representa a mesma informação de luminância, enquanto  $U$  e  $V$  relacionam-se a  $I$  e  $Q$  através das Equações 2.14 e 2.15, nas quais  $\theta$  especifica uma rotação de  $33^\circ$ . A relação entre YUV e RGB é apresentada pela Equação 2.16.

$$I = -U \sin \theta + V \cos \theta \quad (2.14)$$

\* N.T.S.C.: *National Television Systems Committee.*

$$Q = U \sin \theta + V \cos \theta \quad (2.15)$$

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,148 & -0,289 & 0,437 \\ 0,615 & -0,515 & -0,100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.16)$$

### 2.3.3.5 HSI e HSV

Os sistemas HSI e HSV representam, em seus sistemas de coordenadas, medidas relativas à intensidade, matiz e saturação de uma cor. Os valores  $H$ ,  $S$  e  $I$  podem ser obtidos diretamente dos valores  $R$ ,  $G$  e  $B$ , como apresentado nas Equações 2.17 a 2.19.

$$\begin{bmatrix} I \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{6} & -2/\sqrt{6} & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.17)$$

$$H = \arctan \frac{V_2}{V_1} \quad (2.18)$$

$$S = \sqrt{V_1^2 + V_2^2} \quad (2.19)$$

Segundo esta definição, a cor azul corresponde à referência inicial (zero) para o matiz.

### 2.3.3.6 $L^*u^*v^*$

Um problema inerente a representações de cores é o estabelecimento de uma medida de diferença entre dois estímulos, quantificando, para um dado par de estímulos que não se emparelham, uma medida de quanto seria necessário para que representassem a mesma informação perceptiva. O sistema  $L^*u^*v^*$  é um formato de representação de cores no qual modificações lineares em uma das componentes representam alterações em intensidade similar no atributo perceptivo sendo modificado.

A obtenção deste sistema a partir do XYZ é apresentado pelas Equações 2.20 a 2.22 [Wyszecki & Stiles (1982)], sendo que as últimas utilizam-se das relações da Equação 2.23. Os valores indexados por zero indicam o estímulo correspondente à cor branca no sistema XYZ, usado para calcular, também,  $u_0'$  e  $v_0'$ .

$$L^* = \begin{cases} 116\sqrt[3]{\frac{Y}{Y_0}} - 16, & Y/Y_0 > 0,008856 \\ 903,3\frac{Y}{Y_0}, & \text{caso contrário} \end{cases} \quad (2.20)$$

$$u^* = 13L^*(u' - u_0') \quad (2.21)$$

$$v^* = 13L^*(v' - v_0') \quad (2.22)$$

$$\begin{cases} u' = \frac{4X}{X + 15Y + 3Z} \\ v' = \frac{9Y}{X + 15Y + 3Z} \end{cases} \quad (2.23)$$

### 2.3.3.7 $L^*a^*b^*$

$L^*a^*b^*$  é o segundo sistema da C.I.E. estruturado para ser uniforme em relação a uma medida de distância entre estímulos. O obtenção dos valores  $L^*$ ,  $a^*$  e  $b^*$  a partir do sistema XYZ é apresentado nas Equações 2.24 a 2.26. [Pratt (1991)]

$$L^* = 116\sqrt[3]{\frac{Y}{Y_0}} - 16 \quad (2.24)$$

$$a^* = 500 \left( \sqrt[3]{\frac{X}{X_0}} - \sqrt[3]{\frac{Y}{Y_0}} \right) \quad (2.25)$$

$$b^* = 200 \left( \sqrt[3]{\frac{X}{X_0}} - \sqrt[3]{\frac{Z}{Z_0}} \right) \quad (2.26)$$

### 2.3.3.8 $I_1I_2I_3$ (ou espaço de Ohta)

Em 1980, Ohta, Kanade e Sakai [Ohta *et al.* (1980)] propuseram um método de segmentação por regiões através de divisões sucessivas dos dados da imagem\*. Dentro das investigações sobre a aplicação da segmentação, vários espaços de representação de cores foram definidos através de combinações lineares entre as componentes do espaço RGB. Alguns destes espaços, em particular o espaço  $I_1I_2I_3$ , acaba-

\* A segmentação é similar à desenvolvida por Ohlander *et al.* [Ohlander *et al.* (1978)], apresentada no início da seção 3.2.2.

ram por se mostrar mais adequados [Ohta *et al.* (1980)] para a aplicação da metodologia de segmentação para um dado conjunto de imagens. As novas componentes “primárias”  $I_1$ ,  $I_2$ , e  $I_3$  podem ser obtidas a partir de  $R$ ,  $G$  e  $B$  conforme a Equação 2.27.

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & -1/2 & 0 \\ 1/2 & -1/4 & -1/4 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.27)$$

## 2.4 Comentários finais

Este capítulo descreve os principais conceitos básicos envolvidos na visão em cores, especialmente em relação aos mecanismos presentes na visão de primatas, que são considerados possuidores do mais sofisticado sistema de reconhecimento de padrões entre os animais [Costa *et al.* (1994)]. São cobertos, além de um pequeno resumo histórico, os mecanismos dos cones, que permitiram estabelecer a teoria tricromática, e os de oposição cromática, os quais permitem conhecer melhor atributos visuais como contraste simultâneo e constância cromática.

A colorimetria, como área que estuda as medidas quantitativas da visão, é também introduzida, permitindo verificar, sob um ponto de vista de um modelo mais exato (matemático), as representações da visão em cores e suas representações. Também os principais espaços de representação de cores são sucintamente descritos.

O conhecimento destes conceitos permite, em suma, não somente uma melhor compreensão da visão, mas também proporciona elementos para serem aplicados em sistemas artificiais de visão, através do entendimento de seus mecanismos na busca de soluções em problemas de segmentação e reconhecimento, por exemplo, além da representação e codificação das informações visuais.

*“Todas as coisas já foram ditas, mas como  
ninguém escuta é preciso sempre recomeçar.”*  
André Gide

### 3 Segmentação de imagens coloridas: uma revisão

Até os dias atuais, as técnicas para a segmentação de imagens compreendem, em sua grande maioria, o tratamento de imagens em níveis de cinza. A utilização da informação de cor, agora disponível de maneira mais difundida, era bastante modesta nos trabalhos iniciais em segmentação e se apresenta cada vez mais crescente nestes últimos anos. [Pal & Pal (1993); Skarbek & Koschan (1994)]

Imagens coloridas comportam, em princípio, maior quantidade de informação que suas versões monocromáticas, e esta fonte de informação pode ser explorada para a obtenção de melhores segmentações. O sistema visual humano, por exemplo, pode perceber divisões do espectro visível com cores distintas (dos 380 aos 700nm) em aproximadamente 150 passos. Variando as intensidades destas ondas e também sua saturação, podem ser discriminadas mais de 7 milhões de cores diferentes. [Goldstein (1989)]

De forma análoga à vasta literatura sobre segmentação de imagens monocromáticas [Pal & Pal (1993)], a qual divide, *grosso modo*, as técnicas em duas grandes vertentes\*, também o tratamento de imagens coloridas pode ser tratado em dois grupos principais. O primeiro é baseado na detecção de descontinuidades (bordas) no espaço de representação cromática, com detecção de bordas usando, por exemplo, gradientes. O segundo grupo se baseia na identificação de continuidade neste espaço pela determinação de regiões homogêneas, podendo ser caracterizado, basicamente, por um processo de classificação dos pontos da imagem em classes ou grupos.

Este capítulo apresenta as principais técnicas nestas duas categorias, bordas e regiões, procurando caracterizá-las através de exemplos específicos. Técnicas baseadas em bordas são apresentadas na seção 3.1 e as baseadas em regiões na seção 3.2.

---

\* Segmentação baseada em bordas e baseada em regiões. [Gonzales & Woods (1992)]

Além destas grandes vertentes, são expostas, em separado, abordagens baseadas em aprendizado competitivo e redes neurais (seção 3.3), além de técnicas que interpretam a formação da cena segundo um modelo físico de propagação da luz (seção 3.4). Comentários gerais sobre a revisão e as técnicas são apresentados na seção 3.5.

### 3.1 Técnicas baseadas em bordas

No caso monocromático, as bordas caracterizam-se pela descontinuidade na intensidade do sinal. Quando a imagem possui mais de uma banda (componente), a descontinuidade a ser considerada para a determinação das bordas são atributos neste espaço vetorial, no qual a cada ponto da imagem está associado um valor vetorial com três componentes. A utilização da informação das bandas pode ser feita através de várias maneiras, sendo que se optou, neste texto, por sua divisão segundo três pontos de vista com características próprias. A caracterização mais específica e exemplos são apresentados na seqüência (seções 3.1.1 a 3.1.3), após uma caracterização sumária.

As três abordagens compreendem, então:

- a) *Definição de uma métrica ou redução no espaço de cor.* Um exemplo é a utilização de somente uma das componentes cromáticas ou o uso da distância entre as cores em um dado espaço de representação. Esta abordagem caracteriza-se pela redução do espaço vetorial para o espaço escalar, sendo as bordas detectadas por descontinuidades neste espaço reduzido. A consequência é que há, assim, uma redução para um espaço escalar, ou seja, deixa-se de realizar a manipulação das bandas passando-se a operar sobre um espaço equivalente a uma representação monocromática.
- b) *Realização do processamento separadamente em cada banda, unindo os resultados parciais.* Esta abordagem, embora se apresente como uma solução natural para tratar o problema, padece da indefinição de como as bordas encontradas isoladamente podem ser combinadas no final, não sendo esta combinação trivial. Alternativamente é possível que uma técnica para a detecção de bordas para o caso monocromático seja modificada para que, atuando de forma direta sobre cada banda, obedeça a uma limitação imposta a todas as bandas simultaneamente que restrinja a caracterização de um ponto como sendo um ponto de borda.
- c) *Consideração do espaço de cor como um todo, agindo nas bandas simultaneamente.* Esta última forma de abordagem é mais complexa e completa do ponto de vista conceitual, baseando-se na localização de descontinuidades no espaço vetorial definido pelas cores. Pode ser desenvolvido, por exemplo, um gradiente vetorial, o qual pode identificar as variações existentes neste espaço, atuando de modo conjunto nas diversas bandas.

#### 3.1.1 Detecção de bordas por métrica ou redução no espaço de cor

Entre as três diferentes abordagens para detecção de bordas apresentadas, a primeira é a mais direta. Uma forma simplificada de considerar uma borda cromática é

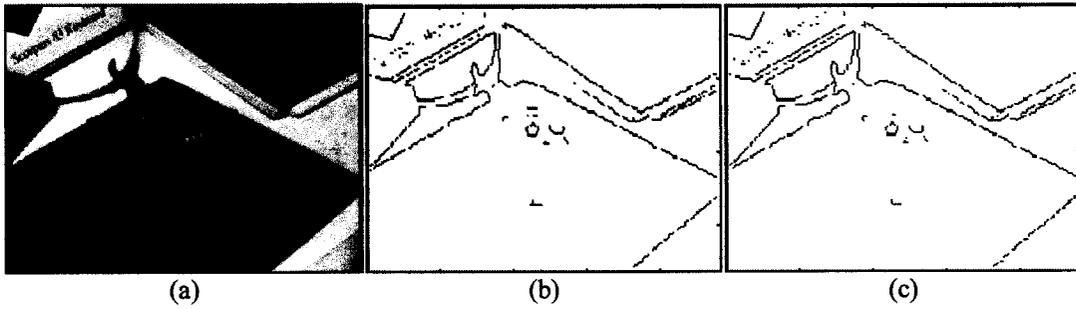


Figura 3.1. Resultados da detecção de bordas pelo operador de Prewitt. (a) Imagem original. (b) Aplicação à versão monocromática da imagem de testes; (c) Aplicação ao módulo do vetor (R,G,B) calculado ponto a ponto sobre a versão original. Notam-se diferenças mínimas, sendo os resultados praticamente os mesmos.

realizar a detecção de bordas somente na componente relacionada à luminância, desconsiderando matiz e saturação. Uma alternativa considera, por exemplo, a determinação dos módulos dos vetores de cor em cada ponto da imagem e posterior determinação das discontinuidades sobre estes módulos. Sob outro ponto de vista, corresponderia a uma determinação de um valor escalar equivalente à intensidade luminosa de cada ponto, de forma que não podem ser esperados resultados muito superiores aos da detecção de bordas na versão monocromática da imagem\*. A Figura 3.1 exemplifica o resultado obtido pela detecção de bordas com o operador de Prewitt [Gonzales & Woods (1992)] (cujas máscaras de convolução são apresentadas na Equação 3.1), respectivamente, para a versão monocromática e a versão obtida pelo módulo do vetor determinado pelas componentes  $R$ ,  $G$  e  $B$  em cada posição da imagem. Na Figura 3.1(b) está representado o resultado para a versão monocromática da imagem, ou seja, de sua componente de intensidade. O mapa de bordas da Figura 3.1(c) resulta não da intensidade, mas considerando-se cada posição da imagem como um vetor tridimensional, cujas coordenadas são especificadas pelas componentes de cor, sendo o gradiente aplicado sobre o módulo (comprimento) dos vetores. Os resultados são praticamente os mesmos, não sendo possível indicar um como melhor que o outro. O número de pontos diferentes entre as duas imagens é de  $185^\dagger$ , contendo cada imagem 180 linhas e 240 colunas. O número total de pontos de borda na Figura 3.1(b) é de 1305, enquanto na Figura 3.1(c) é de 1280.

$$P_h = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, P_v = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad (3.1)$$

\* Em termos estritos, a detecção é realmente feita uma sobre a versão monocromática da imagem, mas não necessariamente a usual (transformação para níveis de cinza segundo intensidades perceptivas).

† O valor se refere não à diferença entre o número de pontos de borda considerados, mas à quantidade de pontos presentes em um mapa de bordas e ausentes no outro, posição a posição.

### 3.1.2 Detecção de bordas através do tratamento isolado das componentes cromáticas

A segmentação por detecção de bordas em imagens de várias bandas pode ser considerada através da aplicação de operadores projetados para o caso monocromático em cada banda separadamente. O problema imediato é a obtenção de um número de mapas de bordas equivalente ao número de bandas da imagem. A combinação destes resultados não se apresenta como trivial, uma vez que não é simples determinar se há algum sentido inerente na combinação destas bordas individuais. Poder-se-iam considerar como pontos de borda, ao final, todos as posições identificadas em qualquer das bandas (realizando-se um *OU* com os mapas) ou somente aquelas presentes em todas as bandas (E dos mapas de bordas), além de outras alternativas que também poderiam ser consideradas. Assim, uma face desta abordagem [Pratt (1991)] é a consideração da presença de uma borda, por exemplo, quando houver uma descontinuidade em qualquer uma das bandas individuais.

O obtenção de mapas de bordas é possível através da combinação das intensidades dos gradientes monocromáticos. Esta combinação pode ser feita, por exemplo, através da soma simples das amplitudes (Equação 3.2) ou pela obtenção do módulo do gradiente resultante (Equação 3.3) [Pratt (1991)], ou ainda obtendo-se um gradiente resultante pela seleção do maior valor entre os gradientes individuais (Equação 3.4) [Robinson (1977)]. Nestas equações  $G_s(i, j)$ ,  $G_m(i, j)$  e  $G_M(i, j)$  indicam, respectivamente, as intensidades do gradiente resultante para cada alternativa de combinação, enquanto  $G_k(i, j)$  denota a intensidade do gradiente para cada banda  $k^*$ . O mapa de bordas pode, então, ser obtido estabelecendo-se um limiar apropriado para o gradiente resultante. Intensidades altas para o gradiente estão diretamente relacionadas à presença de descontinuidades.

$$G_s(i, j) = G_1(i, j) + G_2(i, j) + G_3(i, j) \quad (3.2)$$

$$G_m(i, j) = \sqrt{G_1^2(i, j) + G_2^2(i, j) + G_3^2(i, j)} \quad (3.3)$$

$$G_M(i, j) = \max_{k=1,2,3} \{G_k(i, j)\} \quad (3.4)$$

Uma forma alternativa de combinação de operadores é a proposta por Shiozaki [Shiozaki (1986)] com o operador de entropia, que é um filtro espacial não-linear. Considerando-se  $a_1, a_2, \dots, a_n$  como as intensidades dos  $n$  pontos em uma vizinhança com centro  $(i, j)$ , a entropia da intensidade é dada pela Equação 3.5, com valores de  $p_i$  dados pela Equação 3.6.

---

\* As bandas usadas para o cálculo dos gradientes podem ser quaisquer, desde o espaço RGB até  $L^*a^*b^*$  ou XYZ.

$$H = -\sum_{k=1}^n \frac{p_k \log p_k}{\log(n+1)} \quad (3.5)$$

$$p_i = \frac{a_i}{\sum_{k=1}^n a_k} \quad (3.6)$$

Para imagens coloridas, o conceito de entropia é aplicado definindo-se a entropia de cada componente  $H_R$ ,  $H_G$  e  $H_B$ . Denotando-se  $a_R$ ,  $a_G$  e  $a_B$  como as intensidades de cada banda na posição  $(i, j)$ , a entropia geral é definida pela expressão da Equação 3.7, na qual  $q_k = a_k / (a_R + a_G + a_B)$ , para  $k = R, G, B$ .

$$H = q_R H_R + q_G H_G + q_B H_B \quad (3.7)$$

Assim como para o caso monocromático, a entropia definida é invariante à rotação\* e variações abruptas nas intensidades resultam em valores baixos para  $H$ , enquanto regiões homogêneas têm entropia alta. Associando-se variações locais (de baixa entropia) a bordas, pode-se obter um mapa de bordas pela definição de um limiar para a imagem filtrada pelo operador de entropia. O resultado é atraente, mas a sensibilidade a ruído é grande.

É possível, ainda, processar cada banda como um conjunto independente de dados, mas impondo uma condição de uniformidade para guiar de forma simultânea a detecção das bordas. Na abordagem de Nevatia [Nevatia (1977)] é proposta a definição de um detector de bordas baseado no operador de Hueckel [Hueckel (1973)], estendendo-o para o caso de cores. O espaço de cores utilizado é obtido a partir das componentes RGB através das expressões da Equação 3.8, na qual  $r$  e  $g$  são as componentes cromáticas e  $Y$  é a intensidade, ponderada por constantes apropriadas.

$$\begin{aligned} r &= \frac{R}{R+G+B} \\ g &= \frac{G}{R+G+B} \\ Y &= c_R R + c_G G + c_B B \end{aligned} \quad (3.8)$$

O princípio do operador de Hueckel é ajustar, em uma vizinhança circular, uma borda ideal sobre a imagem, determinando a posição e a orientação da borda, além de sua amplitude. O ajuste da borda ideal ao sinal é feito através da expansão do sinal e da borda nos 9 primeiros termos de uma série de funções ortogonais. Para cada

---

\* A posição dos pontos não interfere no cálculo da entropia, que depende unicamente dos valores das posições vizinhas.

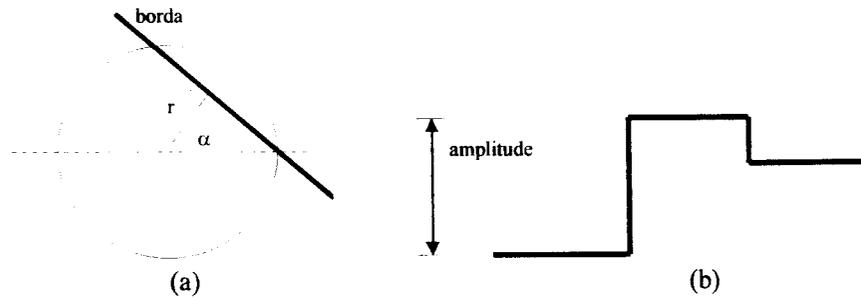


Figura 3.2. Borda ideal de Hueckel. (a) Posição e orientação da borda; (b) Perfil da luminância.

termo  $i$  ( $0 \leq i \leq 8$ ), são calculados os coeficientes da expansão  $a_i$  e  $s_i$ , da borda ideal e do sinal, respectivamente. Os parâmetros da borda ideal podem ser encontrados minimizando-se a Equação 3.9. O resultado da minimização é derivado em termos da direção  $\alpha$  (Figura 3.2).

$$N^2 = \sum_{i=0}^8 (a_i - s_i)^2 \quad (3.9)$$

A presença de uma borda depende da amplitude da borda ideal estar além de um limiar e o ajuste ser bom o suficiente (valor baixo de  $N$  em relação à amplitude).

A extensão para o caso colorido consiste em realizar a aplicação do operador em cada banda isoladamente, acrescentando uma restrição simultaneamente às três componentes. Cada componente cromática pode ser considerada independente e a restrição impõe que a orientação seja a mesma em todas elas em cada ponto  $(i, j)$ .

A borda ideal tem três componentes  $S_1$ ,  $S_2$  e  $S_3$  nas bandas de cor. O ajuste para cada banda  $j$  ( $1 \leq j \leq 3$ ) pode ser dado pela Equação 3.10, com  $a_{ij}$  e  $s_{ij}$  sendo os coeficientes das expansões, respectivamente, do sinal da borda ideal em cada componente. A borda ideal ótima é então definida de forma que  $N^2 = N_1^2 + N_2^2 + N_3^2$  seja mínimo. Minimizar  $N$  equivale, neste caso, a minimizar cada  $N_j$  separadamente.

$$N_j^2 = \sum_{i=0}^8 (a_{ij} - s_{ij})^2 \quad (3.10)$$

O uso mais direto do gradiente para a detecção de bordas é através do estabelecimento de um limiar para a amplitude do gradiente para a geração do mapa de bordas. A escolha do limiar apropriado traz, entretanto, grande dificuldade. Ultré e Macaire [Ultré & Macaire (1996)] propõem, assim, o uso de um algoritmo de relaxação para enfatizar as probabilidades de bordas no intuito de auxiliar no processo de segmentação.

A amplitude do gradiente é obtida para cada banda isoladamente. Quanto maior o valor da amplitude em uma posição, maior a probabilidade desta posição pertencer à classe relativa ao atributo borda. São definidas quatro classes para a relaxação: uma para as bordas (relativas às amplitudes) de cada uma das três componentes es-

pectrais da imagem e uma para a classe não-borda. A relaxação é realizada através de iterações sucessivas até que, para cada posição da imagem, a probabilidade de não-borda tenda ou para zero ou para um. A segmentação final (mapa de bordas) considera como pontos de borda as posições cuja probabilidade de pertencer à classe não-borda esteja próxima a zero. A probabilidade da classe não-borda é calculada em função das probabilidades das outras três classes, dado que a probabilidade total tem que somar 1.

O processo de relaxação, entretanto, por exigir muitas iterações, acaba sendo caro em termos computacionais. Além disso, como o cálculo das iterações usa medidas locais, o algoritmo tende a unir bordas próximas.

### 3.1.3 Detecção de bordas no espaço vetorial das cores

A consideração do espaço de cor como uma única entidade vetorial, ou seja, com a imagem bidimensional assumindo valores múltiplos (três componentes), permite o tratamento através da extensão de operadores simples, como o operador de Prewitt e de Sobel (máscaras de convolução na Equação 3.11).

$$\mathbf{S}_h = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad \mathbf{S}_v = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (3.11)$$

Uma extensão específica do operador de Sobel foi proposta por Yang e Halme [Yang & Halme (1994)], que definem a diferença entre dois pontos coloridos considerando o espaço de cor HSI. Através das projeções dos pontos no plano cromático é possível definir a diferença cromática pura  $d_c$  entre duas cores, além da diferença simples de intensidade  $d_i$ . A diferença cromática absoluta é, então, definida conforme a Equação 3.12, na qual os fatores  $p_i$  e  $p_c$  são pesos arbitrários para as partes luminância e cromática, respectivamente. Tem-se  $p_i > 0$  e  $p_c > 0$ , além de  $p_i + p_c = 1$ .

$$|DIF(i, j)| = \sqrt{p_i d_i^2(i, j) + p_c d_c^2(i, j)} \quad (3.12)$$

Para se estabelecer o operador, define-se, na direção vertical, a diferença entre os pontos  $(i-1, j)$  e  $(i+1, j)$  como  $DIF_v(j)$ , dada pela Equação 3.13.

$$DIF_v(j) = \begin{cases} |DIF_v(j)| & \text{se } I(i+1, j) > I(i-1, j) \\ -|DIF_v(j)| & \text{caso contrário} \end{cases} \quad (3.13)$$

Considerando-se uma janela de dimensão  $3 \times 3$ , define-se o operador de borda vertical pela Equação 3.14 e, por um procedimento análogo, o operador horizontal da Equação 3.15. O resultado do operador de borda final, composição das componentes vertical e horizontal, é dado pela Equação 3.16.

$$B_v(i, j) = DIF_v(j-1) + 2DIF_v(j) + DIF_v(j+1) \quad (3.14)$$

$$B_h(i, j) = DIF_h(i-1) + 2DIF_h(i) + DIF_h(i+1) \quad (3.15)$$

$$B(i, j) = |B_v(i, j)| + |B_h(i, j)| \quad (3.16)$$

O operador  $B(i, j)$  se mostra sensível a alterações de matiz e saturação, mesmo que a intensidade sobre uma região se mostre bastante constante.

Scharcanski e Venetsanopoulos [Scharcanski & Venetsanopoulos (1997)], considerando o espaço RGB, propõem operadores horizontal e vertical similares aos de Prewitt e Sobel, como descritos na Equação 3.17. Os operadores propostos são operadores de convolução e produzem como resultado vetores que são médias locais das cores. Os símbolos  $H_-$  e  $H_+$  indicam submatrizes (de dimensão  $(2w+1) \times w$ ) com valores -1 e 1, respectivamente, enquanto  $\theta$  indica uma coluna de  $2w+1$  zeros, de forma que  $\Delta H$  tenha dimensão  $(2w+1) \times (2w+1)$ . O mesmo se aplica a  $\Delta V$ , considerando a direção vertical. O valor  $w$  é um número inteiro determina a dimensão final das matrizes  $\Delta H$  e  $\Delta V$  (Figura 3.3).

$$\Delta H = [H_- \quad \theta \quad H_+] \quad \text{e} \quad \Delta V = \begin{bmatrix} V_- \\ \theta \\ V_+ \end{bmatrix} \quad (3.17)$$

Assim, para um ponto  $(i, j)$  da imagem, que assume valor (vetorial)  $c(i, j)$ , as expressões na Equação 3.18 fornecem as cores médias para cada parte dos operadores.

$$\begin{aligned} h_+(i, j) &= \frac{1}{2(2w+1)} \sum_{k=j-w}^{j+w} \sum_{l=i+1}^{i+w} c(l, k) \\ h_-(i, j) &= \frac{1}{2(2w+1)} \sum_{k=j-w}^{j+w} \sum_{l=i-1}^{i-w} c(l, k) \\ v_+(i, j) &= \frac{1}{2(2w+1)} \sum_{k=j+1}^{j+w} \sum_{l=i-w}^{i+w} c(l, k) \\ v_-(i, j) &= \frac{1}{2(2w+1)} \sum_{k=j-1}^{j-w} \sum_{l=i-w}^{i+w} c(l, k) \end{aligned} \quad (3.18)$$

As medidas  $\Delta h(i, j) = h_+(i, j) - h_-(i, j)$  e  $\Delta v(i, j) = v_+(i, j) - v_-(i, j)$  estimam a variação local nas direções horizontal e vertical, respectivamente. As grandezas  $\|\Delta h(i, j)\|$  e  $\|\Delta v(i, j)\|$  dão a taxa de variação nas direções ortogonais, ou seja,

$$\begin{array}{|cccc|c|cccc}
 \hline
 \eta_{-}(1,1) & \eta_{-}(1,2) & \cdots & \eta_{-}(1,w) & 0 & \eta_{+}(1,1) & \eta_{+}(1,2) & \cdots & \eta_{+}(1,w) \\
 \eta_{-}(2,1) & \eta_{-}(2,2) & & \eta_{-}(2,w) & 0 & \eta_{+}(2,1) & \eta_{+}(2,2) & & \eta_{+}(2,w) \\
 \vdots & & \ddots & \vdots & \vdots & \vdots & & \ddots & \vdots \\
 \eta_{-}(2w+1,1) & \eta_{-}(2w+1,2) & \cdots & \eta_{-}(2w+1,w) & 0 & \eta_{+}(2w+1,1) & \eta_{+}(2w+1,2) & \cdots & \eta_{+}(2w+1,w) \\
 \hline
 \end{array}$$

Submatriz  $\Delta H_{-}$ 
Submatriz  $\theta$ 
Submatriz  $\Delta H_{+}$

(a)

$$\begin{array}{|ccccc|ccccc}
 \hline
 -1 & -1 & 0 & 1 & 1 & -1 & -1 & -1 & -1 & -1 \\
 -1 & -1 & 0 & 1 & 1 & -1 & -1 & -1 & -1 & -1 \\
 -1 & -1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
 -1 & -1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
 -1 & -1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
 \hline
 \end{array}$$

(b)

Figura 3.3. Representação das matrizes e submatrizes dos operadores de Scharcanski e Venetsanopoulos [Scharcanski & Venetsanopoulos (1997)].  $H_{\pm} = \eta_{\pm}(i, j)$ . (a) Matriz genérica para a direção horizontal; (b) Exemplo para uma matriz  $5 \times 5$  ( $w=2$ ), direções horizontal e vertical, respectivamente.

uma medida do contraste local. A magnitude  $b(i, j)$  que dá a variação máxima em  $(i, j)$  e a direção desta variação são estimadas pelas Equações 3.19 e 3.20.

$$b(i, j) = \sqrt{\|\Delta h(i, j)\|^2 + \|\Delta v(i, j)\|^2} \quad (3.19)$$

$$\theta(i, j) = \arctan \frac{\Delta v'(i, j)}{\Delta h'(i, j)} \quad (3.20)$$

$$\Delta v' = \begin{cases} \|\Delta v(i, j)\| & \text{se } \|\Delta v_{+}(i, j)\| \geq \|\Delta v_{-}(i, j)\| \\ -\|\Delta v(i, j)\| & \text{caso contrário} \end{cases} \quad (3.21)$$

$$\Delta h' = \begin{cases} \|\Delta h(i, j)\| & \text{se } \|\Delta h_{+}(i, j)\| \geq \|\Delta h_{-}(i, j)\| \\ -\|\Delta h(i, j)\| & \text{caso contrário} \end{cases} \quad (3.22)$$

Como não há sinal de direção, convencionou-se atribuir a  $\Delta h'(i, j)$  e  $\Delta v'(i, j)$  valores positivos se a energia luminosa aumenta e negativos caso diminua. A energia luminosa pode ser estimada pelos valores  $\|\Delta h_{+}(i, j)\|$  e  $\|\Delta h_{-}(i, j)\|$  na direção horizontal e  $\|\Delta v_{+}(i, j)\|$  e  $\|\Delta v_{-}(i, j)\|$  na direção vertical.

O vetores de cor podem ser abordados segundo o enfoque de ordenação vetorial. Trahanias e Venetsanopoulos [Trahanias & Venetsanopoulos (1993)] propõem uma classe de detectores de bordas baseados neste princípio. Ao contrário de grandezas escalares, para as quais não há dificuldades em se estabelecer um critério de ordenação, grandezas vetoriais não podem ser comparadas diretamente. Para um conjunto de vetores são calculadas as distâncias entre eles, acumulando para cada um a distância aos demais. Os vetores são ordenados crescentemente segundo as distâncias acumuladas. O primeiro vetor da ordenação é a mediana do conjunto, com menor distância acumulada aos demais, e os do final da seqüência os mais distantes desta mediana. Um operador baseado nesta ordenação pode ser definido conforme a Equação 3.23, na qual  $\mathbf{x}^{(1)}$  é o primeiro vetor da lista e  $\mathbf{x}^{(n)}$  é o último. O operador terá valores maiores quando  $\mathbf{x}^{(n)}$  e  $\mathbf{x}^{(1)}$  estiverem em lados diferentes da borda, sendo este o critério para a determinação das bordas.

$$VR(i, j) = \|\mathbf{x}^{(n)} - \mathbf{x}^{(1)}\| \quad (3.23)$$

A presença de ruído gaussiano influencia os primeiros termos da seqüência ordenada, enquanto ruído impulsivo influencia os termos do final. Após a ordenação dos vetores, a eliminação de  $k$  termos do início da seqüência e  $l$  termos de seu final reduzem esta sensibilidade aos ruídos, resultando na redefinição do operador conforme a Equação 3.24. O conjunto de vetores é selecionado ao redor do ponto  $(i, j)$  e os parâmetros  $k$  e  $l$  são determinados arbitrariamente.

$$VR(i, j) = \|\mathbf{x}^{(n-k)} - \mathbf{x}^{(l+1)}\| \quad (3.24)$$

Outra forma de abordagem do espaço vetorial das cores é a determinação do gradiente sobre os vetores. Conforme DiZenzo [DiZenzo (1986)], a imagem passa a ser vista como uma função vetorial  $\mathbf{f}(x, y)$  definida sobre uma região do plano cujo resultado a cada ponto é um vetor  $[\mathbf{f}_1(x, y), \mathbf{f}_2(x, y), \dots, \mathbf{f}_m(x, y)]^t$ . O  $t$  sobrescrito indica a transposição do vetor e  $m$  é a dimensão do espaço vetorial que, para o caso de imagens coloridas, assume valor  $m = 3$ . Para simplificação da notação usa-se  $\mathbf{w} = (x, y)$ , com  $w_1 = x$  e  $w_2 = y$ . Assim, para  $i = 1, 2$ , define-se a expressão da Equação 3.25, considerando-a contínua, assim como sua primeira derivada. A Equação 3.26 mostra os quatro valores  $g_{ij}(\mathbf{w})$ ,  $i, j = 1, 2$ , componentes de um campo tensorial simétrico. O conjunto da Equação 3.27 é a base de um espaço vetorial bidimensional de vetores tangentes em  $\mathbf{f}(\mathbf{w})$ .

$$\mathbf{q}_i(\mathbf{w}) = [\partial \mathbf{f}_1(\mathbf{w}) / \partial w_i, \partial \mathbf{f}_2(\mathbf{w}) / \partial w_i, \dots, \partial \mathbf{f}_m(\mathbf{w}) / \partial w_i]^t \quad (3.25)$$

$$g_{ij}(\mathbf{w}) = \mathbf{q}_i(\mathbf{w}) \cdot \mathbf{q}_j(\mathbf{w}) \quad (3.26)$$

$$S = \{\mathbf{q}_1(\mathbf{w}), \mathbf{q}_2(\mathbf{w})\} \quad (3.27)$$

Desta forma, deseja-se obter a direção na qual  $\mathbf{f}(\mathbf{w})$  apresenta a maior taxa de variação e o valor absoluto desta taxa de variação máxima. Fazendo-se uso da teoria de vetores e de manipulações algébricas, pode-se chegar à relação da Equação 3.28, que indica o valor de  $\theta$  (ângulo feito com o eixo  $w_1$ ) cuja direção indica a variação máxima ou mínima do campo vetorial  $\mathbf{f}(\mathbf{w})$ .

$$\theta = \frac{1}{2} \arctan\left(\frac{2g_{12}(\mathbf{w})}{g_{11}(\mathbf{w}) - g_{22}(\mathbf{w})}\right) \quad (3.28)$$

Além disto, se  $\theta_+$  é solução para a Equação 3.28, então  $\theta_- = \theta_+ + \pi/2$  também é. As direções obtidas como solução relacionam-se às variações máxima e mínima de  $\mathbf{f}(\mathbf{w})$  e o valor  $\theta_+$  é determinado, assim, como a direção da taxa de variação máxima de  $\mathbf{f}(\mathbf{w})$ . O valor  $\sqrt{F(\mathbf{w}, \theta_0)}$ , dado na Equação é a taxa máxima, associado ao conceito de “intensidade de borda”.

$$\sqrt{F(\mathbf{w}, \theta)} = [g_{11}(\mathbf{w}) \cos^2 \theta + g_{12}(\mathbf{w}) \cos \theta \sin \theta + g_{22}(\mathbf{w}) \sin^2 \theta]^{1/2} \quad (3.29)$$

Cumani [Cumani (1991)], propõe uma forma alternativa para a obtenção dos valores de intensidade e direção do gradiente vetorial baseada em autovalores e autovetores. A Equação 3.30 define uma função de contraste na direção do vetor unitário  $\mathbf{n} = (n_1, n_2)$ . Sabendo-se que a forma quadrática apresenta, através da variação de  $\mathbf{n}$ , valores extremos, o máximo e o mínimo podem ser obtidos dos autovalores da matriz  $2 \times 2$   $[g_{ij}(\mathbf{w})]$ , e coincidem com a direção dos autovetores associados.

$$\begin{aligned} S(\mathbf{w}, \mathbf{n}) &= \sum_{i=1}^2 \sum_{j=1}^2 g_{ij}(\mathbf{w}) n_i n_j \\ &= g_{11}(\mathbf{w}) n_1^2 + g_{12}(\mathbf{w}) n_1 n_2 + g_{22}(\mathbf{w}) n_2^2 \end{aligned} \quad (3.30)$$

Portanto os valores de máximo e mínimo, respectivamente  $\lambda_+(\mathbf{w})$  e  $\lambda_-(\mathbf{w})$ , bem como as direções relativas são obtidos pelas expressões na Equação 3.31, nas quais  $\theta_+(\mathbf{w})$  e  $\theta_-(\mathbf{w})$  satisfazem a Equação 3.28.

$$\begin{aligned} \lambda_{\pm}(\mathbf{w}) &= \frac{g_{11}(\mathbf{w}) + g_{22}(\mathbf{w}) \pm \sqrt{[g_{11}(\mathbf{w}) - g_{22}(\mathbf{w})]^2 + 4g_{12}^2(\mathbf{w})}}{2} \\ \mathbf{n}_{\pm}(\mathbf{w}) &= (\cos \theta_{\pm}(\mathbf{w}), \sin \theta_{\pm}(\mathbf{w})) \end{aligned} \quad (3.31)$$

Algoritmo 3-1. Extração de bordas por gradiente cromático [Lee & Cok (1991)].

Suavize a imagem com um filtro gaussiano.  
 Calcule as derivadas parciais em relação a  $x$  e  $y$  usando máscaras apropriadas.  
 Estime as amplitudes e direções do gradiente usando as derivadas parciais.  
 Marque como candidatos a pontos de borda as posições correspondentes a máximos locais do gradiente ao longo das direções deste gradiente.  
 Marque como pontos de borda definitivos os candidatos que apresentarem amplitude do gradiente acima de um limiar inferior e estiverem conectados a pelo menos um ponto com amplitude maior que um limiar superior.

Propriedades interessantes estão associadas a estes resultados [Lee & Cok (1991)]. Dado que  $\lambda_+(\mathbf{w})$  é a amplitude da intensidade do gradiente vetorial ao quadrado e que  $[g_{11}(\mathbf{w}) - g_{22}(\mathbf{w})]^2 + 4g_{12}^2(\mathbf{w}) \geq 0$ , tem-se  $g_{11}(\mathbf{w}) + g_{22}(\mathbf{w}) \geq \lambda_+(\mathbf{w})$ . Dadas ainda as relações da Equação 3.32, obtém-se que a soma dos quadrados dos gradientes escalares  $g_{11}(\mathbf{w}) + g_{22}(\mathbf{w})$  é sempre maior ou igual ao gradiente vetorial  $\lambda_+(\mathbf{w})$ .

$$\begin{aligned} g_{12}(\mathbf{w}) + g_{22}(\mathbf{w}) &= \sum_{k=1}^m \left( \frac{\partial f_k(\mathbf{w})}{\partial w_1} \right)^2 + \sum_{k=1}^m \left( \frac{\partial f_k(\mathbf{w})}{\partial w_2} \right)^2 \\ &= \sum_{k=1}^m \|\nabla f_k(\mathbf{w})\|^2 \end{aligned} \quad (3.32)$$

Denominam-se, assim, a quantidade  $\sqrt{g_{11}(\mathbf{w}) + g_{22}(\mathbf{w})}$  a amplitude do gradiente escalar e  $\sqrt{\lambda_+(\mathbf{w})}$  a amplitude do gradiente vetorial. O gradiente vetorial tem amplitude igual à do gradiente escalar quando todos os gradientes de cada componente tiverem a mesma direção. Para uma imagem colorida, na qual as componentes apresentam, em geral, grande correlação, a situação descrita acima (aproximadamente) é quase sempre verdadeira. Quando houver ruído descorrelacionado, os gradientes das componentes terão direções diferentes e o gradiente vetorial terá valor menor que o escalar. Como o sinal é geralmente mais correlacionado que o ruído, ao final tem-se um aumento da relação sinal/ruído na detecção de bordas, o que leva à conclusão que, para imagens coloridas típicas, o gradiente vetorial é menos sensível ao ruído que o gradiente escalar.

Como exemplo, o Algoritmo 3-1 apresenta de forma geral o procedimento para a detecção de bordas proposto por Lee e Cok [Lee & Cok (1991)].

O gradiente multiespectral pode ser calculado, ainda, em termos de atributos perceptivos [Pujas & Aldon (1997)]. O espaço de representação HSV apresenta dois problemas quanto à componente do matiz  $H$ : a componente tem periodicidade e não está definida quando  $S = 0$ . A implicação do primeiro problema é o cálculo de médias entre matizes, devendo ser definida uma álgebra adequada. O segundo problema ocor-

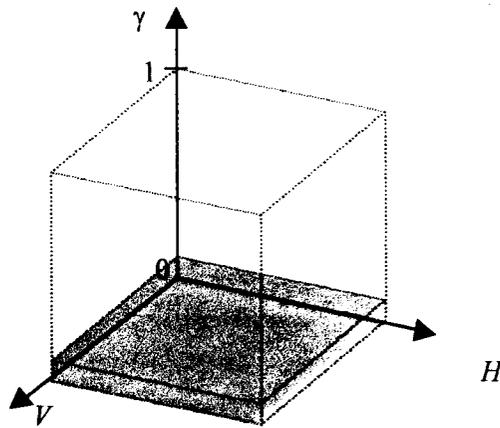


Figura 3.4. Ilustração do novo espaço de representação de cores  $HV\gamma$ . A área hachurada representa a porção do espaço no qual o grau de cromaticidade é mais baixo.

re quando a saturação é próxima de zero, podendo ser resolvido considerando-se uma área acromática, na qual  $H$  não é utilizado (indicando “indefinido” ou “indevidamente definido”), e uma cromática, na qual  $H$  é bem definido. Somente  $S$  e  $V$  definem os atributos na área acromática.

Pujas e Aldon [*ibid.*] propõem, assim, um espaço de representação  $HV\gamma$  (Figura 3.4), mantendo as coordenadas de matiz e intensidade, mas substituindo a saturação pelo “grau de cromaticidade”  $\gamma$ , que é determinado iterativamente através da escolha de pontos representativos da imagem. Os valores que  $\gamma$  assume variam de 0 a 1, e quanto mais próximo a zero, mais acromático é o ponto; quanto mais próximo a 1, maior sua representação cromática. A saturação fica incorporada em  $\gamma$ , uma vez que é função de  $S$  e  $V$ .

Para este espaço, o gradiente pode ser dado pela Equação 3.33, lembrando que  $\mathbf{w}$  representa a posição  $(i, j)$ . O fator  $\gamma$  contribui para o grau em que se consideram os gradientes do matiz e da intensidade.

$$\nabla C(\mathbf{w}) = \gamma \nabla H(\mathbf{w}) + (1 - \gamma) \nabla V(\mathbf{w}) + \nabla \gamma(\mathbf{w}) \quad (3.33)$$

Na representação tensorial, a estimativa  $\hat{g}(\mathbf{w})$  pode ser dada conforme a Equação 3.34. O tensor  $\hat{g}(\mathbf{w})$  representa a métrica do espaço perceptivo definido.

$$\hat{g}(\mathbf{w}) = \begin{bmatrix} \hat{g}_{11}(\mathbf{w}) & \hat{g}_{12}(\mathbf{w}) \\ \hat{g}_{21}(\mathbf{w}) & \hat{g}_{11}(\mathbf{w}) \end{bmatrix} \quad (3.34)$$

$$\hat{g}_{11}(\mathbf{w}) = \gamma \left( \frac{\partial H(\mathbf{w})}{\partial w_1} \right)^2 + (1 - \gamma) \left( \frac{\partial V(\mathbf{w})}{\partial w_1} \right)^2 + \left( \frac{\partial \gamma(\mathbf{w})}{\partial w_1} \right)^2 \quad (3.35)$$

$$\begin{aligned}\hat{g}_{12}(\mathbf{w}) &= \hat{g}_{21}(\mathbf{w}) = \\ &= \gamma \frac{\partial H(\mathbf{w})}{\partial w_1} \frac{\partial H(\mathbf{w})}{\partial w_2} + (1-\gamma) \frac{\partial V(\mathbf{w})}{\partial w_1} \frac{\partial V(\mathbf{w})}{\partial w_2} + \frac{\partial \gamma(\mathbf{w})}{\partial w_1} \frac{\partial \gamma(\mathbf{w})}{\partial w_2}\end{aligned}\quad (3.36)$$

$$\hat{g}_{22}(\mathbf{w}) = \gamma \left( \frac{\partial H(\mathbf{w})}{\partial w_2} \right)^2 + (1-\gamma) \left( \frac{\partial V(\mathbf{w})}{\partial w_2} \right)^2 + \left( \frac{\partial \gamma(\mathbf{w})}{\partial w_2} \right)^2 \quad (3.37)$$

Esta abordagem apresenta como vantagem a característica que, em baixas saturações, os gradientes do matiz são relevados, diminuindo a segmentação excessiva em áreas de matiz homogênea. Por outro lado, a determinação de  $\gamma$  é interativa e depende do conjunto de dados escolhido como “representativo” da imagem.

### 3.2 Técnicas baseadas em regiões

Segmentar uma imagem considerando as regiões significa encontrar um critério que permita agrupar pontos que tenham características semelhantes sob um mesmo rótulo. Em contraposição à detecção de bordas, na qual são consideradas as variações abruptas, o princípio destas abordagens é a similaridade.

As abordagens baseadas em regiões podem ser agrupadas em duas vertentes gerais: uma baseada em pontos (atributos) e a outra abordando a conectividade de regiões (área). A primeira vertente é relativa aos pontos e seus atributos individuais, tendo como fundação a classificação dos pontos proporcionando rótulos específicos independentes da distribuição espacial. Já a abordagem considerando áreas opera sobre princípios que avaliam se uma região não é homogênea e deve ser dividida ou se pontos conectados são similares e devem ser agrupados em uma mesma região\*.

#### 3.2.1 Segmentação por classificação e agrupamento

A classificação é utilizada para associar a cada ponto na imagem um rótulo, usualmente considerando que cada região (objetos na cena, ou partes deles) tenha características mais ou menos uniformes em relação a um conjunto de atributos. Os atributos dos pontos da imagem são então considerados e sua separação segundo um conjunto de rótulos resulta na segmentação. O uso de agrupamentos atinge um resultado final similar, mas a rotulação é obtida através da determinação de subconjuntos dentro do conjunto de dados.

Uma forma de se fazer a classificação é através de um esquema no qual as classes sejam conhecidas e o objetivo seja identificar quais pontos na imagem estão relacionados a cada classe. A técnica de classificação  $k$  vizinhos mais próximos [Duda & Hart (1973)] utiliza este princípio e geralmente fornece resultados bastante robustos.

\* As regiões podem ser formadas pela adição sucessiva de novos pontos ou ainda pela junção de duas ou mais regiões similares.

tos. O procedimento consiste em escolher um certo número de amostras para cada classe, rotulando-as adequadamente. Para cada dado a ser classificado é feita uma comparação de distância (usando uma métrica arbitrária) do dado a cada uma das amostras, ordenando estas últimas crescentemente pela distância. As  $k$  amostras mais próximas, com  $k$  também arbitrário, são verificadas e a classe que tiver maior número de representantes entre estas amostras determina a rotulação do dado. Havendo empate, um critério de amostra mais próxima pode ser usado para o desempate. Sua aplicação para a segmentação de imagens apresentou resultados robustos, considerando sua aplicação a imagens coloridas, sendo que tanto os dados quanto as amostras consideravam os atributos de cromaticidade (normalização das componentes RGB) [Moreira & Costa (1995)].

Outra forma também supervisionada para classificação se dá através da classificação bayesiana, que se baseia nas distribuições de probabilidades das classes. Sua aplicação pode ser restringida considerando que as distribuições são próximas à distribuição normal. Esta restrição não é irreal dado que os dados são relacionados à cores dos objetos (ou parte de objetos), cada cor simples compondo uma classe. Mesmo na presença de ruído e distorções introduzidas na captura, cada cor é considerada aproximadamente constante ao longo da superfície do objeto, com variações significativas apenas em intensidade. Distribuições para cada cor em torno de uma média são, portanto, esperadas. A Equação 3.38 apresenta a forma geral de uma densidade normal multivariada, na qual  $\mathbf{x}$  é um vetor coluna de  $d$  componentes,  $\bar{\mu}$  é o vetor média,  $\Sigma$  é a matriz de covariância  $d \times d$ ,  $(\mathbf{x} - \bar{\mu})'$  é a transposta de  $(\mathbf{x} - \bar{\mu})$ ,  $\Sigma^{-1}$  e  $|\Sigma|$  são, respectivamente, a inversa e o determinante de  $\Sigma$  [Duda & Hart (1973)].

$$P(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{x} - \bar{\mu})' \Sigma^{-1} (\mathbf{x} - \bar{\mu})} \quad (3.38)$$

Para dados de cromaticidade  $d$  assume valor 2, já que apenas duas cores normalizadas descrevem completamente os atributos relativos a matiz e saturação, sendo a terceira componente dependente das outras duas.

A classificação é preparada com a seleção inicial de amostras de cada classe, juntamente com uma probabilidade *a priori*. Este conjunto de amostras deve ser suficiente para proporcionar a informação necessária para determinar as funções de densidade, de forma que o cálculo das probabilidades *a posteriori* seja possível. As amostras são selecionadas (e rotuladas) diretamente dos dados normalizados da imagem. Um pré-processamento para a redução de ruído pode ser realizado. A Equação 3.39 mostra a regra de Bayes para o cálculo da probabilidade *a posteriori*  $P(C_j|\mathbf{x})$ , ou seja, a probabilidade da classe  $j$  dado que se dispõe do valor  $\mathbf{x}$ . Na equação,  $p(\mathbf{x}|C_j)$  é função de densidade condicional para  $\mathbf{x}$  e  $P(C_j)$  é a probabilidade da *a priori* da classe  $j$  ( $1 \leq j \leq c$ ) [Duda & Hart (1973)].

Algoritmo 3-2. Procedimento básico do algoritmo  $k$  médias. (Adaptado de Duda e Hart [Duda & Hart (1973)].)

Determine um conjunto de  $k$  centróides iniciais arbitrários.  
 Repita  
   Classifique os dados em relação ao centróide mais próximo.  
   Recalcule os novos centróides (como a média das amostras) para cada agrupamento.  
 Até que não haja alteração dos centróides.

$$P(C_j|\mathbf{x}) = \frac{p(\mathbf{x}|C_j)P(C_j)}{\sum_{j=1}^c p(\mathbf{x}|C_j)P(C_j)} \quad (3.39)$$

Dado que todos os parâmetros estão disponíveis, os dados da imagem são processados para rotular cada ponto da imagem segundo sua classe, ou seja, associá-lo à classe com probabilidade *a posteriori* máxima, minimizando, assim, a probabilidade de erro.

As técnicas de agrupamento, por sua vez, usualmente operam sobre um conjunto de dados determinando qual a melhor maneira de separá-los em classes de forma a minimizar um critério de erro global. A segmentação é dada pela rotulação dos pontos segundo as classes (grupos) determinadas.

Coleman e Andrews [Coleman & Andrews (1979)] apresentam uma técnica que é uma variação do algoritmo de agrupamento  $k$  médias [Duda & Hart (1973)], o qual ajusta o conjunto de dados iterativamente, relacionado-os a  $k$  centróides (ou médias). A operação se dá de forma não supervisionada e determina o número ótimo de grupos no conjunto de dados. O procedimento básico da classificação por  $k$  médias é apresentado no Algoritmo 3-2. Em cada iteração é feita a classificação dos dados, associando-os ao centróide mais próximo segundo um critério de distância. Pode ser feito, então, o cálculo dos novos centróides de cada agrupamento. O processo é encerrado quando os grupos se tornam estáveis e os centróides não mais se modificam, tendendo estes últimos para uma condição que minimiza uma função de erro quadrático.

Sendo que o número  $k$  de médias adotado é arbitrário e imutável até o final do procedimento, a correta estimativa do número de classes presentes determina a qualidade da classificação. Na abordagem de Coleman e Andrews [*ibid.*], um critério é adicionado para que se compare a classificação para número distintos de centróides, permitindo escolher um número “ótimo” de agrupamentos a serem utilizados.

Coleman e Andrews [*ibid.*] propõem, desta forma, que considerando-se  $\mathbf{x}$  como um vetor de atributos,  $\mathbf{x} = (x_1, x_2, \dots, x_n)^t$ , o objetivo do agrupamento é a obtenção de  $K$  regiões mutuamente exclusivas que compreendam todos os dados. A determinação do melhor número de grupos é feita através de um parâmetro baseado

Algoritmo 3-3. Segmentação por agrupamento de Coleman e Andrews [Coleman & Andrews (1979)].

Selecione dois centróides iniciais.  
 Classifique todos os dados em relação a estes centróides (usando um critério de distância).  
 Repita  
   Crie um novo centróide, selecionando o vetor mais distante de seu centróide.  
   Classifique novamente os dados em função do conjunto atual de centróides.  
   Calcule o valor do parâmetro de controle  
 Até que o parâmetro comece a decrescer  
 Classifique os dados segundo o valor máximo do parâmetro de controle (antes da descendente).

nas matrizes de espalhamento intragrupos e intergrupos, especificadas pelas Equações 3.40 e 3.41, respectivamente [Duda & Hart (1973)]. O vetor  $\bar{\mu}_k$  é a média e  $M_k$  é o número de elementos para um dado grupo  $k$ ;  $S_k$  é o conjunto de pontos pertencentes ao grupo  $k$  e  $\bar{\mu}_0$  especifica o vetor média compreendendo todos os  $M$  dados.

$$S_w = \frac{1}{K} \sum_{k=1}^K \frac{1}{M_k} \sum_{x_i \in S_k} (x_i - \bar{\mu}_k)(x_i - \bar{\mu}_k)^t \quad (3.40)$$

$$S_b = \frac{1}{K} \sum_{k=1}^K (\bar{\mu}_k - \bar{\mu}_0)(\bar{\mu}_k - \bar{\mu}_0)^t \quad (3.41)$$

$$\bar{\mu}_0 = \frac{1}{M} \sum_{i=1}^M x_i \quad (3.42)$$

Para a definição do parâmetro de controle  $\beta$  do número de grupos faz-se uso da Equação 3.43, na qual  $\text{tr}(\cdot)$  é o traço da matriz. Na situação de um único agrupamento de dados,  $\text{tr } S_w = \sigma^2$  (a variância dos dados) e  $\text{tr } S_b = 0$ , de forma que  $\beta = 0$ ; se o número de grupos for  $M$ , o número total de amostras, então  $\text{tr } S_w = 0$  e  $\text{tr } S_b = \sigma^2$ , e novamente  $\beta = 0$ . Os valores de  $\beta$  são iguais a zero em ambos os limites e positivos dentro do intervalo, de forma que há, necessariamente, um ponto de máximo (ou vários). Este ponto de máximo indica um balanceamento entre  $S_w$  e  $S_b$ , o que é utilizado como critério para a determinação do melhor número de grupos.

$$\beta = \text{tr } S_w \cdot \text{tr } S_b \quad (3.43)$$

O procedimento para a aplicação da técnica é apresentado no Algoritmo 3-3. Avaliações subjetivas permitem considerar bons os resultados da segmentação, os quais são usualmente melhorados se for feita a decorrelação dos atributos. A escolha

Algoritmo 3-4. Segmentação por agrupamento no espaço  $L^*a^*b^*$ . [Celenk (1990)]

Calcule os histogramas unidimensionais para cada componente.  
 Selecione em um dos histogramas dois agrupamentos principais (naquele em que for mais proeminente).  
 Usando a informação obtida, avalie os histogramas das outras duas componentes, usando somente a porção do espaço relativa ao grupo encontrado. Determine os picos das outras componentes.  
 Considerando o modo principal do histograma e seu vizinho, determine a função de discriminação de Fischer para projeção e divisão por limiar unidimensional.  
 Verifique os pontos restantes para a determinação de modos, repetindo o processo. Caso não se verifiquem mais modos significativos, extraia um novo conjunto de atributos, submetendo ao método novamente.

dos centróides iniciais, entretanto, tem influência significativa sobre a classificação final.

Em vez de usar os vetores em si, é possível obter um ganho computacional considerável na realização da segmentação se o histograma multidimensional for utilizado para a obtenção dos grupos. Considerando uma abordagem de “escalar picos” no histograma, Khotanzad e Bouarfa [Khotanzad & Bouarfa (1990)] propõem uma abordagem para a segmentação de imagens baseada em duas etapas. A primeira etapa é a obtenção do histograma de dimensão  $d$  (número de bandas), no qual cada célula guarda o número de vetores dentro de uma dada quantização. O agrupamento é realizado “escalando” os picos do histograma através da verificação, entre células vizinhas\*, das relações de ascendência (transições de valores menores para maiores). Uma célula para a qual é possível determinar uma vizinha com valor maior é uma célula filha que tem uma célula pai associada. Um pico, então, pode ser definido como a célula de maior densidade em uma região, ou seja, a que não tem pai (isto é, as células vizinhas são todas filhas ou nulas).

A construção dos agrupamentos é iniciada definindo cada pico encontrado como um grupo. As células vizinhas com ligação a um dado pico são localizadas e incorporadas ao grupo. O procedimento prossegue através da localização e incorporação de cada célula com ligação estabelecida a uma das já agrupadas, sucessivamente até que todas estejam associadas a um dos picos iniciais. Como todas as células não vazias estão, ao final, associadas a um dos grupos, também estão os vetores associados a cada célula.

Em vez do espaço RGB, Celenk [Celenk (1990)] usa o espaço  $L^*a^*b^*$ , o qual aproximadamente satisfaz a propriedade de uniformidade† [Wyszecki & Stiles (1982)]. Para a representação, definem-se as coordenadas cilíndricas  $L^*$ ,  $H^0$  e  $C^*$ , que se relacionam, respectivamente, às propriedades psicofísicas de luminância, matiz e croma, conforme a Equação 3.44.

\* A vizinhança é 8 para um histograma bidimensional, 26 para um tridimensional e assim por diante. Ou seja, verifica-se cada vizinho da hipercélula em todas as dimensões.

† Veja a descrição do espaço  $L^*a^*b^*$  na seção 2.3.3.5, caracterizando a uniformidade de percepção.

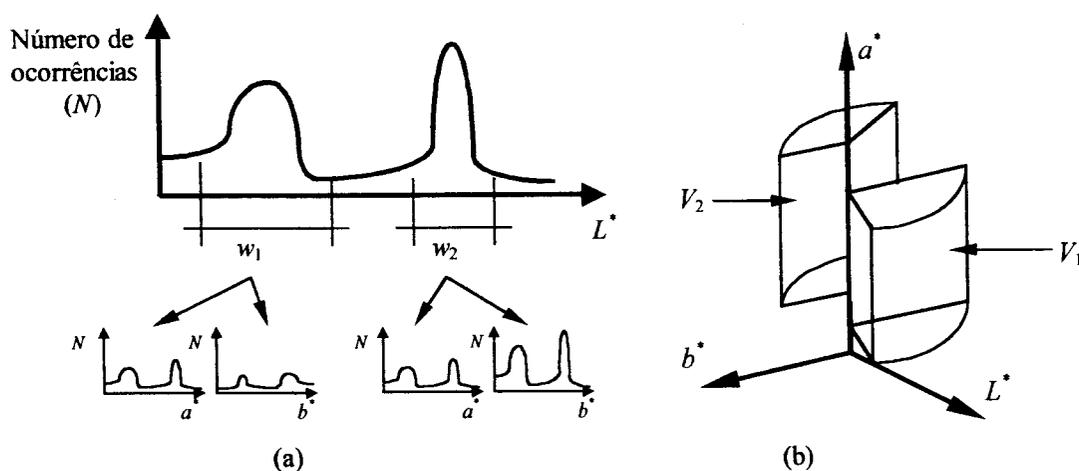


Figura 3.5. Estabelecimentos dos agrupamentos no espaço  $L^*a^*b^*$ . (a) Os modos mais proeminentes ( $w_1$  e  $w_2$ ) são detectados em um dos histogramas unidimensionais — por exemplo  $L^*$ . Para cada modo é identificado o melhor pico nos outros dois histogramas — componentes  $a^*$  e  $b^*$ , neste exemplo. (b) Os intervalos definidos pelos três picos nos histogramas delimitam um volume circular cilíndrico no espaço de cor.  $V_1$  é definido em relação a  $w_1$  e  $V_2$  em relação a  $w_2$ .

$$\begin{aligned}
 L^* &= L^* \\
 H^0 &= \arctan \frac{b^*}{a^*} \\
 C^* &= \sqrt{(a^*)^2 + (b^*)^2}
 \end{aligned} \tag{3.44}$$

O procedimento proposto para a segmentação através do agrupamento dos dados é apresentada no Algoritmo 3-4. A determinação dos volumes é feita, inicialmente, através da avaliação dos histogramas unidimensionais de cada componente do espaço de cores. Aquele que apresentar modos (picos) mais proeminentes é selecionado, permitindo a definição de dois intervalos correspondentes ao modo principal ( $w_1$ ) e seu vizinho ( $w_2$ ), como esquematizado na parte superior da Figura 3.5(a).

Para o modo  $w_1$  é feita a avaliação dos histogramas unidimensionais das componentes restantes (restrita pela abrangência estabelecida pelo modo), verificando pelos seus modos e restringindo um intervalo para cada componente (parte inferior da Figura 3.5(a)). As restrições efetuadas definem em  $L^*a^*b^*$  um volume circular cilíndrico  $V_1$ , caracterizando a estimativa do primeiro agrupamento. O mesmo procedimento é aplicado a  $w_2$  para a obtenção de  $V_2$  (Figura 3.5(b)).

A separação dos dados dos dois volumes é feita através da projeção dos dados em uma reta, cuja determinação é estabelecida pelo discriminante de Fischer [Duda & Hart (1973)]. Feita a projeção, procede-se à avaliação do histograma dos dados projetados, permitindo a realização efetiva do agrupamento do modo principal. O processo é aplicado em sua íntegra, então, aos pontos ainda não incluídos em agrupamentos, repetindo-se até que não mais sejam obtidos modos significativos nos histogramas de qualquer componente. Neste último caso, um novo conjunto de atributos é derivado para os dados, gerando um novo espaço a ser submetido ao algoritmo descrito.

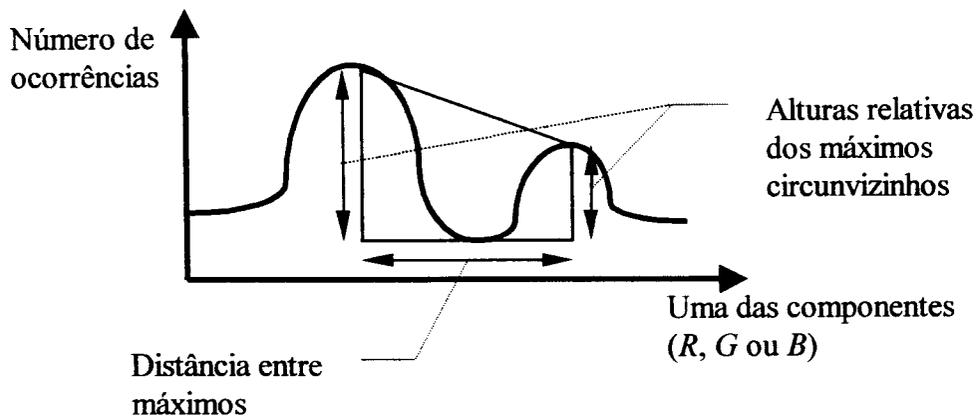


Figura 3.6. Para uma dada resolução do histograma, são estabelecidas proporções entre a altura do vale e picos no histograma, além de ser considerada uma distância mínima entre estes picos.

Desenvolvida para a segmentação de mapas coloridos, a abordagem de Hedley e Yan [Hedley & Yan (1992)] também opera sobre histogramas unidimensionais, usando o espaço RGB. O processo para a segmentação baseada nos histogramas adota várias escalas (refletidas na resolução dos histogramas). A avaliação do histograma é feita em busca de vales com máximos significativos em ambos os lados. São estabelecidos critérios de altura mínima, dada por um fator de proporção em relação à altura do mínimo sendo avaliado em relação aos máximos circunvizinhos, e distância mínima entre máximos (Figura 3.6). Tais critérios são usados recursivamente a partir de uma resolução grossa até que se determine o “melhor” mínimo do histograma, por rastreamento, para uma resolução mais fina. Baseado nesta análise recursiva uma árvore de decisão é montada, sendo os dados classificados. Interessantemente este procedimento é somente aplicado aos pontos “bem comportados” da imagem, de maneira que um critério espacial é utilizado para a seleção dos dados a serem processados. Considerando-se a aplicação (mapas), regiões de transição de cor (em marcações de rodovias, rios etc.) geram instabilidade nas cores, dada tanto pelo processo de aquisição quanto pelo registro imperfeito da impressão do mapa, proporcionando diferenças no que se refere às cores principais. Estes dados “imprecisos” são removidos do processamento por uma seleção do gradiente calculado sobre a imagem. Regiões de amplitude de gradiente alta, associada a bordas, são desconsideradas e somente os pontos restantes são classificados pela análise dos histogramas. Os pontos de gradiente alto são posteriormente unidos às regiões definidas pelo agrupamento realizado através do crescimento de tais regiões, até que todos os pontos pertençam a uma ou outra região.

A classificação, usada como ferramenta para a segmentação de imagens, é bastante difundida [Pal & Pal (1993)], sendo que para imagens coloridas a utilização

de classificação ou agrupamento difusos\* é bastante abrangente. O algoritmo  $c$  médias difuso é o mais empregado, às vezes com sugestões de modificações. O objetivo é a partição da totalidade do conjunto de dados ( $n$  vetores) em  $c$  grupos distintos. Os vetores são indicados por  $x_k$ ,  $k$  variando de 1 a  $n$  e o número de grupos  $c$  varia de 2 a  $n$ .

A Equação 3.45 indica o critério de partição, incluindo as restrições especificadas na Equação 3.46.  $V_{cn}$  é o conjunto de matrizes reais  $c \times n$ ,  $U \in V_{cn}$  tem cada linha  $i$  indicando o grau de participação†  $u_i$  dos vetores  $x_k$ . Cada  $u_{ik}$  indica o grau de participação do vetor  $x_k$  na classe  $c_i$ , assumindo valores no intervalo  $[0, 1]$ . A primeira restrição na Equação 3.46 especifica que cada vetor  $x_k$  tenha grau de participação total em todas as classes igual a 1. A segunda restrição impõe que nenhuma partição seja vazia.

$$M = \{U \in V_{cn} | u_{ik} \in [0,1], \forall i, k\} \quad (3.45)$$

$$\begin{aligned} \sum_{i=1}^c u_{ik} &= 1, \forall k \\ 0 < \sum_{k=1}^n u_{ik} < n, \forall i \end{aligned} \quad (3.46)$$

A função de avaliação do algoritmo de partição difusa pode ser como a da Equação 3.47, na qual  $U$  é a matriz de participação nos grupos ( $U = [u_{ik}]$ ),  $m$  é o “índice de difusão” ( $1 \leq m \leq \infty$ ) e  $V = \{v_1, v_2, \dots, v_c\}$  é o conjunto de centróides dos grupos. O símbolo  $\|\cdot\|$  indica uma norma, como a distância euclidiana. Esta expressão é a função do mínimo erro quadrático para uma partição difusa.

$$J_m(U, V) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m \|x_k - v_i\|^2 \quad (3.47)$$

Para  $x_k \neq v_i$ , o valor local mínimo para  $J_m(U, V)$ , para qualquer  $i$  e  $k$ , é dado por  $(U, V)$ , considerando os graus de participação e centróides dados na Equações 3.48 e 3.49.

$$u_{ik} = \frac{\|x_k - v_i\|^{\frac{2}{m-1} - 1}}{\sum_{j=1}^c \|x_k - v_j\|^{\frac{2}{m-1} - 1}} \quad (3.48)$$

\* O termo “difuso” é usado como tradução para *fuzzy*, embora o termo “nebuloso” também seja largamente empregado.

† O grau de participação é representado por um valor entre 0 e 1, que fornece a proporção em que um vetor se associa a um grupo. Assim, um vetor pode se associar com grau 0,80 a um grupo A, 0,15 a um grupo B e 0,05 a um grupo C, indicando sua ligação mais acentuada (80%) ao grupo A.

Algoritmo 3-5. Procedimento básico para o agrupamento  $c$  médias difuso.

Selecione os centróides iniciais.  
 Repita  
   Calcule as participações (Equação 3.48).  
   Calcule os novos centróides usando os novos valores de participações (Equação 3.49).  
 Até que a Equação 3.50 seja satisfeita.

$$v_i = \frac{\sum_{k=1}^n u_{ik}^m x_k}{\sum_{k=1}^n u_{ik}^m} \quad (3.49)$$

No Algoritmo 3-5 é apresentada a repetição iterativa para a determinação das partições difusas. A Equação 3.50 mostra o critério de convergência, sendo que  $\hat{v}_i$  indica os centróides antes do passo atual e  $\varepsilon$  é o erro tolerado.

$$\|v_i - \hat{v}_i\| < \varepsilon, \quad \forall i \in [1, c] \quad (3.50)$$

Uma aplicação do algoritmo utiliza número de classes  $c$  fixo em 4 e índice de difusão  $m$  com valor 2 [Huntsberger *et al.* (1985)]. Após cada associação difusa dos pontos às classes, aqueles que tiverem associação acima de um limiar têm sua rotulação definida. Os pontos restantes permanecem no processo, que é repetido até que todos (ou uma porcentagem especificada) sejam classificados. Em cada estágio um conjunto de 2400 pontos são aleatoriamente escolhidos da imagem. Centróides muito próximos podem ser fundidos em uma única classe, usando como critério a expressão da Equação 3.51, na qual  $v_i^{(k)}$  indica o  $k$ -ésimo atributo de  $v_i$ .

$$2 \frac{|v_1^{(k)} - v_2^{(k)}|}{v_1^{(k)} + v_2^{(k)}} \leq 0,075, \quad \forall k \quad (3.51)$$

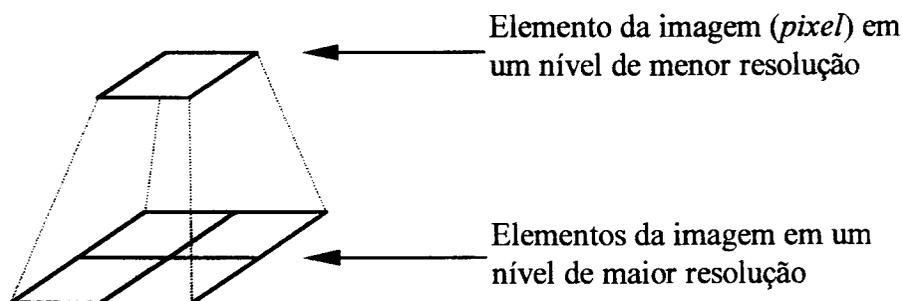


Figura 3.7. Representação da pirâmide de resolução. Cada elemento de imagem de um nível superior da pirâmide compreende a média de quatro elementos do nível imediatamente inferior. Ou seja, a cada nível a imagem tem seu tamanho reduzido pela metade tanto em altura quanto em largura.

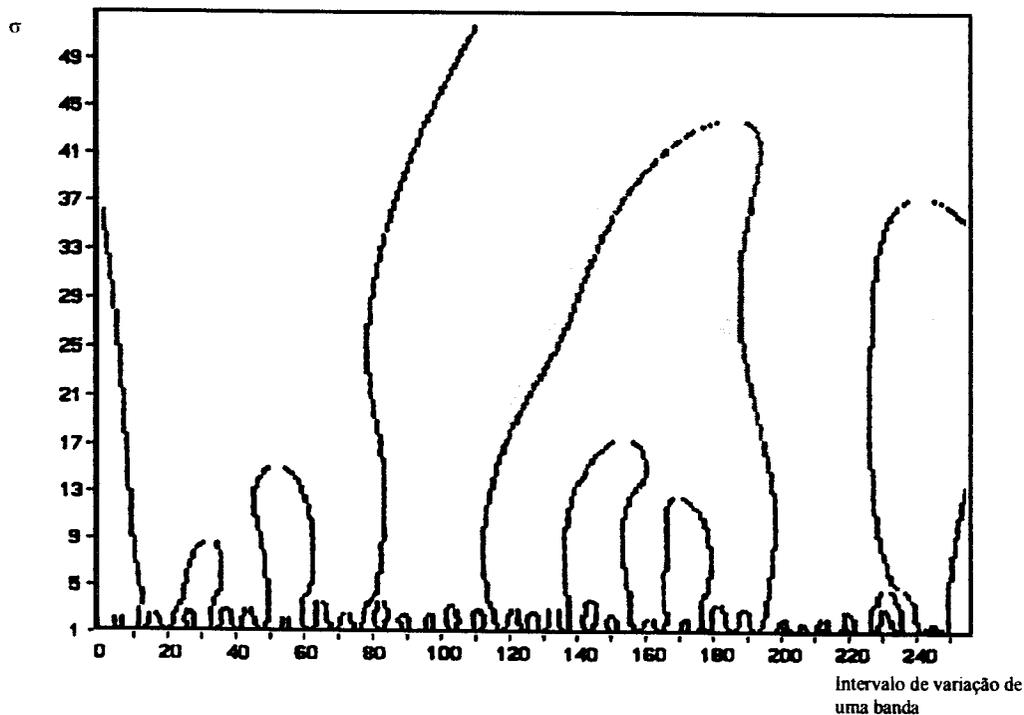


Figura 3.8. As linhas indicam os cruzamentos em zero da derivada do histograma, filtrado com variâncias distintas. O retângulo indica uma das regiões definidas por pares de cruzamentos em zero considerados significativos.

Trivedi e Bezdec [Trivedi & Bezdec (1986)] propõem uma forma de segmentação para imagens multiespectrais usando agrupamento difuso com uma pirâmide de resolução para a redução do custo de processamento. Cada ponto de um dado nível  $l$  da pirâmide compreende a média de uma janela  $2 \times 2$  do nível imediatamente inferior (Figura 3.7). O conjunto de vetores relativos a cada elemento de resolução da pirâmide é então agrupado em  $c_l$  grupos, usando o algoritmo  $c$  médias difuso. Elementos não homogêneos, ou seja, classificados como pertencentes a duas classes ou mais, são divididos em elementos menores em um nível de maior resolução, de forma a serem novamente enquadrados em  $c_{l+1}$  grupos. O processo termina quando nenhum dos elementos da pirâmide falhar no teste de homogeneidade. Tal critério de homogeneidade usado em cada nível da pirâmide, considerando um elemento  $e_i$  no nível  $l$ , é  $\max\{u_1(x_k), u_2(x_k), \dots, u_{c_l}(x_k)\} \geq T$ . Se a condição não for satisfeita, o elemento é subdividido em um nível  $l+1$ . Embora o processo de agrupamento difuso não seja supervisionado, exige-se que sejam estabelecidos os números de grupos para cada nível.

Uma técnica para a redução do custo do processo de segmentação pela classificação difusa foi proposta por Lim e Lee [Lim & Lee (1990)] usando uma abordagem em duas etapas. A primeira, mais “grossa”, é baseada na divisão dos histogramas unidimensionais por um limiar. A divisão é obtida por uma análise espaço-escala dos histogramas, no qual os histogramas são filtrados por gaussianas de variâncias distintas. Cruzamentos por zero da segunda derivada do histograma filtrado são buscados

Algoritmo 3-6. Algoritmo de classificação difusa  $c$  médias com seleção automática do número de classes [Kamei-Harada *et al.* (1994)].

Determine o número inicial de grupos e os limiares de fusão e divisão de grupos.  
 Selecione os centróides iniciais.  
 Repita  
   Calcule as partições difusas  $u_{ik}$  (Equação 3.48).  
   Remova, sucessivamente, as classes com coeficiente  $u_{ik}$  inferior a 0,6. (Havendo a remoção de classes os valores  $u_{ik}$  são recalculados.)  
   Se o número de grupos for 14 ou inferior, verifique a condição de divisão. (Havendo a remoção de classes os valores  $u_{ik}$  são recalculados.)  
   Calcule os novos centróides  $v_i$  (Equação 3.48).  
   Verifique a condição de fusão. (Havendo a remoção de classes os valores  $u_{ik}$  e  $v_i$  são recalculados.)  
 Até que uma condição de convergência seja satisfeita (Equação 3.50, por exemplo).

para cada escala, definindo linhas no espaço-escala. Linhas mais longas indicam cruzamentos por zero mais significativos e, em torno dessas linhas são definidas regiões retangulares limitadas, na escala, pelo aparecimento de um par de cruzamentos por zero até o aparecimento de um novo par e, no espaço, pelas localizações do par de cruzamentos na menor escala.

A Figura 3.8 ilustra as linhas de cruzamento em zero e as regiões significativas. Regiões cujos limites de escala são mais distantes têm maior significância, sendo possível determinar vales e picos significativos no histograma e conseqüente avaliação da escala “ótima” na qual ocorrem. Os vales são usados para a divisão do histograma por um limiar, mas como os valores são obtidas de formas suavizadas, somente os dados mais próximos aos picos são considerados. Em torno dos vales, como há ambigüidade em relação à partição do histograma, os dados vinculados a estas regiões são processados posteriormente. Os picos significativos são encontrados para as três componentes, o que define hexaedros no histograma tridimensional. Cada hexaedro que acumule número de pontos acima de um limiar é selecionado e define uma partição inicial. A segunda etapa, mas refinada, utiliza o algoritmo de classificação difusa  $c$  médias para os pontos restantes, rotulando-os segundo as classes preexistentes.

Uma das principais características do algoritmo de agrupamento difuso  $c$  médias é a necessidade do conhecimento prévio do número de classes às quais os dados serão associados. Estimativas de número podem ser feitas frente ao conhecimento do tipo da imagem e aplicação em questão, ou através de uma pré-avaliação dos dados da imagem (como na abordagem de Lim e Lee [*ibid.*]). Alternativamente Kamei-Harada *et al.* [Kamei-Harada *et al.* (1994)] propõem uma modificação do algoritmo a qual incorpora mecanismos de fusão de classes similares e divisão de classes heterogêneas durante as iterações. O Algoritmo 3-6 apresenta, dentro do contexto do algoritmo original, as modificações introduzidas.

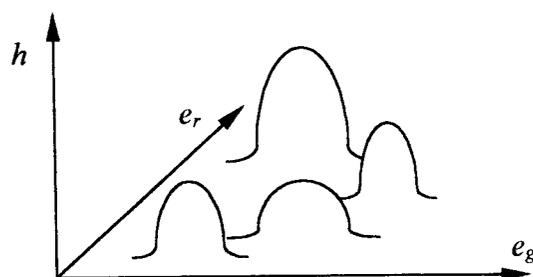


Figura 3.9. Espaço definido pelas componentes cromáticas  $e_r$  e  $e_g$ , além da componente  $h$ , a qual corresponde ao número de ocorrências das outras duas componentes.

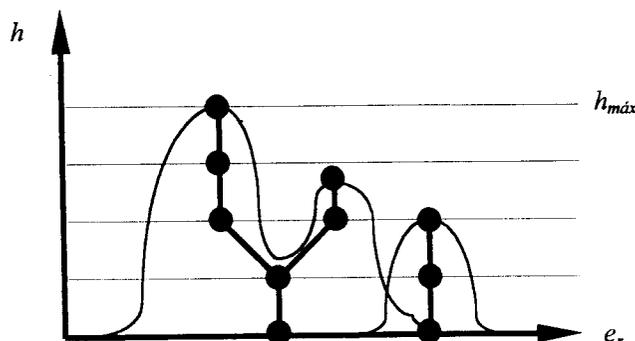


Figura 3.10. Estrutura de árvore estabelecida segundo vários planos paralelos ao plano  $e_r, e_g$ . A ilustração mostra uma visão na direção do eixo  $e_g$ .

O procedimento de divisão é baseado no histograma, usando o espaço  $L^*u^*v^*$ , procurando por picos nos quais o número de pontos multiplicado por  $u_{ik}$  é máximo dentro da classe. Verifica-se, ainda dentro de uma mesma classe, um segundo pico máximo. Se este segundo pico da classe distar do primeiro acima de um limiar preestabelecido, passa a constituir o centróide de uma nova classe, efetivando o desmembramento de uma classe em duas.

Por outro lado, grupos cujos centróides apresentem uma distância inferior a um limiar de fusão são unidos em um só grupo, cujo novo centróide assume o valor médio entre os centróides originais. A convergência é um dos principais problemas da abordagem, pois não é garantida. Para valores menores do coeficiente de difusão  $m$  a convergência é obtida; para valores maiores, o algoritmo pode não convergir.

Outros algoritmos de classificação de pontos podem, ainda, incorporar ao agrupamento difuso estimadores estatísticos robustos, sendo aplicáveis a dados ruidosos [Frigui & Krishnapuran (1996)].

A utilização de uma estrutura de árvore para a obtenção dos agrupamentos e conseqüente segmentação é proposta por Uchimura [Uchimura (1994)], a qual utiliza como espaço de representação uma derivação do espaço RGB. A nova representação é obtida normalizando-se as componentes RGB, convertendo-as para um vetor unitário na forma  $e_c = (e_r, e_g, e_b)$ , com  $e_c^2 = e_r^2 + e_g^2 + e_b^2$ . Como esta representação cria dependência entre as componentes, somente as duas primeiras são utilizadas. Um terceiro eixo ortogonal  $h$  é acrescido, correspondendo ao número de pontos da imagem

**Algoritmo 3-7. Simplificação da árvore de distribuição cromática [Uchimura (1994)].**

Remova da árvore todas as folhas conectadas diretamente a um nó de ramificação (Figura 3.11(b)).

Para cada nó de ramificação, traga-o para próximo à raiz, de forma que as ramificações fiquem no segundo nível da árvore (Figura 3.11(c)).

Reduza cada ramificação até o segundo nível da árvore (Figura 3.11(d)).

para cada  $(e_r, e_g)$ , na forma de histograma. Este novo espaço (ilustrado na Figura 3.9) mostra a distribuição de matiz da imagem, sendo que picos indicam ocorrências mais frequentes e vales indicam poucas ocorrências. Uma estrutura de árvore é definida neste espaço, iniciando-se no maior valor do eixo  $h$ ,  $h_{\max}$ . A partir deste valor, o eixo  $h$  é quantizado em  $n$  níveis, criando planos  $p(k)$ . As interseções do espaço com o plano definem contornos (curvas de nível). Para cada plano e cada contorno, um nó da árvore é criado, identificando o contorno. As conexões dos nós de planos adjacentes definem a estrutura da árvore, com apresentado na Figura 3.10.

Considerando que o espaço  $h, e_r, e_g$  representa a distribuição cromática da imagem, a simplificação da árvore associada ao gráfico equivale a uma segmentação ou agrupamento neste espaço. A simplificação da árvore adota os passos apresentados no Algoritmo 3-7, com ilustrações das simplificações apresentadas na Figura 3.11. Os nós remanescentes da simplificação indicam as divisões na imagem. Divisões excessivas, controladas por um limiar, são fundidas, reduzindo regiões muito pequenas.

### 3.2.2 Segmentação por técnicas de divisão/fusão e crescimento de regiões

Uma técnica baseada na divisão de regiões a partir da análise dos histogramas unidimensionais é proposta por Ohlander *et al.* [Ohlander *et al.* (1978)]. O princípio é, recursivamente, dividir regiões em regiões menores se for possível detectar, nos histogramas, a presença de características distintas (modos em pelo menos um dos

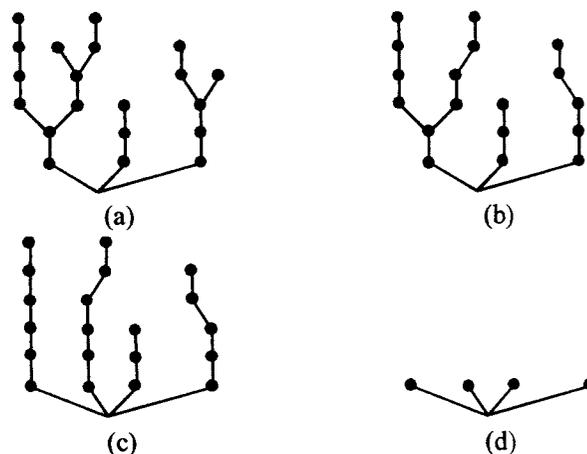


Figura 3.11. Redução da árvore. (a) Árvore inicial. (b) Após a remoção das folhas conectadas às ramificações; (c) “Redução” do nível dos nós de ramificação; (d) Redução das ramificações (árvore final).

Algoritmo 3-8. Segmentação por divisão de regiões através da análise de histogramas [Ohlander *et al.* (1978)].

```

Repita
  Selecione a próxima região a ser dividida.
  Calcule os histogramas unidimensionais.
  Se há um melhor pico entre os histogramas então:
    Aplique a divisão por limiar.
    Selecione regiões conectadas
  Até que não haja mais regiões divididas.

```

histogramas). Os histogramas são calculados para os dados da região, para um conjunto de atributos envolvendo as componentes  $R$ ,  $G$  e  $B$ ,  $H$ ,  $S$  e  $I$ , além de  $Y$ ,  $I$  e  $Q$ . Cada histograma é analisado em busca do melhor pico, caracterizado por mínimos bem definidos em ambos os lados. Se houver modos característicos, os dados da região podem ser, então, divididos segundo um valor de limiar determinado pela separação de picos, gerando duas sub-regiões. As novas regiões são tratadas para a eliminação de pequenos buracos e conexões finas por um operador de suavização ou operações de expansão e contração\* da região. Ambas as novas regiões são incluídas na lista de regiões a serem analisadas para futuras divisões, considerando-se que um critério de tamanho mínimo para cada região deva ser estabelecido. O Algoritmo 3-8 apresenta os passos para a divisão das regiões.

Utilizando o mesmo procedimento, Ohta *et al.* [Ohta *et al.* (1980)] utilizam, para os histogramas, três novos componentes derivados a partir da análise de componentes principais (pela transformada de Karhunen-Loève [Gonzales & Woods (1992)]) aplicada às componentes do espaço de cor utilizado. Adicionalmente os autores apresentam resultados do método comparando vários espaços de representação de cor†.

Outro algoritmo que considera a fusão e divisão de regiões, dado que se disponha das distribuições de probabilidade para algumas classes conhecidas, é proposto por Gorte-Kroupnova & Gorte [Gorte-Kroupnova & Gorte (1996)]. A partir das informações disponíveis, que incluem as densidades de probabilidade condicional  $P(x | C_i)$  mas não as probabilidades incondicionais  $P(x)$  ou *a priori*  $P(C_i)$ , são estimadas estas probabilidades por meio de uma pré-classificação pelo algoritmo dos  $k$  vizinhos mais próximos. Estimadas as probabilidades para as classes conhecidas é criada a classe “desconhecida” (sobre a qual não se dispõe de informação prévia), cuja probabilidade é definida de forma que as somas das probabilidades das classes seja 1. Pela regra de Bayes podem ser estimadas, também, as probabilidades *a posteriori*. Para cada região são calculadas as probabilidades *a posteriori* para cada classe e um critério de confiança definido pela diferença entre as duas maiores probabilidades. A

\* Como operadores morfológicos de dilatação e erosão.

† A seção 2.3.3.8 apresenta informações complementares sobre espaços de cor relativos aos resultados da análise de segmentação realizada.

fusão de regiões ocorre para regiões de alta confiança que compartilhem maior probabilidade para a mesma classe. Se as maiores probabilidades forem relativas à classe desconhecida ou se houver baixa confiança, um critério baseado em características espectrais (usando o espaço de cor HSI) é empregado para decidir sobre a fusão.

Mesmo não sendo exatamente uma abordagem de crescimento de regiões ou utilizando divisão e fusão de regiões, Hurlbert [Hurlbert (1989)] propõe um “homogeneização” das cores usando bordas detectadas em cada banda como restrição no processo. O espaço utilizado contém as duas componentes de cromaticidade  $r$  e  $g$ , além de uma de intensidade  $I$ , dada pela soma das componentes  $R$ ,  $G$  e  $B$ . O algoritmo a ser aplicado à componente cromática  $r$ , por exemplo, é iterativo e dado pela Equação 3.52, na qual  $N(r_{i,j}^t)$  é o conjunto dos  $n(N^t)$  pontos entre os vizinhos de  $(i, j)$  que diferem de  $r_{i,j}^t$  em menos que um limiar e que não são cruzados por uma borda em qualquer dos mapas de bordas.

$$r_{i,j}^{t+1} = \frac{1}{n(N^t)} \sum_{l,m \in N(r_{i,j}^t)} r_{l,m}^t \quad (3.52)$$

Em outras palavras, em cada iteração o valor do ponto é substituído pela média de sua circunvizinhança, desde que esta vizinhança não ultrapasse uma borda. Os mapas de bordas são obtidos através de algoritmos convencionais aplicados às imagens  $r$ ,  $g$  ou  $I$ . Um ponto da vizinhança contribui se é um dos 4 vizinhos de  $(i, j)$ , se tem o mesmo rótulo de borda que o ponto central, se seu valor é não nulo e se seu valor está dentro de uma faixa próxima do ponto central.

Como efeitos do algoritmo há, primeiramente, a suavização das cores após muitas iterações, deixando as regiões mais homogêneas. Adicionalmente as bordas consideradas evitam que se faça a média de uma cor com outra de outro lado da borda, impedindo a borradura nestas regiões. Finalmente, os cálculos sucessivos das médias locais acabam reduzindo bastante as variações causadas por reflexos especulares. Porém, através do número elevado de iterações, o custo do algoritmo é também alto.

### 3.3 Segmentação por aprendizado competitivo e redes neurais

O aprendizado competitivo também pode ser empregado para a divisão dos dados em grupos [Uchiyama & Arbib (1994)], podendo-se chegar a um conjunto de vetores que correspondem à quantização do espaço dos dados (e aos centróides dos grupos). O algoritmo de determinação dos vetores quantizados (Algoritmo 3-9) gera unidades até um máximo de  $n$ , que é o número máximo de grupos. A notação utilizada define  $N_R$  como o contador do número de repetições,  $N_{máx}$  como o número máximo de repetições,  $v_t$  como o limiar de vezes e  $\gamma$  a taxa de aprendizado.

$$\Delta w_{vencedor} = \gamma(x - w_{vencedor}) \quad (3.53)$$

## Algoritmo 3-9. Segmentação por aprendizado competitivo [Uchiyama &amp; Arbib (1994)].

Inicie o número de repetições em zero. Defina uma unidade, com vetor de peso  $w$  de valor igual à média de todos os dados e um campo de contador, que indica o número de vezes que a unidade é vencedora no processo.

Realize o aprendizado competitivo seguindo os passos:

- Selecione aleatoriamente dos dados um vetor  $x$ , selecionando (por um critério de distância euclidiana) a unidade com peso mais próximo, indicando-a como vencedora.
- Atualize o peso da unidade vencedora usando a Equação 3.53 e incremente seu contador de 1.
- Caso o contador atinja valor  $v_i$  e o número de unidades seja menor que  $n$ , então crie uma nova unidade, com vetor de peso igual ao da unidade vencedora. Atribua zero aos contadores de ambas unidades.
- Incremente  $N_R$  de 1 e repita o processo.

O agrupamento (classificação) é feito de forma conforme as expressões da Equação 3.54, criando os grupos  $C_i$ .

$$x_j \in C_i \text{ se } \|x_j - w_i\| < \|x_j - w_k\|, \forall k, k \neq i \quad (3.54)$$

A determinação de  $n$  é feita previamente. A relação  $N_{\max} \geq (2n - 3)v_i$  garante que o algoritmo gere pelo menos  $n$  unidades. (A derivação da expressão pode ser encontrada em [Uchiyama & Arbib (1994)].) Quanto maior o valor de  $v_i$ , maior a acurácia proporcionada mas um número maior de repetições é exigido.

Nesta linha de aprendizado competitivo, a segmentação pode ser obtida pelo sistema neural conhecido como “*learning vector quantizer*” (LVQ) [Anderson (1995); Fausett (1994); Hertz *et al.* (1991)], que também opera de forma não supervisionada. Nesta abordagem, o número de nós da rede é determinado previamente, tendo a seus pesos internos atribuídos valores aleatórios. Um dado conjunto de amostras é extraído dos dados da imagem, cada uma delas alimentando a rede na fase de treinamento. Para cada amostra, é localizado o nó cujo peso interno seja mais próximo dos atributos do vetor de entrada, sendo denominado nó vencedor e tendo seu peso modificado para que se aproxime do dado de entrada (de forma similar à expressa na Equação 3.53). Após o treinamento, que também assume que a taxa de aprendizado representada por  $\gamma$  decai de forma monótona com o tempo, cada nó da rede está associado a um padrão extraído do conjunto de dados. A classificação se dá conforme a resposta da rede para um padrão de entrada na fase de reconhecimento. A melhor resposta é considerada aquela relativa ao nó de resposta máxima, ou seja, aquele cuja distância de seu peso interno ao dado de entrada seja a menor entre todos os nós. Auge e Pitas [Auge & Pitas (1995)] propõem a utilização do LVQ utilizando como atributos as componentes RGB extraídas dos pontos de uma região  $3 \times 3$  para cada posição da imagem submetida como dado para aprendizado, perfazendo um vetor com 27 dados. Vetores criados desta forma alimentam os nós de um primeiro nível de uma rede de dois níveis, os quais, por sua vez, alimentam o segundo nível com 16

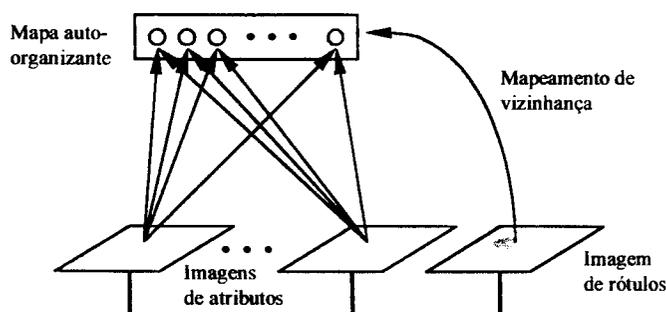


Figura 3.12. Esquema de rede neural para segmentação de imagens coloridas (adaptado de Zheng [Zheng (1995)]).

nós, ou seja, com 16 classes finais. Um critério baseado na variância é utilizado nesta abordagem para fundir classes similares.

Mapas auto-organizáveis podem ser considerados extensões das redes LVQ com a inclusão do princípio múltiplos nós vencedores e serão abordados com mais detalhes no Capítulo 4. São utilizados por Zheng [Zheng (1995)] para a segmentação de imagens coloridas, particionando-as em regiões. O paradigma proposto abrange duas etapas: o agrupamento através de mapas auto-organizáveis e uma descrição voltada para modelo. A segmentação, então, procura determinar regiões conexas de pontos que apresentem cores e texturas similares. As regras de agrupamento envolvem:

- a) Pontos com cores similares tendem a ser agrupados (atributo de similaridade);
- b) Pontos com texturas similares tendem a ser agrupados (similaridade);
- c) Pontos geometricamente próximos a uma dada região tendem a ser agrupados a esta região (proximidade e continuidade);

As duas primeiras regras são consideradas na característica auto-organizável da rede neural empregada, enquanto a terceira regra é incorporada através da definição de uma imagem de rótulos, a qual considera uma pequena região em torno da posição do ponto que está sendo classificado.

A Figura 3.12 apresenta a estrutura da rede neural empregada na segmentação. As imagens de atributos são compostas, cada uma, de atributos relevantes da imagem colorida original e a imagem de rótulos indica a rotulação associada a cada ponto em particular. O processamento é “sincronizado”, isto é, os atributos e rótulos considerados em um dado momento são relativos a uma mesma posição da imagem.

O mapa auto-organizável é composto de nós, a cada qual está associado um vetor peso relacionado aos atributos extraídos da imagem. Há uma conexão de cada dado de entrada para todos os nós do mapa, os quais são ativados de forma competitiva. Uma ligação entre a imagem de rótulos e os nós é utilizada para indicar quais as unidades que podem competir para um determinado dado. Somente nós que compartilham o mesmo rótulo dos pontos situados em uma pequena região ao redor das coordenadas sendo consideradas podem ser disparados no processo de competição. Os atributos considerados são as componentes do espaço de Ohta [Ohta *et al.* (1980)] e

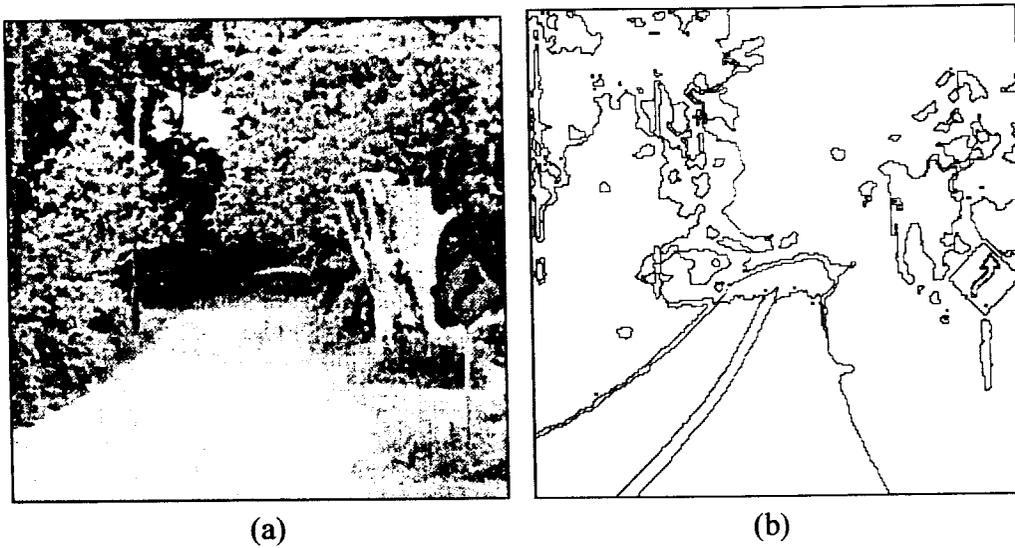


Figura 3.13. Segmentação de imagem colorida. (a) Versão monocromática da imagem colorida. (b) Resultado da segmentação considerando cor e textura. [Zheng (1995)]

medidas locais de textura (entropia e energia) aplicadas a duas das componentes espectrais.

O treinamento da rede é feito apresentando intensivamente os atributos da imagem e, a cada entrada, o nó vencedor tem seus pesos internos atualizados, tornando-o mais próximo ao vetor de entrada. Seu rótulo é também atualizado com o mesmo rótulo da posição correspondente na imagem de rótulos. Após o treinamento é possível associar a cada ponto da imagem o rótulo apropriado. A Figura 3.13 mostra um exemplo de resultado de segmentação através deste procedimento.

Outras considerações abrangendo redes neurais compreendem a segmentação de imagens de células para identificação de células cancerígenas através de redes de Hopfield [Rachid *et al.* (1997)] e a identificação de cores de impressão (padrão CMYK\*) por segmentação da imagem impressa utilizando redes hierárquicas modulares [Verikas *et al.* (1997)]. Ambas as abordagens são, entretanto, muito específicas quanto a aplicação, o que limita sua utilização de forma mais geral. Não são, desta forma, detalhadas neste texto. O processamento de imagens coloridas pode também ser realizado por meio do uso de redes neurais celulares [Lee & Gyvez (1996)].

No que concerne aos mapas auto-organizáveis tem-se ainda recentemente o aparecimento de novas aplicações e variações (sejam aperfeiçoamentos ou expansões) [Kohonen (1996); Sirosh & Miikkulainen (1997); Babu (1997)], indicando sua importância.

### 3.4 Abordagens baseadas em propriedades físicas da cena

Uma abordagem bastante específica para o tratamento de imagens coloridas é através do estabelecimento de um modelo de iluminação, através do qual se podem

\* CMYK: *cyan, magenta, yellow e black.*

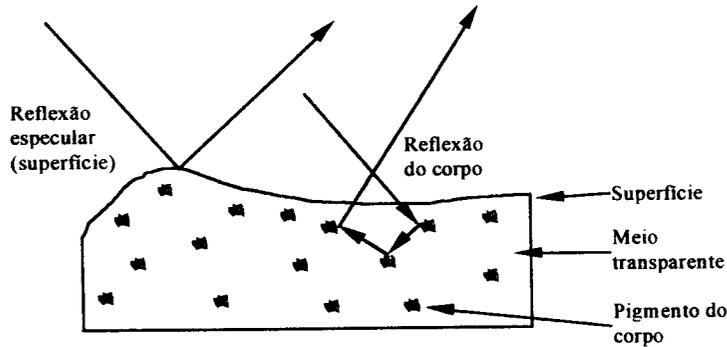


Figura 3.14. As duas componentes de reflexão do modelo de reflexão dicromática: especular e do corpo.

levantar hipóteses sobre como cada parte da imagem foi formada. Baseado nas propriedades físicas de reflexão dos materiais dielétricos\*, o “modelo de reflexão dicromática” foi proposto por Shafer [Shafer (1985)]. O modelo assume que cada objeto reflete a luz incidente em duas componentes distintas. A primeira é relativa às características materiais do corpo e a segunda relativa à um reflexo especular da superfície (Figura 3.14), ambas dependentes da natureza do objeto. A luz refletida  $L(\lambda, i, e, g)$  é uma composição entre a luz do corpo do objeto  $L_c(\lambda, i, e, g)$  e da reflexão da superfície  $L_s(\lambda, i, e, g)$ . O comprimento de onda é indicado por  $\lambda$  e os parâmetros  $i$ ,  $e$  e  $g$  são relativos aos ângulos de luz incidente, luz emitida e fase, respectivamente. É feita uma separação entre as propriedades geométricas e distribuições espectrais, na forma apresentada na Equação 3.55.

$$\begin{aligned} L(\lambda, i, e, g) &= L_s(\lambda, i, e, g) + L_c(\lambda, i, e, g) \\ &= m_s(i, e, g)c_s(\lambda) + m_c(i, e, g)c_c(\lambda) \end{aligned} \quad (3.55)$$

Para um objeto, cuja cor possa ser descrita pela combinação das componentes espectrais  $c_s(\lambda)$  para a superfície e  $c_c(\lambda)$  para o corpo ao longo de sua extensão, há valores diversos para  $m_s(i, e, g)$  e  $m_c(i, e, g)$ . Verificando os agrupamentos formados por essas duas últimas componentes é possível identificar cores homogêneas e separá-las de seus reflexos especulares [Klinker *et al.* (1990)]. Pelas restrições do modelo, os objetos devem ser necessariamente convexos e compostos de um único material de cor uniforme. Maxwell e Shafer [Maxwell & Shafer (1993); Maxwell & Shafer (1995)] propuseram uma representação expandida do modelo, incorporando representações por pequenas áreas e permitindo a interpretação de objetos compostos de regiões de cores diferentes (com várias cores). Tanto Klinker *et al.* como Maxwell e Shafer abordam a segmentação através da geração de hipóteses segundo o modelo proposto, avaliando quais as mais prováveis para a representação dos dados da imagem e indicando como interpretar (e segmentar) os objetos constituintes. Como em

\* Incluem tintas, plásticos, acrílicos, cerâmicas e papel [Maxwell & Shafer (1993)].

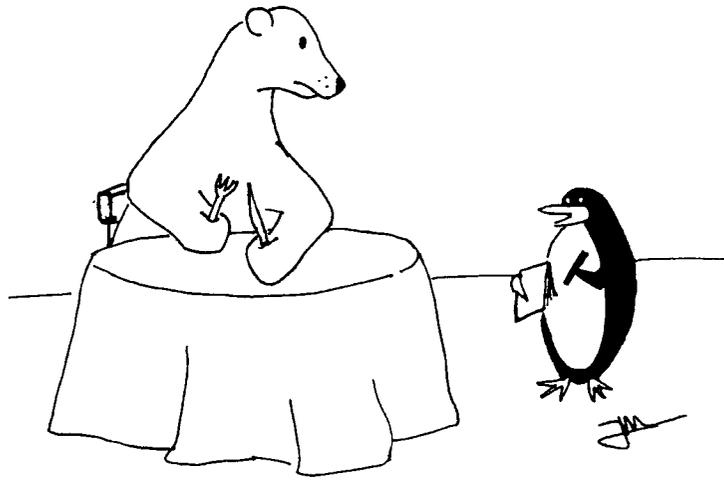


Figura 3.15. O urso polar e o pingüim.

geral a quantidade de hipóteses é grande e muitas delas serem plausíveis ao mesmo tempo, é empregada a redução das possíveis hipóteses e diminuição das classes de imagens.

### 3.5 Os pingüins e os ursos polares

Às vezes, as pessoas se deparam com charadas formadas por perguntas de aparente simplicidade, mas cuja solução nem sempre é direta ou rápida. Seja, por exemplo, a seguinte questão: “Habitando as regiões mais frias do globo, ursos polares encontram-se plenamente adaptados a sobreviver nestas áreas inóspitas. Para os pingüins, que também habitam mares e costas geladas, esta adaptação é também evidente. Por que, entretanto, focas fazem parte do cardápio dos ursos polares mas os pingüins são excluídos?” O raciocínio direto normalmente leva a questionamentos sobre os pontos mais óbvios, como levantar hipóteses sobre os ursos não gostarem da carne dos pingüins, destes terem gosto ruim ou serem espertos o suficiente para não serem surpreendidos. Através da quebra desta forma de raciocínio (*insight*) [Stenberg & Davidson (1995)] é possível estabelecer, após algum tempo (nem sempre curto), que a razão é a distância. Enquanto ursos polares são habitantes do pólo norte, pingüins habitam o pólo sul; hábitats similares com o restante do globo os separando.

Da mesma forma podem ser vistos os métodos e técnicas do vasto conjunto abordado neste capítulo. Embora a organização deste elenco possa ser estabelecida em termos de técnicas voltadas a bordas, voltadas a agrupamentos ou outras a classificação, não são estabelecidas, em geral, as inter-relações entre elas. Técnicas pingüins e técnicas ursos aplicam-se sobre os dados, mas não “interagem”. Não se permite, assim, estabelecer conexões entre as técnicas, que se comportam como elementos isolados em seu próprio ambiente, “ignorando” a existência dos outros. Desta forma não há uma metodologia trivial que permita comparar tais técnicas, avaliando seus desempenhos relativos, seja em termos de qualidade da segmentação, seja em termos de custo de cômputo.

*"Sua deficiência em cavalarias de armamento leve os privou, naqueles dias desprovidos mapas, da obtenção da tão necessária informação."*  
T. Hodgkin — "Dynasty of Theodosius"

## 4 Segmentação de imagens por mapas auto-organizáveis

A segmentação, essencial nas primeiras etapas do processo de visão [Goldstein (1989)], está representada no sistema de visão *Cyvis-1* em várias modalidades, desde detecção de bordas em níveis de cinza até segmentação de regiões por textura e cor. Dentro desta perspectiva, a segmentação de imagens utilizando cor se compõe de um dos módulos de processamento do sistema, cujo objetivo básico é produzir um mapa de regiões dos objetos presentes na cena (ou partes deles). A metodologia de segmentação proposta (Figura 4.1) faz uso de uma rede neural não supervisionada para a identificação das principais componentes cromáticas\* da imagem, gerando um conjunto representativo dos dados, os quais são separados em grupos, proporcionando a identificação das classes presentes na imagem. Dadas as classes, os pontos da imagem são associados a elas, o que proporciona uma segmentação em regiões. Um mapa de bordas pode ser derivado, então, a partir dos limites entre as regiões.

A rede neural utilizada, estruturada na forma de um mapa auto-organizável, é parte de uma classe de redes neurais dentre cujas características se destaca, no âmbito deste trabalho, a exclusão da supervisão no processo de aprendizado. Características adicionais deste tipo de rede envolvem a conexão lateral entre suas unidades e a influência de um dado de entrada sobre unidades vizinhas, proporcionando ao mapa uma estrutura que preserva uma relação topológica (isto é, de vizinhança) entre os pesos internos das unidades da rede. A ativação de uma das unidades faz com que unidades vizinhas também sejam ativadas, mas em menor intensidade, até que, a partir de uma dada distância, a influência do estímulo seja insignificante. Além da excitação de uni-

---

\* No contexto, *componente cromática* refere-se à composição da cor, em termos de seu matiz, sua saturação e sua intensidade.

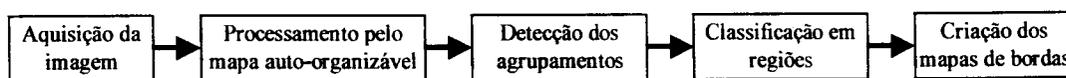


Figura 4.1. Esquema de processamento.

dades vizinhas, também pode haver inibição, em conformidade com a forma com que um estímulo afeta a rede.

Este capítulo descreve a metodologia adotada para a obtenção de elementos estruturais básicos (regiões e contornos) a partir de imagens coloridas, apresentando em detalhes a metodologia estabelecida para a segmentação da imagem e os mapas auto-organizáveis, com sua estruturação e características, além de seu emprego na identificação dos atributos da imagem. A técnica de agrupamento empregada é também exposta, com sua caracterização, seguida da forma de obtenção do mapa de regiões e conseqüente mapa de bordas. O mapa auto-organizável é apresentado na seção 4.1.1, enquanto a técnica de agrupamento é exposta na seção 4.1.2. Os resultados obtidos, assim como as considerações pertinentes, são apresentados na seção 4.2. A seção 4.3 discorre sobre as possibilidades da aplicação do método proposto. A comparação de desempenho do método proposto em relação à classificação bayesiana e à classificação por  $k$  vizinhos mais próximos [Duda & Hart (1973)] encontra-se na seção 4.4. Os comentários finais estão apresentados na seção 4.5.

#### 4.1 Caracterização da metodologia

O método de segmentação estabelecido é baseado na classificação dos pontos individuais da imagem para a obtenção das regiões presentes. Assume-se que as imagens sejam formadas por objetos ou partes de objetos, cujos atributos permaneçam aproximadamente constantes na extensão de sua superfície. Através da identificação dos pontos com atributos semelhantes, são estabelecidas sobre a imagem regiões associadas ao atributo e a conseqüente determinação das bordas das regiões (e dos objetos) se torna viável. [Moreira & Costa (1996)]

##### 4.1.1 Mapas auto-organizáveis

O cérebro é composto, em sua essência\*, por neurônios e suas interconexões. A interconexão completa entre todos os neurônios não é viável em termos físicos, de modo que a organização cerebral estabelece áreas especializadas, conectadas entre si. Na parte responsável pela visão (parte posterior do cérebro [Levine (1985); Goldstein (1989)]), existe, além de um mapeamento topológico entre a retina e as células no núcleo lateral geniculado ou do córtex, também variações suaves ao longo de superfícies corticais que respondem a um mesmo atributo. Isso ocorre, por exemplo, com a resposta a orientação. Neurônios vizinhos tendem a apresentar respostas

\* O cérebro também apresenta, em sua composição, as células gliais, responsáveis não somente pela sustentação, mas também pela absorção de neurônios mortos e fornecimento de substâncias nutritivas aos neurônios.

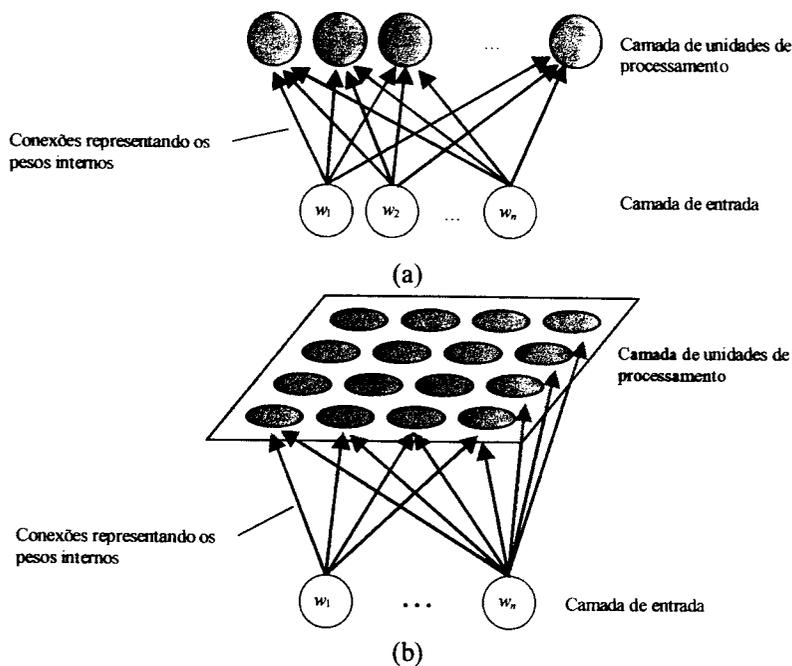


Figura 4.2. Esquema de interconexão de entradas com os neurônios. A interconexão é completa. (a) Unidimensional; (b) bidimensional (com as conexões parcialmente representadas).

similares, mas não iguais, a uma mesma orientação. As respostas, geralmente, variam de forma suave ao longo de sua superfície segundo mudanças na orientação do estímulo [Anderson (1995)]. Esta característica é incorporada na concepção dos mapas auto-organizáveis, representada pelas conexões entre unidades (neurônios) vizinhas.

A representação de atributos, os quais podem incorporar dimensionalidade alta, em redes usualmente bidimensionais implica, além da organização topográfica, em uma redução desta dimensionalidade, que pode ser associada à redução e representação mais abstrata da informação [Kohonen (1984); Anderson (1995)]. A presença desta característica está relacionada à interconexão lateral entre os neurônios, que compartilham suas informações com os vizinhos [Kohonen (1984)].

#### 4.1.1.1 Aspectos gerais

Os mapas auto-organizáveis [Kohonen (1984)], como já mencionado, compõem uma classe de redes neurais artificiais sem supervisão. Sua estruturação corresponde a uma rede linear ou planar, com uma única camada. Cada unidade da rede mantém seus pesos internos, relativos aos atributos de entrada. A Figura 4.2 ilustra a completa conexão das entradas para a rede com todas as células da única camada do mapa, independentemente do arranjo das unidades em uma ou duas dimensões. Cada unidade de processamento da rede, então, incorpora pesos internos relativos a cada uma das dimensões da entrada, usualmente escalares e contínuas. Além destas conexões com as entradas, as unidades mantêm conexões com suas vizinhas, as quais permitem que a ativação de uma unidade afete não somente ela própria, mas também

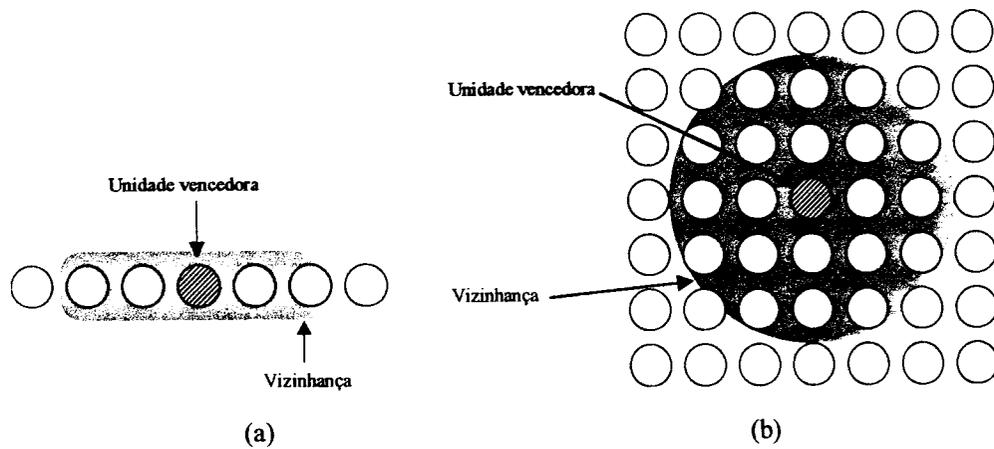


Figura 4.3. Vizinhança em mapas auto-organizáveis, indicada pela região escura envolvendo os nós. O nó hachurado representa o nó vencedor; os nós em cinza claro indicam os nós pertencentes à vizinhança; os nós em branco não são afetados pela ativação do nó vencedor. (a) Estruturação unidimensional do mapa; (b) Estruturação bidimensional com grade regular e vizinhança circular.

uma vizinhança ao seu redor. À medida que a distância da célula central aumenta, reduz-se o efeito causado.

Durante a fase de aprendizado, as amostras são submetidas à rede, que adapta seus pesos internos conforme a nova informação. Esta adaptação é realizada identificando-se, inicialmente, qual unidade apresenta a maior resposta ao dado de entrada, ou seja, a unidade cujo peso interno se aproxima mais do valor da entrada. Esta unidade, denominada *unidade vencedora*, terá seu valor interno modificado para que se aproxime do valor submetido à rede. Esta modificação pode ser realizada conforme a Equação 4.1, na qual  $w$  indica o peso interno de um nó (nos passos anterior e atual) e  $x$  corresponde a um vetor de entrada;  $\alpha$  é uma taxa de aprendizado, que varia entre 0 e 100%.

$$w_{\text{atual}} = w_{\text{anterior}} + \alpha(x - w_{\text{anterior}}) \quad (4.1)$$

A Equação 4.1 mostra que, cada vez que uma unidade é eleita vencedora, seus pesos se aproximam do dado de entrada, tornando-se cada vez mais representativos daquele valor. Em outras palavras, se um conjunto de dados de entrada varia em torno de um valor  $x$ , a submissão destes valores para o aprendizado da rede tende a fazer com que uma das unidades acabe por representar um valor médio deste conjunto, se ela for a vencedora para este conjunto de dados. Além disto, a rede incorpora também uma influência de vizinhança, de forma que unidades vizinhas propaguem seus pesos internos a outras unidades próximas. Desta forma, um conjunto de dados similares tende a incidir sobre uma região da rede, a qual passa a representar este conjunto em particular. Além disto, os mapas apresentam transições suaves de atributos no caminho que liga áreas que representam atributos distintos.

A regra de atualização dos pesos não se aplica, assim, somente às unidades vencedoras, mas também a uma vizinhança ao redor desta unidade, como a Figura 4.3

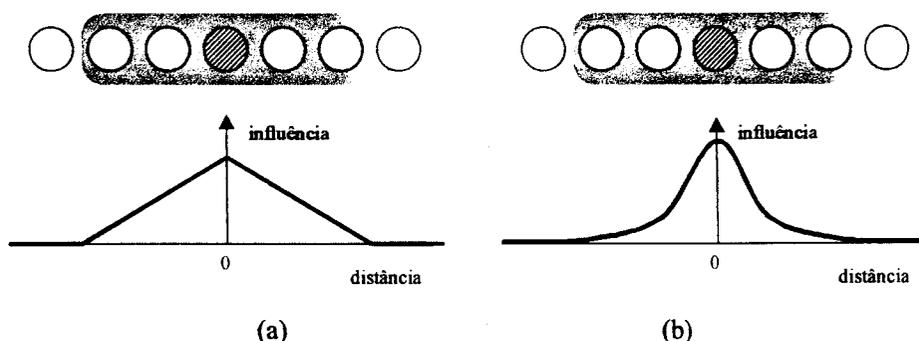


Figura 4.4. Esquemas de influência conforme a distância do nó vencedor. Decaimentos segundo uma função (a) linear e (b) gaussiana.

ilustra para os casos unidimensional e bidimensional. Quanto maior a distância da unidade vencedora, menor será a influência propagada por ela. Duas formas, entre várias possíveis, de se estabelecer a influência conforme a distância são ilustradas na Figura 4.4. Uma função que decai linearmente conforme a distância do nó é representada na Figura 4.4(a), enquanto na Figura 4.4(b) a influência decai segundo uma função gaussiana. A influência, a partir de uma dada distância, é nula. A extensão para o caso bidimensional é direta, correspondendo a um cone para a função linear e a uma gaussiana bidimensional para o segundo exemplo. As variâncias das gaussianas determinam sua maior ou menor abertura ou, em outras palavras, a restrição ou não da influência à região central da vizinhança. Outra função que também pode ser utilizada é a do laplaciano da gaussiana (o “chapéu mexicano”) [Jain *et al.* (1995)].

Considerando ainda o caso bidimensional, a vizinhança não necessita ser, forçosamente, circular, podendo-se adotar quaisquer outros critérios de distância, como por exemplo a distância do quarteirão ou *chessboard*. Quando a vizinhança atinge a borda do mapa, podem ser adotadas as estratégias de truncar a influência ou de se criar um “hipermapa” no qual a vizinhança se propaga para a borda oposta, tornando-o “fechado” em todas as direções, como ilustrado na Figura 4.5.

O efeito causado pela influência de vizinhanças é o estabelecimento da relação topológica sobre o mapa, no qual atributos similares tendem a estar próximos no espaço. Quanto mais próximos dois vetores de entrada, mais próximos estarão os nós da rede que responderão a eles.

Para que se obtenha a convergência do mapa para um estado estável durante a fase de aprendizado, utiliza-se a estratégia de reduzir sistematicamente tanto a vizinhança de influência quanto a taxa de aprendizado  $\alpha$ , introduzidos na Equação 4.1. Assim, o tamanho da vizinhança (denotada por  $N$ ) e a taxa de aprendizado tornam-se variáveis com o tempo (Equação 4.2). [Kohonen (1984)]

$$\begin{aligned}\alpha &= \alpha(t) \\ N &= N(t)\end{aligned}\tag{4.2}$$

Algoritmo 4-1. Fase de aprendizado para a formação do mapa auto-organizável. (Adaptado de Fausett [Fausett (1994)].)

```

Ajuste os valores iniciais dos pesos dos nós.
Ajuste os valores iniciais para a vizinhança.
Ajuste a taxa de aprendizado inicial.
Repita
  Determine o nó  $j$  que proporciona maior resposta à entrada  $x$ .
  Para cada nó dentro da vizinhança, ajuste os valores internos se-
  gundo a Equação 4.1.
  Atualize a taxa de aprendizado.
  Reduza a vizinhança.
Até que a condição de parada seja satisfeita.
  
```

Ambos os parâmetros, vizinhança e taxa de aprendizagem, devem ser decrescidos de forma gradativa e monotônica, até que a vizinhança compreenda somente o nó vencedor e a taxa de aprendizado se aproxime de zero. A escolha da função de decaimento é arbitrária, podendo ser linear, exponencial ou outra que mantenha as características citadas. Um possível critério para determinar a convergência do mapa é a verificação se, entre dois passos sucessivos (duas unidades de tempo), a diferença geral dos pesos é menor que um dado valor arbitrário de erro  $\epsilon$ , determinando sua estabilidade.

O Algoritmo 4-1 apresenta uma visão geral do procedimento de aprendizado. Os valores iniciais para os pesos internos das unidades são usualmente aleatórios, mantendo-os dentro do intervalo de variação dos dados de entrada. Havendo conhecimento prévio da estrutura dos agrupamentos, em termos de expectativa, é possível a atribuição de valores especiais, iniciando com antecedência uma tendência na distribuição topológica do mapa.

São determinadas, então, a taxa de aprendizado inicial, bem como o raio inicial para a vizinhança. Dentro da vizinhança há, ainda, uma diferenciação entre as unidades, que são influenciadas distintamente conforme sua distância do nó vencedor; quanto mais próximas à periferia da vizinhança, menor o grau de ajuste dos pesos. Há,

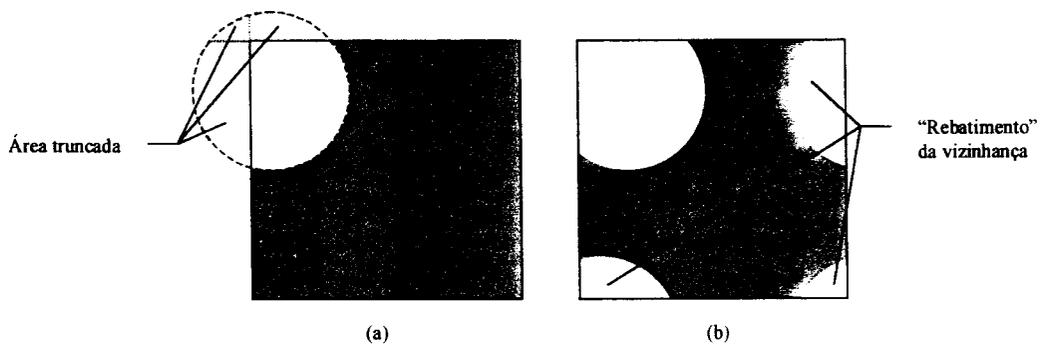


Figura 4.5. Vizinhança circular (em branco): (a) influência truncada além das bordas do mapa; (b) estruturada na forma de um hipermapa.

finalmente, o estabelecimento da condição de parada para o algoritmo, que pode ser feita através de um critério de convergência ou ainda através da determinação de um máximo para o número de entradas que podem ser submetidas no aprendizado.

Neste último caso não há qualquer garantia de convergência em relação à dependência da taxa de aprendizado e tamanho da vizinhança, mas resultados experimentais mostram que a formação básica do mapa é rápida em relação ao tempo até a convergência. Em outras palavras, embora a chegada ao estado de convergência seja feita lentamente, o mapa tende a adquirir uma forma topológica próxima da forma final rapidamente. A partir desta estruturação inicial, embora cada nova entrada seja significativa, a disposição espacial dos atributos representados nos pesos das unidades tende a não se modificar.

A rede é alimentada por meio da submissão um a um dos dados, os quais podem ser escolhidos ordenadamente, aleatoriamente, uma ou várias vezes cada amostra. A utilização de todos os dados, usando um critério de ordenação ou não, implica, para um grande número de amostras (imagens facilmente atingem 65.000 amostras\*), em um tempo de processamento elevado. Uma opção interessante que não privilegia qualquer organização preexistente dos dados é a escolha aleatória. Considerando que a seleção (pseudo)aleatória não apresente vícios, a probabilidade de que qualquer ponto seja escolhido é a mesma e, assim, todas as classes podem ter seus dados submetidos à rede.

Embora o resultado, após a convergência, seja a distribuição dos atributos no mapa, respeitando suas inter-relações, a rede em si não realiza um procedimento de agrupamento. Não ficam determinados quais são os grupos presentes nem seu número. A determinação dos grupos (ou classes) presentes requer um passo adicional de processamento, descrito na seção 4.1.2.

#### **4.1.1.2 Aspectos de implementação**

Embora os mecanismos envolvidos na estruturação dos mapas auto-organizáveis não sejam tão complexos, a implementação exige a escolha de determinados procedimentos para a rede neural, definindo seu padrão de comportamento.

A implementação está especificamente direcionada para segmentação de imagens coloridas através de suas características de cromaticidade. Dentro deste enfoque, a rede foi estruturada em um arranjo planar bidimensional completo (Figura 4.6), com unidades de processamento espalhadas uniformemente segundo uma grade regular. As dimensões da rede são um parâmetro do sistema, não sendo exigido que a altura ou largura do arranjo sejam iguais. Permite-se, assim, que o mapa seja unidimensional linear, definindo-se apenas uma linha ou apenas uma coluna na rede.

Os pesos internos a cada nó representam, em princípio, as componentes de cromaticidade envolvidas ( $r$ ,  $g$  e  $b$ , normalizadas a partir das componentes do espaço RGB), de modo que cada nó da rede armazena um vetor de conexão de três dimen-

\* Uma imagem quadrada com 256 pontos de lado perfazem 65.536 amostras individuais.

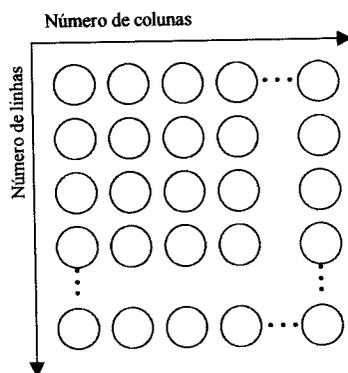


Figura 4.6. Estruturação bidimensional regular da rede que representa o mapa auto-organizável.

sões. Como as componentes são normalizadas (Equação 4.3), somente duas delas são necessárias, visto que a terceira pode ser obtida a partir das outras duas ( $g = 1 - r - b$ , por exemplo).

$$\begin{cases} r = \frac{R}{R + G + B} \\ g = \frac{G}{R + G + B} \\ b = \frac{B}{R + G + B} \end{cases} \quad (4.3)$$

A forma escolhida para implementar a vizinhança é a de um círculo cujo centro se situa na posição do nó vencedor. O parâmetro para a vizinhança é seu raio, que determina a extensão de influência de uma determinada amostra considerando-se as conexões laterais. O decaimento do raio obedece uma taxa predefinida, estabelecida por uma porcentagem que é subtraída a cada unidade de tempo, conforme a Equação 4.4, na qual  $\rho$  é o raio da vizinhança e  $p_\rho$  é a taxa de decaimento [Kohonen (1984)]. Esta taxa de decaimento é exponencial, como se exemplifica na Figura 4.8, na qual o valor inicial é 1 e a taxa de redução é de 1% a cada unidade de tempo. Valores usuais para as taxas usadas na rede situam-se na faixa de 0,05 a 1,2%, tornando a redução

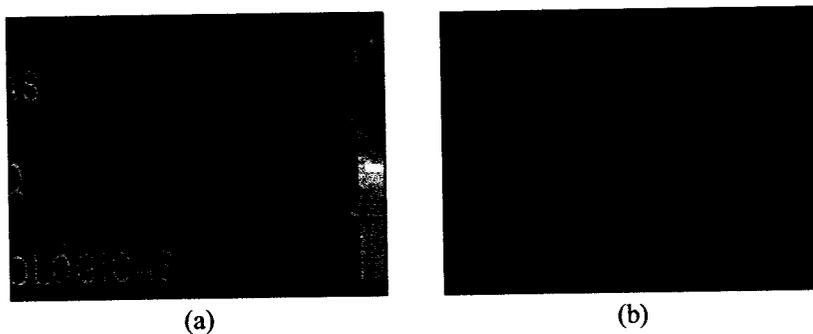


Figura 4.7. (a) Imagem original. (b) Representação da imagem com cores normalizadas (cromaticidade).

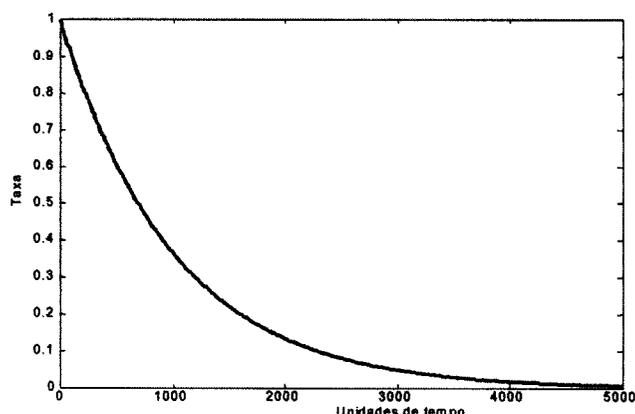


Figura 4.8. Ilustração do decaimento percentual considerando-se uma taxa de redução de 1%, ao longo de 5000 unidades de tempo.

bastante lenta tanto para o raio da vizinhança quanto para a taxa de aprendizado. A influência da taxa de aprendizado diminui conforme aumenta a distância do centro do círculo da vizinhança, variando linearmente, conforme ilustra o perfil apresentado na Figura 4.4(a).

$$\rho_{t+1} = (1 - p_p)\rho_t \quad (4.4)$$

#### 4.1.1.3 Exemplos

O comportamento dos mapas auto-organizáveis depende de um grande conjunto de parâmetros, que variam desde as dimensões da rede até as taxas de decréscimo adotadas para o tamanho da vizinhança e taxa de aprendizado, passando neste intermédio pela escolha dos atributos que representam os dados de entrada. Esta seção tem como objetivo exemplificar o comportamento geral destes mapas, ilustrando alguns pontos relevantes.

A Figura 4.9 mostra, em alguns instantes do processo de aprendizado, o estado do mapa. Como conjunto de dados utilizado para alimentar a rede foram escolhidos os pontos que compõem a imagem da Figura 4.9(a), obtida em laboratório. O mapa foi dimensionado com tamanho de  $10 \times 10$  unidades, cada uma com pesos iniciados com valores aleatórios. A camada de entrada é configurada para aceitar os valores RGB das amostras. A cada passo do processo, um ponto da imagem é escolhido aleatoriamente e submetido à rede, a qual ajusta seus pesos internos. A Figura 4.9 (b) mostra o estado das unidades da rede após 100 passos, sendo que cada unidade é representada por um quadrado na cor equivalente ao valor RGB armazenado em suas conexões.

A evolução do estado da rede para tempos 500, 1000, 4000, 5000, 10000 e 24460 está apresentada nas Figuras 4.9(c) a (h), respectivamente. O último estado indica os valores após o critério de convergência ter sido satisfeito, isto é, após a modificação total da rede entre passos sucessivos ter sido menor que  $\epsilon = 0,001$ . Alguns

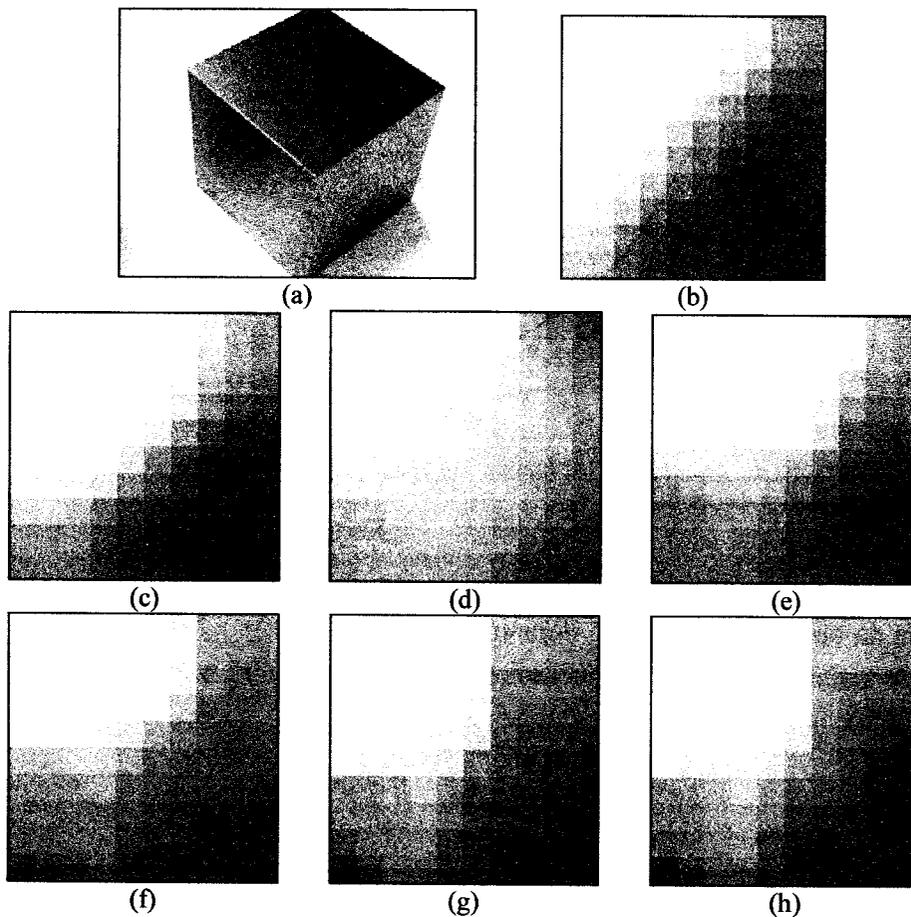


Figura 4.9. Evolução da configuração de um mapa auto-organizável de tamanho  $10 \times 10$ . Cada unidade da rede armazena como pesos os valores RGB, representada graficamente pela cor correspondente. (a) Imagem  $180 \times 240$  pontos. (b)-(g) Evolução dos valores do mapa para tempos 100, 500, 1000, 4000, 5000 e 10000, respectivamente. (h) Mapa após a convergência (24460 passos).

dos parâmetros principais da rede foram: vizinhança inicial com raio de valor 8, taxa de aprendizado inicial de 60%, ambos com taxa de decaimento de 0,02% a cada iteração.

A distribuição espacial (topológica) dos atributos já pode ser notada no início do processo de aprendizado, na Figura 4.9(b), após os 100 primeiros passos. As posições das cores branca, azul, amarela e verde não mais se alteram, modificando-se apenas seus limites, embora as unidades que representam os valores azuis ainda não estejam muito bem definidas. No restante das representações do mapa é possível notar os ajustes topológicos, fazendo com que a rede progrida para um estado estável.

A Figura 4.10 ilustra o mesmo processo de aprendizado para outra imagem, com mapa com estruturação de atributos similares, usando um mapa de dimensões menores, com  $5 \times 5$  unidades. Os valores das taxas de decaimento, taxa inicial de aprendizado, condição de convergência, entre outros, são iguais aos do exemplo anterior, com exceção do raio inicial da vizinhança, ajustado para 4. A imagem tem as mesmas dimensões e mostra uma cena externa, com maior complexidade de informa-

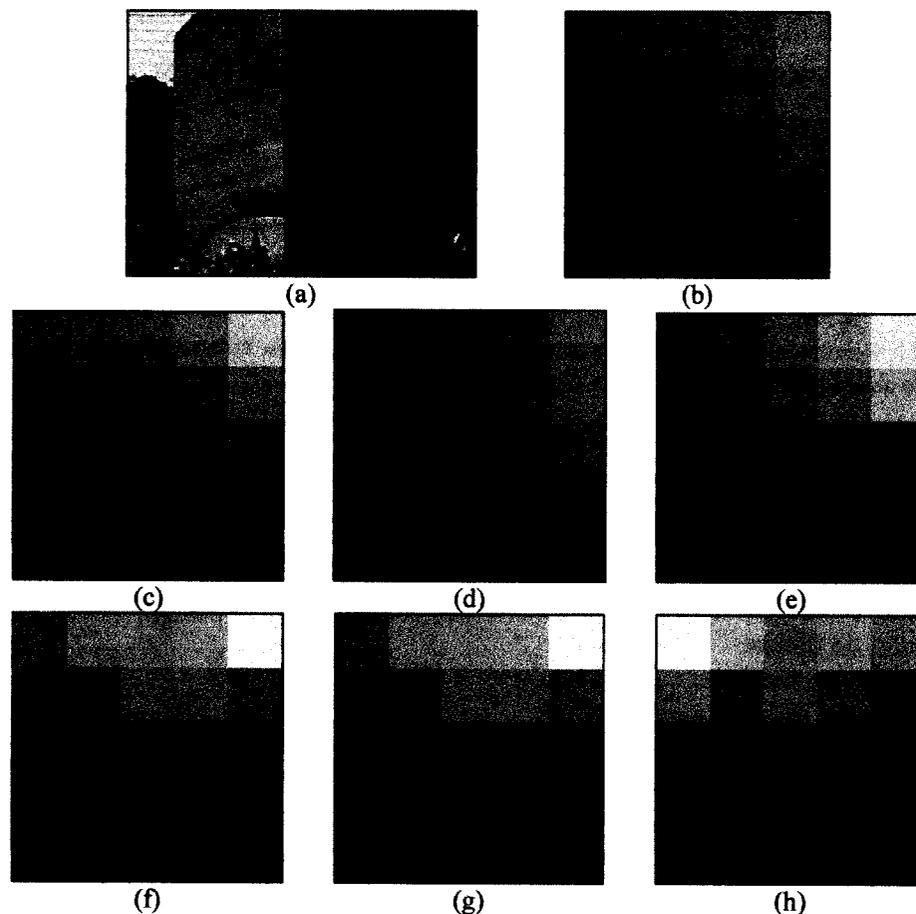


Figura 4.10. Evolução da configuração de um mapa auto-organizável de tamanho  $5 \times 5$ , com representação similar à Figura 4.9. (a) Imagem utilizada com tamanho de  $180 \times 240$ . (b)-(g) Evolução dos valores do mapa para tempos 100, 500, 1000, 4000, 10000 e 28042, respectivamente. (h) Estado final para outra execução do procedimento (29632 passos até a convergência).

ção e maior variação de cores. A contribuição do azul do céu da Figura 4.10(a) somente pode ser distinguida com mais nitidez na Figura 4.10(f), após 10000 amostras.

Detalhes mais específicos também são mais evidentes no final do processo, quando se podem notar alguns tons de verde, o azul acinzentado e alguns tons alaranjados, presentes como ruído na parte mais clara da parede, mostrado na Figura 4.10(g). Partindo-se de outra configuração inicial, a Figura 4.10(h) mostra o resultado final para uma segunda execução do procedimento. Pode ser observado que a maioria das cores representadas no mapa da Figura 4.10(g) também estão aparentemente representadas nesta nova configuração, apesar das posições não terem sido mantidas. Deve-se, ainda, tomar o cuidado de não afirmar que os valores finais são os mesmos, dado que a percepção humana para a distinção visual das cores tem suas limitações.

A Figura 4.11 mostra exemplos de mapas unidimensionais. Cada mapa tem 16 unidades e foi treinado a partir dos dados RGB selecionados aleatoriamente da imagem correspondente. Nota-se, nos mapas, que a progressão dos atributos internos (as cores, no caso) ocorre de forma suave ao longo de nós adjacentes, além dos atributos similares tenderem a ficar próximos em termos de distribuição espacial. Uma utiliza-

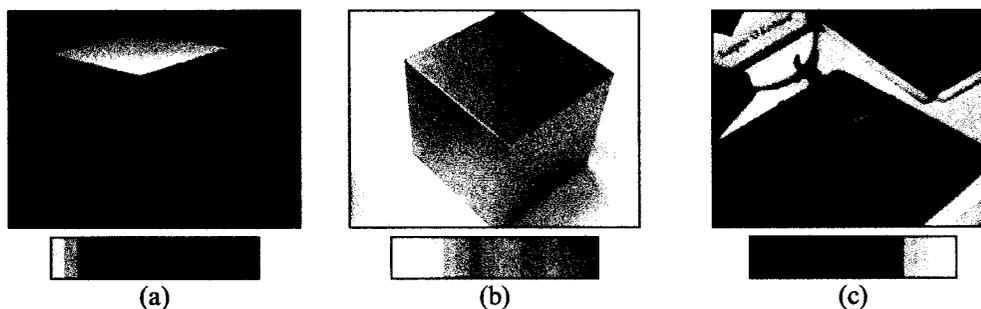


Figura 4.11. Resultados finais para mapas lineares, com 16 unidades, considerando as componentes RGB como dados para os pesos internos dos nós.

ção direta para mapas simples é a redução do número de cores de uma imagem para efeito de compressão, por exemplo [Auge & Pitas (1995)]. Cada ponto da imagem pode ser substituído pelo conteúdo da unidade que mais se assemelha a ele. Godfrey e Attikiouzel [Godfrey & Attikiouzel (1992)] também utilizam redes neurais para a compressão de dados.

#### 4.1.2 Técnica de agrupamento

Os mapas auto-organizáveis não realizam, por si, detecção ou separação de grupos, mas organizam uma representação dos dados submetidos à rede [Kangas *et al.* (1990)]. O mapa se estrutura para representar os dados significativos submetidos; quanto maior o número de amostras de uma determinada classe presente no conjunto de dados, melhor será o estabelecimento dos pesos internos da rede que a espelha. Dados pouco significativos, por sua vez, acabam diluídos e eliminados. Há, assim, não somente a redução para as dimensões do mapa, mas também a redução da quantidade de dados, estes últimos podendo ser considerados como representativos do conjunto total.

Neste novo conjunto de dados é aplicado um procedimento de agrupamento, o qual tem a propriedade de avaliar um número “ótimo” de grupos. O algoritmo de segmentação em questão é basicamente o proposto por Coleman e Andrews [Coleman & Andrews (1979)], já apresentado na seção 3.2.1 e considerado de forma mais detalhada nesta seção.

O algoritmo proposto, inspirado no procedimento de agrupamento denominado  $k$  médias [Duda & Hart (1973)], consiste na avaliação, em diversas iterações, das variâncias (representadas por matrizes de espalhamento) entre os grupos existentes e internamente a cada grupo. Em cada passo os centróides são recalculados, de modo a variarem dinamicamente conforme mudam as composições dos grupos. Comparativamente ao algoritmo  $k$  médias, a proposta de Coleman e Andrews é capaz de modificar o número de grupos, enquanto o  $k$  médias mantém o mesmo número estabelecido no início do processo (o que requer a disponibilidade de um conhecimento prévio). A cada iteração é acrescentado um novo centróide, ao qual está associada uma nova classe.

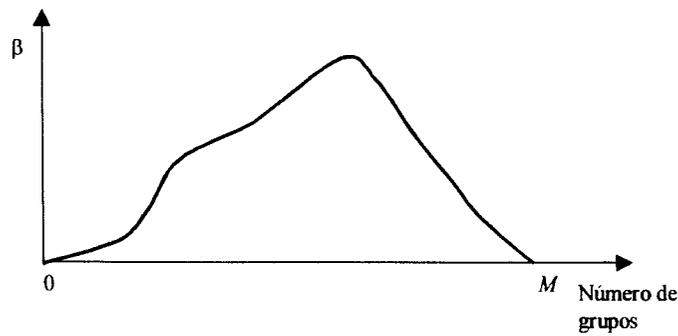


Figura 4.12. Ilustração de um comportamento hipotético para parâmetro  $\beta$ , segundo a variação do número de agrupamentos em que os dados são divididos.

A natureza dinâmica do número de grupos requer o estabelecimento de um critério que determine a situação na qual há um número “ótimo” de grupos em relação ao conjunto de dados. Para este fim é definido um *parâmetro de controle*  $\beta$ , cujo comportamento durante o processo iterativo define o término do procedimento e a decisão por uma partição específica dos dados.

A separação dos dados em grupos considera os atributos na forma de um vetor  $\mathbf{x}$ ,  $\mathbf{x} = (x_1, x_2, \dots, x_n)^t$ , para a obtenção de  $K$  regiões mutuamente exclusivas que compreendam a totalidade dos dados. Desta forma, dados não classificados não são permitidos. As matrizes de espalhamento intragrupos e intergrupos são definidas pelas Equações 4.5 e 4.6, respectivamente, sendo que o vetor  $\bar{\mu}_k$  representa a média e  $M_k$  representa o número de elementos para um dado grupo  $k$ .  $S_k$  é o conjunto dos dados associados ao grupo  $k$  e  $\bar{\mu}_0$  especifica o vetor média compreendendo todos os  $M$  dados [Duda & Hart (1972)]. Estas matrizes permitem obter uma caracterização da distribuição dos dados entre os grupos, para uma configuração de grupos específica.

$$\mathbf{S}_w = \frac{1}{K} \sum_{k=1}^K \frac{1}{M_k} \sum_{x_i \in S_k} (\mathbf{x}_i - \bar{\mu}_k)(\mathbf{x}_i - \bar{\mu}_k)^t \quad (4.5)$$

$$\mathbf{S}_b = \frac{1}{K} \sum_{k=1}^K (\bar{\mu}_k - \bar{\mu}_0)(\bar{\mu}_k - \bar{\mu}_0)^t \quad (4.6)$$

$$\bar{\mu}_0 = \frac{1}{M} \sum_{i=1}^M \mathbf{x}_i \quad (4.7)$$

O parâmetro de controle  $\beta$  é definido pela Equação 4.8, na qual  $\text{tr}(\cdot)$  é o traço da matriz [Coleman & Andrews (1979)]. Por esta definição, na situação de haver um único agrupamento englobando todos os dados,  $\text{tr} \mathbf{S}_w = \sigma^2$ , valor que equivale à variância dos dados, e  $\text{tr} \mathbf{S}_b = 0$ . Esta ocorrência resulta em  $\beta = 0$ . Se, por outro lado,

Algoritmo 4-2. Segmentação por agrupamento de Coleman e Andrews [Coleman & Andrews (1979)].

Defina os dois centróides iniciais.  
 Classifique todos os dados em relação a estes centróides (usando um critério de distância).  
 Determine o valor inicial para o parâmetro de controle.  
 Repita  
   Crie um novo centróide, selecionando o vetor mais distante do centróide de seu grupo.  
   Classifique novamente os dados em função do conjunto atual de centróides.  
   Calcule o novo valor do parâmetro de controle.  
 Até que o parâmetro comece a decrescer.  
 Classifique os dados segundo o valor máximo do parâmetro de controle (antes da descendente).

cada amostra individual constituir um grupo isolado, em um total de  $M$  grupos, então  $\text{tr } S_w = 0$  e  $\text{tr } S_b = \sigma^2$ , resultando novamente em  $\beta = 0$ . Os valores para o parâmetro  $\beta$  são nulos em ambas as situações limites e assumem valores positivos nas demais configurações. Isto implica na existência de pelo menos um ponto de máximo na função descrita por  $\beta$  conforme a variação do número de grupos (Figura 4.12). Este ponto de máximo pode ser interpretado como um balanceamento entre as matrizes de espalhamento  $S_w$  e  $S_b$ , de maneira a ser utilizado como critério para a determinação do melhor número de grupos.

$$\beta = \text{tr } S_w \cdot \text{tr } S_b \quad (4.8)$$

O uso das matrizes  $S_w$  e  $S_b$  define um critério de agrupamento que utiliza um esquema de pesos implícito. Cada item das matrizes de espalhamento intergrupos e intragrupos tem seus termos com peso baseado na frequência relativa dos dados em cada grupo. Considerando-se a relação da Equação 4.9,  $C$  é a soma da variância total dos dados e tem valor constante.

$$C = \text{tr } S_w + \text{tr } S_b \quad (4.9)$$

Pode-se, então, expressar  $\beta$  em relação a  $C$ , conforme a Equação 4.10, a qual, se derivada e igualada a zero, proporciona solução  $\text{tr } S_w = C/2$ . Isto significa que o máximo de  $\beta$  é atingido quando  $\text{tr } S_w$  tem valor igual à metade de  $\text{tr } \phi$ , sendo  $\phi$  a matriz de covariância total dos dados.

$$\beta = \text{tr } S_w \cdot \text{tr } S_b = (C - \text{tr } S_w) \text{tr } S_w \quad (4.10)$$

Além disto, a razão entre os traços das matrizes de espalhamento intragrupos e intergrupos, quando  $\text{tr } S_w = C/2$ , tem valor 1, como ilustrado nas Equações 4.11 e 4.12.

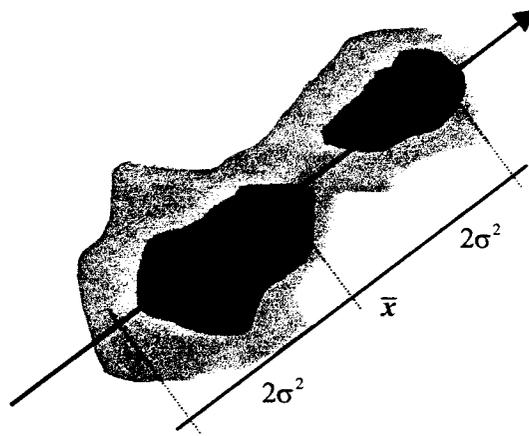


Figura 4.13. Representação do eixo principal de uma distribuição de dados (duas dimensões), com indicação da média e primeiros centróides, diametralmente opostos ao longo do eixo e equidistantes em relação à média. Áreas mais escuras representam maior densidade de dados.

$$\gamma = \frac{\text{tr } S_w}{\text{tr } S_b} = \frac{C - \text{tr } S_w}{\text{tr } S_w} \quad (4.11)$$

$$\gamma = \frac{C - C/2}{C/2} = 1 \quad (4.12)$$

O Algoritmo 4-2 esquematiza os passos gerais do processo de agrupamento e classificação. A determinação dos centróides iniciais é arbitrária e estabelece uma divisão inicial para os dados e os primeiros valores para as matrizes de espalhamento. Entretanto, esta escolha inicial tem influência severa sobre os passos seguintes e o conseqüente resultado final. Desta forma, a escolha dos centróides que iniciam o processo se impõe como uma etapa importante. É interessante que os centróides separem os dois grupos da melhor maneira possível, o que permite colocar que a determinação do eixo principal dos dados se apresenta como uma forma adequada para determinar este alinhamento. Através de uma análise de componentes principais [Castleman (1996)] é determinado o eixo ao longo do qual o “espalhamento” dos dados é maior. Ambos os centróides iniciais são determinados sobre este eixo, um de cada lado em relação à média dos valores (Figura 4.13). Dado que o eixo principal representa o eixo no qual há a melhor separabilidade das classes, os centróides são determinados sobre ele. A variância é usada como parâmetro para restringir os centróides a valores dentro da variação dos valores dos dados. No algoritmo, os centróides são então recalculados a partir das médias dos dados já classificados, tornando-se maior sua relação com os dados das classes determinadas.

A necessidade de se estabelecer uma metodologia para esta determinação inicial dos dois centróides da fase de iniciação do algoritmo se deu em virtude das classificações finais dependerem bastante desta configuração inicial. Assim, para determinações aleatórias, considerando-se um mesmo conjunto de dados, o agrupamento

diferia não somente no número de classes, mas também na pertinência dos dados às classes, quando o número de classes era o mesmo para execuções distintas.

Os dois agrupamentos iniciais, definidos deste modo, são compostos pela associação a cada dado individual da classe relativa ao centróide mais próximo. Disponibilizados estes dois grupos, seus centróides reais podem então ser calculados, caracterizando-os adequadamente. É incluída, também, a determinação do valor do parâmetro de controle  $\beta$  para estes grupos.

Desta maneira, procede-se então a repetição que adiciona um novo centróide como tentativa para a obtenção de uma melhor distribuição de grupos. O valor do novo centróide é determinado pelo valor do dado cuja distância do centróide de sua classe é o maior entre todas as classes. Novamente são recalculados os centróides, agora incrementados de um em número, e o parâmetro  $\beta$ . A cada repetição o valor atual de  $\beta$  é comparado ao valor anterior. É avaliado o perfil da “curva” que este parâmetro descreve à medida que o número de grupos aumenta (Figura 4.12), e quando este valor começa a decrescer a repetição é interrompida. O “melhor” valor para o número de grupos é aquele que detém o maior valor de  $\beta$ , ou seja, o que foi obtido imediatamente antes da interrupção do laço de repetição.

#### 4.1.3 Segmentação em regiões e extração de bordas

O processo de segmentação da imagem é feito pela classificação dos pontos segundo as classes detectadas pela classificação dos dados do mapa auto-organizável. A classificação é simples, utilizando os centróides dos agrupamentos como referência e um critério de distância euclidiana para associar cada ponto a uma classe. Cada ponto da imagem é, assim, associado à classe do centróide mais próximo. Antes da classificação, entretanto, um filtro de média é aplicado a cada banda da imagem colorida, utilizando uma janela  $3 \times 3$  centrada em cada ponto, suavizando em parte o ruído presente.

Assume-se que a imagem é formada por objetos, os quais compartilham (ao menos em partes definidas) um mesmo atributo cromático. Pontos da imagem que compartilham um atributo estão, deste modo, formando regiões aproximadamente contínuas. A associação dos pontos destas regiões a uma dada classe permite estabelecer as regiões na imagem. Quando a classificação atribui um ou outro ponto dentro de uma região homogênea a outra classe, a aplicação de um filtro da mediana pode ser feita para remover estes pontos esparsos. Há, entretanto, um maior comprometimento dos limites das bordas entre as regiões.

Os resultados da segmentação da imagem da Figura 4.14(a) são apresentados, identificando, inicialmente, as regiões detectadas na Figura 4.14(b), cada classe representada por um nível de cinza distinto. O mapa auto-organizável em seu estado final é apresentado na Figura 4.14(c), na qual as cores são uma representação da cromaticidade armazenada nos pesos internos dos nós, modificada unicamente com o fim de visualização. O mapa tem dimensões de  $16 \times 16$  unidades e os limites entre as regiões detectadas encontram-se delineados. A Figura 4.14(d) apresenta um detalhe ampliado

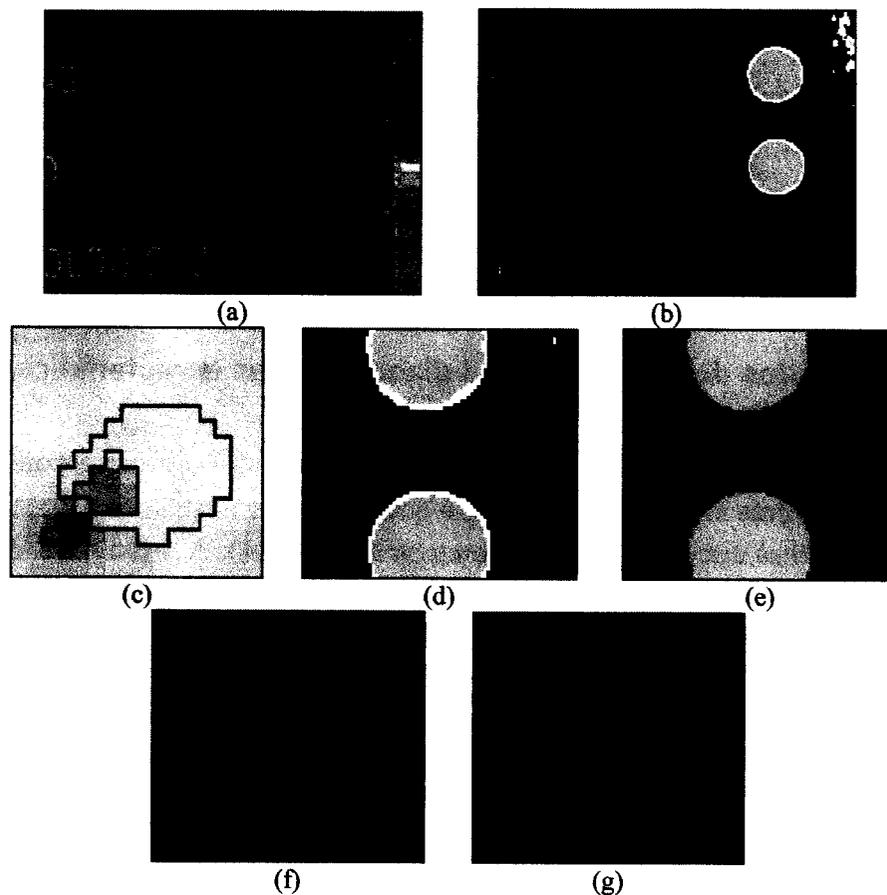


Figura 4.14. Exemplo da classificação através dos centróides detectados pela metodologia proposta. (a) Imagem original; (b) imagem resultante da classificação dos pontos; (c) representação do mapa auto-organizável após a convergência; (d) detalhe da borda entre duas regiões; (e) mesmo detalhe após a aplicação do filtro da mediana; (f) e (g) Detalhes de outra parte da imagem, respectivamente antes e após a aplicação do filtro da mediana.

da imagem de regiões, na qual podem ser observadas regiões pequenas. Após a aplicação de um filtro da mediana de raio 3, o mesmo detalhe da segmentação deixa de apresentar as regiões pequenas bem como a região fina em branco; detalhes da borda, como a reentrância entre as lâminas da placa, entretanto, são perdidos, como se pode notar na parte central direita da Figura 4.14(e). Embora localmente o resultado tenha sido interessante, em outros pontos da imagem, como mostrado nas Figuras 4.14(f) e (g), o filtro acaba removendo detalhes que poderiam ser considerados importantes.

A obtenção do mapa de bordas a partir da imagem segmentada por regiões é feita determinando-se os pontos que constituem a interface entre regiões. Adota-se como critério uma máscara  $3 \times 3$  (Figura 4.15), centrada no ponto a ser analisado, que é deslizada sobre a imagem. Um ponto é considerado como ponto de borda caso qualquer um de seus vizinhos nas direções N, O, S ou L cubra uma classe distinta da do centro da máscara. São assim determinadas bordas em ambos os lados da fronteira entre duas regiões, como pode ser observado na Figura 4.16.

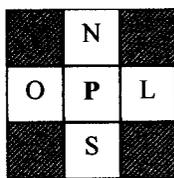


Figura 4.15. Máscara usada para determinar pontos de borda. P: Centro da máscara, sobre o ponto sendo analisado. N, O, S e L: vizinhos nas direções cardeais. Áreas hachuradas não são utilizadas.

## 4.2 Resultados de desempenho e discussão das características do método proposto

Os resultados conseguidos pela aplicação da metodologia corresponderam a algumas expectativas iniciais, principalmente quanto à parte relativa ao uso da cromaticidade (o que será discutido posteriormente na seção 4.2.4). A seção 4.2.1 descreve a estratégia para uma avaliação quantitativa do desempenho das classificações. Nas seções seguintes são apresentados os resultados considerando-se variações nos parâmetros do mapa, envolvendo suas dimensões (seção 4.2.2) e taxas de decaimento para o raio da vizinhança e taxa de aprendizado (seção 4.2.3). Finalmente, a seção 4.3 apresenta algumas considerações sobre a aplicabilidade da classificação por mapas auto-organizáveis.

### 4.2.1 Avaliação da classificação

A análise visual das segmentações obtidas permite considerar os resultados adequados sob um ponto de vista subjetivo. As regiões representam as áreas da imagem cobertas pelos objetos, as diferentes cores levam, em geral, a classes diferentes e as principais classes são, na maioria dos casos, detectadas conforme esperado. Em alguns casos, porém, pode ocorrer a não detecção de uma das classes esperadas *a priori*, bem como a classificação subdividir uma das classes, considerada única inicialmente, em duas ou mais.

Uma medida quantitativa de avaliação, em contraposição à avaliação puramente subjetiva, foi também considerada, abrangendo a taxa de classificação realizada

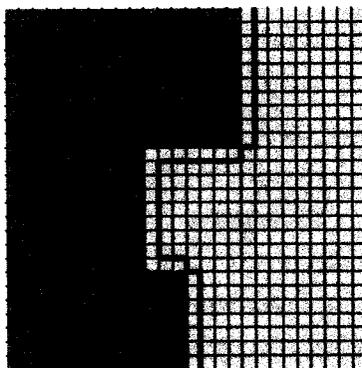


Figura 4.16. Esquema da fronteira entre duas regiões, indicadas por tons de cinza diferentes. Contornos pela linha mais espessa se encontram os pontos considerados como bordas entre elas.

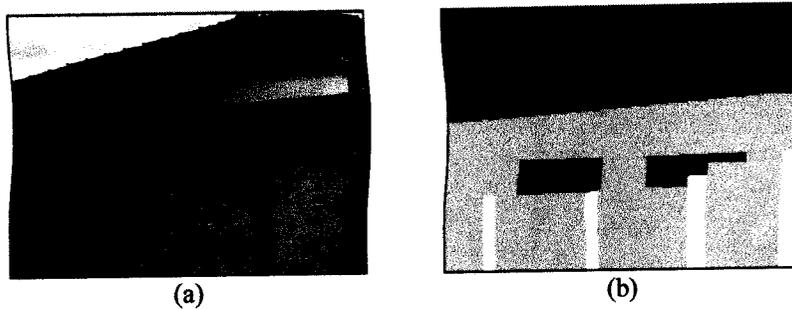


Figura 4.17. Exemplo de rotulação para a imagem House06. Níveis de cinza indicam classes distintas.

corretamente. A imagem é segmentada manualmente, utilizando critérios não somente específicos, como bordas e limites entre objetos, mas também avaliações de mais alto nível de conhecimento, que levam em consideração o discernimento do operador, que é capaz de reconhecer os objetos presentes, além de inferir seus limites mesmo quando a imagem não apresenta características locais suficientes. Esta imagem segmentada resulta na chamada *imagem de regiões de referência*. Para que se pudesse seguir um padrão geral de comparação, a segmentação manual é realizada considerando algumas diretrizes gerais, que incluem:

- a) A imagem é segmentada em regiões, atribuindo a todos os pontos da imagem um rótulo (classe);
- b) Os principais elementos cromáticos (regiões mais significativas) da imagem são considerados para determinar o número de classes;
- c) São desconsiderados os reflexos especulares ou reflexos de objetos em objetos;
- d) Sombras e outras “bordas” causadas por diferença de intensidade não são consideradas relevantes;
- e) O conhecimento de nível alto (identificação dos objetos) é utilizado para resolver rotulações que são duvidosas quando se dispõe apenas dos valores locais dos pontos.

A utilização deste critério não permite, entretanto, uma segmentação completamente objetiva, porém direciona para que a escolha das classes e suas delimitações não seja totalmente subjetiva. Em particular, a segmentação manual, como delineada, permite uma avaliação que considere a perspectiva da percepção visual humana. Esta opção por uma caracterização mais abstrata que física das regiões leva em consideração um grau mais elevado de interpretação da cena, de forma que se possa avaliar o comportamento da classificação utilizando-se a cromaticidade, dadas suas características de quase independência das variações da intensidade luminosa (por exemplo, sombras).

A Figura 4.17(a) mostra a imagem da fachada de uma casa e sua rotulação manual. Nota-se, na Figura 4.17(b), que a determinação da classificação dos pontos desconsidera as sombras e alguns detalhes, como os frisos das janelas, o padrão das telhas. Da mesma forma, foi ignorada a distinção entre as nuvens e o céu, assim como

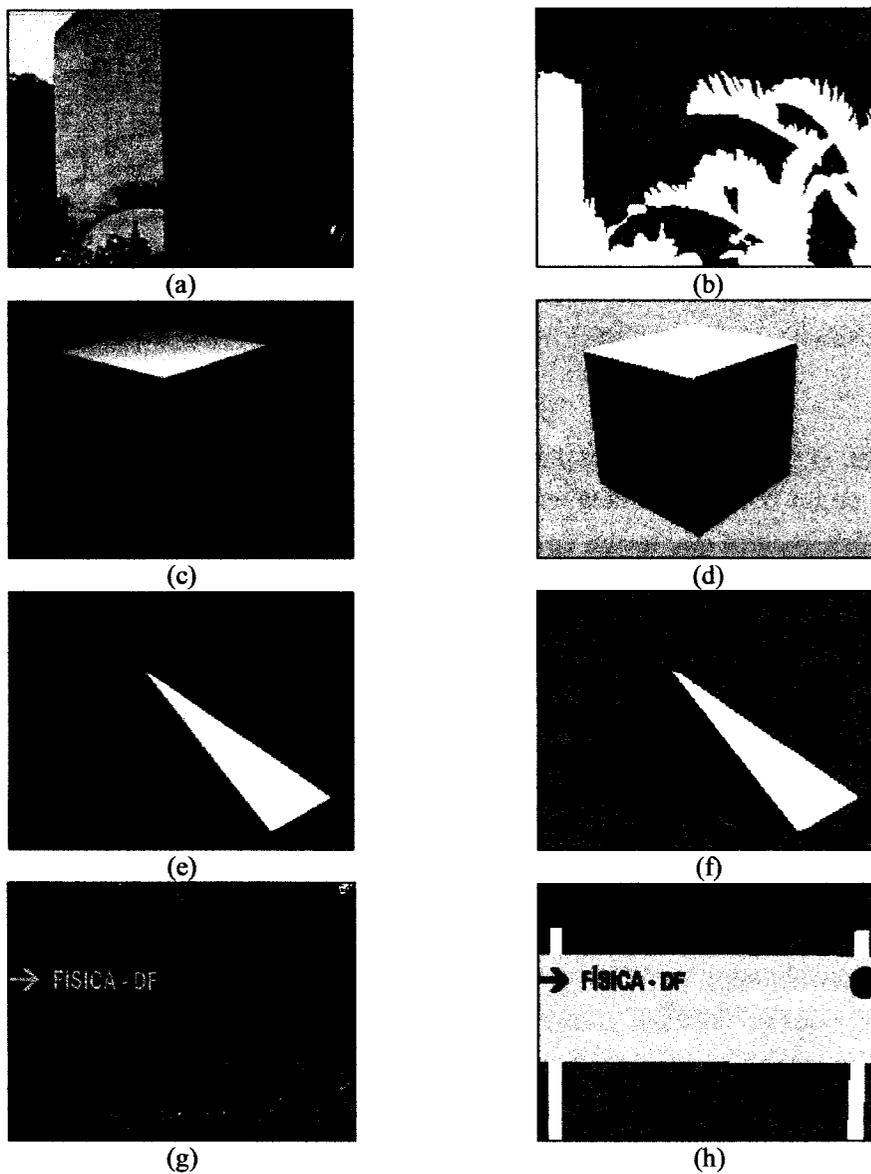


Figura 4.18. Imagens para os exemplos da segmentação para diferentes tamanhos de mapas. (a) Imagem original e (b) imagem de referência para imagem House03. (c) Imagem original e (d) imagem de referência para imagem Cube014. (e) Imagem original e (f) imagem de referência para imagem Pyramid01. (g) Imagem original e (h) imagem de referência para imagem Sign04.

a diferença de iluminação entre o muro e a parede, que são considerados da mesma cor.

Define-se como avaliação quantitativa, desta forma, a razão entre o número de pontos da imagem classificados em consonância com a classificação manual. Esta razão é expressa pela Equação 4.13, na qual  $M_i$  é a taxa de acerto,  $C_i$  é o número de pontos classificados como pertencentes à classe  $i$  e  $N_i$  é o número de pontos atribuído manualmente à classe  $i$ .

$$M_i = \frac{C_i}{N_i} \quad (4.13)$$

Considerando-se, porém, a forma como estão delimitadas as regiões obtidas manualmente, uma taxa próxima a 100% não é esperada em regiões menos homogêneas ou com a presença de detalhes ignorados, como é o caso das telhas descrito para o exemplo da Figura 4.17.

O método de segmentação sofre influência de um grande número de parâmetros, tais como as dimensões do mapa, os fatores de decaimento do tamanho da vizinhança e da taxa de aprendizado, bem como seus valores iniciais e o erro do critério usado para determinar a convergência. Os experimentos permitiram identificar que vários destes parâmetros apresentam uma tolerância grande, permitindo ajustar valores que se comportam bem para uma gama de imagens bastante variada.

O mapa auto-organizável retém os parâmetros variáveis do sistema, dado que o processo de determinação dos grupos de dados se impõe como unicamente dependente dos valores dos dados. Não são incluídos, assim, parâmetros que monitorem ou influenciem o processo de agrupamento.

As imagens usadas para os testes são as apresentadas no Apêndice D, as quais se compõem de cenas em ambiente interno (laboratório) e externo, tomadas sob condições variáveis de iluminação. Um conjunto razoavelmente genérico de imagens foi, assim, construído.

#### 4.2.2 Tamanho e dimensão do mapa auto-organizável

A Figura 4.18 mostra as imagens usadas para produzir os gráficos que ilustram o comportamento da classificação conforme variam as dimensões do mapa. As classes consideradas para a imagem *House03* são *céu*, *parede* e *vegetação*, e para a imagem *Cube011*, *azul*, *verde*, *amarela* (para as faces do cubo) e *fundo*. Na imagem *Pyram01*, as classes são *verde*, *amarela* (faces) e *fundo*, enquanto na imagem *Sign04* existem *placa*, *grama* (incluindo a vegetação mais clara, na parte superior), *código* (identificação por cor\*), *suporte* (postes que fixam a placa) e *árvores* (vegetação escura, próximo ao cinza, na parte superior). As classes são consideradas conforme a segmentação manual descrita na seção 4.2.1.

As Figuras 4.19 a 4.22 ilustram o comportamento da metodologia de segmentação de acordo com a variação do número de unidades e de sua disposição no mapa auto-organizável. A estruturação foi organizada de forma linear (dimensão  $1 \times n$ ) e quadrada ( $n \times n$ ), para  $n$  assumindo os valores 5, 10, 20, 40, 60, 80 e 100 unidades. Os resultados apresentados consistem nos valores médios para quatro execuções do programa. Observa-se, ainda, que para execuções distintas com os mesmos parâmetros, a variação dos resultados é bastante pequena. Os demais parâmetros, utilizados para todos os testes desta seção, são: vizinhança inicial correspondente a 80% da maior dimensão do mapa, reduzindo-se com uma taxa de 0,06 a cada passo; peso inicial de 0,6 para a unidade vencedora, decaindo com uma taxa de 0,06 a cada passo;

---

\* Os círculos identificam Centros e Unidades da Universidade Federal de São Carlos através de cores específicas.

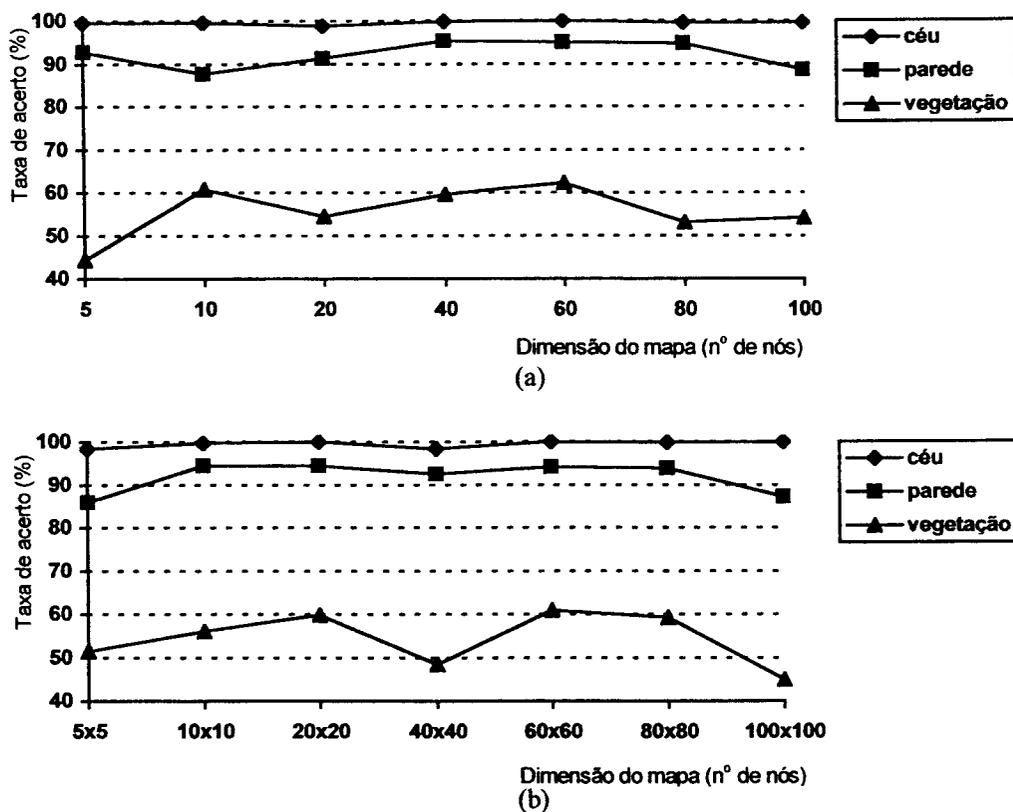


Figura 4.19. Taxa de acerto na classificação (em porcentagem) conforme variação do tamanho usado para mapa auto-organizável: (a) linear e (b) quadrado. Imagem *House03*.

peso invariável para a periferia da vizinhança de 0,01; e critério de convergência com  $\epsilon = 0,0001$ . A taxa inicial de aprendizado para a unidade vencedora é definido em 0,6, que é um valor usual na literatura [Anderson (1995)].

A Figura 4.19(a) mostra, para mapas lineares, os resultados da segmentação da imagem *House03* quando comparados à imagem de regiões de referência da Figura 4.18(b). É possível observar que a taxa é elevada e constante para a classe *céu*, oscilando em torno de 98%. Para a classe *parede*, a taxa de acerto é um pouco menor, variando aproximadamente entre 89 e 95%; apresenta também uma variação um pouco maior conforme varia o comprimento do mapa. A classe *vegetação* apresenta uma taxa de acerto substancialmente menor, em grande parte devido à textura da própria vegetação (folhas das árvores), que acaba por induzir a classificação a criar uma classe adicional nesta região. Esta classe “extra” abrange certas partes da área manualmente designada como uma única classe. Em virtude desta “sobressegmentação”, a quantidade de pontos efetivamente associados à classe decresce, assim como a consequente taxa de acerto. A taxa de acerto oscila entre 50 e 60%, aproximadamente, com exceção para a rede pequena com 5 unidades, que produziu uma taxa abaixo de 50%. Redes pequenas, em geral, apresentam uma separação de classes mais drástica, dado que são poucas as unidades da rede e a transição dos atributos entre unidades vizinhas não se permite ser suave.

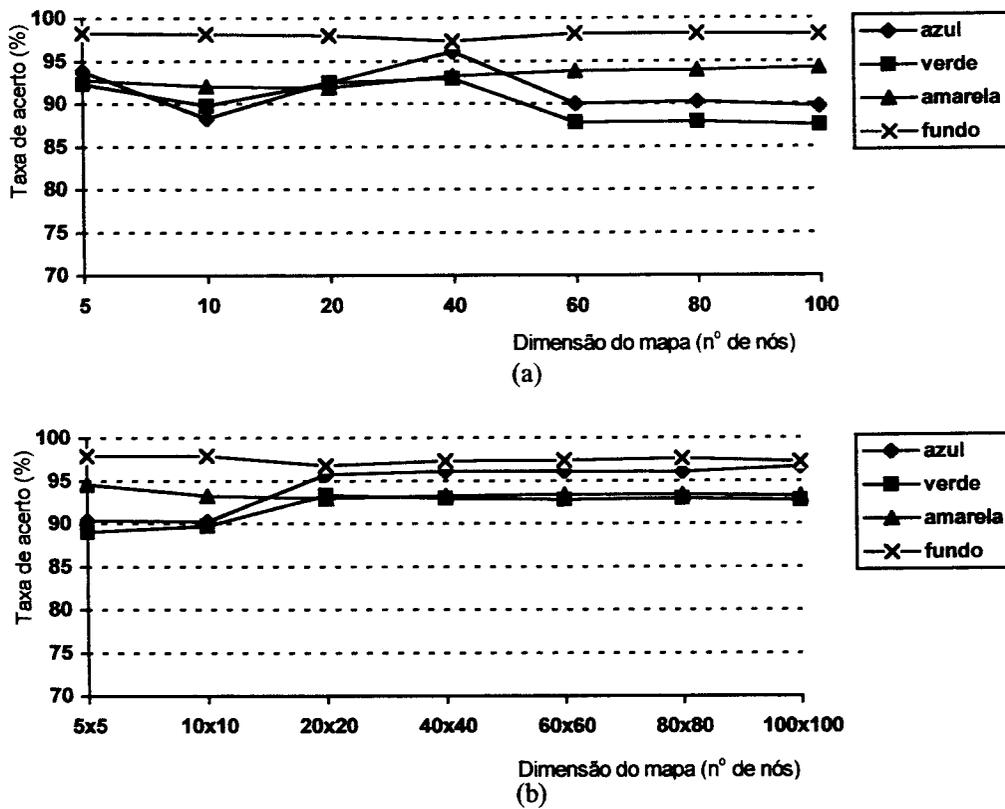


Figura 4.20. Taxa de acerto na classificação (em porcentagem) conforme variação do tamanho usado para mapa auto-organizável: (a) linear e (b) quadrado. Imagem Cube011.

Na Figura 4.19(b) são mostradas as taxas de acerto médias para redes quadradas, também em relação à imagem *House03*. Os resultados são bastante similares às das redes lineares, com taxas próximas às dos mapas lineares, mas levemente superiores, no caso geral. Nota-se, particularmente para os mapas de tamanhos 40×40 e 100×100, que a taxa de acerto para *vegetação* foi inferior a 50%. Não houve, entretanto, nenhuma particularidade no comportamento de convergência dos mapas ou nos dados que pudessem justificar ou prever tal comportamento, embora, do ponto de vista da textura apresentada para a região, possa-se considerar tal variação tolerável.

Para a imagem da Figura 4.18(c), denominada *Cube011*, e respectiva imagem de regiões de referência, na Figura 4.18(d), são apresentadas na Figura 4.20 as taxas de acerto da classificação. Sem depender severamente das dimensões do mapa estruturado linearmente, as taxas determinadas foram todas acima de 85% de acerto. A classe mais estável foi a *fundo*, com 96 a 99%. *Azul* manteve-se entre 90 e 95%, com melhor desempenho para o mapa com 40 unidades. As classes *azul* e *verde* apresentaram uma taxa de acerto menor para mapas mais longos (acima de 60 unidades), causado pela classificação de alguns pontos nas áreas escuras do cubo como pertencentes ao fundo. Para o caso de mapas bidimensionais, a Figura 4.20(b) ilustra que o comportamento da classificação se estabiliza para mapas acima de 20 unidades de lado, mantendo-se com grau de acerto maior de 90% para todas as classes e, de modo mais

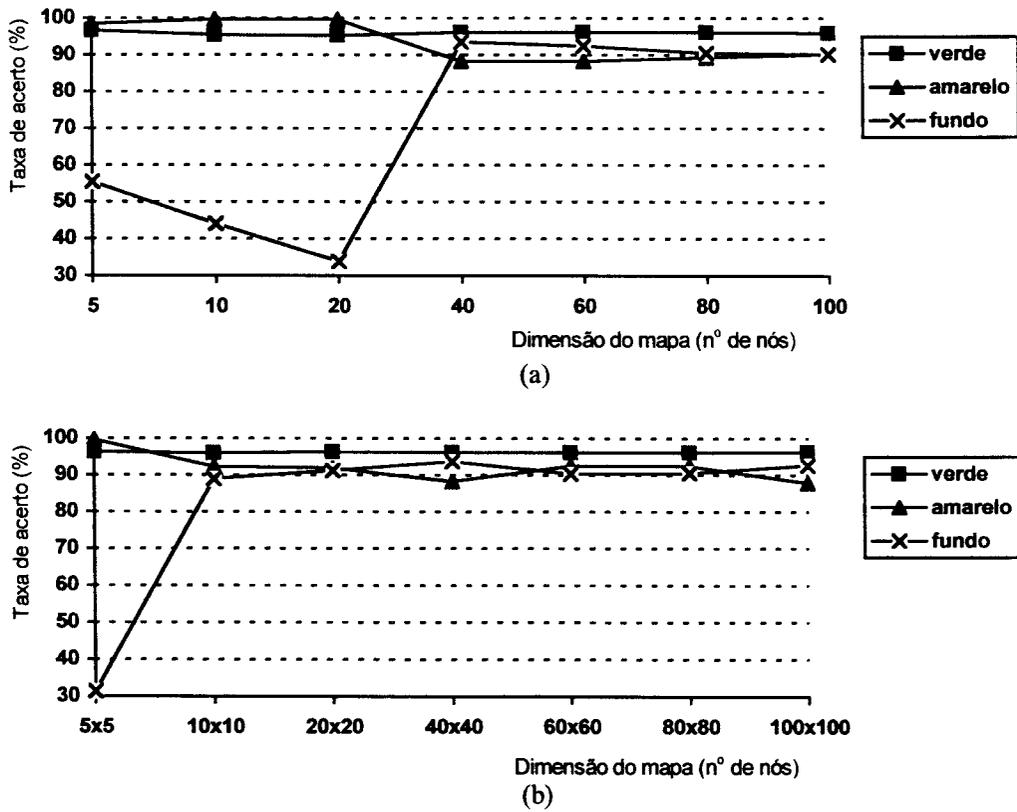


Figura 4.21. Taxa de acerto na classificação (em porcentagem) conforme variação do tamanho usado para mapa auto-organizável: (a) linear e (b) quadrado. Imagem *Pyram01*.

específico, acima de 95% para as classes *fundo* e *azul*. Este desempenho é particularmente interessante pelo aumento do reconhecimento correto das áreas da imagem cobertas por sombras (parte inferior esquerda da área verde e inferior direita da área azul). As regiões correspondentes à face amarela e ao fundo permaneceram praticamente constantes.

Gráficos similares estão representados na Figura 4.21, com as taxas de acerto para a imagem *Pyram01*. Na Figura 4.21(a) estão as taxas de acerto para mapas organizados linearmente. A partir do mapa com comprimento 40 unidades a correta separação das classes é obtida. Com comprimentos menores, o fundo é reconhecido como composto por duas classes distintas: partes menos escuras, com tom azulado, e regiões mais escuras, em preto. Estas diferenças de cor foram proporcionadas por reflexos nas ondulações do tecido preto que compõe o fundo da imagem. As taxas de acerto baixas correspondem, neste caso, a detecções parciais da classe *fundo*. Mapas mais extensos permitem a determinação adequada do fundo como uma única classe, proporcionando um desempenho de acerto acima de 90% na maioria dos casos, embora em alguns casos a classe *amarela* ficasse com taxa em torno de 87%. A Figura 4.21(b), para mapas quadrados, mostra resultados bastante similares, com as mesmas variações e valores para as taxas de acerto. Uma característica importante é o fato da correta separação do fundo em uma única classe já ocorrer para mapas com dimensão

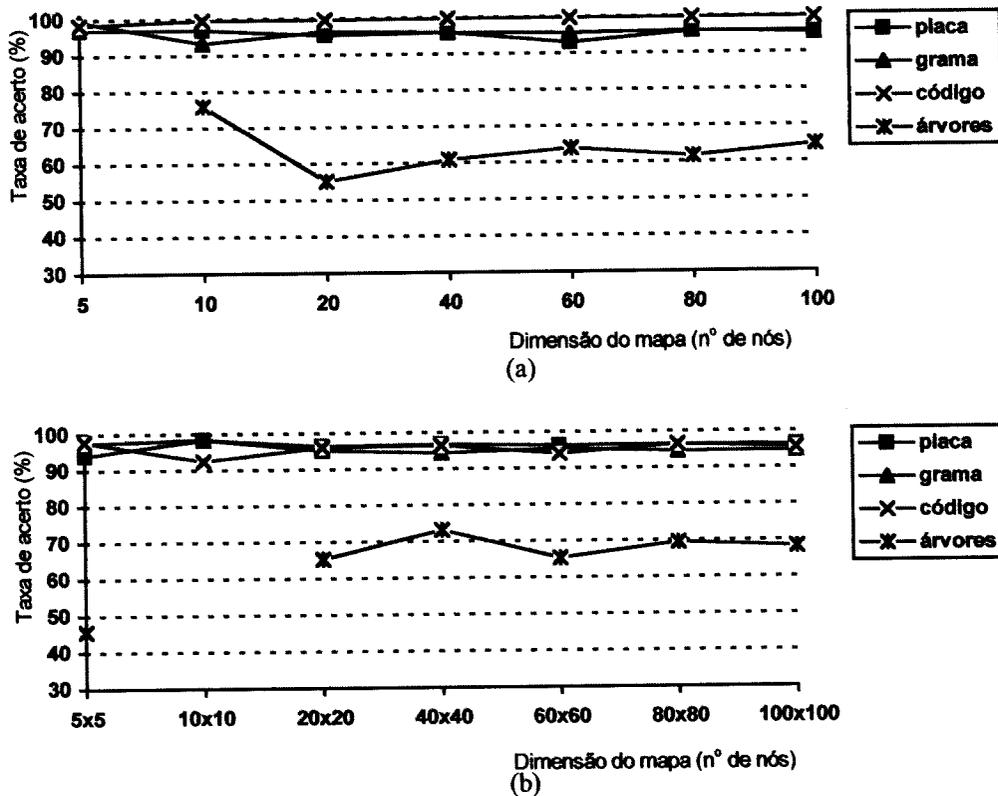


Figura 4.22. Taxa de acerto na classificação (em porcentagem), conforme variação do tamanho usado para mapa auto-organizável: (a) linear e (b) quadrado. Imagem Sign04.

10×10. Apenas no mapa de tamanho 5×5 unidades houve a separação do fundo em duas classes distintas (com acerto em torno de apenas 30% para uma das regiões detectadas como fundo).

A imagem Sign04 tem sua segmentação para mapas lineares mostrada na Figura 4.22(a). Este gráfico apresenta uma particularidade para mapas com poucas unidades: a classe *árvores* não pôde ser identificada. Como a quantidade de unidades é pequena, a influência lateral entre vizinhos acabou levando o método de agrupamento a considerar como uma única classe grupos de pontos considerados distintos na segmentação manual. Este caso está representado na ausência da classe *árvores* para o mapa de comprimento 5.

Particularmente a esta classe, novamente se atribui à textura mais intensa\* das árvores a baixa taxa de acerto, mesmo quando a classe é corretamente identificada. Vários pontos da área designada manualmente como classe *árvores* são classificados como pertencente à classe *grama*, principalmente nas regiões mais claras nas copas das árvores. Esta característica da imagem faz com que a taxa permaneça um pouco acima de 60% para redes de comprimentos diferentes, com acertos específicos em torno de 76% para a rede com 10 unidades e 55% para 20 unidades. Para as demais

\* Mais intensa quando comparada à grama, que apresenta uma textura visualmente mais suave.

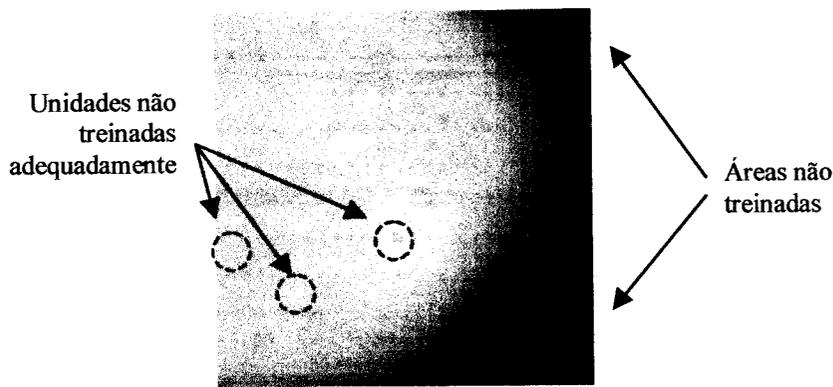


Figura 4.23. Representação de um mapa auto-organizável  $32 \times 32$ , processado com taxa de decaimento muito alta para o raio da vizinhança.

classes, mais suaves do ponto de vista de textura, as taxas permanecem em torno de 95%, usualmente acima deste valor. Especificamente para a classe *código*, a taxa é praticamente de 100% de acerto, principalmente para mapas mais longos.

Mapas quadrados, cujas taxas de acerto são apresentadas na Figura 4.22(b), apresentam comportamento semelhante, com acertos em torno de 95% para todas as classes. Novamente a classe *árvores* não é detectada para o mapa de tamanho  $10 \times 10$  e tem sua pior classificação para o tamanho  $5 \times 5$  (aproximadamente 45%). Para os mapas bidimensionais, a taxa para a região de árvores oscila entre 65 e 70%. Destaca-se, para esta imagem, ainda, que uma classe inicialmente esperada não é identificada em nenhuma das configurações de tamanho do mapa. As áreas da imagem formada pelos postes que sustentam a placa (classe *suporte*) são ignoradas no processo de agrupamento dos dados do mapa auto-organizável, provavelmente por seus pontos não poderem ser considerados relevantes para compor uma nova classe, de acordo com os critérios descritos em 4.1.2. Usualmente os pontos pertencentes a estas regiões são classificados como *árvores* ou *grama*.

Observa-se, como um caso geral, uma constância entre os resultados obtidos em termos de taxa de acerto para as classes individuais. Resultados problemáticos aparecem, na maioria dos casos, para mapas de tamanhos pequenos, como pode ser observado nos exemplos apresentados. Os desempenhos para mapas bidimensionais têm se mostrado, também, mais apropriados em contraposição aos mapas lineares. A utilização de uma dimensão para o mapa da ordem  $16 \times 16$  a  $40 \times 40$  proporciona resultados apropriados para o elenco de imagens usadas como conjunto de testes, o que levou à adoção de uma rede  $32 \times 32$  como padrão inicial a ser aplicado a uma segmentação qualquer. Tamanhos menores de rede têm desempenho interessante para o caso da ausência de texturas na imagem. Uma calibração para um conjunto de imagens específico pode ser determinado para uma aplicação específica.

O tempo de execução para mapas maiores é mais elevado, dado principalmente o maior raio inicial para a vizinhança, o que leva a retardar a convergência. O tempo é também influenciado pelo maior número de unidades afetadas a cada dado

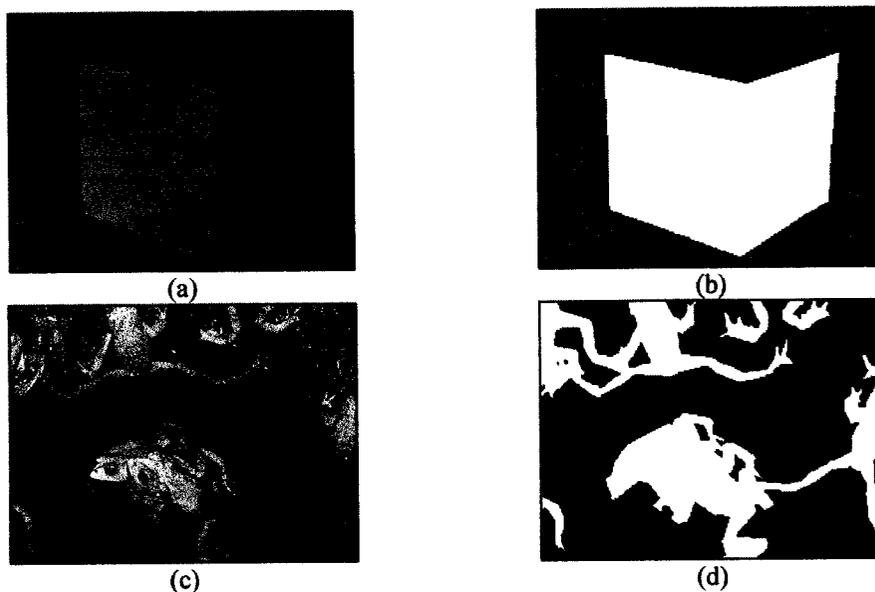


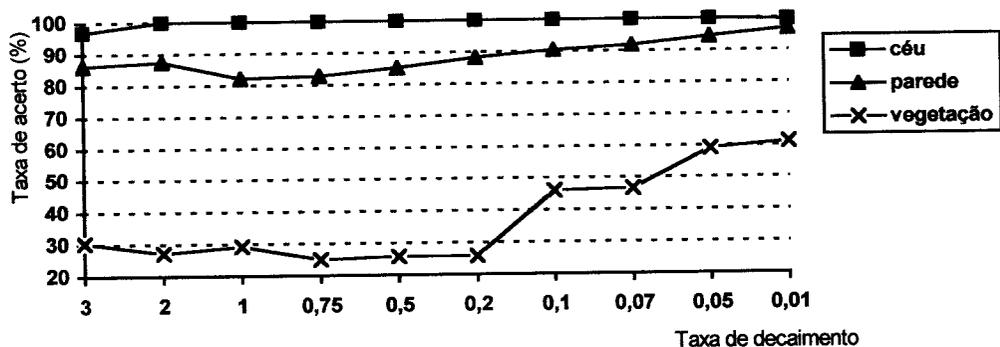
Figura 4.24. Figuras usadas nos experimentos dos gráficos das Figuras 4.26 e 4.27. (a) e (b): Imagem Cube104 e sua imagem de regiões (manual), respectivamente. (c) e (d): Idem para a imagem Frogs01.

submetido à rede durante o treinamento. A escolha da dimensão padrão foi, também, influenciada por fatores de desempenho de tempo.

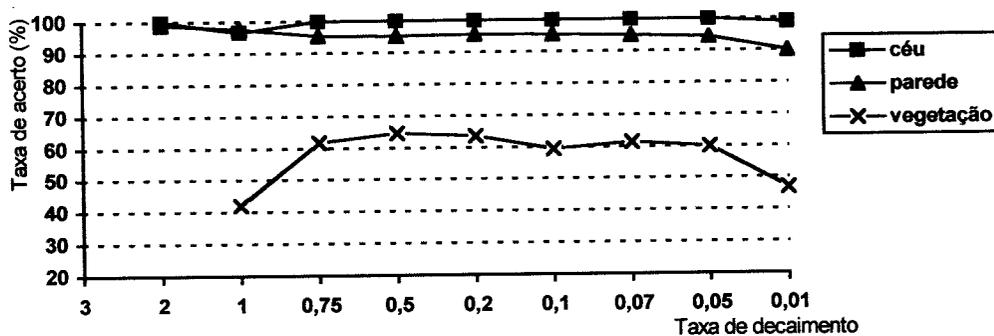
### 4.2.3 Taxas de decaimento

Outros dois parâmetros a serem considerados no comportamento da classificação são as taxas de decaimento para o tamanho da vizinhança e para a taxa de aprendizado. Para que haja uma convergência correta, o decaimento de ambos deve ser bastante lento [Kohonen (1984); Anderson (1995)], o que garante a suavidade dos valores à medida que se desloca de uma unidade para outra.

As Figuras 4.25, 4.26 e 4.27 ilustram como o resultado da segmentação, expresso pela porcentagem de acerto para as diferentes classes na imagem, varia para taxas distintas de decaimento. A Figura 4.25(a) mostra as taxas de acerto para a imagem House03, apresentada na Figura 4.18(a), de acordo com diferentes valores utilizados como taxa de decaimento para o raio da vizinhança. Valores mais altos, como a taxa de 3%, até menores são apresentados, cobrindo os valores 0,01, 0,07, 0,1, 0,2, 0,5, 0,75, 1, 2 e 3%. Verifica-se uma certa constância nos resultados, apesar da utilização de taxas díspares. Melhores resultados, entretanto, são observados para as pequenas taxas, abaixo de 1%, que permitem a lenta adaptação dos pesos internos das unidades do mapa. No caso de taxas grandes, como 2 e 3%, ocorre a não adaptação de parte das unidades do mapa, pois o “alcance” da influência das informações é abruptamente reduzido logo nos primeiros dados submetidos à rede. A Figura 4.23 mostra uma representação de um mapa treinado com taxa de 4%. Em alguns casos, devido à identificação de áreas não corretamente adaptadas do mapa em classes, houve a incorreta separação também das classes esperadas, principal causa do baixo desempenho para a classe *vegetação*, subdividida em mais classes. Para este gráfico o



(a)



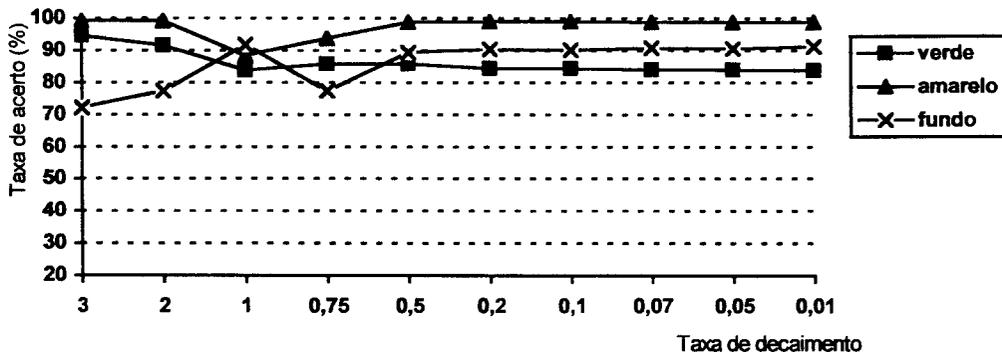
(b)

Figura 4.25. Taxa de acerto na classificação (em porcentagem) conforme variação da taxa de decaimento utilizada (a) para o raio da vizinhança e (b) para a taxa de aprendizado. Decaimento expresso em porcentagem. Imagem House03.

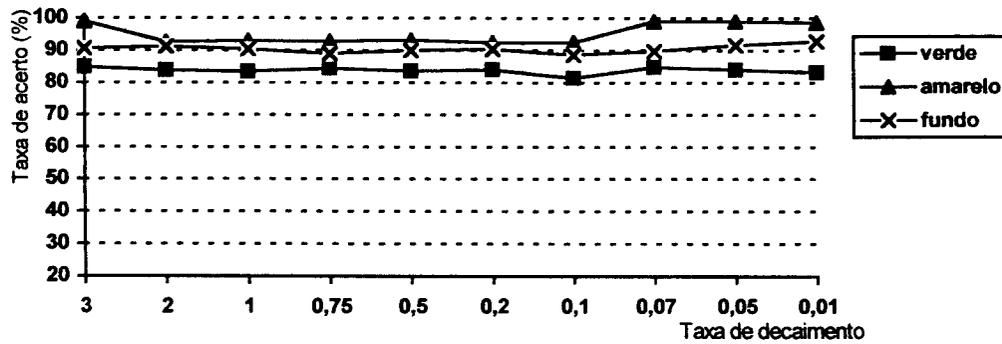
mapa utilizado foi dimensionado em  $32 \times 32$  unidades, utilizando raio inicial de 80% do lado do mapa (arredondado para 26), taxa de aprendizado inicial de 0,6 para a unidade vencedora, com taxa de decaimento de 0,05% a cada passo. A convergência utilizada ficou definida com erro de 0,0001.

Para o gráfico da Figura 4.25(b) foram utilizados os mesmos parâmetros, diferenciando apenas para a taxa de decaimento do raio da vizinhança, fixado em 0,05. Os valores testados para a taxa de decaimento da influência do aprendizado foram os mesmos utilizados na Figura 4.25(a). Para o caso particular da taxa de 3% não houve a correta identificação das classes presentes. A queda brusca da taxa de aprendizado a cada novo dado submetido à rede não permitiu uma convergência adequada, ficando muitas das unidades modificadas apenas superficialmente\*. Notadamente pôde ser observada a influência dos valores aleatórios utilizados na iniciação dos pesos internos. Para a taxa de 2%, também, a classe *vegetação* não foi discriminada, sendo que

\* Várias unidades não chegaram a ter seus pesos internos adequados para refletir os dados de cromaticidade. Foram notadas, também, regiões do mapa sem suavidade entre unidades vizinhas, bem como cromaticidades similares presentes em áreas distintas do mapa. (A análise foi feita verificando-se os valores e observando uma representação visual colorida do mapa. Veja-se a Figura 4.23 como exemplo destas falhas no treinamento.)



(a)



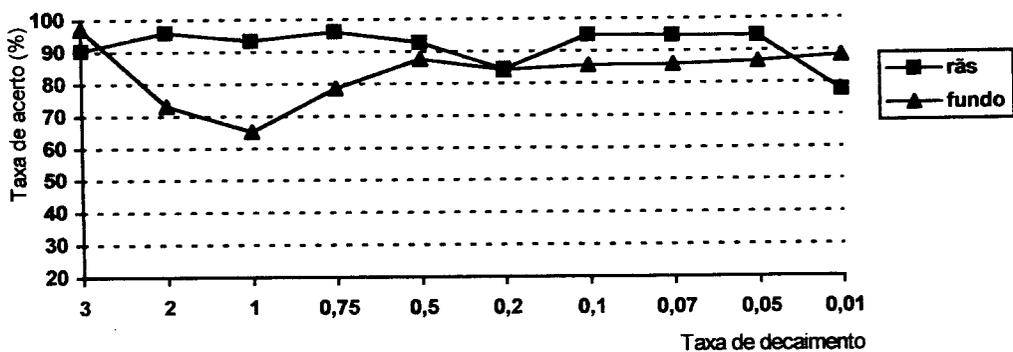
(b)

Figura 4.26. Taxa de acerto na classificação (em porcentagem) conforme variação da taxa de decaimento utilizada (a) para o raio da vizinhança e (b) para a taxa de aprendizado. Decaimento expresso em porcentagem. Imagem Cube104.

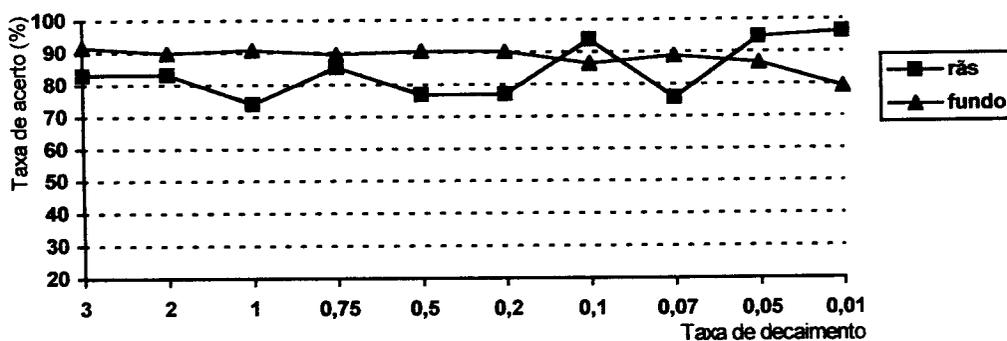
seus dados foram classificados na classe *céu*. Os resultados mais adequados permaneceram em torno da taxa de decaimento de 0,05%.

A Figura 4.26 mostra o mesmo experimento aplicado à imagem da Figura 4.24(a), a qual tem suas regiões de referência apresentadas na Figura 4.24(b). As classes selecionadas manualmente compreendem as faces amarelas (classe *amarelo*), a face superior (*verde*) e o fundo azul (*fundo*). Embora sejam duas faces amarelas, sob o ponto de vista da cor (e cromaticidade), considera-se haver uma única classe. A classificação dos pontos se mostrou mais estável em termos das taxas de acerto para as classes, com *fundo* e *amarelo* acima de 90% e *verde* acima de 80% de acerto. Mesmo assim, ocorrências de áreas não treinadas e unidades mal treinadas se manifestaram para as taxas de decaimento da vizinhança com valores acima de 1%, apresentadas na Figura 4.24(a). Para o caso da Figura 4.24(b), que considera taxas diferentes para o decaimento da taxa de aprendizado, várias unidades permaneceram com treinamento inadequado (insuficiente) para taxas altas, o que é mais visível para taxas maiores que 1%. Apesar destes fatos, o agrupamento foi feito separando-se as classes de forma apropriada, provendo resultados consistentes com as expectativas das classes manuais.

O último experimento está apresentado nas Figuras 4.27(a) e (b), com variações, respectivamente, para os decaimentos do raio da vizinhança e da taxa de apren-



(a)



(b)

Figura 4.27. Taxa de acerto na classificação (em porcentagem) conforme variação da taxa de decaimento utilizada (a) para o raio da vizinhança e (b) para a taxa de aprendizado. Decaimento expresso em porcentagem. Imagem *Frogs01*.

dizado. A imagem utilizada, *Frogs01*, é a mostrada na Figura 4.24(c), com respectiva segmentação manual na Figura 4.24(d). Os mesmos comentários feitos para os experimentos anteriores se aplicam às segmentações desta imagem. Destaca-se, neste exemplo, uma oscilação um pouco maior das taxas de acerto para as duas classes presentes: *rãs* e *fundo*. Devido à existência de reflexos espaciais no fundo e uma cor bem mais saturada nas bordas dos corpos das rãs, a identificação de uma terceira classe pela metodologia foi freqüente, manifestando-se em todas as imagens, em maior ou menor grau. Por vezes a terceira classe abrangia em maior quantidade os pontos do fundo; em outros casos, abrangia a periferia das áreas das rãs. Não pôde ser identificada nenhuma tendência para esta terceira classe que levasse a julgar que a segmentação manual tivesse sido inadequada ao estabelecer somente duas classes para a imagem.

Para que uma convergência adequada dos pesos internos do mapa auto-organizável aconteça, tanto o raio da vizinhança quanto a taxa de aprendizado devem decrescer lentamente. Caso contrário, instabilidades são inseridas nos pesos internos e o agrupamento (e conseqüente segmentação) fica comprometido. Então, a cada passo, ou seja, a cada novo dado usado no treinamento do mapa, taxas entre 0,05 e 0,09% foram adotadas para reduzir cada um dos parâmetros analisados. Esta pequena

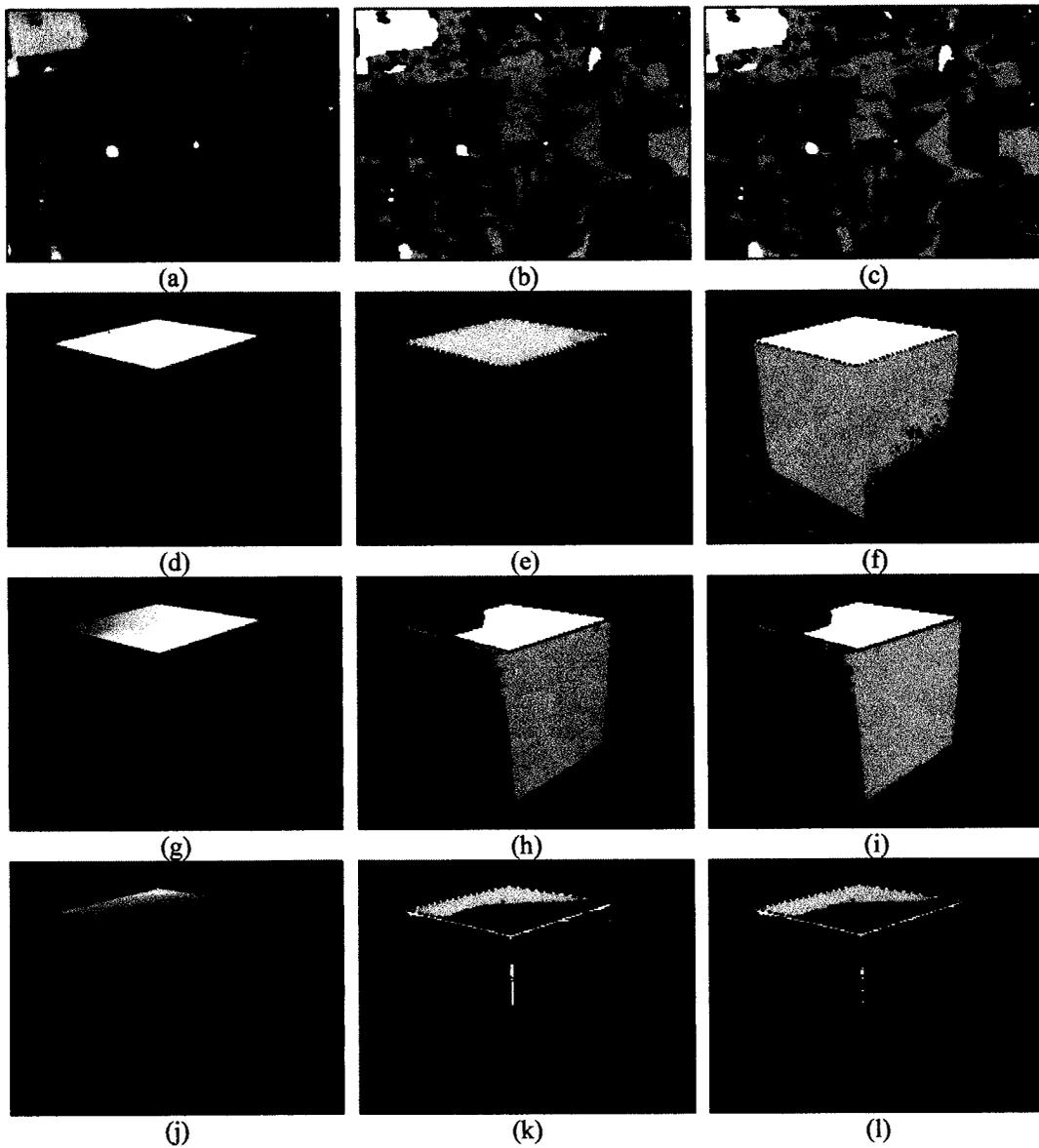


Figura 4.28. Resultados para a segmentação através de mapas auto-organizáveis com incorporação da intensidade luminosa. Imagens originais: (a): Aerial101; (d): Cube009; (g): Cube010; (j): Cube014; (m): Cube102; (p): Cube104; (s): Cube105; (v): Cube203; (y): House02; (ab): House03; (ae): House05; (ah): Pyram01; (ak): Pyram04; (an): Pyram06; (aq): Donna; (at): Flowers. (b), (e), (h), (k), (n), (q), (t), (w), (z), (ac), (af), (ai), (al), (ao), (ar), (au): Imagens segmentadas utilizando intensidade e cromaticidade, relativas às imagens da primeira coluna, respectivamente. (c), (f), (i), (l), (o), (r), (u), (x), (aa), (ad), (ag), (aj), (am), (ap), (as), (av): Imagens segmentadas utilizando intensidade, relativas às imagens da primeira coluna, respectivamente.

faixa de valores tornou-se o padrão para o processamento das demais imagens e raramente modificações são necessárias.

#### 4.2.4 Uso da cromaticidade

Como já discutido anteriormente, o uso da cromaticidade foi adotado como normalização dos dados para evitar a luminosidade. Procura-se evitar, assim, sombras e outras variações de iluminação. Uma comparação do método de segmentação usan-

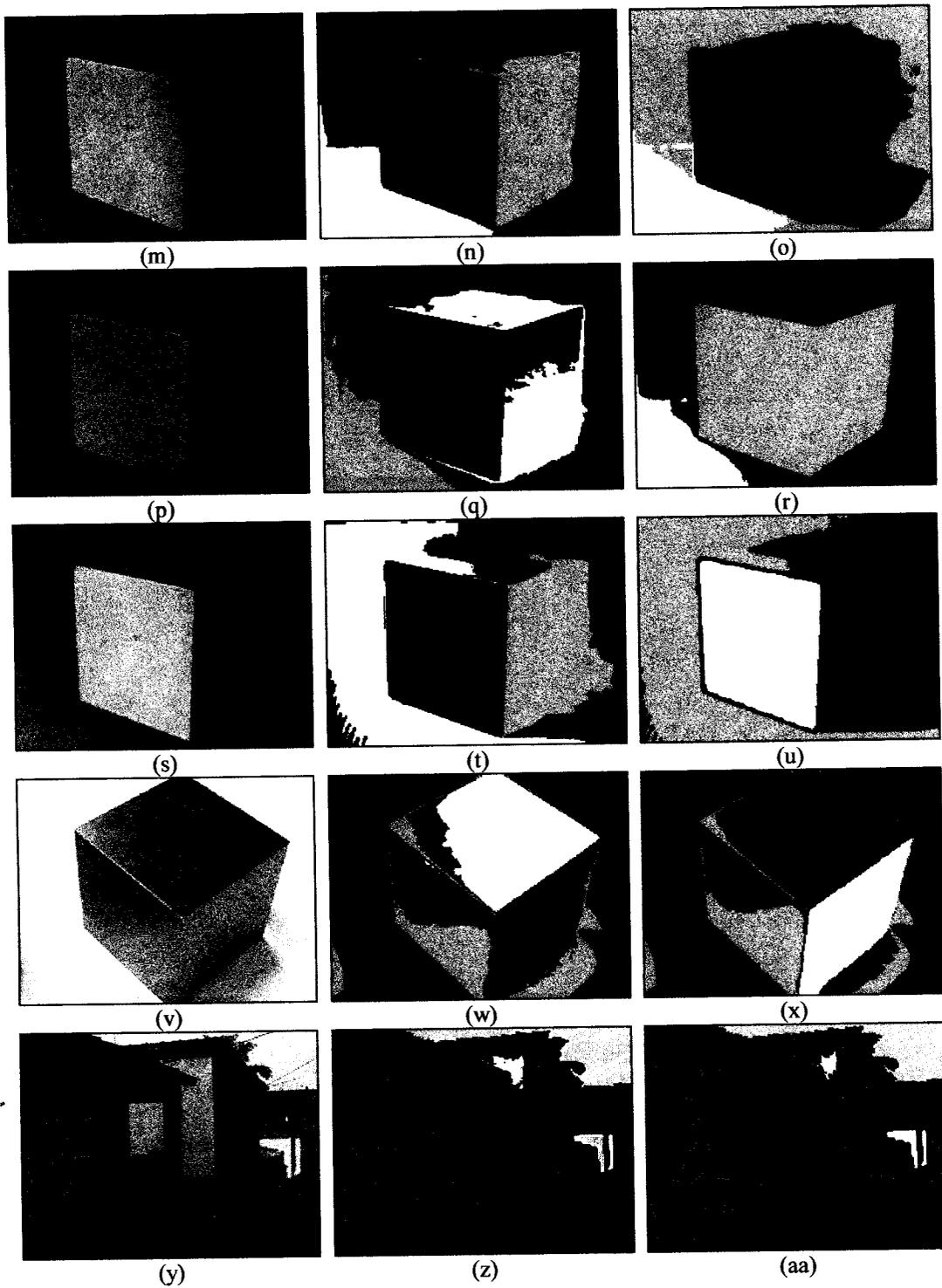


Figura 4.28. *Continuação.* (m): Cube102; (p): Cube104; (s): Cube105; (v): Cube203; (y): House02. (n), (q), (t), (w), (z): Imagens segmentadas utilizando intensidade e cromaticidade, relativas às imagens da primeira coluna, respectivamente. (o), (r), (u), (x), (aa): Imagens segmentadas utilizando intensidade, relativas às imagens da primeira coluna, respectivamente.

do a cromaticidade com sua aplicação usando também a iluminação da imagem é necessária para identificar suas diferenças básicas.

Nesta comparação optou-se por duas formas na inclusão da componente de iluminação. A primeira é a aplicação do método diretamente às componentes RGB da

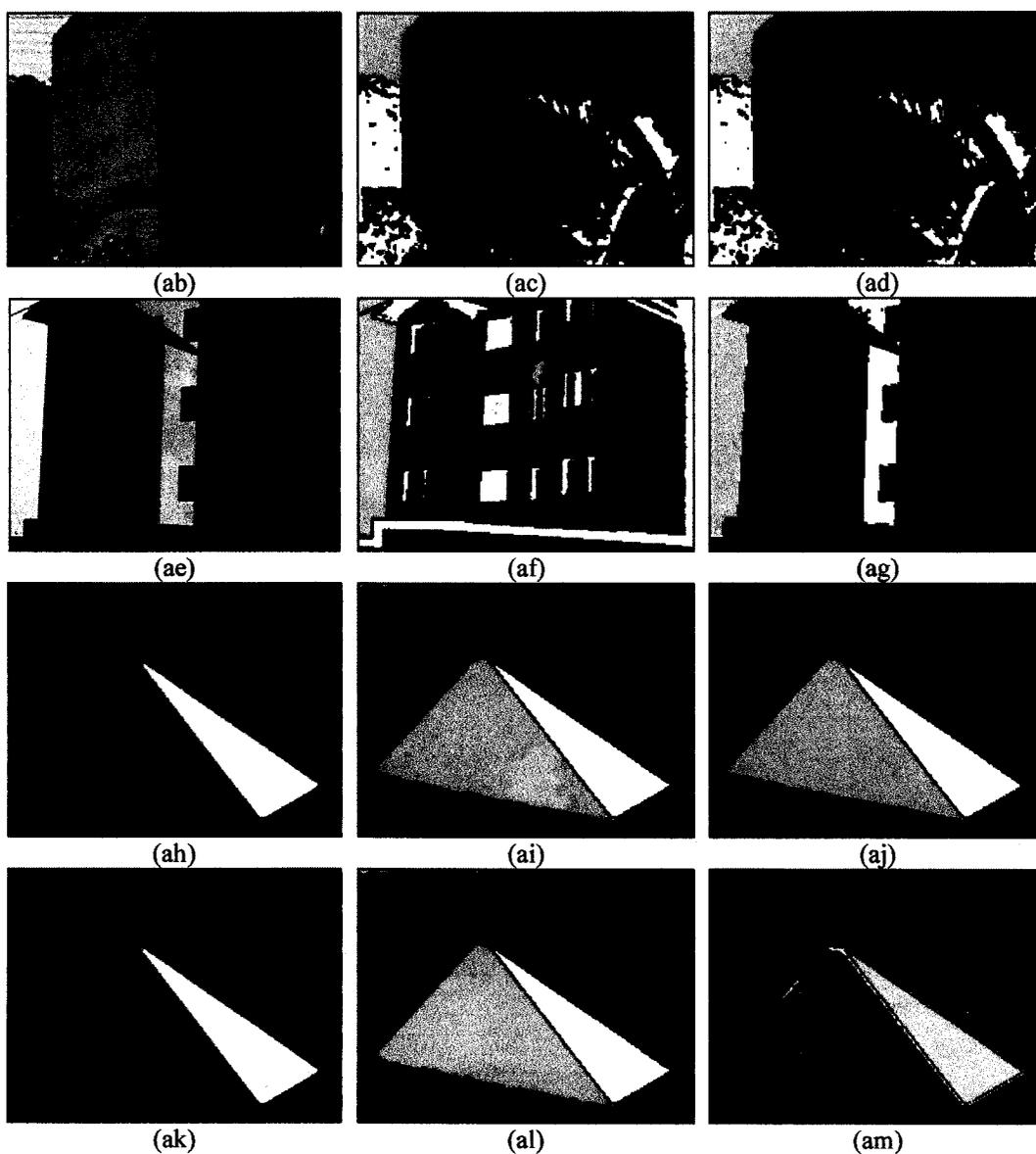


Figura 4.28. *Continuação.* (ab): House03; (ae): House05; (ah): Pyram01; (ak): Pyram04. (ac), (af), (ai), (al): Imagens segmentadas utilizando intensidade e cromaticidade, relativas às imagens da primeira coluna, respectivamente. (ad), (ag), (aj), (am): Imagens segmentadas utilizando intensidade, relativas às imagens da primeira coluna, respectivamente.

imagem, sem sua normalização. Outra é a utilização, no mapa, de três atributos: as componentes de cromaticidade  $r$  e  $g$  (lembrando que  $b$  é correlacionada) e a componente de iluminação.

A diferença básica entre as duas alternativas é que a primeira utiliza a influência da iluminação em todas as componentes diretamente, ao passo que na segunda a iluminação é considerada um atributo “separado” das componentes de cromaticidade. Com a mudança da estruturação dos atributos, altera-se também sua distribuição e o conseqüente comportamento da metodologia de segmentação.

Em virtude das classes identificadas, quando usada a iluminação, não se comportarem conforme a expectativa das classes de referência, uma comparação do de-

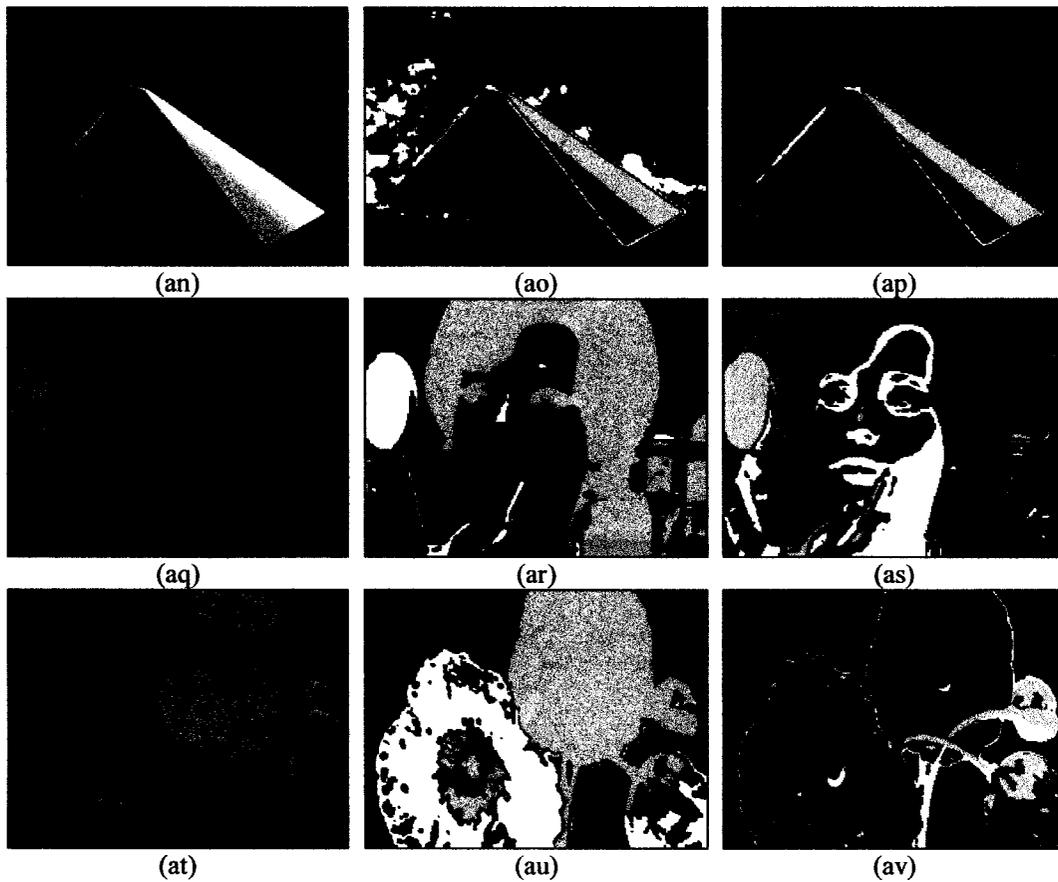


Figura 4.28. *Continuação*. (an): Pyram06; (aq): Donna; (at): Flowers. (ao), (ar), (au): Imagens segmentadas utilizando intensidade e cromaticidade, relativas às imagens da primeira coluna, respectivamente. (ap), (as), (av): Imagens segmentadas utilizando intensidade, relativas às imagens da primeira coluna, respectivamente.



Figura 4.29. Imagem House06.

sempenho em termos de taxa de acerto não é adequada, o que leva a uma discussão em termos mais subjetivos. Os parâmetros utilizados no mapa para as segmentações que incorporam intensidade foram semelhantes aos utilizados para o caso da cromaticidade, com mesmos tamanho do mapa, taxas de decaimento, erro de convergência e valores iniciais para raio e taxa de aprendizado.

A Figura 4.28 apresenta alguns dos resultados da segmentação com a inclusão da intensidade luminosa, indicando que a diferenciação mais evidente, quando se considera a componente de iluminação, é sua grande influência na segmentação. Esta

situação é bastante clara, por exemplo, nas imagens *Aerial01*, *Cube009*, *Cube010*, *Cube014*, *Cube104*, *Cube203*, *House03* e *Pyram06*, apresentadas na Figura 4.28(a), (d), (g), (j), (p), (v) e (an), respectivamente. Considerando-se uma iluminação mais uniforme e bem comportada, a utilização da iluminação pode levar a segmentações interessantes, como é o caso de *Cube104*, *Cube203*, *House02*, *House03*, *House05*, *Pyram04*, *Donna* e *Flowers*, Figura 4.28(p), (v), (y), (ab), (ae), (ak), (aq) e (at), respectivamente, nas quais as regiões, embora subdividam uma região visualmente homogênea em vários casos, apresentam bastante definição nas áreas dos limites entre regiões. Algumas destas segmentações são bastante interessantes, como por exemplo a da Figura 4.28(ah), *Pyram01*, na qual as classes foram identificadas corretamente, apresentando taxas de acerto de 99,6% para a face verde, 88,9% para a face amarela e 99% para o fundo. Por outro lado, a separação em classes pode ser desviada bastante do esperado, como para *Cube010*, *Cube102* ou *Pyram06*, como pode ser observado na Figura 4.28(g), (h) e (i), (m), (n) e (o), (an), (ao) e (ap), respectivamente.

De um modo geral, a utilização exclusiva da cromaticidade condiz mais proximamente com as expectativas das classes de referência, com taxas de acerto relativamente altas para vários casos. A inserção da iluminação como atributo, seja nas três componentes cromáticas, seja na cromaticidade, usualmente leva à segmentação excessiva, embora separe geralmente de forma coerente os dados da imagem\*. Outra característica importante da iluminação é nas bordas entre regiões. Nos locais onde se esperam bordas entre objetos, a coincidência com limites entre regiões é maior que a obtida usando cromaticidade. Apesar disto, em várias ocorrências, limites entre regiões não podem ser associados a bordas (subdivisão em virtude de sombras, por exemplo).

A utilização de uma das três alternativas ainda é dependente dos dados a serem processados, caracterizados pelo tipo de cena a ser capturada, considerando fatores como condições gerais de iluminação, sombras, presença de texturas etc.

### 4.3 Aplicação

A segmentação utilizando cromaticidade e mapas auto-organizáveis, embora tenha se mostrado eficiente em termos dos resultados produzidos, apresenta algumas deficiências que limitam uma aplicação muito geral. Uma delas é a sensibilidade à cromaticidade em áreas muito escuras da imagem, nas quais não há “cor” (veja-se o fundo na imagem da Figura 4.28(an), *Pyram06*). A tolerância a texturas é também baixa, como visto nas áreas de telhados das imagens *House02* e *House06*, Figuras 4.28(y) e 4.29, respectivamente, e da vegetação de *House03*, Figura 4.28(ab).

Para ambientes mais controlados e que permitam a aquisição das imagens com uma iluminação mais adequada, entretanto, o método apresenta boa robustez, permitindo a determinação automática das classes.

---

\* Com exceções evidentes, como *Cube105* para a utilização das componentes RGB.

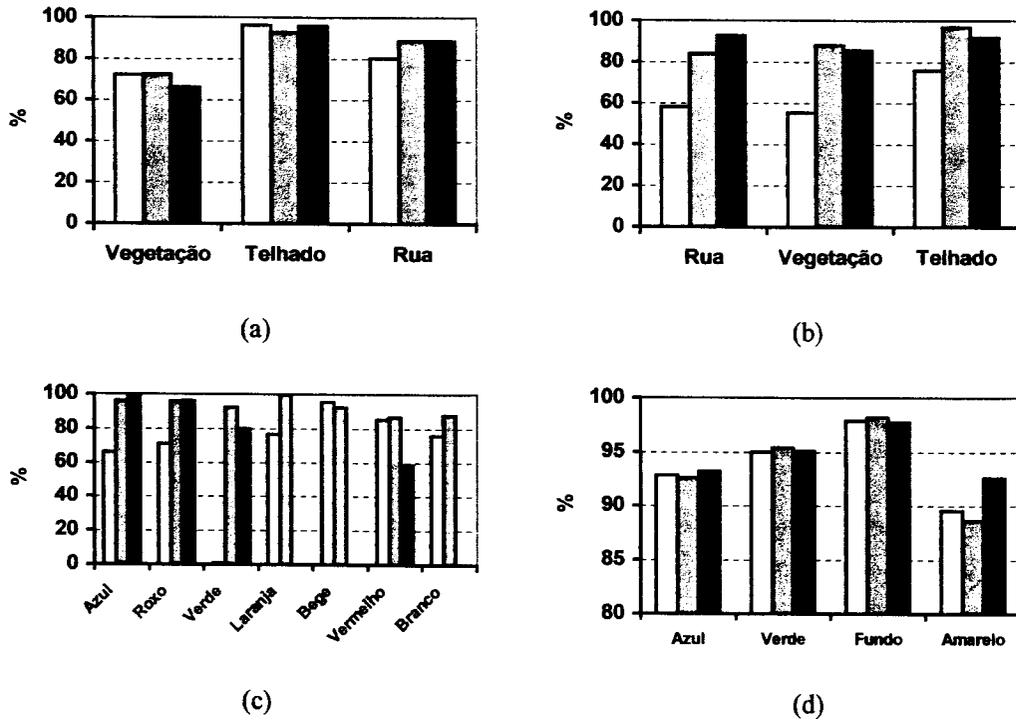


Figura 4.30. Comparação das taxas de acerto para as classes das figuras do conjunto de testes, apresentadas por imagem. Em branco, a taxa obtida usando a segmentação proposta; em cinza médio, usando a classificação bayesiana; e em cinza escuro as taxas produzidas pela classificação com  $k$  vizinhos mais próximos. (a) Aerial01. (b) Aerial02. (c) Balloons. (d) Cube009. (e) Cube010. (f) Cube011. (g) Cube014. (h) Cube015. (i) Cube101. (j) Cube102. (k) Cube104. (l) Cube105. (m) Cube202. (n) Cube203. (o) Cube204. (p) Flowers. (q) Flower02. (r) Frogs02. (s) House01. (t) House02. (u) House03. (v) House05. (w) House06. (x) House07. (y) Pyram01. (z) Pyram03. (aa) Pyram04. (ab) Pyram06. (ac) Sign01. (ad) Sign02. (ae) Sign03. (af) Sign04. (ag) Toy. (ah) Wall01. (ai) Wall02.

#### 4.4 Comparação da classificação com mapa auto-organizável com as classificações bayesiana e por $k$ vizinhos mais próximos

Para avaliar o desempenho da metodologia, os resultados relativos às taxas de acerto para as classes nas imagens de testes foram comparados aos resultados produzidos pela segmentação bayesiana e pela classificação pelos  $k$  vizinhos mais próximos.

A classificação bayesiana, apresentada na seção 3.2.1, utiliza-se da matriz de covariância, a qual representa as variâncias e covariâncias dos vários atributos utilizados para o cálculo da probabilidade *a posteriori*. Permite-se, pela regra de Bayes [Duda & Hart (1973)], que a probabilidade *a posteriori* seja avaliada pela Equação 4.14\*. Esta classificação é ótima no sentido estatístico.

$$P(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\bar{\mu})' \Sigma^{-1}(\mathbf{x}-\bar{\mu})} \quad (4.14)$$

\* Repetição da Equação 3.38, descrita na página 41.

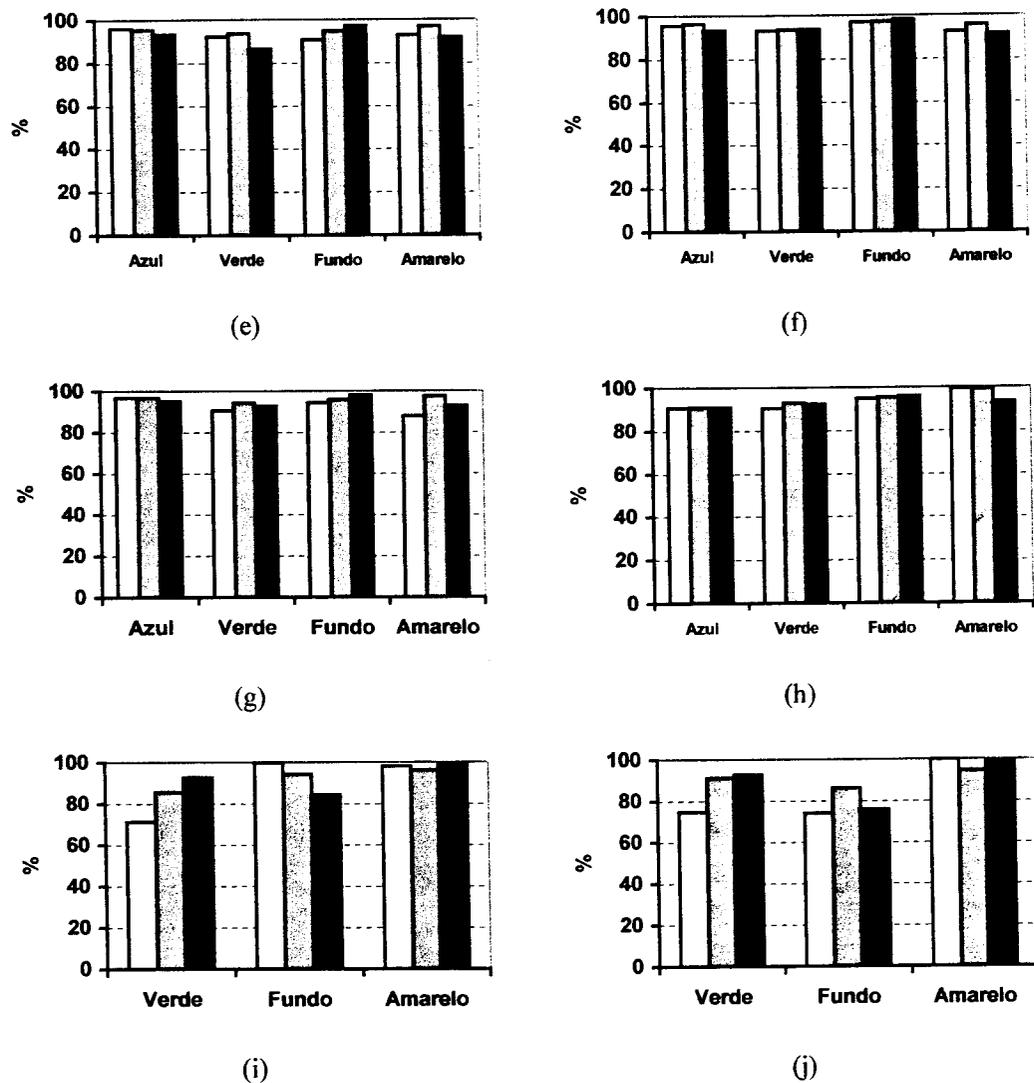


Figura 4.30. *Continuação.* (e) Cube010. (f) Cube011. (g) Cube014. (h) Cube015. (i) Cube101. (j) Cube102.

As probabilidades *a priori* especificadas para os testes foram distribuídas igualmente entre as classes, para que não houvesse favorecimentos arbitrários. A partir de um conjunto de dados extraídos da imagem, associados previamente a classes, é possível estimar a matriz de covariância e definir as funções de densidade. São estimadas, assim, as probabilidades *a posteriori* de um dado ponto pertencer a uma das classes. A classe que proporcionar a maior probabilidade determina a rotulação deste ponto. Para os testes, foram selecionados arbitrariamente 25 amostras de cada classe para determinar os parâmetros da classificação. Os resultados das classificações estão apresentados no Apêndice D, juntamente com outros resultados. As classes utilizadas são as mesmas representadas na imagem de regiões de referência.

Outra forma de classificação é o método dos  $k$  vizinhos mais próximos (abreviada aqui para  $k$  vizinhos) [Duda & Hart (1973)]. Pelo método, um conjunto de amostras é previamente rotulado, definindo amostras para todas as classes. Um dado ponto a ser classificado é comparado, por um critério de distância, a cada uma das

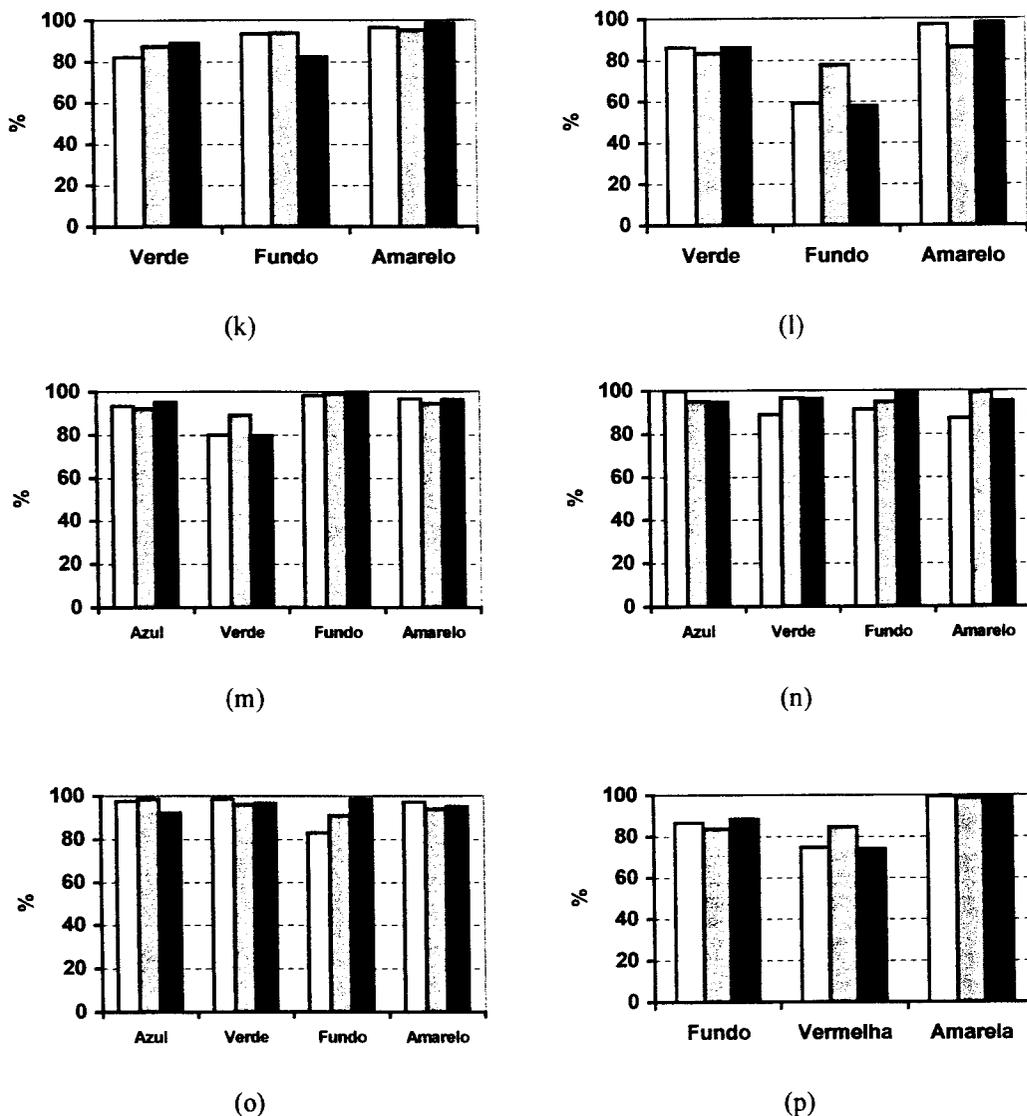


Figura 4.30. *Continuação.* (k) Cube104. (l) Cube105. (m) Cube202. (n) Cube203. (o) Cube204. (p) Flowers.

amostras e são escolhidas as  $k$  (arbitrário) mais próximas. A classe associada ao ponto é a que tiver maioria entre os  $k$  valores mais próximos. Em caso de empate, pode-se optar por um critério, como proximidade a uma amostra em particular, para definir a classe resultante. Nos experimentos realizados foram escolhidos 25 amostras de cada classe e  $k$  foi definido como 20 vizinhos, em função da extensão da análise deste método realizada anteriormente [Moreira & Costa (1995)].

A Figura 4.30 apresenta as taxas de acerto para as classes de várias das imagens do Apêndice D, ilustrando comparativamente os resultados da classificação proposta com as duas outras técnicas de classificação descritas. Uma análise geral dos histogramas apresentados permite verificar um desempenho aproximadamente equivalente para as três formas de classificação, com uma concordância qualitativa da taxa de acerto. Entretanto algumas discrepâncias ocorrem. A classificação bayesiana mostrou-se a mais robusta sob o aspecto de não falhar na identificação de uma das clas-

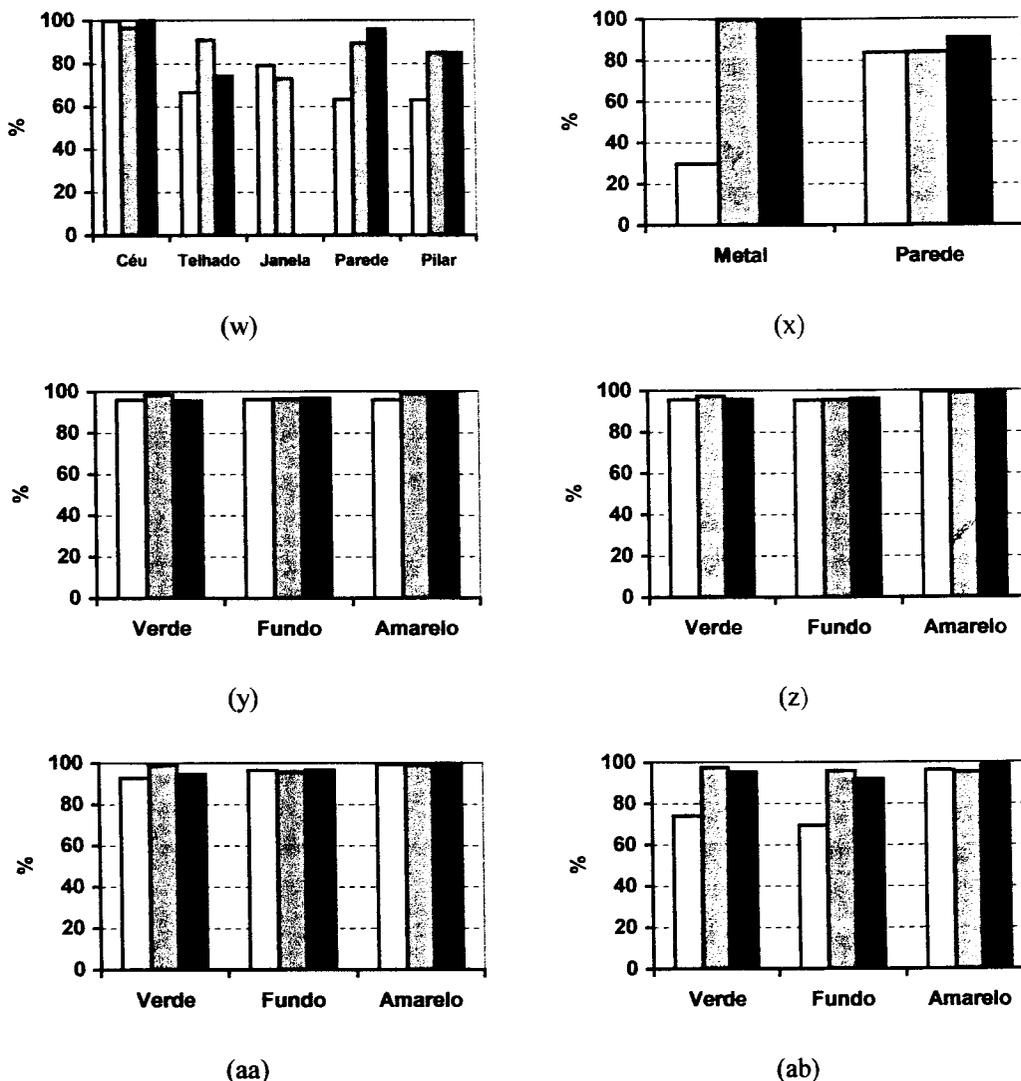


Figura 4.30. *Continua  o*. (w) House06. (x) House07. (y) Pyram01. (z) Pyram03. (aa) Pyram04. (ab) Pyram06.

ses. Para imagens mais complexas, como Balloons, House01, Sign02, Sign04 e Toy, por exemplo, tanto a metodologia proposta quanto a classifica  o por *k* vizinhos n o reconheceram uma ou mais classes.

No gr fico para a imagem Toy, Figura 4.30(ag), por exemplo, a classifica  o com o mapa auto-organiz vel n o separou os dados do mapa para identificar os pontos relativos   classe azul (anel azul do brinquedo) nem relativos   classe base (base do brinquedo), esta  ltima “incorporada”   classe metal (objeto met lico acobreado). Da mesma forma, o anel amarelo n o foi corretamente identificado, e a taxa de acerto (em torno de 16%) foi meramente casual, uma vez que a regi o identificada cobre mais a parte superior do anel vermelho que o anel amarelo. Esta classifica  o pode, ent o, ser desconsiderada. A classifica  o por *k* vizinhos tamb m n o se mostrou robusta neste aspecto para as imagens mais complexas. Usando a mesma imagem exemplo, por *k* vizinhos tamb m n o se puderam detectar as classes azul, base e amarelo,

Jander Moreira

INSTITUTO DE PESQUISA E DESENVOLVIMENTO EM CI NCIAS E TECNOLOGIA

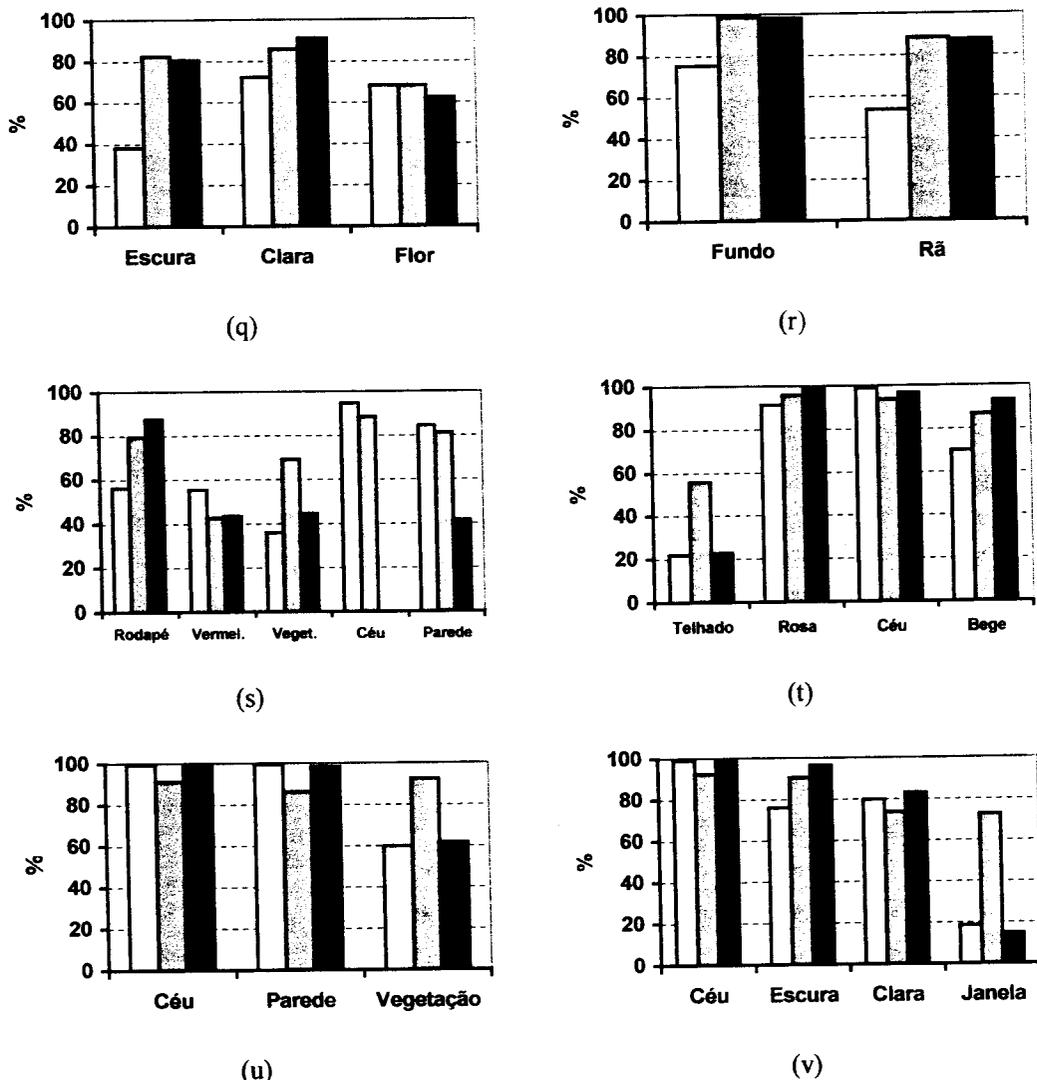


Figura 4.30. *Continuação.* (q) Flower02. (r) Frogs02. (s) House01. (t) House02. (u) House03. (v) House05.

cujos pontos foram associados a outras classes. Nas demais imagens com falhas de identificação de classes, situações semelhantes ocorreram.

As principais discrepâncias entre os três métodos ocorreram em situação particulares, como presença de sombras muito escuras ou regiões pouco homogêneas. Os métodos de classificação bayesiana e por  $k$  vizinhos, que são supervisionados na ocasião da determinação das classes, tendem a ter um comportamento mais eficiente nas sombras pesadas, como é o caso da imagem *Pyram06*, para a qual a sombra pesada na face verde modificou a cromaticidade o suficiente para induzir o mapa auto-organizável a considerar uma quarta classe, cobrindo esta sombra. Esta situação é evitada pelos outros dois métodos, dado que a escolha das classes, bem como das amostras utilizadas para caracterizá-las, é arbitrária e fixa. Nesta situação, a rotulação dos pontos é levada a umas das classes existentes, o que, para um caso geral, tende a ser a classe esperada, mesmo que os atributos difiram das amostras iniciais. Esta tendência é nítida para a classificação bayesiana, a qual, determinada uma classe para

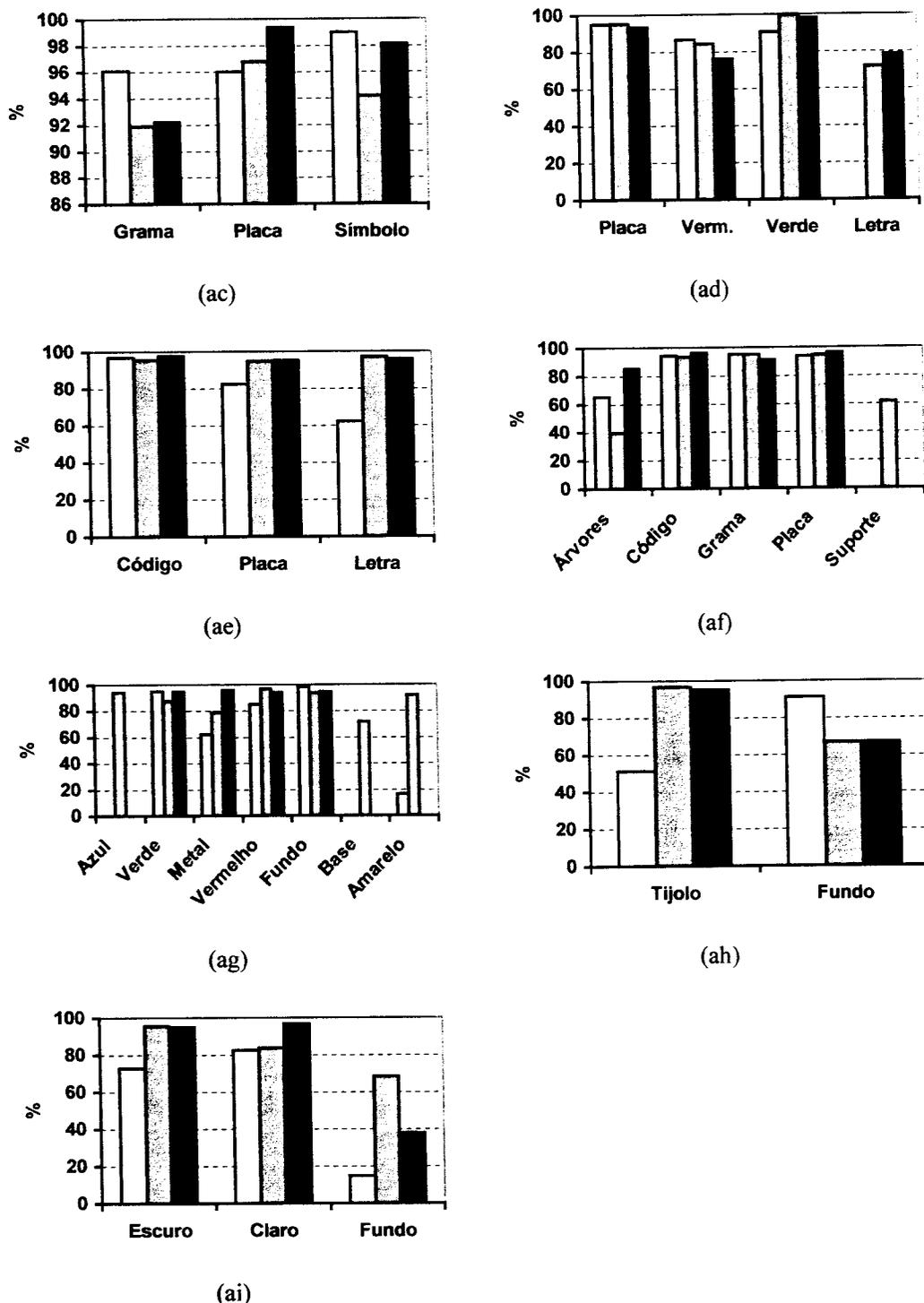


Figura 4.30. Continuação. (ac) Sign01. (ad) Sign02. (ae) Sign03. (af) Sign04. (ag) Toy. (ah) Wall101. (ai) Wall102.

alguns pontos da imagem, permite associar a esta classe pontos semelhantes, como é o caso da classe *suporte* na imagem *Signal04*, que corresponde aos postes de suporte da placa.

De forma geral, o desempenho da metodologia de segmentação proposta se mostrou bastante robusto para aplicações genéricas, com configuração não crítica dos parâmetros do mapa. O desempenho é satisfatório quando comparada às técnicas su-

pervisionadas apresentadas, com desvantagem para o caso de subdivisões de regiões não esperadas. Mesmo nestas situações, a análise visual da imagem segmentada permite confirmar a coerência entre as regiões encontradas e o conteúdo das imagens.

Um último comentário é relativo às imagens *Donna* e *Lenna*, excluídas da análise de classes. Para estas imagens, a segmentação manual das imagens acabou utilizando critérios de alto nível para a definição das classes, ignorando diferenças fortes na iluminação que não puderam ser desprezadas pelas três formas de segmentação. Por exemplo, na imagem *Lenna*, a forte diferença de iluminação no chapéu acabou sendo ignorada ao se estabelecer uma classe, o que ocorreu, também, na face. A segmentação manual, neste caso, impôs uma classe cujos pontos da imagem que a compõem não podem ser considerados aproximadamente constantes em relação aos atributos de cromaticidade. A mesma situação ocorre para a imagem *Donna* (por exemplo, nos vidros de perfume e na face, com menor intensidade).

Analisar as classes para estas imagens não se mostrou produtivo, dado que a correspondência entre as classes detectadas automaticamente e as da região de referência seriam arbitrárias, influenciando nos resultados. Para a imagem *Lenna*, por exemplo, parte do chapéu e parte do rosto foram identificadas como uma única classe. Determinar se a comparação desta região seria com a classe *chapéu* ou a classe *rosto* exige uma arbitrariedade que não permite uma análise mais objetiva dos resultados. Subjetivamente, porém, as segmentações são coerentes como conteúdo das cenas, embora não da forma esperada.

#### 4.5 Comentários finais

Os resultados da segmentação foram satisfatórios na maioria dos casos aplicados, como pode ser observado nas imagens de regiões do Apêndice D. A confiabilidade do método proposto é alta em termos de resultados similares para imagens similares, bem como resultados similares foram obtidos em execuções diferentes para a mesma imagem. Mesmo havendo um fator aleatório na escolha das amostras para o treinamento do mapa, sua adaptação gradual e suas características de ajuste para os pesos internos absorvem bem as características de cromaticidade extraídas da imagem.

A utilização da cromaticidade em vez das componentes RGB diretamente, reduzindo drasticamente a influência da iluminação, permitiu uma sensibilidade menor a sombras mais leves (que não escurecessem demais a região, alterando sua cromaticidade). Como consequência, os resultados se mostraram coerentes quase que na totalidade dos casos, com as expectativas iniciais das classes para as imagens.

Uma das grandes contribuições que a metodologia proposta apresenta é a ausência de supervisão para a determinação das classes relativas às regiões da imagem. Isso elimina uma etapa manual no processo de segmentação, o qual exige que um operador tenha acesso prévio às classes existentes. A versatilidade de um esquema automático é, assim, uma característica importante, principalmente considerando-se um sistema de visão com uma gama de aplicações variada.

Em termos de esforço computacional, a implementação do processo usando um *hardware* específico pode reduzir drasticamente o tempo de execução, que pode ser elevado (da ordem de alguns minutos) para mapas grandes e taxas de decaimento muito pequenas. Atualmente, a segmentação completa de uma imagem consome entre 10 e 90 segundos, dependendo dos dados da imagem, considerando-se a escolha das configurações iniciais e as estruturas definidas como padrão. O equipamento utilizado consiste em um microcomputador padrão PC, com processador Pentium II (333 MHz) da Intel Corp., e memória de 128 Mbytes.

## 5 Estudo comparativo de técnicas de segmentação

Importantes em sistemas de visão, as bordas, como caracteres de nível baixo, representam um passo inicial para a compreensão do conteúdo de uma imagem, carregando informações relevantes sobre ela [Schalkoff (1989)]. O *Cyvis-1*, que incorpora esta representação de nível mais baixo no processo de análise, tem como uma de suas propostas a utilização de informações de vários subsistemas. Da seleção adequada destas informações depende o funcionamento adequado do sistema. Assim, a comparação de mapas de bordas diferentes pode ser utilizada para fundir informações de subsistemas distintos, promovendo a cooperação e integração dentro de um sistema de visão com maior abrangência.

O estudo de técnicas cujo objetivo é a detecção das bordas presentes na imagem tem sido constante na comunidade de processamento de imagens [Pal & Pal (1993); Skarbek & Koschan (1994)]. A possibilidade de comparar os diferentes métodos, entretanto, tem praticamente se restringido a uma avaliação mais subjetiva. A indicação de bons detectores de bordas, em geral, considera o exame visual dos resultados produzidos e a verificação se as bordas obtidas correspondem às esperadas [Heath *et al.* (1997); Demigny & Kamlé (1997)]. Da mesma forma, cada autor apresenta seus resultados ressaltando suas vantagens e sua aplicabilidade a um determinado conjunto característico de imagens ou problema. A comparação com outros métodos se dá, em geral, através da confrontação visual entre os vários mapas de bordas.

Segundo Heath *et al.* [Heath *et al.* (1997)], as avaliações dos resultados obtidos por detectores de bordas podem ser classificadas em dois grandes grupos: teórico e experimental. As avaliações teóricas incluem modelos matemáticos, desconsiderando a aplicação real do algoritmo à imagem. A entrada do algoritmo é caracterizada matematicamente e o desempenho é avaliado analiticamente ou por simulações. As avaliações experimentais, caracterizadas pela avaliação dos mapas de bordas efetivamente produzidos, podem abranger duas vertentes principais: as que utilizam e as que

não utilizam a posição real das bordas. A análise das bordas produzidas pode ser feita, no primeiro caso, comparando-se o mapa produzido às posições reais das bordas e, no segundo, considerando-se somente a “qualidade” das bordas, em termos de espessura ou continuidade, por exemplo.

Como alternativa de avaliação, Heath *et al.* [Heath *et al.* (1997); Heath *et al.* (1998)] propõem que a comparação das técnicas diferentes seja realizada de forma subjetiva, através de uma análise visual. Mapas de bordas, produzidos em condições diversas (métodos distintos e variações nos parâmetros de cada método), são julgados através de notas por um conjunto de pessoas, as quais qualificam com grau maior os que permitem identificar o melhor objeto presente na imagem. Características dos mapas como exatidão da posição das bordas e bordas irrelevantes\* são desconsideradas, levando a uma comparação extremamente subjetiva.

Avaliar resultados produzidos por métodos diferentes, sob um prisma mais quantitativo, não é uma tarefa simples. Este capítulo descreve uma proposta básica para uma metodologia de comparação dos mapas de bordas de origens diversas, apresentando uma forma de combinar os resultados de modo a minimizar um critério geral de erro. Assim, uma caracterização dos tipos de bordas produzidas por diferentes métodos pode ser obtida, bem como a avaliação de sua qualidade em termos dos atributos locais da imagem. A partir da análise, optar por uma ou outra técnica se torna viável, considerando melhorar o mapa de bordas como um todo.

Dentro deste contexto, este capítulo apresenta a estruturação deste modelo comparativo, indicando e caracterizando também um conjunto de atributos locais selecionados para esta investigação (seção 5.1.1). A seção 5.1 disserta sobre o modelo de comparação utilizado e os resultados obtidos e discussões pertinentes são apresentados nas seções 5.2 e 5.3, acrescentadas as conclusões pertinentes. Uma proposta de combinação dos resultados da avaliação é apresentada na seção 5.4 e seus principais resultados são mostrados e discutidos na seção 5.5. As principais conclusões do modelo e da combinação são abordadas na seção 5.6.

## 5.1 Modelo comparativo

Em imagens, bordas estão diretamente associadas a descontinuidades em alguns de seus atributos locais, usualmente induzidas pelos limites dos elementos que compõem a cena [Pratt (1991); Gonzales & Woods (1992); Castleman (1996)]. Esta ligação estreita das bordas com as descontinuidades faz com que os dados locais da imagem, que caracterizam as descontinuidades, apresentem-se como elementos de especial importância para a compreensão do comportamento dos resultados produzidos por métodos de segmentação variados.

---

\* As imagens, em níveis de cinza, além de manipuladas no brilho e contraste para produzir uma imagem visualmente melhor, têm seu objeto de atenção centrado. Regiões fora da área central (e as bordas ali presentes) são, provavelmente, ignoradas no processo de avaliação.

Assim, considerando imagens coloridas com atenção especial, o estudo comparativo aqui proposto analisa a qualidade das bordas produzidas por alguns métodos de segmentação, examinando a forma com que determinam as bordas segundo um conjunto de atributos locais.

Ainda dentro destas considerações, somente técnicas voltadas à detecção de bordas são tradicionalmente incluídas [Heath *et al.* (1997); Demigny & Kamlé (1997)]. As segmentações voltadas a regiões, sumariadas na seção 3.2 e acrescidas de uma nova proposta no Capítulo 4, embora não direcionadas à detecção de descontinuidades, como é o caso das técnicas que localizam bordas, acabam realizando este trabalho de forma indireta. Ao determinar quais regiões da imagem compartilham um atributo, ou seja, ao delimitar as regiões que compartilham certa homogeneidade de conteúdo, consideram também as modificações nas fronteiras destas regiões, as quais caracterizam exatamente as descontinuidades.

Assim, como apresentado na seção 5.1.2, através da determinação dos pontos que delimitam as regiões é possível o estabelecimento do mapa de bordas correspondente àquela segmentação. Torna-se viável, então, utilizar mapas de bordas provenientes não somente da detecção de descontinuidades, mas também das segmentações por regiões. Exemplos destes mapas estão apresentados no Apêndice D, juntamente com algumas segmentações originalmente direcionadas a bordas.

Para que diferentes técnicas possam ser comparadas umas às outras, em termos de seus mapas de bordas, um modelo que permite avaliar seus desempenhos relativos é proposto. O princípio adotado foi a investigação do comportamento dos métodos ao rotularem um ou outro ponto como um ponto de borda, tomado em função de um conjunto de atributos calculados localmente. Assim, este estudo é baseado em uma análise do comportamento de um conjunto de técnicas de segmentação que produzem mapas de bordas, fazendo uso das variâncias definidas na seção 5.1.1 como atributos. A associação das bordas com as variações dos atributos locais é uma característica que fundamenta esta abordagem.

### 5.1.1 Atributos locais

Como vários atributos locais podem ser considerados, procurou-se determinar um elenco, embora limitado, que distinguísse alguns dos principais atributos cromáticos da imagem. Deste modo, quatro atributos foram destacados, três deles considerando matiz, saturação e iluminação e outro, isoladamente, considerando apenas a iluminação. Levar em consideração variações dos níveis de cinza foi julgado importante não somente pela predominância deste tipo de representação de imagens [Pal & Pal (1993)] e pela consideração de métodos exclusivos para imagens em cinza, mas também por considerar a importância da iluminação na percepção de descontinuidades [Levine (1985); Tovée (1996)].

O primeiro atributo, designado por  $g$ , é calculado como a variância dos pontos da imagem em uma vizinhança, utilizando somente a componente de intensidade luminosa. Considera-se um círculo centrado em uma posição da imagem, e a variância

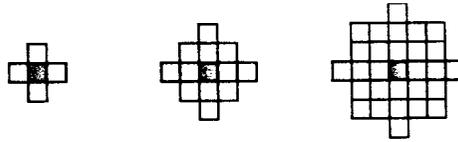


Figura 5.1. Círculos discretos de raios 1, 2 e 3.

(Equação 5.1) é calculada utilizando-se os dados circunscritos até uma distância  $l$ .  $N_C$  indica o número de pontos dentro do círculo discreto  $C$ ,  $i$  representa cada um dos pontos circunscritos,  $I_i$  é o valor da iluminação do ponto  $i$  e  $\bar{I}$  é a esperança dos valores  $I_i$ .

$$g = \sigma_g^2 = \frac{1}{N_C - 1} \sum_{i \in C} (I_i - \bar{I})^2 \quad (5.1)$$

As variâncias são calculadas para cada um dos pontos da imagem, excluídas as bordas, nas quais os círculos cobririam áreas não definidas. Os círculos definem quais os pontos vizinhos que contribuem para o cálculo, incluindo aqueles cuja distância ao centro seja inferior ou igual ao raio. Estabelecem-se assim, vizinhanças circulares discretas (Figura 5.1).

Os demais atributos consideram as cores e são calculados sobre os dados RGB das imagens. O atributo  $c$  corresponde à variação dos dados RGB em relação à sua média. Cada tríade de coordenadas RGB é usada para definir o ponto final de um vetor em três dimensões, cuja outra extremidade está na origem. Para se disponibilizar uma medida escalar foi calculada a variância das distâncias das cores (sob o ponto de vista dos vetores tridimensionais) em relação à sua média. Este valor representa um “espalhamento” dos vetores que representam as cores para uma determinada posição da imagem. As Equações 5.2 a 5.4 ilustram a obtenção destes valores, sendo  $D_i$  o valor da distância para uma posição  $i$  dentro do círculo,  $R_i$ ,  $G_i$  e  $B_i$  as componentes RGB do ponto  $i$ ,  $R_j$ ,  $G_j$  e  $B_j$  as componentes RGB do ponto central e  $\bar{D}$  a distância média dentro desta vizinhança.

$$c = \sigma_c^2 = \frac{1}{N_C - 1} \sum_{i \in C} (D_i - \bar{D})^2 \quad (5.2)$$

$$D_i = \sqrt{(R_i - R_j)^2 + (G_i - G_j)^2 + (B_i - B_j)^2} \quad (5.3)$$

$$\bar{D} = \frac{1}{N_C} \sum_{i \in C} D_i \quad (5.4)$$

As Equações 5.5 e 5.6 mostram a obtenção das outras duas variâncias, calculadas sobre o atributo de oposição cromática nos canais RG (vermelho e verde) e YB

(amarelo e azul). Os valores  $S_i$  representam as diferenças entre as componente  $R$  e  $G$ , enquanto  $T_i$  representa a oposição entre  $Y$  e  $B$ , sendo  $Y$  calculado pela mistura\* de  $R$  e  $G$ , como mostrado na Equação 5.7.

$$rg = \sigma_{rg}^2 = \frac{1}{N_C - 1} \sum_{i \in C} (S_i - \bar{S})^2 \quad (5.5)$$

$$yb = \sigma_{yb}^2 = \frac{1}{N_C - 1} \sum_{i \in C} (T_i - \bar{T})^2 \quad (5.6)$$

$$S_i = R_i - G_i, \quad T_i = \frac{R_i + G_i}{2} - B_i \quad (5.7)$$

$$\bar{S} = \frac{1}{N_C} \sum_{i \in C} S_i, \quad \bar{T} = \frac{1}{N_C} \sum_{i \in C} T_i \quad (5.8)$$

A variância  $g$  representa a alteração de contraste relativa à intensidade luminosa em cada posição da imagem. Como variações nos níveis de cinza estão relacionadas a bordas na imagem, também há uma associação com o contraste, dado que, se objetos diferentes apresentam intensidades diferentes, também o contraste nos pontos que fazem a fronteira entre os objetos é maior. Esta afinidade intrínseca entre contraste e bordas permitem considerar a relevância deste atributo local na análise das bordas. E da mesma forma que contrastes de iluminação estão relacionados às bordas, igualmente estão as variâncias relativas às cores. Considerando-se as imagens coloridas, bordas entre objetos de cores distintas apresentam uma variação nos atributos cromáticos. Estas medidas de contraste local relacionadas à cor e às oposições cromáticas representam modificações nos atributos dos pontos da imagem e podem ser associadas a bordas entre elementos da cena.

A Figura 5.2 mostra a representação das variâncias calculadas para a imagem *Aerial01*. A versão monocromática da iluminação da imagem é apresentada na Figura 5.2(b). As Figuras 5.2(c) a (f) mostram, respectivamente, o mapa de variâncias calculadas para o contraste monocromático  $g$  e para os coloridos  $c$ ,  $rg$  e  $yb$ . Para melhorar a visualização, foi aplicada uma função logarítmica† a cada ponto para ressaltar os valores menos intensos e, na seqüência, foi feita a mudança de escala dos valores para 256 níveis.

O relacionamento entre as bordas dos elementos presentes na imagem (ruas, casas etc.) e as variâncias mais altas é bastante nítido no exemplo apresentado. Algu-

\* Mistura aditiva de cores, introduzida na seção 2.2 e representada na Figura 2.2.

† A função utilizada foi  $\log(v + 1)$ , sendo  $v$  o valor da variância em cada ponto. [Gonzales & Woods (1992)]

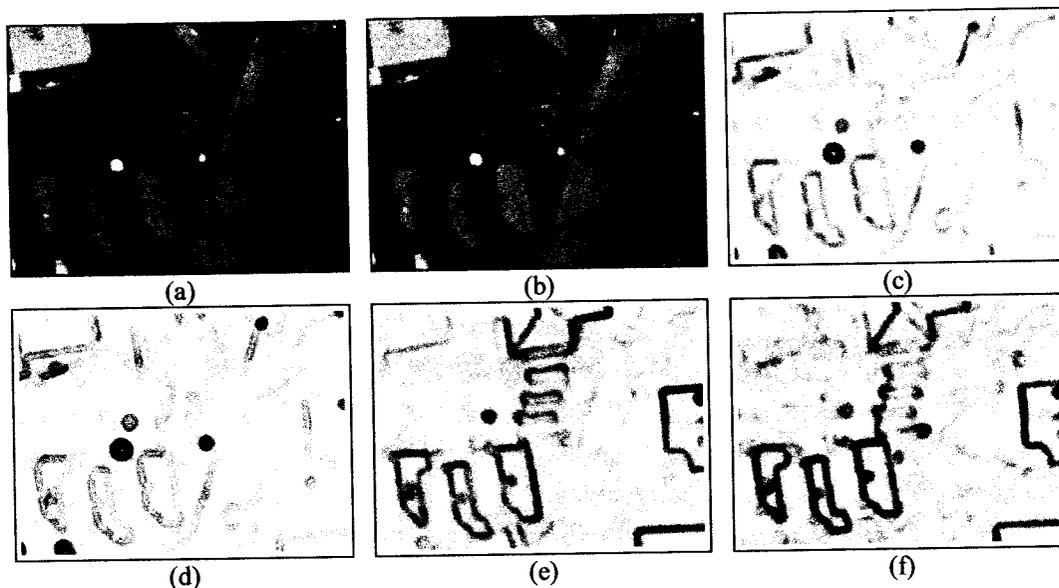


Figura 5.2. Exemplo das variâncias locais. (a) Imagem original *Aerial01*. (b) Versão em níveis de cinza. (c) Mapa da variância dos níveis de cinza. (d) Mapa da variância das cores RGB. (e) Mapa de variância da oposição *R* e *G*. (f) Mapa de variância da oposição *Y* e *B*. Valores mais altos são representados por tons mais escuros, após a aplicação de um logaritmo e a escalação para o intervalo  $[0, 255]$ . O raio utilizado para o cálculo foi de 3 pontos.

mas diferenças entre seus comportamentos também podem ser observados, como na região da rua (um pouco à direita do centro da imagem) na qual pode ser notada a sombra da casa, por exemplo. A sombra representa um contraste mais relevante na variância  $g$  e pode ser observado nos mapas das variâncias  $c$  e  $y_b$ . No caso do mapa da variância  $rg$ , na Figura 5.2(e) os limites entre a sombra e o restante da rua se mostra muito pouco significativo. Outra área em que as diferenças são mais visíveis é a do círculo branco, próxima ao centro da imagem, à esquerda, com menor representatividade nas variâncias  $rg$  e  $y_b$ .

### 5.1.2 Obtenção das bordas a partir de regiões

Estabelecido inicialmente para a avaliação dos mapas de bordas produzidos por técnicas de segmentação por regiões (em especial a proposta que utiliza o mapa auto-organizável), o modelo para comparação considera qualquer mapa de bordas descrito por uma imagem binária, permitindo comparações entre quaisquer técnicas. Assim, para esta avaliação, foram consideradas segmentações baseadas em regiões, como a por mapas auto-organizáveis, a produzida pela classificação bayesiana e a pela classificação por  $k$  vizinhos mais próximos. Para cada uma delas foi extraído o mapa de bordas a partir dos limites entre as regiões determinadas. É interessante notar, em particular, que essa classe de bordas é sempre de espessura dupla, com pontos rotulados como bordas em ambos os lados da fronteira (Figura 4.16), uma consequência da

definição de bordas entre fronteiras estabelecida na seção 4.1.3\*. Considerar, arbitrariamente, um ou outro lado como o lado correto para reduzir a espessura da borda foi cogitado, porém esta decisão influenciaria os resultados e nenhuma alternativa de escolha parece plenamente justificável. Por exemplo, poder-se-ia considerar o lado mais escuro como o lado da borda, pois a parte clara representaria um objeto e a parte escura representaria o fundo. Ter-se-ia, então, o contorno externo do objeto. Nem sempre, entretanto, os objetos são mais claros que o fundo, além do contorno externo poder não representar adequadamente o objeto. Como outras alternativas (por exemplo, considerando as cores ou o tamanho das regiões) também podem ser contrapostas, optou-se pela adoção da borda dupla por sua imparcialidade.

### 5.1.3 Outras técnicas comparadas

A comparação, além das técnicas citadas na seção anterior, envolve duas técnicas voltadas à detecção de descontinuidades: a segmentação usando os operadores de Sobel [Sobel (1970)] e através do gradiente vetorial [DiZenzo (1986)].

Ambas as técnicas são utilizadas para avaliar, de forma numérica, o valor do gradiente em cada ponto da imagem. Como as bordas estão associadas a descontinuidades, os pontos que apresentam gradiente com magnitude alta correspondem, em geral, a pontos limítrofes entre objetos. Calculadas as magnitudes do gradiente em cada ponto da imagem, pode-se determinar um limiar que selecione os valores maiores da magnitude, gerando o mapa de bordas. Pontos com valores acima do limiar estabelecido, então, são rotulados como bordas.

A principal diferenciação entre o gradiente de Sobel e o vetorial é o domínio da aplicação. As máscaras usadas para avaliar o gradiente propostas por Sobel são aplicadas aos níveis de cinza da imagem, quantificando, nas direções horizontal e vertical, as derivadas parciais. A magnitude do gradiente na posição  $(i, j)$ , dada na Equação 5.9, pode, assim, ser obtida calculando-se o módulo do vetor resultante, considerando-se  $G_h(i, j)$  e  $G_v(i, j)$  os resultados após a aplicação das máscaras horizontal e vertical, respectivamente.

$$G(i, j) = \sqrt{G_h^2(i, j) + G_v^2(i, j)} \quad (5.9)$$

No caso do gradiente vetorial, o domínio não é escalar e o gradiente é derivado em função do espaço vetorial descrito pelas componentes que formam a imagem colorida. Pode-se, assim, utilizar as componentes  $R$ ,  $G$  e  $B$ , definidas para cada ponto da imagem, na forma de um vetor  $(R, G, B)$ , ao qual a derivação do gradiente apresentada na seção 3.1.3 é aplicada. Obtém-se, desta forma, o valor da magnitude do gradiente considerando-se os dados de cor.

---

\* Deve ser notado, ainda, que em virtude da estruturação da comparação entre os mapas de bordas, a representação da borda entre pontos da imagem ("subpixel") [Ballard & Brown (1982)] não foi considerada.

As implementações de ambas as técnicas apresentam alguns detalhes, apresentados abaixo. Além destes, alternativas para segmentação através de mapas auto-organizáveis são também apresentadas.

### 5.1.3.1 Sobel

O gradiente gerado com as máscaras de Sobel é calculado sobre a versão monocromática da imagem. Uma convolução das máscaras com as intensidades dos pontos é realizada para cada posição da imagem, obtendo-se as componentes parciais do gradiente que são combinadas para estimar a magnitude do gradiente.

O mapa de bordas é produzido escolhendo-se o limiar que produz o melhor resultado visual. Esta escolha é feita para cada imagem em particular e o limiar varia conforme seu conteúdo. Como não é sempre nítida a melhor opção para o limiar, subjetivamente se escolhe aquele que produz um mapa de bordas adequado. A utilização de um ou outro valor como limiar é estritamente arbitrária.

Como esta arbitrariedade pode induzir um favorecimento da técnica pela opção das “melhores” bordas, procurou-se também incluir nas avaliações formas mais objetivas para a determinação do limiar. Foram incluídas, deste modo, duas opções para a avaliação automática do limiar: o limiar adaptável local e o limiar adaptável global.

O limiar adaptável local, como proposto por Robinson [Robinson (1977)], considera o valor do limiar como dependente da intensidade original da imagem no ponto, fazendo com que o limiar seja reduzido onde a imagem é mais escura e aumentado nas áreas mais claras. O limiar  $LAT(i, j)$ , para cada posição  $(i, j)$  da imagem, pode ser calculado pela Equação 5.10, na qual  $I^*(i, j)$  é o valor do nível de cinza após uma suavização e  $G(i, j)$  é o valor do gradiente na posição.

$$LAT(i, j) = \frac{I^*(i, j)}{G(i, j)} \quad (5.10)$$

Os valores do gradiente, após calculados, são escalados para o intervalo  $[0, 255]$  e, então, o limiar local é calculado. Para suavizar a imagem é feita a média entre os pontos vizinhos, através de uma convolução com uma máscara de tamanho  $11 \times 11$ . Esta dimensão para a máscara foi escolhida experimentalmente e foi mantida para todo o conjunto de imagens. Alterações nesta máscara, entretanto, não induzem mudanças significativas nos resultados obtidos.

O gradiente com limiar adaptável global considera a forma usual do histograma dos gradientes de uma imagem. Em geral, há uma certa quantidade de gradientes com valores menores, os quais correspondem às áreas homogêneas da imagem, e uma quantidade usualmente menor de gradientes mais elevados, os quais corresponderiam a pontos de bordas. A Figura 5.3 mostra os histogramas para os valores já escalados do gradiente das imagens *Cube011* e *Flower02*. Pode-se observar que a maioria dos gradientes, em ambos os histogramas, concentra-se na faixa de valores menores, cor-

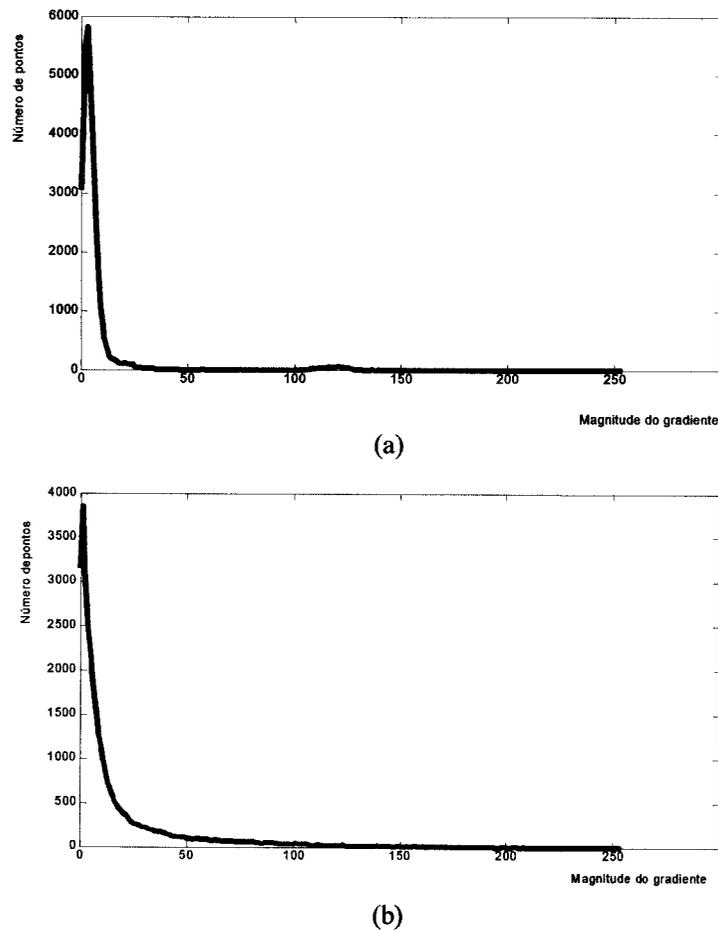
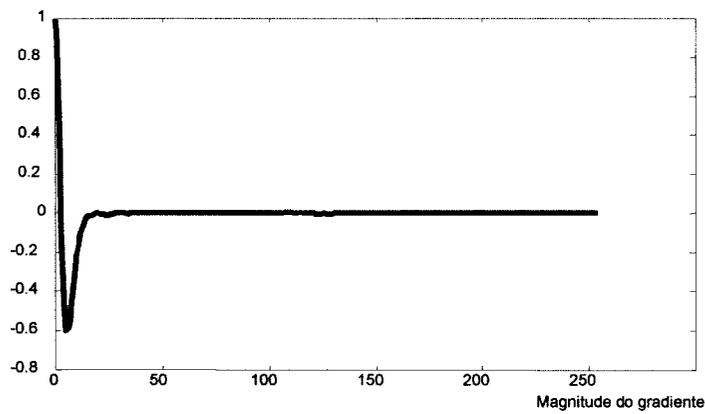


Figura 5.3. Histogramas dos valores da magnitude do gradiente calculados para as imagens (a) Cube011 e (b) Flower02. Os valores da magnitude foram escalados para o intervalo  $[0, 255]$ .

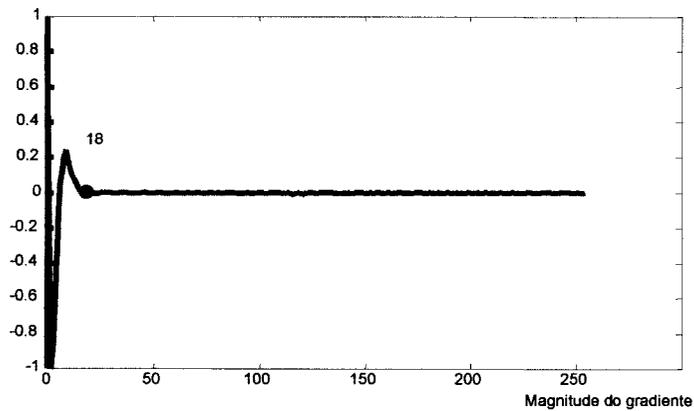
respondentes às áreas sem bordas da imagem. Gradientes mais altos são sensivelmente em menor número.

Propõe-se, assim, utilizar a informação de que a maioria dos valores de gradiente correspondem, normalmente, a pontos que não são bordas e determinar, no histograma, um valor adequado para servir como limiar na discriminação entre bordas e não-bordas. Desta forma, considera-se que o modo no início do histograma deva ser isolado e, para isso, localiza-se na curva descrita o valor do eixo no qual ela se “estabiliza”, aproximando-se de zero. São ignorados outros picos, sejam sutis ou mais significativos, após esta primeira tendência assintótica.

Para esta seleção do valor, a curva do histograma é suavizada com uma gaussiana discreta, sendo então calculada suas derivadas de primeira e segunda ordens. Escalar as curvas para o intervalo  $[-1, 1]$  simplifica o processo de escolher o ponto de inflexão da curva do histograma, que é feito estabelecendo-se limiares numéricos nas derivadas. Assim, o limiar é determinado pelo primeiro ponto que tem derivada “zero” (ou seja, com módulo abaixo do valor arbitrário  $l_1$ ) e segunda derivada também “zero” (com valor absoluto menor que  $l_2$ ). Permite-se, assim, verificar onde a curva do histograma se horizontaliza (com a primeira derivada tendendo a zero), mas evi-



(a)



(b)

Figura 5.4. (a) Primeira derivada para a curva do histograma da Figura 5.3(a). (b) Segunda derivada da mesma curva. O ponto indicado (18) representa o ponto de inflexão no histograma. Ambas as curvas foram escaladas para o intervalo  $[-1, 1]$ .

tando os pontos de máximo e mínimos locais do início do histograma, o que é determinado exigindo-se a proximidade dos pontos de máximo e mínimo com os pontos de inflexão. A Figura 5.4 apresenta, para imagem *Cube011*, as curvas da primeira e segunda derivadas para a curva do histograma. No gráfico para a segunda derivada é apresentado o primeiro ponto no qual a condição conjunta de proximidade a zero das derivadas de primeira e segunda ordem ocorre (valor 18).

Os valores para  $l_1$  e  $l_2$  são determinados experimentalmente e, aplicados às curvas restritas ao intervalo  $[-1, 1]$ , foram estabelecidos, respectivamente, em 0,008 e 0,02. Pequenas variações são toleradas, causando variação também pequena no valor do limiar. Uma vez escolhidos os valores, foram utilizados os mesmos parâmetros para todo o conjunto de imagens. Os resultados obtidos estão apresentados junto a outros no Apêndice D.

Uma avaliação mais extensiva desta abordagem ainda deve ser realizada, mas os resultados preliminares se mostram satisfatórios quando analisados de um ponto de vista subjetivo. Optou-se, desta forma, por incluir este tipo de mapa de bordas na comparação.

### 5.1.3.2 *Gradiente vetorial*

Apresentado na revisão, na seção 3.1.3, o gradiente colorido foi implementado utilizando-se a forma alternativa proposta por Cumani [Cumani (1991)], através da qual o cálculo da magnitude do gradiente pode ser obtido pelos autovalores da matriz  $[g_{ij}(\mathbf{w})]$ , definida na Equação 3.25. Assim, utilizando-se a função de contraste definida pela Equação 3.30, obtém-se as magnitudes dos gradientes através dos autovalores  $\lambda_+(\mathbf{w})$ , mais especificamente utilizando a Equação 5.11.  $\sqrt{\lambda_+(\mathbf{w})}$  é o valor da magnitude do gradiente na direção da maior variação [Cumani (1991); Lim & Lee (1990)].

$$\sqrt{\lambda_+(\mathbf{w})} = \sqrt{\frac{g_{11}(\mathbf{w}) + g_{22}(\mathbf{w}) + \sqrt{[g_{11}(\mathbf{w}) - g_{22}(\mathbf{w})]^2 + 4g_{21}^2(\mathbf{w})}}{2}} \quad (5.11)$$

Uma vez calculada a magnitude do gradiente, a operação de encontrar um limiar é análoga ao caso monocromático. Um limiar arbitrário, escolhido em função da análise visual do mapa de borda produzido foi utilizado para gerar os mapas de bordas a partir dos gradientes vetoriais, apresentados no Apêndice D. O emprego das abordagens para se encontrar automaticamente um limiar não puderam, entretanto, ser utilizadas. Pelo fato dos limiares, para o caso colorido, apresentarem usualmente valores muito baixos (valores entre 4 e 10, para a maioria das imagens), a determinação automática deixava de produzir resultados aceitáveis, já que valores maiores eram determinados e o número de pontos de borda acabavam por ser drasticamente reduzidos. Deste modo, somente a determinação manual do limiar na segmentação foi utilizada para o gradiente usando cores.

### 5.1.3.3 *Segmentação por mapas auto-organizáveis: Alternativas*

Como o interesse deste estudo é relativo a bordas, além da segmentação pelo mapa auto-organizável utilizando a cromaticidade também foram incluídos outros dois mapas de bordas. Proporcionados pela aplicação da metodologia proposta no Capítulo 4 e utilizando como dados a intensidade das três componentes RGB diretamente e também a combinação da iluminação com a cromaticidade (descritas na seção 4.2.4) são definidos os mapas I (intensidade das componentes) e IC (iluminação em cinza e cromaticidade), respectivamente. Considerando-se que seus desempenhos, em termos de borda, apresentam uma precisão satisfatória quando analisadas visualmente, também estes mapas de bordas foram incluídos na comparação aqui proposta.

### 5.1.4 *Descrição do modelo de comparação*

O conjunto de imagens da base de dados apresenta uma diversidade de imagens com características relativamente variadas. Origens distintas para grupos de imagens, bem como o não estabelecimento de condições prévias para a captura, fazem

deste conjunto uma coleção interessante para ser avaliada segundo as segmentações produzidas.

A comparação é feita sobre o desempenho dos métodos em relação à posição “real” das bordas, determinada através da segmentação manual da imagem de regiões de referência. São consideradas duas medidas básicas, ambas determinadas em termos das razões entre números de pontos entre as imagem de bordas de referência e a obtida por uma das técnicas: a taxa de coincidência e a taxa de excesso. A taxa de coincidência corresponde ao número de pontos de borda definidos manualmente que foram efetivamente rotulados como bordas por uma técnica de segmentação (Equação 5.12). Como a análise é feita considerando-se os atributos locais,  $v$  indica o centro de um intervalo no eixo das variâncias para o qual a razão é aplicada. Para este intervalo,  $T_a(v)$  é a taxa de coincidência,  $S_a(v)$  é o número de pontos de borda detectados e  $N(v)$  é o número de pontos de borda de referência\*.

$$T_a(v) = \frac{S_a(v)}{N(v)} \quad (5.12)$$

De forma similar, a taxa de excesso  $T_e(v)$  é dada pela Equação 5.13, na qual  $S_e(v)$  representa o número de pontos de borda detectados pelo método que não encontram contrapartida na imagem de bordas de referência, também considerando um intervalo nas variâncias centrado em  $v$ .

$$T_e(v) = \frac{S_e(v)}{N(v)} \quad (5.13)$$

Como  $T_a(v)$  representa a taxa de acerto, seus valores estão restritos ao intervalo  $[0, 1]$ , com acerto máximo de 100%. Por outro lado, a relação estabelecida para  $T_e(v)$  pode proporcionar taxas superiores a 100%.

Para o tratamento conjunto das imagens foram adotadas duas estratégias, ambas procurando avaliar o melhor ou pior desempenho de cada método dentro do conjunto de imagens. A primeira verifica o número de vezes em que cada método teve desempenho superior ao dos outros, segundo os critérios das taxas estabelecidas. A outra elabora estima as “taxas médias”, as quais são usadas para caracterizar o desempenho médio dos métodos.

A contagem do número de vezes em que um método tem desempenho superior aos outros é realizada, também, para faixas de variância de um dos atributos selecionados. Para cada intervalo é estipulado um contador para o número de vezes que o método apresenta a maior taxa de coincidência e outro para o número de vezes que apresenta menor taxa de excesso. Calculados os totais, a estimativa de um índice de qualidade leva em conta não somente o quanto o método maximiza a taxa de coinci-

\* Para simplificar a notação,  $v$ , ao invés de  $\sigma^2$ , simboliza o centro de um dos intervalos para as variâncias  $g$ ,  $c$ ,  $rg$  e  $y_b$ .

dência, mas também o quanto ele minimiza a taxa de excesso. Estabelece-se, assim, um coeficiente de desempenho geral considerando uma proporção entre estes extremos. Este coeficiente, denominado *coeficiente acumulativo* e designado por  $H_a$ , é definido em conformidade com a Equação 5.14, na qual  $p$  é a proporção da combinação, que assume valores no intervalo  $[0, 1]$ , e  $V_a(v)$  e  $V_e(v)$  são os resultados relativos da contagem para as taxas de coincidência e de excesso, respectivamente. O valor  $v$  indica o centro de um intervalo em uma das variâncias. Permite-se, assim, privilegiar as coincidências ou privilegiar a ausência de excessos.

$$H_a(v) = pV_a(v) + (1 - p)V_e(v) \quad (5.14)$$

A estimativa das taxas médias é realizada, por sua vez, utilizando as médias das taxas de coincidência e de excesso, considerando-se um elenco de imagens. Da mesma forma que para o coeficiente acumulativo, o *coeficiente proporcional*  $H_p(v)$  é definido pela média da ponderação entre as taxas de coincidência e excesso, como introduzido na Equação 5.15.  $N_I$  é o número de imagens utilizadas na comparação,  $p$  é a proporção da combinação de ambas as taxas e  $T_{a_i}(v)$  e  $T_{e_i}(v)$  são as taxas de coincidência e excesso, respectivamente, obtidas para a  $i$ -ésima imagem. Manipulada adequadamente, a Equação 5.15 pode ser rescrita no formato da Equação 5.16, mais adequada para utilização, com as taxas médias descritas na Equação 5.17.

$$H_p(v) = \frac{1}{N_I} \sum_{i=1}^{N_I} [pT_{a_i}(v) - (1 - p)T_{e_i}(v)] \quad (5.15)$$

$$H_p(v) = p\bar{T}_a(v) - (1 - p)\bar{T}_e(v) \quad (5.16)$$

$$\bar{T}_a(v) = \frac{1}{N_I} \sum_{i=1}^{N_I} T_{a_i}(v), \quad \bar{T}_e(v) = \frac{1}{N_I} \sum_{i=1}^{N_I} T_{e_i}(v) \quad (5.17)$$

A combinação utiliza, apesar da taxa  $p$ , o sinal negativo ao incorporar a contribuição da taxa de excesso ao coeficiente proporcional, pois quanto maior esta taxa, menor deve ser o coeficiente calculado. Nota-se, ainda, que a combinação entre as duas taxas não é balanceada, visto que  $T_e(v)$  pode assumir valores maiores que 100%. A escolha da proporção de combinação  $p$  deve compensar esta diferença de magnitude. Para manter os valores dos coeficientes no intervalo  $(0, 1)$ , as taxas de excesso são normalizadas antes do cálculo de  $H_p(v)$ . Esta normalização leva em consideração todos os valores  $\bar{T}_e(v)$ , de forma que o maior assumia valor 1 e os demais sejam es-

---

\* O valor é, na realidade, a proporção das vezes que o método obteve o melhor desempenho para um conjunto de imagens.

Tabela 5-I. Síntese das medidas de desempenho propostas.

Medida de desempenho	Símbolo	Descrição
Taxa de coincidência	$T_a(v)$	Medida da proporção do número de pontos de borda que uma técnica de segmentação detectou e que concordam com as bordas da segmentação manual de referência. É condicionada a um intervalo nas variâncias centrado em $v$ e definida pela Equação 5.12.
Taxa de excesso	$T_e(v)$	Medida da proporção do número de pontos de borda que uma técnica de segmentação detectou e que não coincidem com as bordas da segmentação manual de referência. É condicionada a um intervalo nas variâncias centrado em $v$ e definida pela Equação 5.13.
Coefficiente acumulativo	$H_a(v)$	Coefficiente que considera o número de vezes, para um conjunto de imagens, em que uma técnica de segmentação obteve maiores taxas de coincidência combinada com menores taxas de excesso. O valor é condicionado a um intervalo das variâncias centrado em $v$ e a um fator $p$ , que determina os pesos a serem dados a ambas as taxas. Definido pela Equação 5.14. O valor resultante é proporcional ao número de imagens.
Coefficiente proporcional	$H_p(v)$	Coefficiente que considera, para um conjunto de imagens, a taxa de coincidência média combinada com a taxa de excesso média. O valor é condicionado a um intervalo das variâncias centrado em $v$ e a um fator $p$ , que determina os pesos a serem dados a ambas as taxas. Definido pela Equação 5.16.

calados proporcionalmente. Mantém-se, assim, os desempenhos relativos entre as técnicas, permitindo a comparação.

A Tabela 5-I apresenta uma síntese das medidas de desempenho apresentadas. Procura-se, desta forma, verificar o comportamento das diferentes técnicas em função dos dados locais, estabelecendo parâmetros quantitativos para a comparação das técnicas.

### 5.1.5 Classes de imagens

As imagens que compõem a base de imagens variam em termos de composição espectral (considerada através das componentes RGB) e condições de iluminação (tipo de iluminação, sombras, reflexos etc.), além do próprio conteúdo. Sob este

Tabela 5-II. Classes de imagens.

Classe	Imagens incluídas	Número de imagens
Geral	Todas as imagens	38
Casas	House01, House02, House03, House04, House05, House06, House07	7
Cubos	Cube009, Cube010, Cube011, Cube014, Cube015	5
Externas	House01, House02, House03, House04, House05, House06, House07, Sign01, Sign02, Sign03, Sign04, Wall01, Wall02	13
Laboratório	Cube009, Cube010, Cube011, Cube014, Cube015, Cube101, Cube102, Cube104, Cube105, Cube202, Cube203, Cube204, Pyram01, Pyram03, Pyram04, Pyram06	16

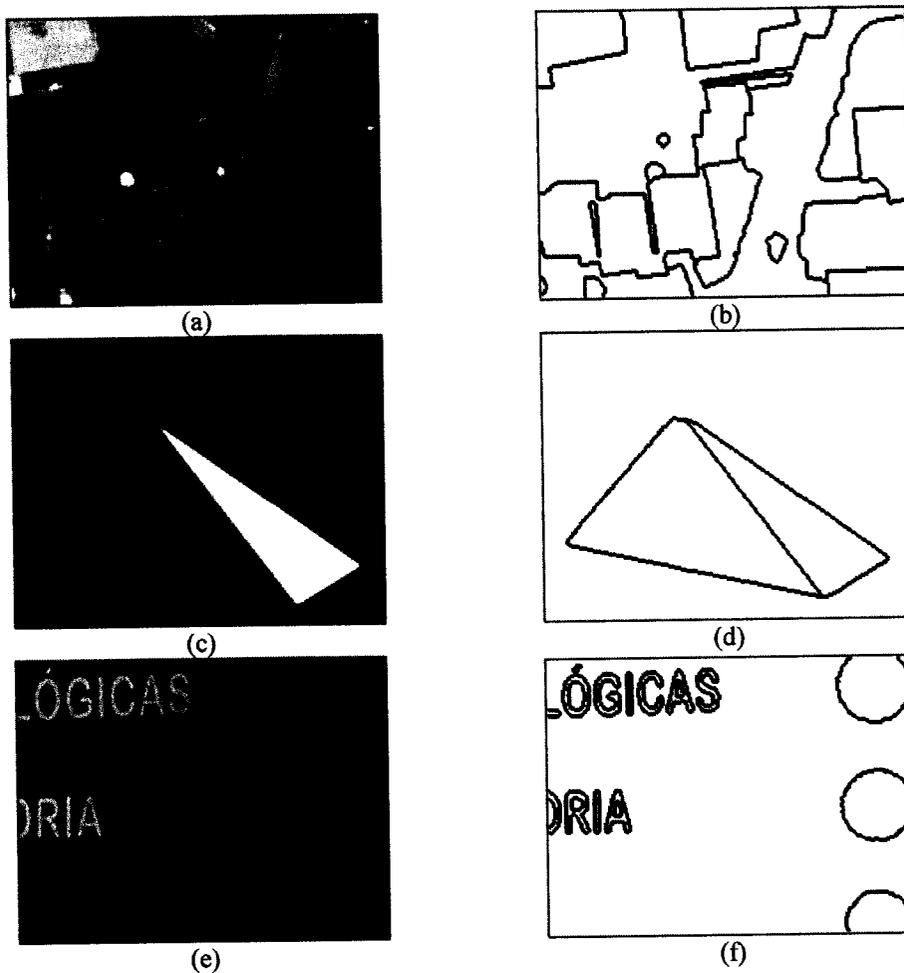


Figura 5.5. Imagens utilizadas nos exemplos das Figuras 5.6 a 5.8. Imagens originais: (a) Aerial01; (c) Pyram04 e (e) Sign03. (b), (d) e (f): Respetivos mapas de bordas de referência.

ponto de vista, a diversificação das imagens é relativamente alta, de forma que, além da avaliação comparativa geral do desempenho, classes de imagens podem ser consideradas. Assim, subconjuntos de imagens são estipulados e as comparações entre as técnicas podem ser realizadas para classes que compartilhem algumas características gerais.

As características comuns consideradas são bastante gerais e subdividem a base de imagens em classes considerando um assunto específico ou uma condição de aquisição em especial. A Tabela 5-II introduz as classes utilizadas, listando as imagens incluídas em cada uma. Incluindo todas as imagens, a classe *geral* considera o comportamento genérico dos métodos; a classe *casas* abrange cenas externas de casas e prédios; a classe *externas* estende a classe *casas* para outras cenas externas; cubos adquiridos em ambiente de laboratório estão representados na classe *cubos*; e *laboratório* é a classe que amplia a classe *cubos* para todas as imagens adquiridas em laboratório. Procurou-se, no estabelecimento das classes, utilizar um mínimo de cinco imagens para que se tivesse uma representatividade mínima.

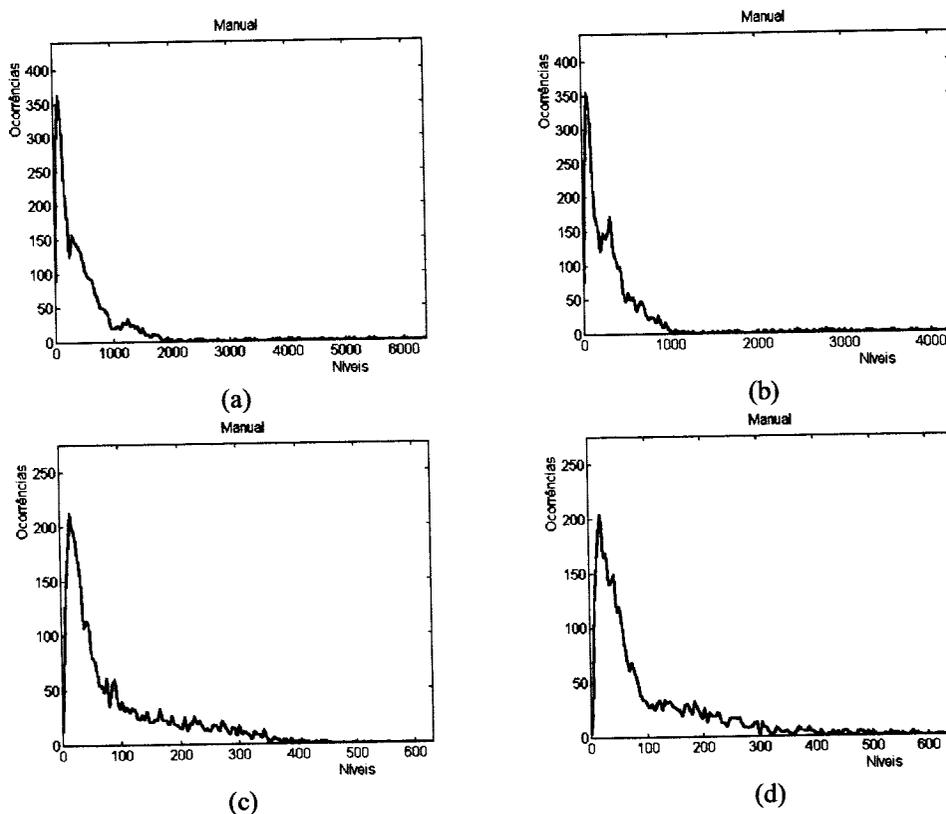


Figura 5.6. Histogramas das variâncias nos pontos de borda definidos manualmente para a imagem Aerial01: (a)  $g$ ; (b)  $c$ ; (c)  $rg$ ; e (d)  $y_b$ .

## 5.2 Resultados da comparação e discussão

As imagens da Figura 5.5 são as utilizadas nos exemplos desta seção. A variância, independentemente do atributo, é considerada nas posições em que foram definidos pontos de borda, em conformidade com a segmentação manual.

A Figura 5.6 ilustra, para os quatro atributos selecionados, os histogramas que representam o número de pontos de borda para faixas da variância divididas em 150 intervalos. Nas Figuras 5.6(a) a (d) estão representadas as distribuições dos pontos de borda para as variâncias  $g$ ,  $c$ ,  $rg$  e  $y_b$ , respectivamente, calculadas para a imagem Aerial01. Pode-se observar que há uma concentração das bordas em regiões com variância menor, principalmente para as variâncias dos níveis de cinza e das cores. Para as variâncias  $rg$  e  $y_b$ , apesar de também haver esta concentração, os pontos têm maior representação até os valores intermediários das variâncias.

Deve-se observar que a presença de pontos de bordas em regiões de variância pequena não é uma contradição por, principalmente, duas razões. Primeiramente a segmentação manual pode seguir, como apresentado, outros critérios além da descontinuidade, como o conhecimento de nível alto para determinar o limite correto e fechado de um objeto da imagem. A outra razão é que, embora haja uma relação entre as variâncias e as bordas, esta não é, necessariamente, a mesma estabelecida na segmentação manual ou a produzida pelas técnicas de segmentação. Para a estimativa

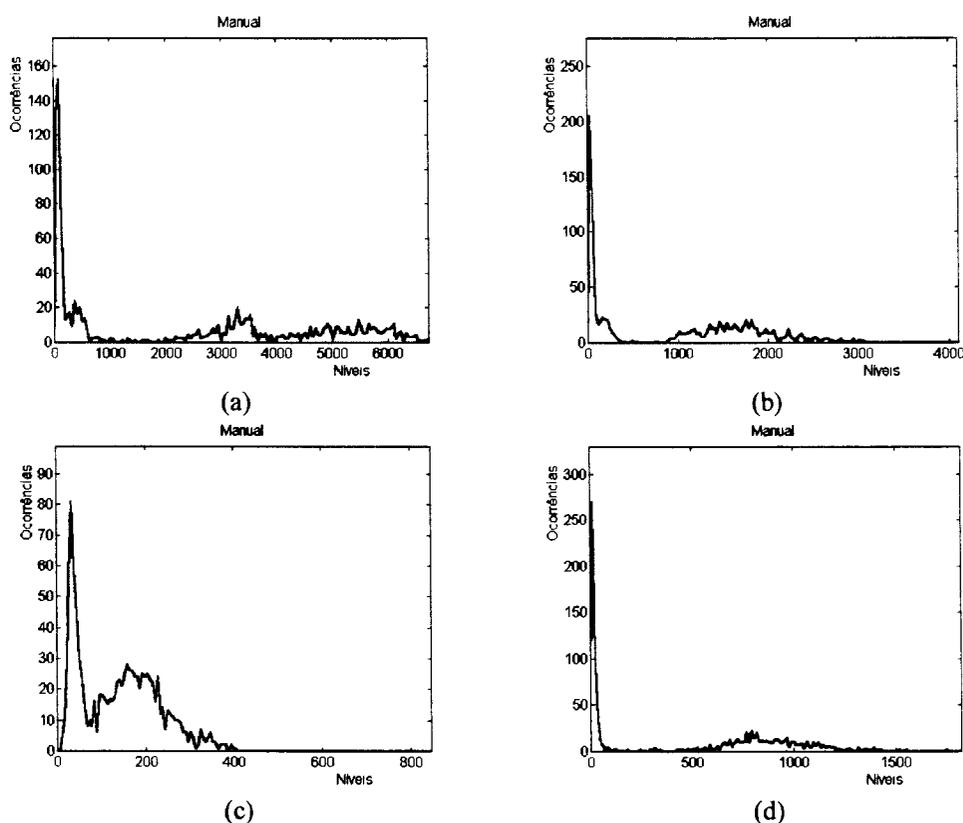


Figura 5.7 Histogramas das variâncias nos pontos de borda definidos manualmente para a imagem *Pyram04*: (a) *g*; (b) *c*; (c) *rg*; e (d) *yb*.

das variâncias locais foi usado um raio de quatro pontos para o círculo discreto. Os atributos foram, também, considerados para raios de três e cinco pontos, mas uma comparação cuidadosa\* dos resultados permitiu considerar que os resultados não se alteravam significativamente. Desta forma, somente as avaliações que utilizam raio quatro são apresentadas neste capítulo, o que não compromete a discussão dos resultados.

As Figuras 5.7 e 5.8 mostram os mesmos dados para as imagens *Pyram04* e *Sign03*, respectivamente. Notam-se distribuições com características diferentes para cada uma das três imagens, especialmente a concentração das bordas em intervalos bastante distintos para as quatro variâncias. Destaca-se, por exemplo, a quase ausência de bordas para as variâncias em níveis de cinza entre 1000 e 2000 da imagem *Pyram04*, como se pode observar na Figura 5.7(a) e a mesma situação em torno da variância 500 para o atributo *c*, mostrado na Figura 5.7(b). Comentários similares podem ser feitos para a imagem *Sign03*, também para as Figuras 5.7(a) e (b). É possível, assim, observar uma grande variação nas distribuições das bordas manuais, que se caracterizam de forma diversa entre as imagens da base.

\* A comparação envolveu analisar, para os raios citados, as tendências dos resultados e não houve alteração que pudesse ser considerada significativa. Entretanto raios maiores não foram considerados, principalmente por se considerarem ultrapassados os limites de localidade dos atributos.

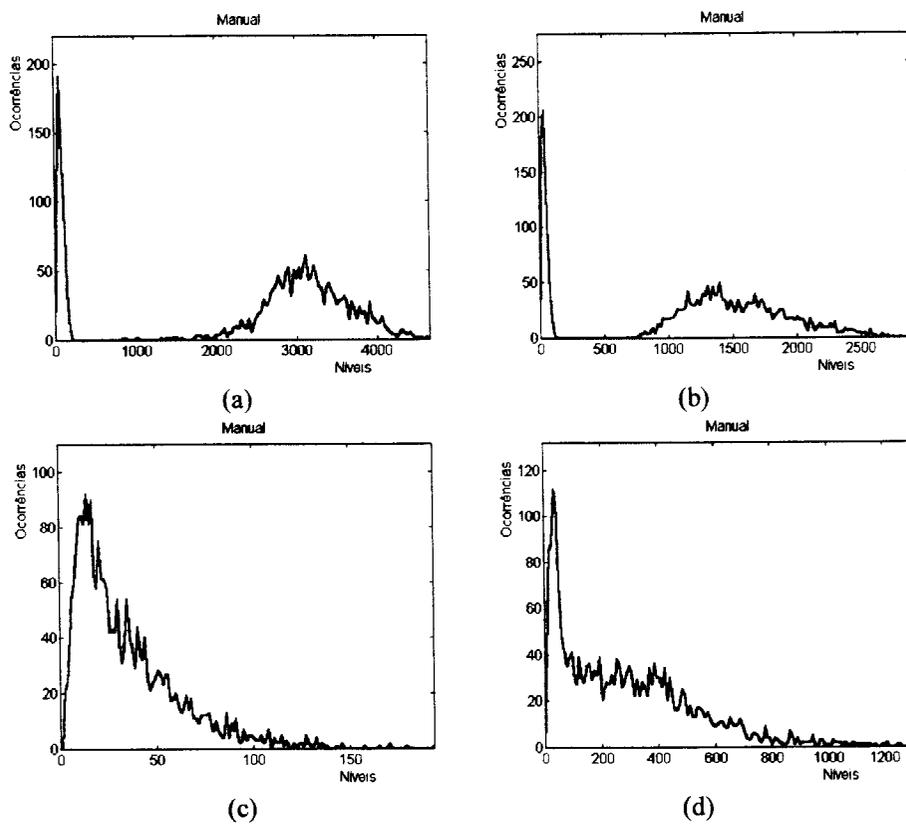


Figura 5.8 Histogramas das variâncias nos pontos de borda definidos manualmente para a imagem *sign03*: (a) *g*; (b) *c*; (c) *rg*; e (d) *yb*.

### 5.2.1 Taxas de coincidência e excesso

Levando em conta o critério de comparação dos os mapas de bordas, a Figura 5.9 mostra os histogramas das taxas de acerto para a imagem exemplo *Aerial01*, segundo as especificações das Equações 5.12 e 5.13. Foram utilizadas 150 divisões entre zero e a variância máxima para cada atributo. O número de intervalos, além de considerado satisfatório para representar a grande diversidade de variâncias, permite uma boa visualização dos histogramas. A apresentação na forma de linhas, ao invés das colunas tradicionais, foi escolhida para evitar o excesso de retas nas figuras, permitindo uma melhor observação dos dados.

Os resultados da comparação das bordas manuais com as bordas produzidas pela aplicação do gradiente de Sobel (limiar escolhido manualmente) estão apresentados na Figura 5.9(a). A taxa de acerto é praticamente de 100% para as variâncias maiores, reduzindo-se para as variâncias menores, com taxas de acerto caindo para próximo de zero. Nas regiões de maior variância, os intervalos para os quais a taxa cai subitamente para zero são relativos a intervalos para os quais a segmentação manual não produziu nenhuma borda; são portanto, desconsiderados nas comparações. Para algumas faixas, entretanto, a taxa de excesso é relativamente grande, chegando a valores em torno de 400% e picos acima de 600%, ou seja, detecção de mais de cinco vezes o número esperado de pontos para um dado intervalo da variância.

Comportamentos similares podem ser observados para as segmentações pelo gradiente vetorial, na Figura 5.9(b), e Sobel com limiares adaptáveis local e global, nas Figuras 5.9(h) e (i), com diferenças mais significativas em uma ou outra faixa de variação. Estas técnicas, em especial, são as voltadas diretamente para detecção de bordas. As demais técnicas, para esta imagem, mostram-se com taxa de acerto sensivelmente menor, oscilando em torno de 40% para o mapa auto-organizável usando a cromaticidade, apresentado na Figura 5.9(c), 50% para a classificação bayesiana, Figura 5.9(d), e abaixo de 40% para  $k$  vizinhos mais próximos, apresentado na Figura 5.9(e). A utilização do mapa auto-organizável utilizando atributos de intensidade (RGB) e de iluminação aliada à cromaticidade, porém, responderam com taxas ligeiramente superiores para as coincidências que as outras técnicas de segmentação por regiões, como pode ser observado nas Figuras 5.9(f) e (g). Notam-se, em vários histogramas, taxas de excesso muito altas em intervalos de variâncias específicos, dados que a detecção proporcionou uma quantidade muito superior à expectativa.

As Figuras 5.10 a 5.12 ilustram, respectivamente, as taxas obtidas para as variâncias  $c$ ,  $rg$  e  $yb$ . Para estes histogramas não há uma tendência significativamente diferente dos resultados apresentados para a variância dos níveis de cinza.

Notam-se taxas de coincidência altas, quase todas 100%, para as variâncias mais altas, relativamente à imagem *Aerial101*. Para estas variâncias, o número de pontos de borda é bastante pequeno, resultando mais facilmente nos extremos coincidência total ou nenhuma coincidência, em um dado intervalo.

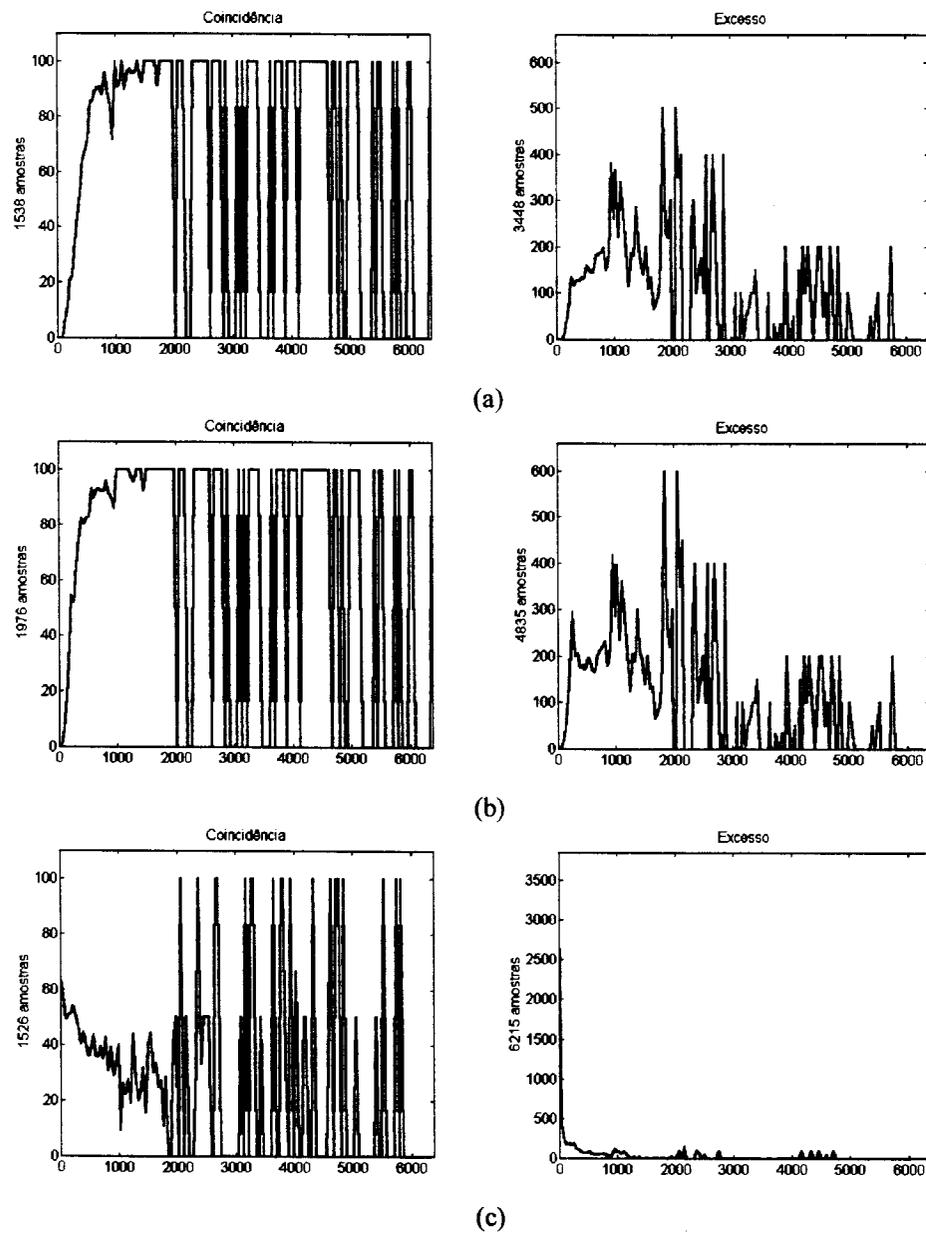


Figura 5.9. Histogramas das taxas de coincidência (à esquerda) e taxas de excesso (à direita) para os mapas de bordas gerados por várias técnicas. A imagem é Aerial01 e o atributo considerado é a variância dos níveis de cinza. (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

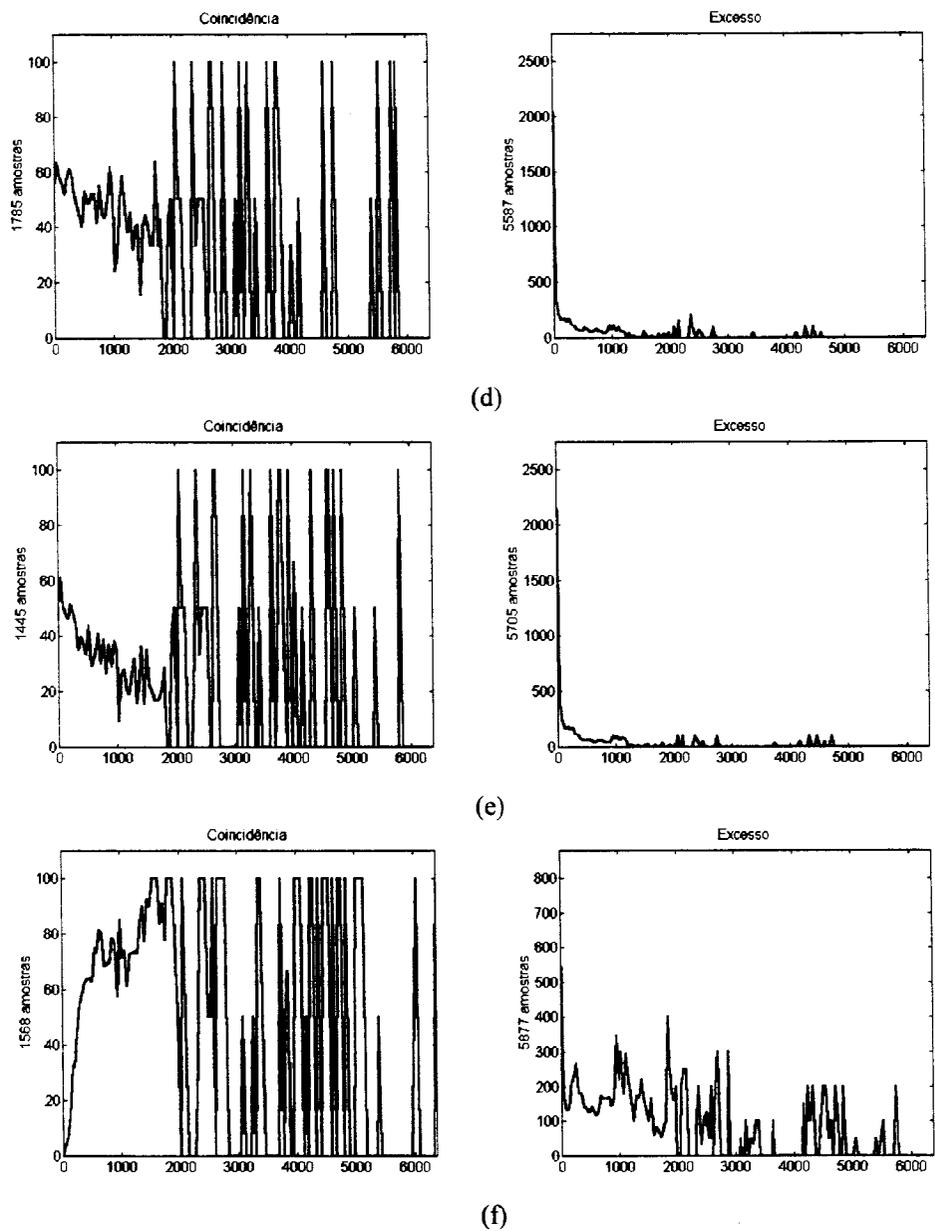


Figura 5.9. Continuação. (d) Classificação bayesiana; (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação.

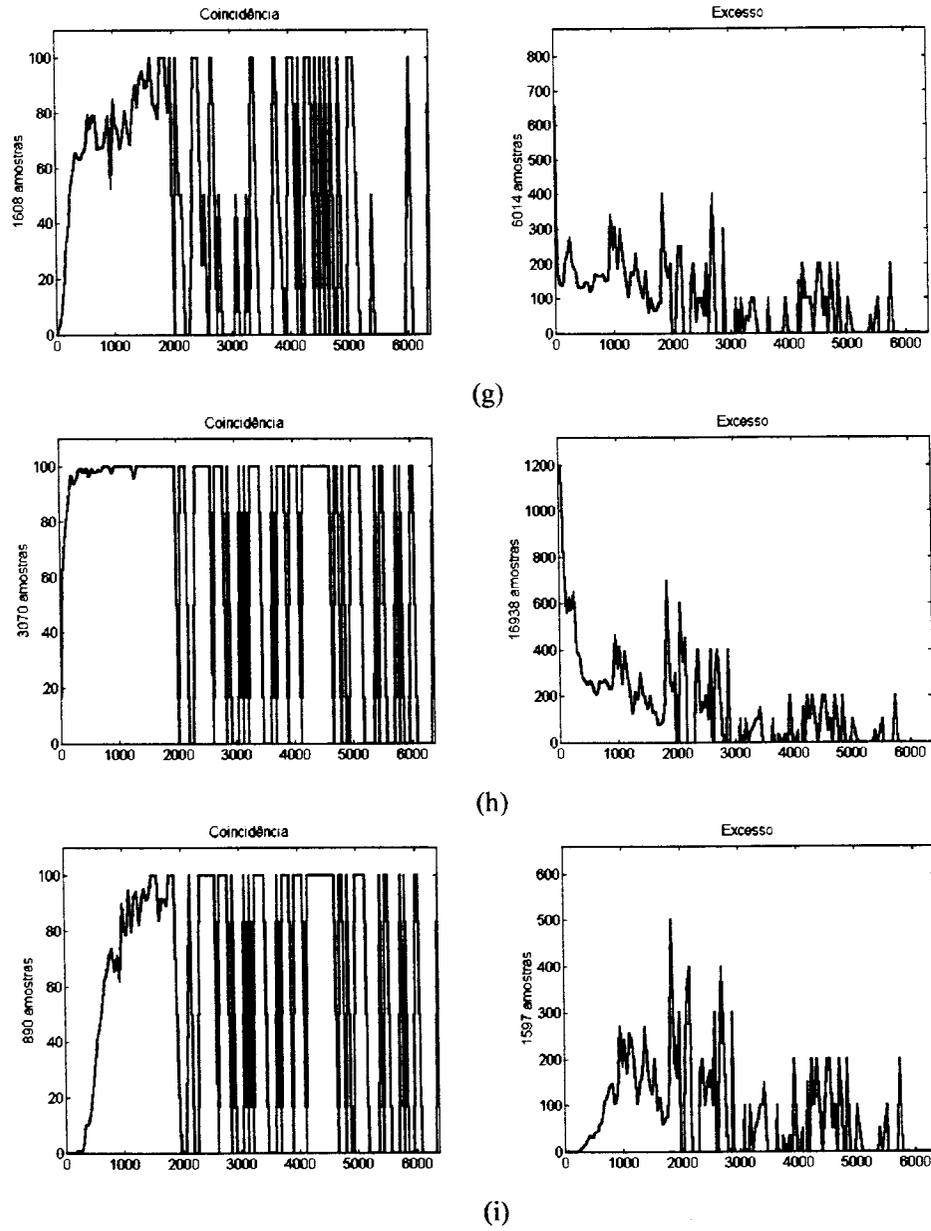


Figura 5.9. *Continuação.* (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

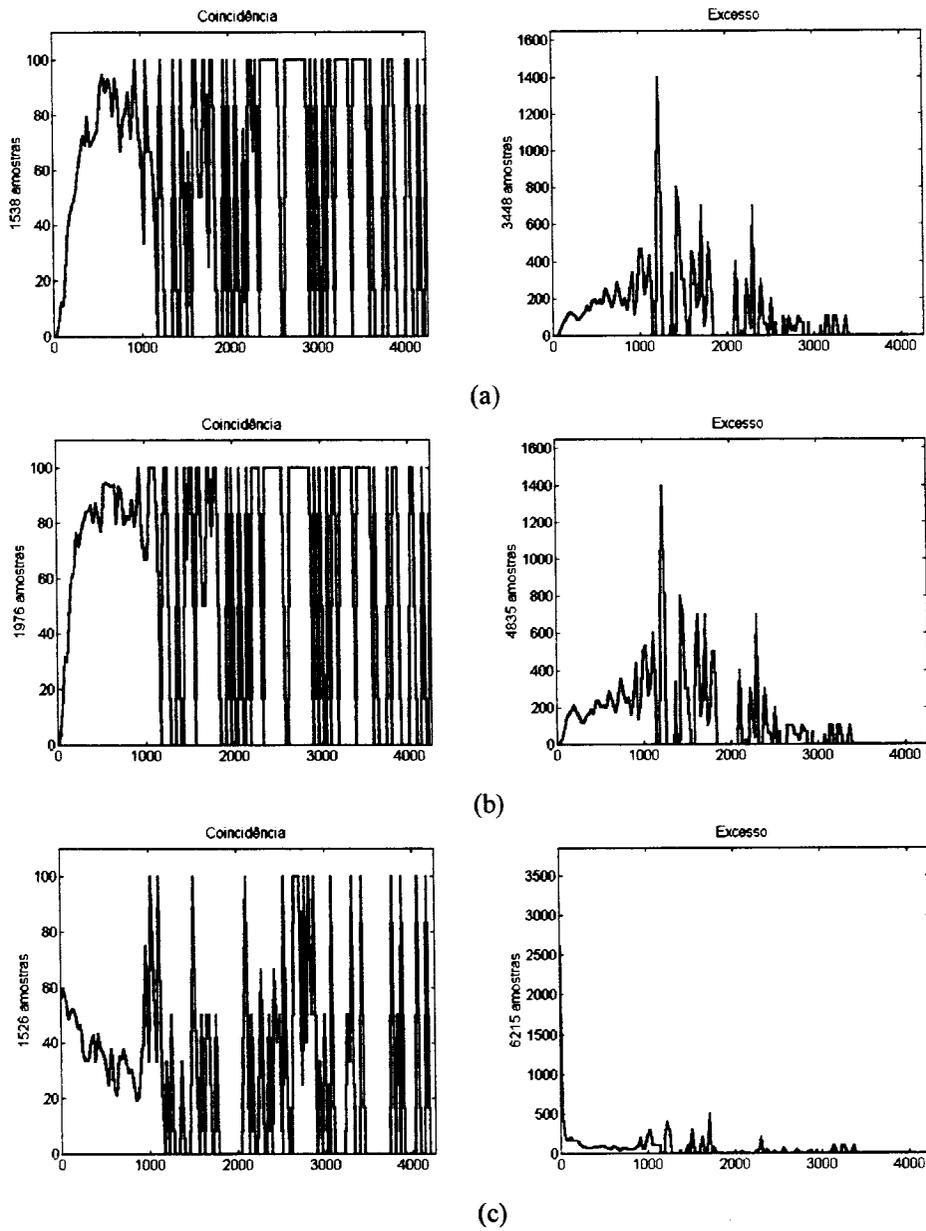


Figura 5.10. Histogramas das taxas de coincidência (à esquerda) e taxas de excesso (à direita) para os mapas de bordas gerados por várias técnicas. A imagem é *Aerial01* e o atributo considerado é a variância dos vetores de cor (atributo *c*). (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por *k* vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

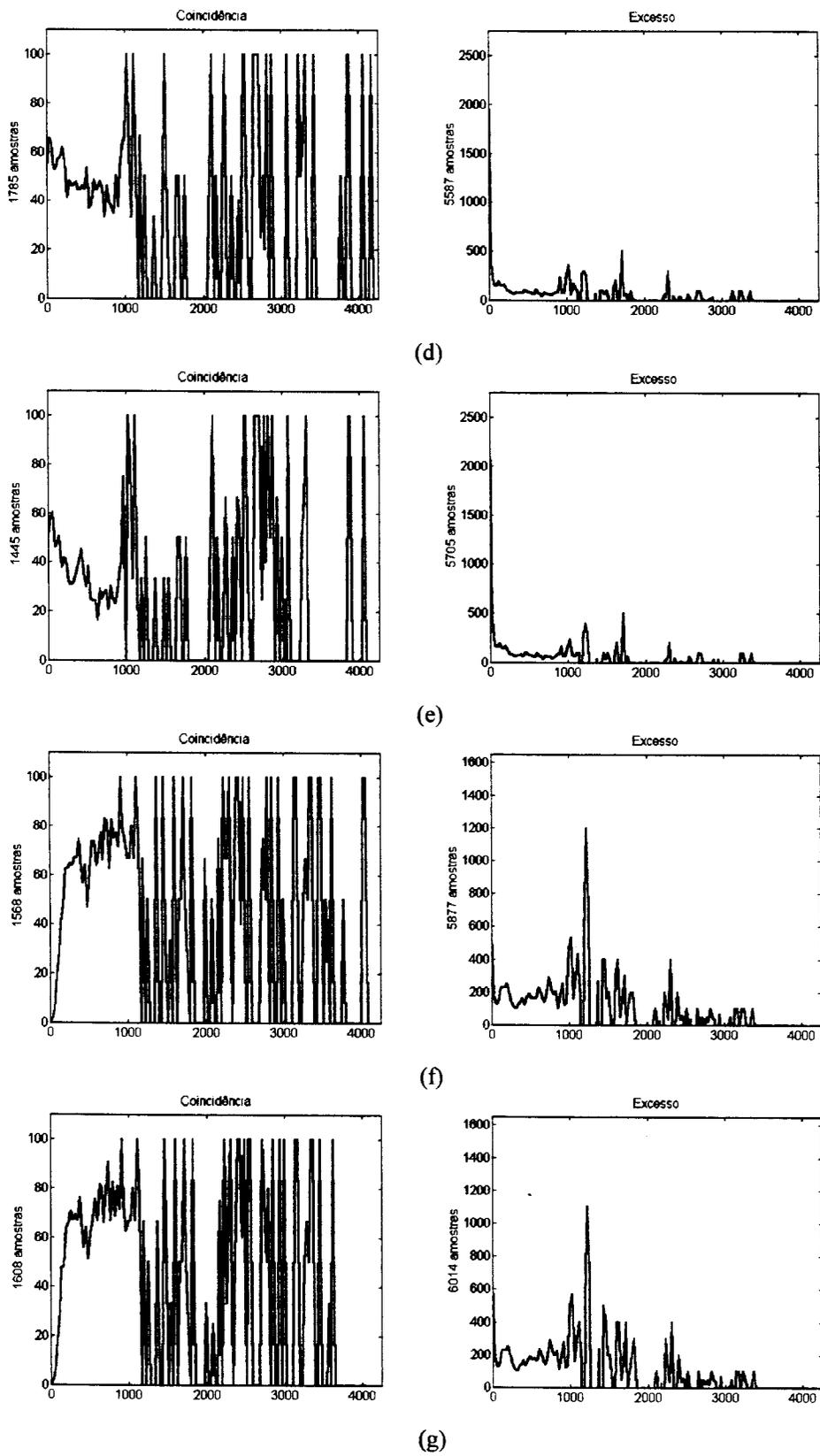


Figura 5.10. *Continuação.* (d) Classificação bayesiana; (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade.

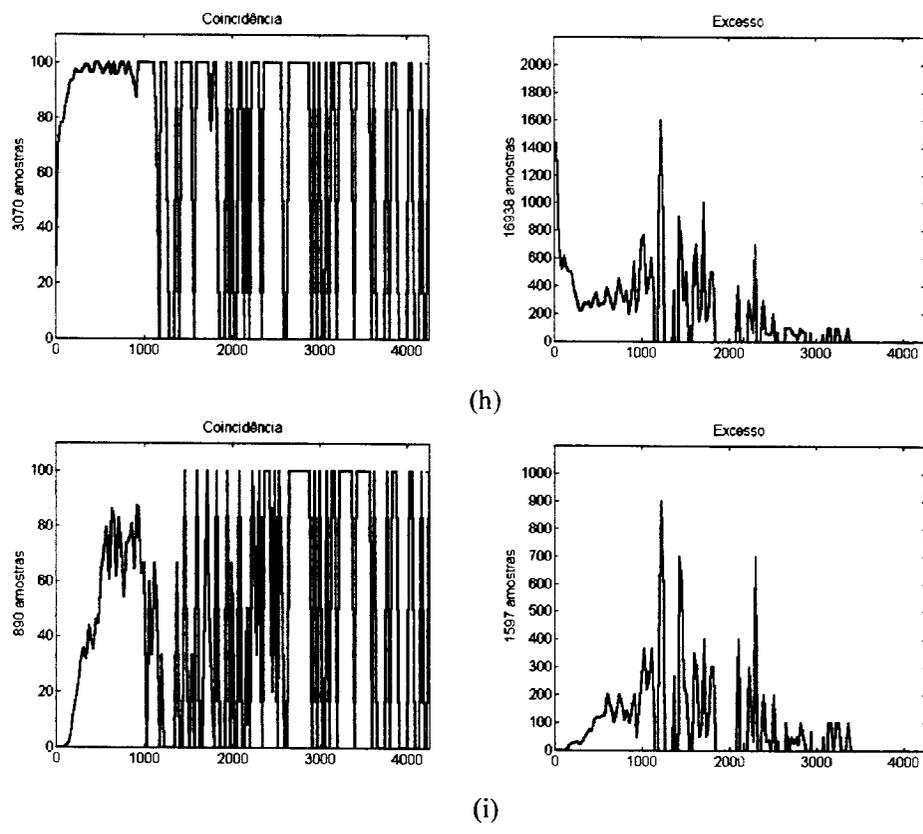


Figura 5.10. *Continuação.* (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global

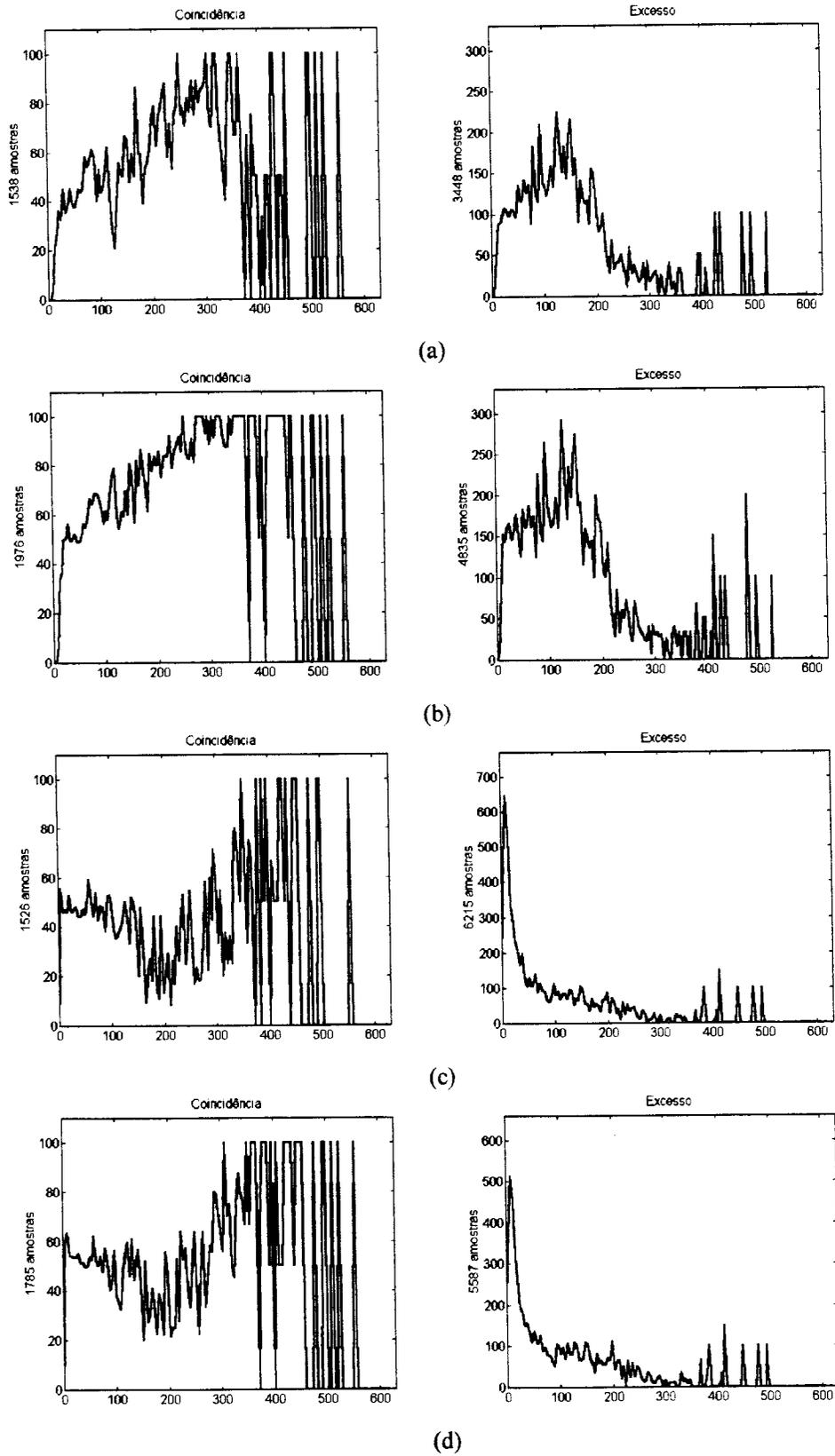


Figura 5.11. Histogramas das taxas de coincidência (à esquerda) e taxas de excesso (à direita) para os mapas de bordas gerados por várias técnicas. A imagem é *Aerial01* e o atributo considerado é a variância da oposição cromática RG. (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

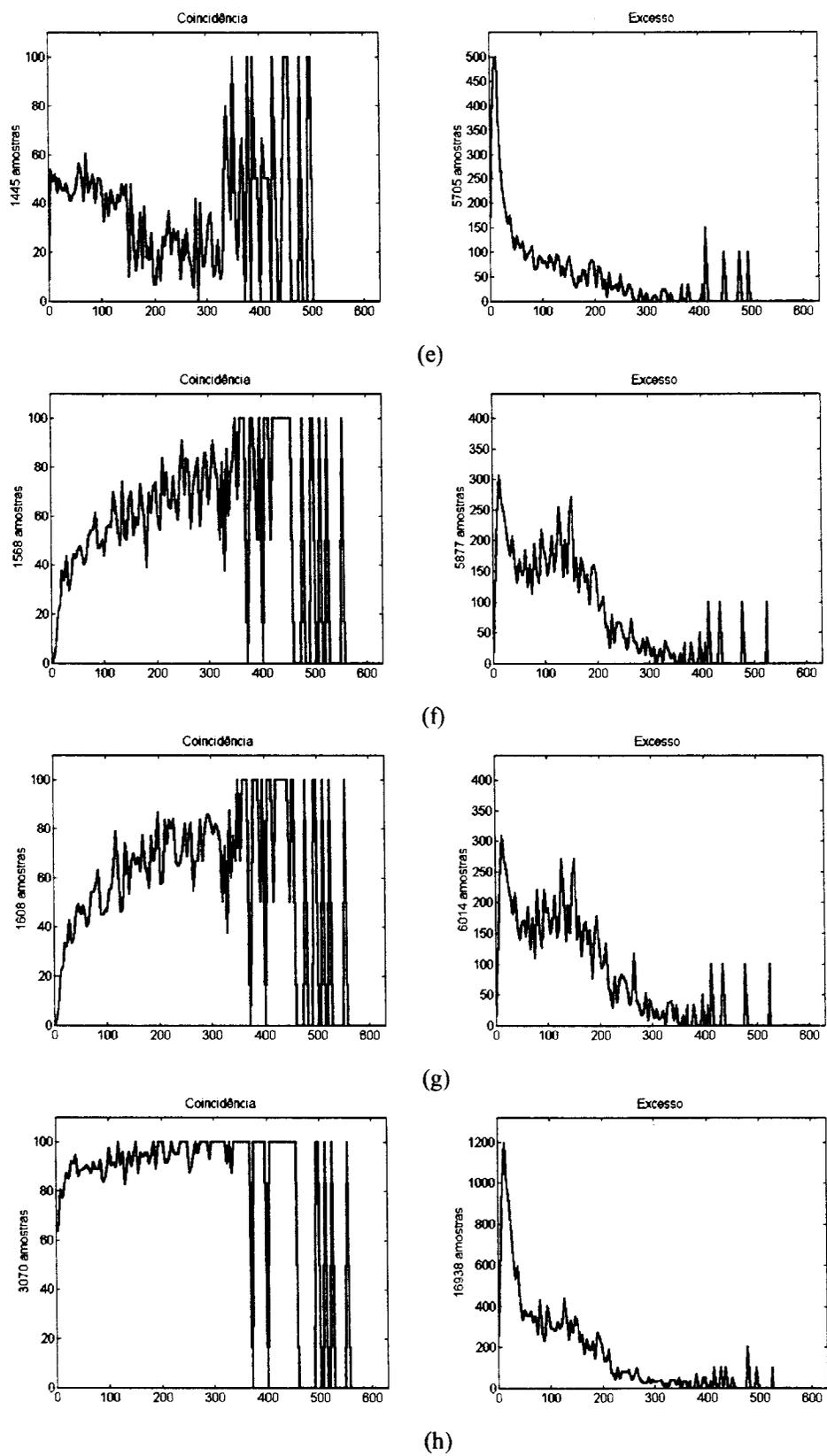


Figura 5.11. *Continuação.* (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local.

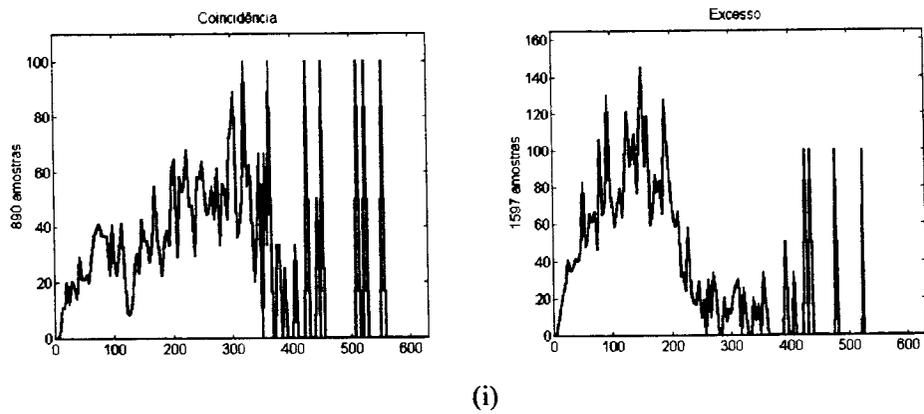


Figura 5.11. *Continuação.* (i) Sobel usando limiar adaptável global

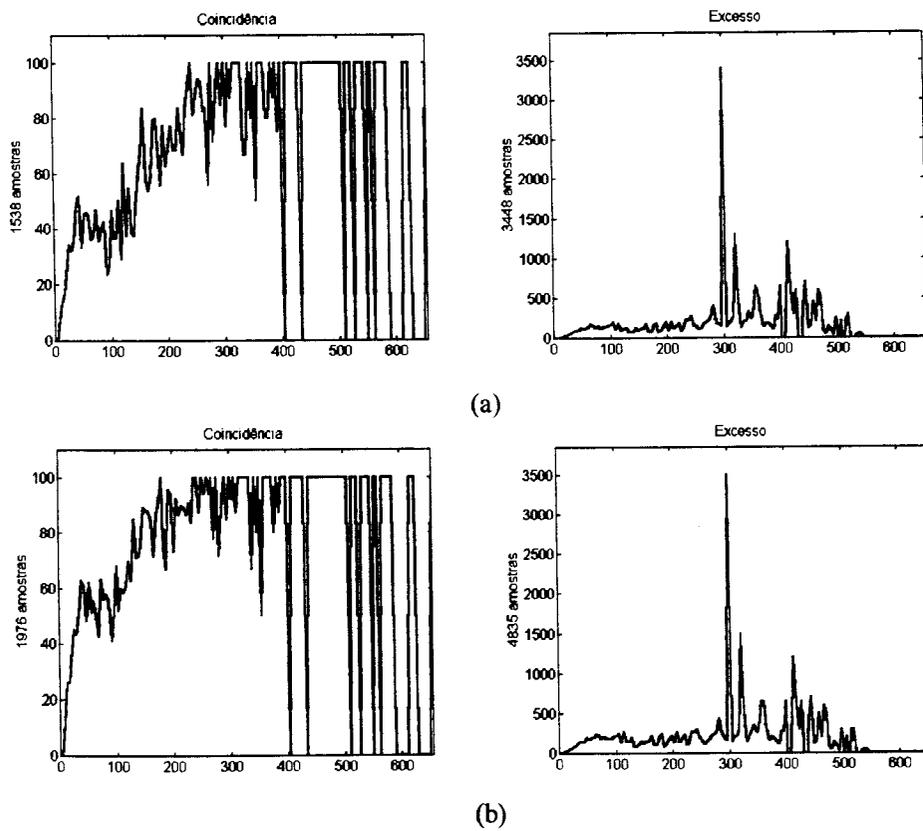


Figura 5.12. Histogramas das taxas de coincidência (à esquerda) e taxas de excesso (à direita) para os mapas de bordas gerados por várias técnicas. A imagem é *Aerial01* e o atributo considerado é a variância da oposição cromática YB. (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

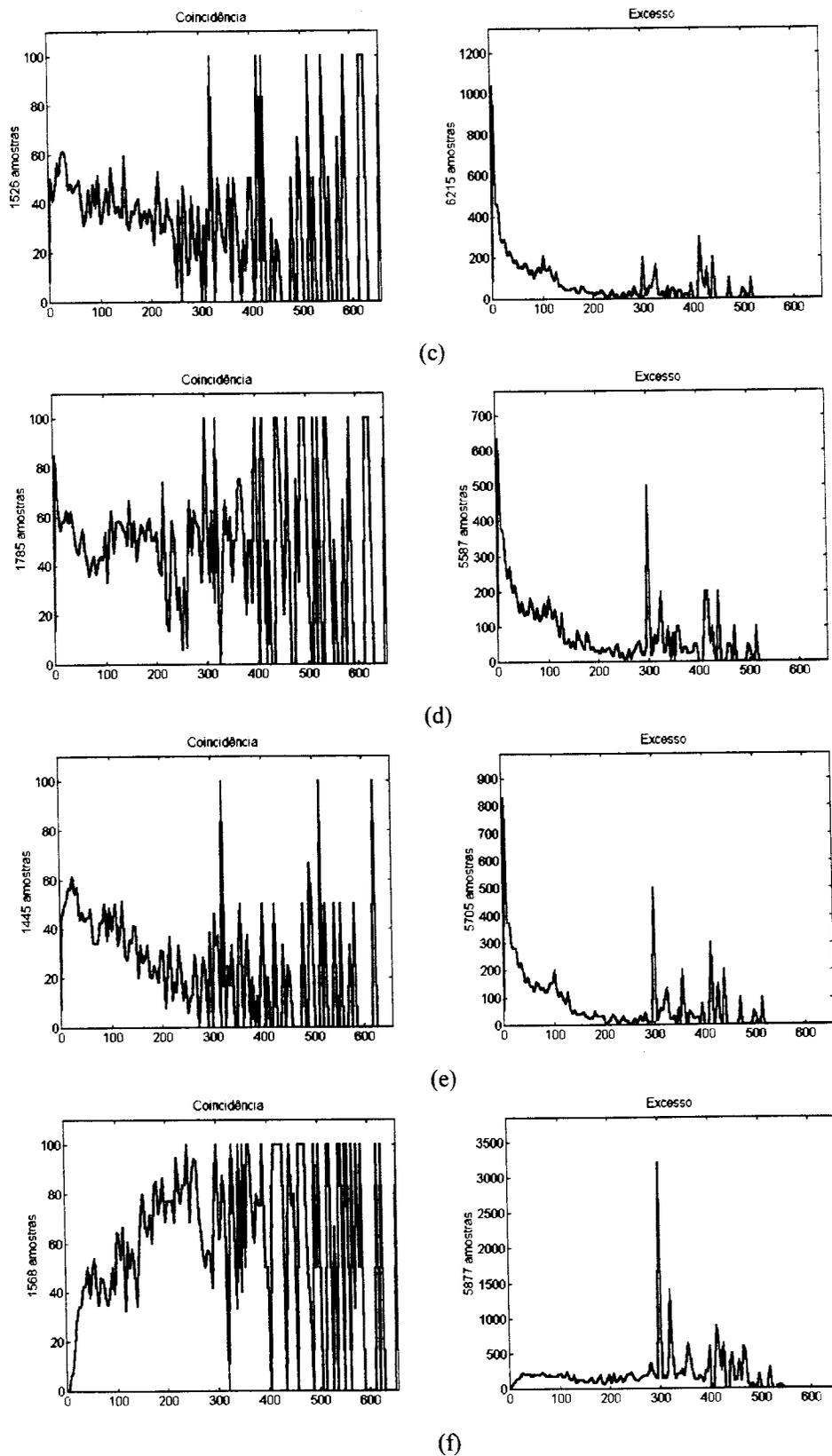


Figura 5.12. *Continuação.* (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação.

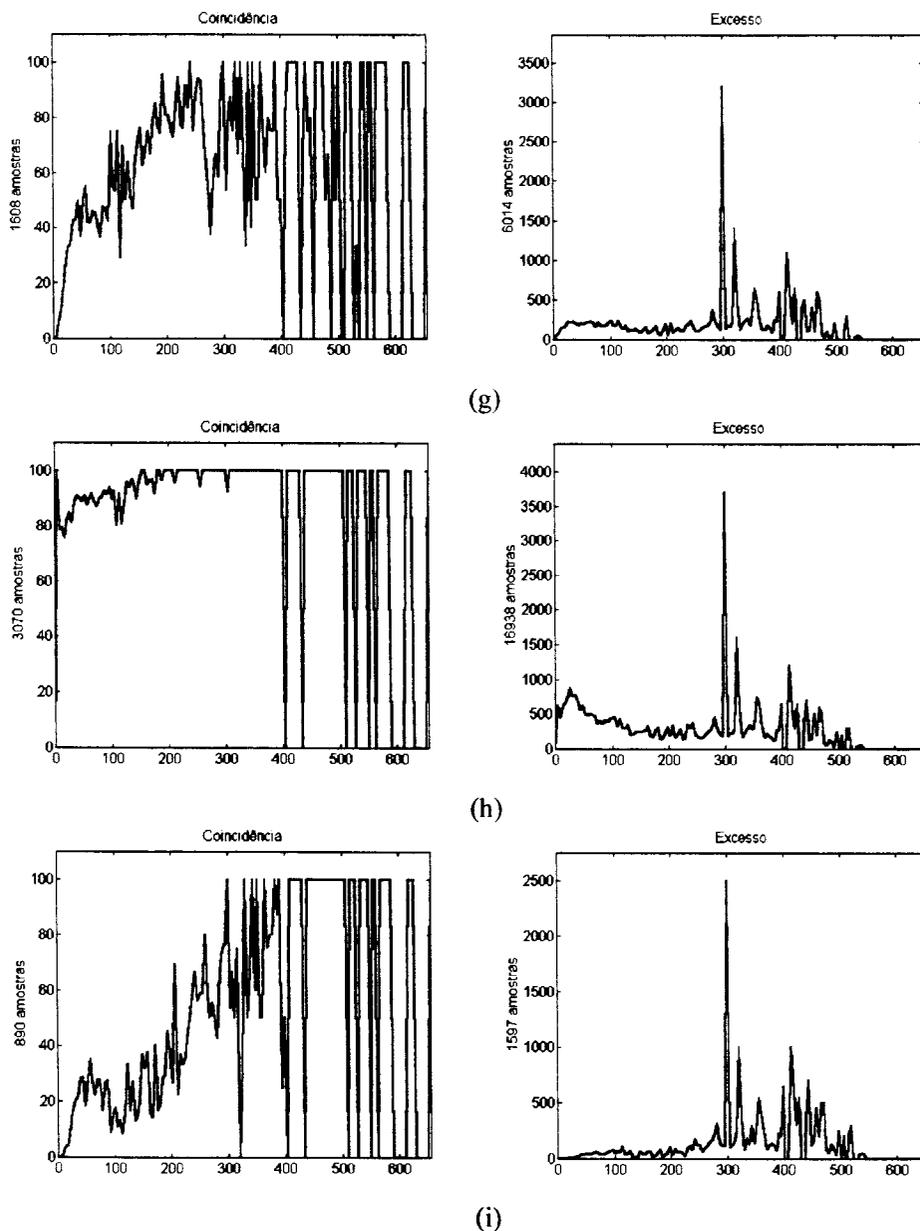


Figura 5.12. *Continuação.* (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

Os desempenhos dos métodos estão apresentados nas Figuras 5.13 a 5.16, respectivamente para os atributos  $g$ ,  $c$ ,  $rg$  e  $yb$ . Cada gráfico representa a uma “medida de qualidade”, na qual 1 indica que, para uma determinada faixa de variação do atributo, a técnica de segmentação apresentou tanto o melhor resultado quanto o que menos detectou bordas erroneamente. O melhor resultado corresponde, neste caso, à maior taxa de coincidência entre todos os métodos e a menor detecção de bordas erradas corresponde à menor taxa de excesso. Valores 0,5, no gráfico, indicam a maior taxa de coincidência ou a menor taxa de excesso, exclusivamente. Valores nulos, por

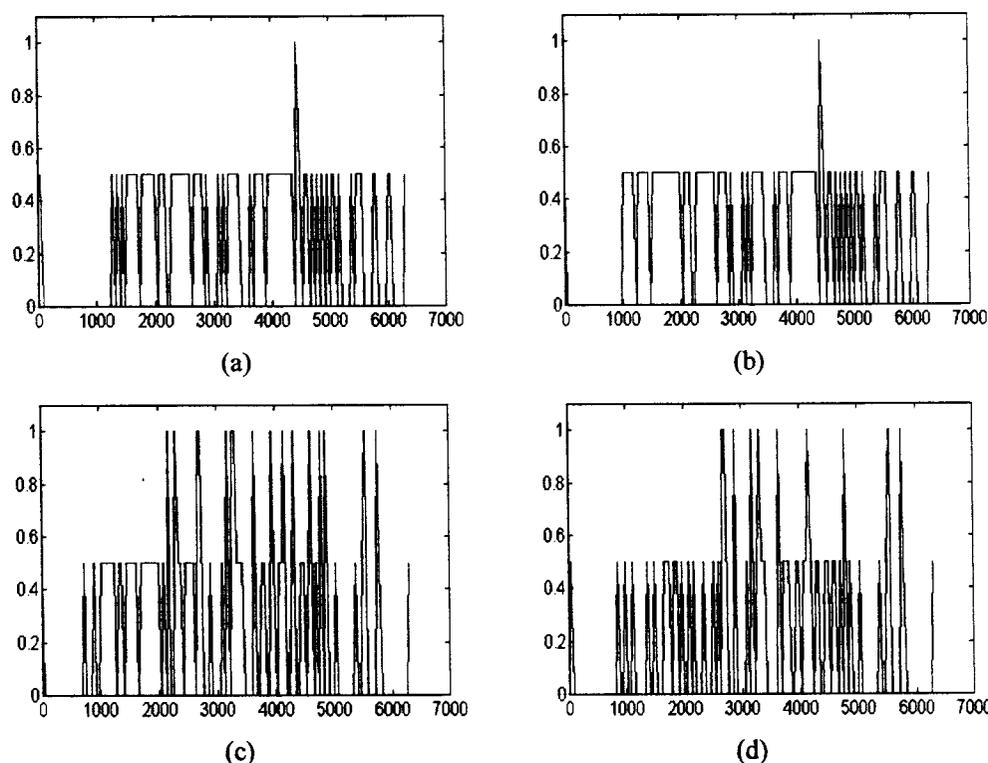
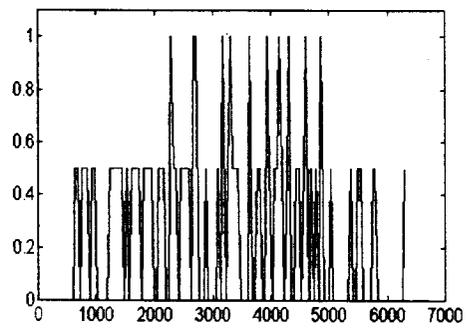
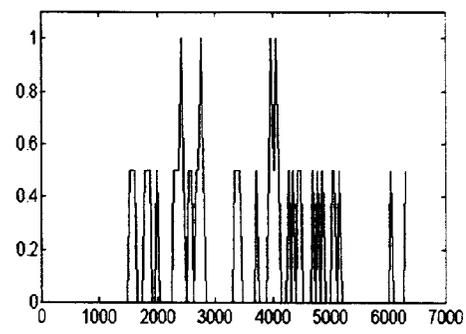


Figura 5.13. Desempenho dos métodos de segmentação para a imagem *Aerial01*, considerando o atributo  $g$ . (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

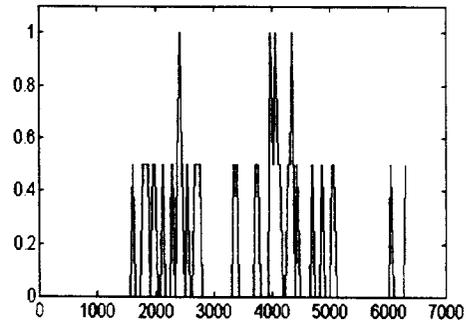
sua vez, indicam intervalos nos quais o método não foi o que apresentou maior taxa de coincidência nem o que apresentou menor taxa de excesso. Esta última condição, entretanto, não significa que o comportamento da técnica foi ruim, apenas que não foi a melhor. Pequenas diferenças nas taxas obtidas por métodos diferentes proporciona apenas um método com melhor desempenho dentro dos parâmetros estabelecidos. Nota-se, em vários casos, que mais de uma técnica é considerada a de melhor desempenho. Estas situações ocorrem, na maior parte dos casos, nos intervalos em que as taxas de coincidência ou taxas de excesso atingem 100% para mais de uma técnica; ou seja, caso haja “empate”, todas as que tiverem maior taxa são consideradas como tendo melhor desempenho.



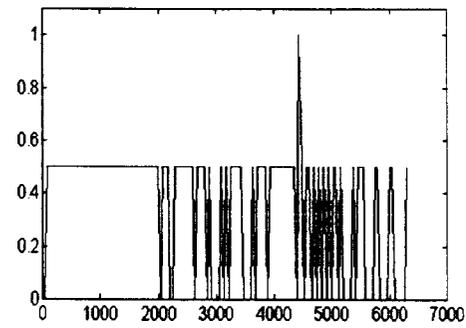
(e)



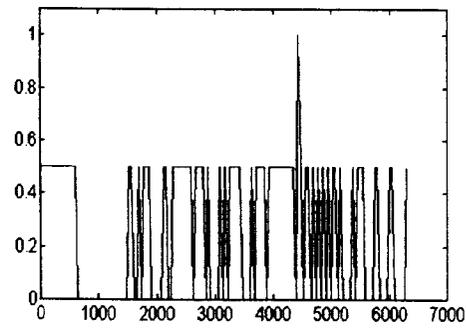
(f)



(g)



(h)



(i)

Figura 5.13. *Continuação.* (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

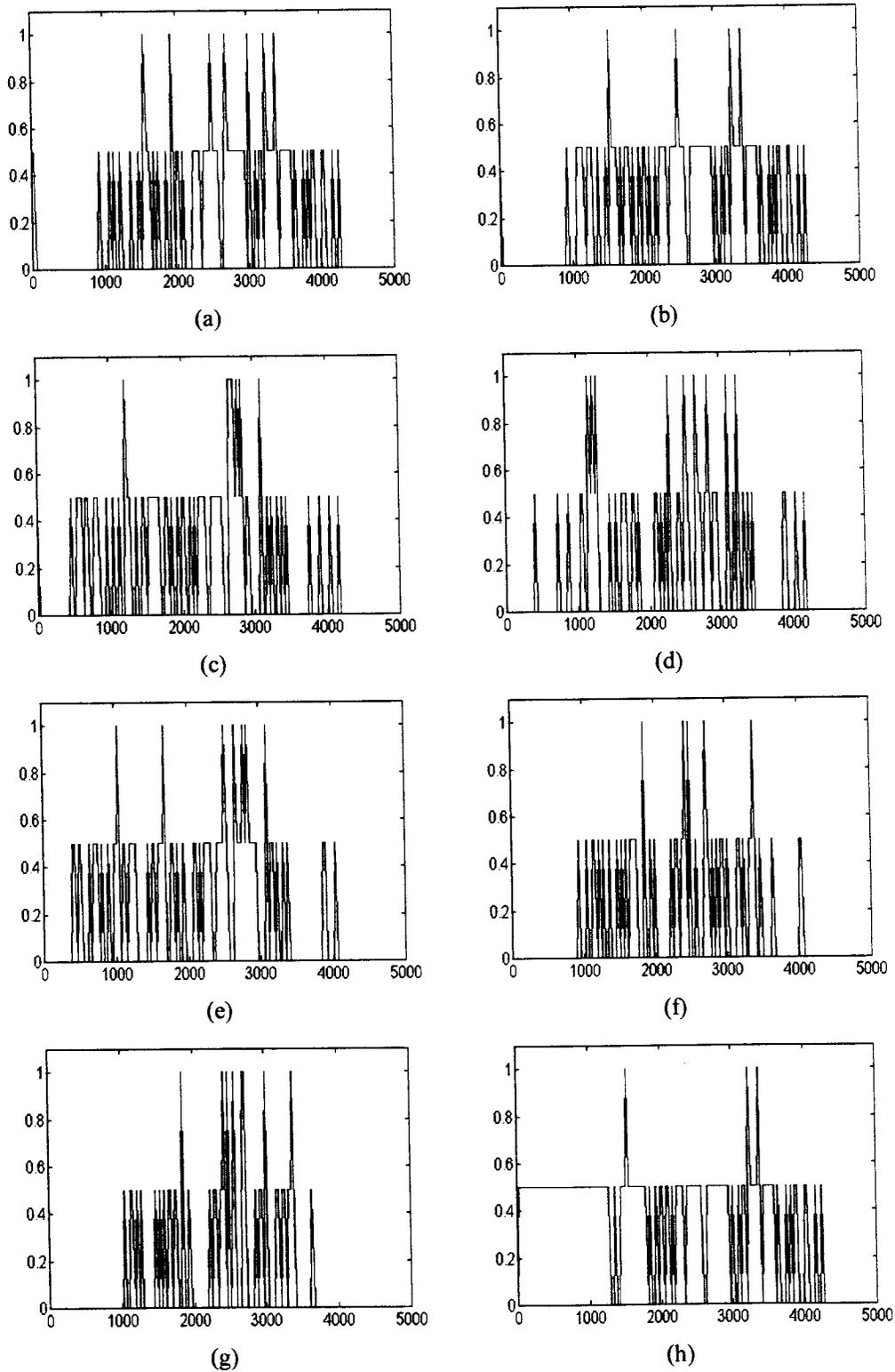


Figura 5.14. Desempenho dos métodos de segmentação para a imagem Aerial01, considerando o atributo  $c$ . (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

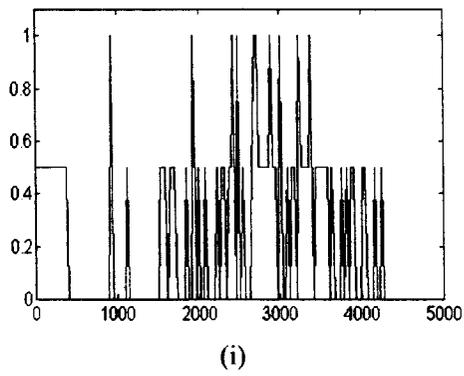


Figura 5.14. *Continuação.* (i) Sobel usando limiar adaptável global.

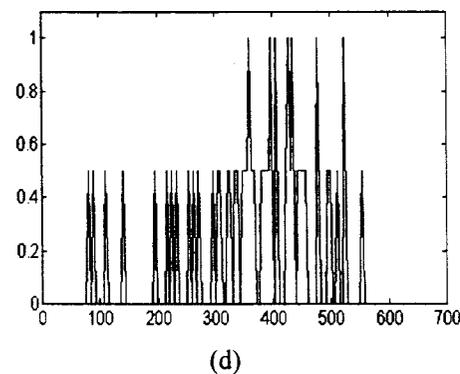
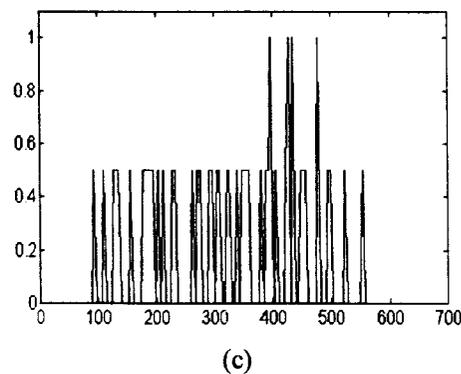
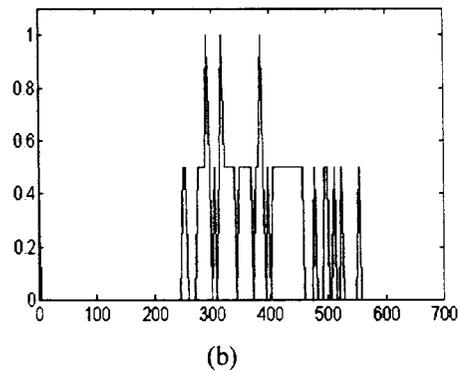
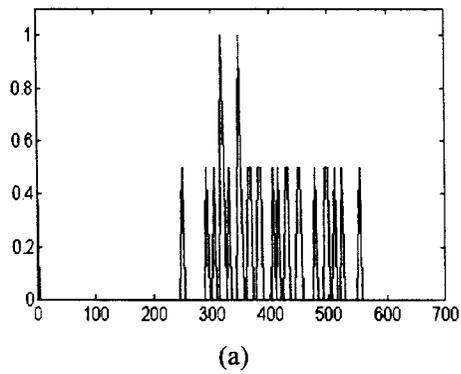


Figura 5.15. Desempenho dos métodos de segmentação para a imagem *Aerial01*, considerando o atributo *rg*. (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

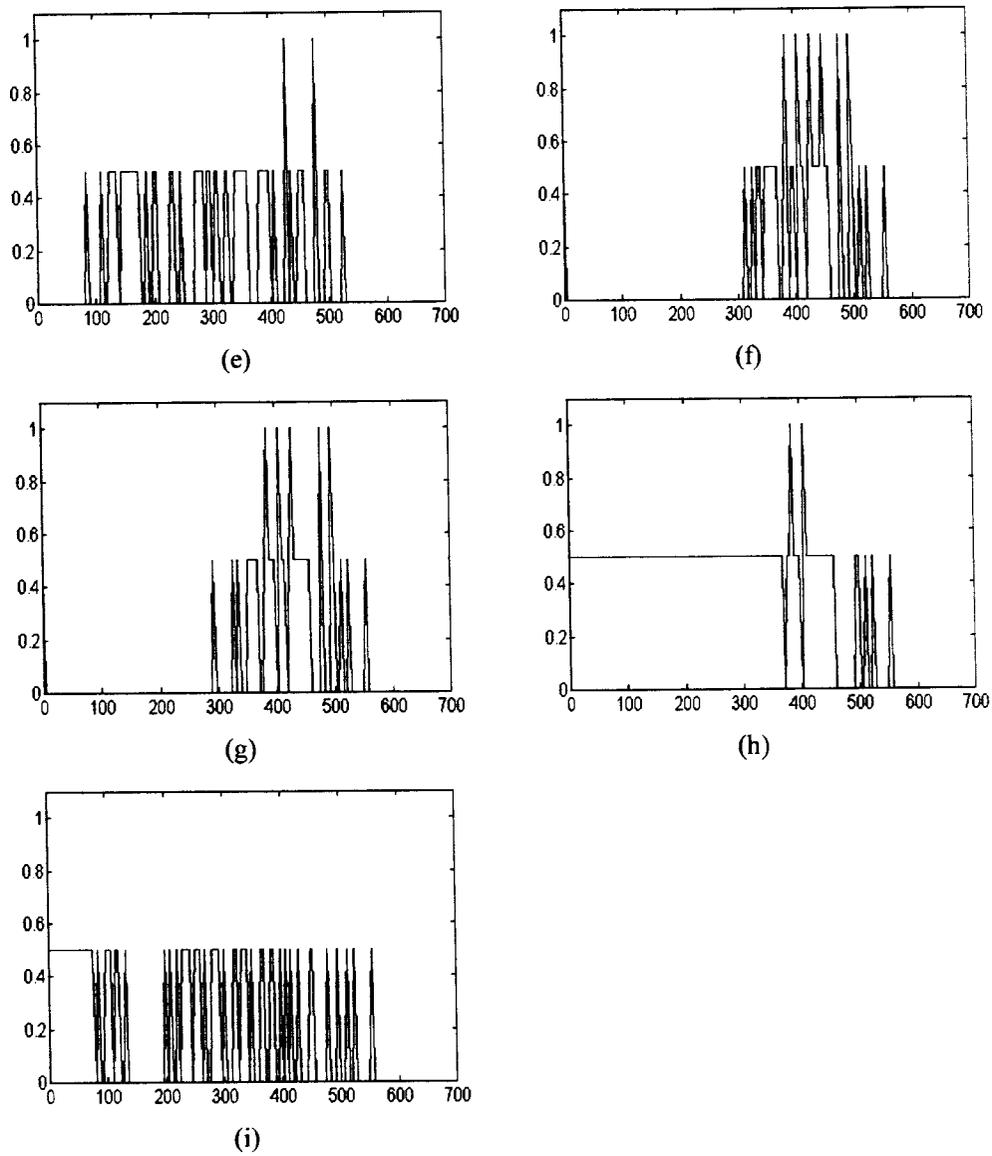


Figura 5.15. *Continuação.* (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

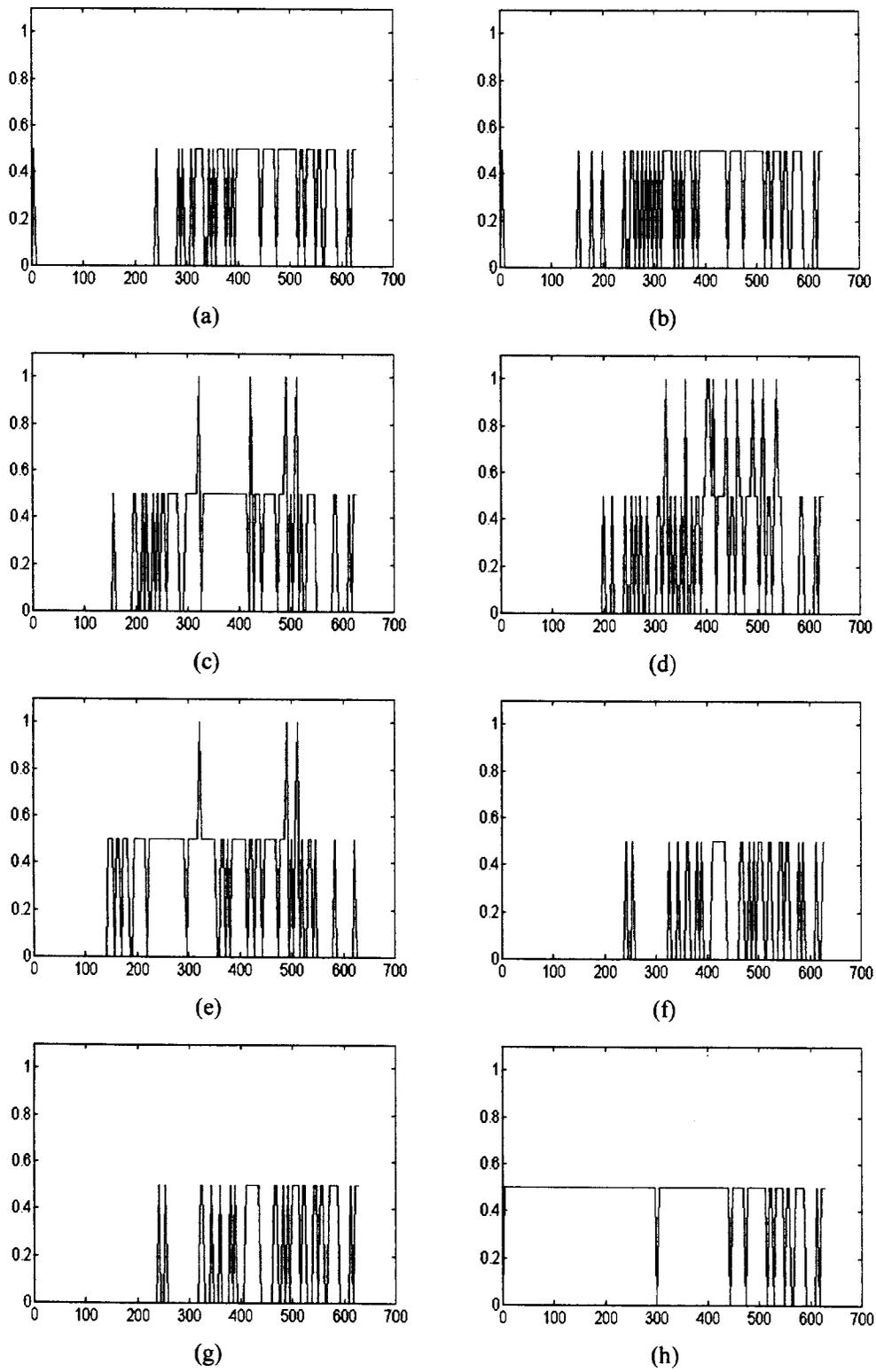


Figura 5.16. Desempenho dos métodos de segmentação para a imagem Aerial01, considerando o atributo  $y_b$ . (a) Sobel manual; (b) Gradiente vetorial; (c) Mapa auto-organizável usando cromaticidade; (d) Classificação bayesiana; (e) Classificação por  $k$  vizinhos mais próximos; (f) Mapa auto-organizável usando iluminação; (g) Mapa auto-organizável usando iluminação e cromaticidade; (h) Sobel usando limiar adaptável local; (i) Sobel usando limiar adaptável global.

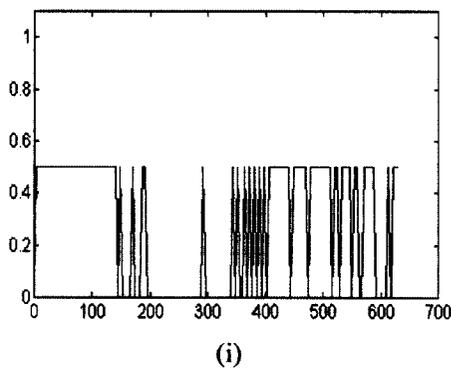


Figura 5.16. *Continuação.* (i) Sobel usando limiar adaptável global.

Para a imagem *Aerial01*, desta forma, a maior parte das técnicas apresenta um comportamento que, de um ponto de vista mais geral, se caracteriza por poucos intervalos “ótimos” (melhor coincidência e menor excesso) e um número maior de intervalos “ruins” (não detém melhor coincidência nem o menor excesso). A maioria dos intervalos, entretanto, apresenta um comportamento médio, como boa taxa de coincidência em alguns intervalos e boa taxa de excesso em outros.

Entre as diferentes variâncias utilizadas, as que mais discriminam os métodos são as de oposição cromática  $rg$  e  $yb$ . Algumas faixas de variação do atributo podem ser, com mais nitidez, vinculadas ao melhor desempenho de um ou outro método. Por exemplo, na Figura 5.15(f) nota-se que outro métodos que não o mapa auto-organizável usando as intensidades das cores apresentam melhor desempenho para o intervalo  $[0, 300]$ , aproximadamente. Melhores resultados, neste intervalo, são apresentados pelo Sobel com limiar adaptável local. Deve-se notar, ainda, que não somente esta última técnica apresentou melhor desempenho para a faixa de variância citada, mas também o mapa auto-organizável usando cromaticidade, na Figura 5.15(c), e a classificação bayesiana, na Figura 5.15(d), mostram bons resultados, embora em número menor. Situações similares podem ser notadas para os resultados apresentados na Figura 5.16, para o atributo  $yb$ .

Estes comentários, porém, aplicam-se à imagem *Aerial01*. Outras imagens tendem, em geral, a apresentar comportamento diferenciado, e comentá-las individualmente não se apresenta como uma alternativa que permita verificar um comportamento médio das técnicas. Este comportamento médio pode ser considerado, então, pelos coeficiente definidos nas Equações 5.14 e 5.16, como é exposto na próxima seção.

### 5.2.2 Coeficiente acumulativo

O número relativo médio dos melhores resultados, como estabelecido na Equação 5.14, é considerado para a classe de imagens *geral*, incluindo todas as imagens da base. As cores que identificam cada método em particular nos gráficos estão referenciadas na Figura 5.17.

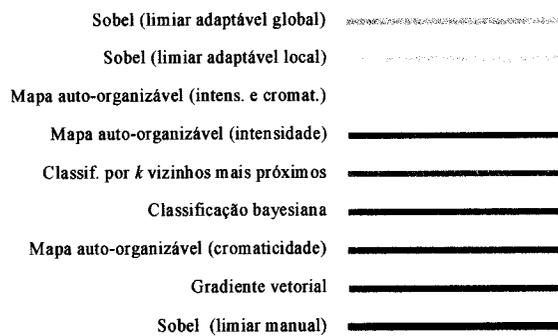


Figura 5.17. Legenda das cores utilizadas na comparação das técnicas.

A Figura 5.18 apresenta, assim, os resultados médios dos desempenhos das técnicas considerando somente a taxa de coincidência. Os valores apresentados correspondem à porcentagem do número de vezes em que uma técnica obteve a maior taxa de coincidência para uma dada imagem e um dado intervalo da variância. Para o caso de duas técnicas obterem a mesma taxa, ambas são consideradas na contagem, de forma que o total das porcentagens para as várias técnicas não totaliza necessariamente 100%. Esta opção deu-se, principalmente, por não se dispor de um critério objetivo que permitisse escolher uma ou outra técnica em caso de empate nas taxas. Nota-se, ainda, nos gráficos, que para as variâncias maiores as taxas são pequenas, o que se deve, na grande maioria dos casos, pelo fato de nem todas as imagens poderem contribuir em todo o intervalo das variâncias. Observe-se, por exemplo, que a imagem *Aerial01*, usada como exemplo para os gráficos anteriores, apresenta valores máximos para as variâncias  $g$ ,  $c$ ,  $rg$  e  $yb$ , respectivamente, em torno de 6400, 4300, 630 e 630, o que não permite que contribua para as contagens acima destes valores. Apesar dos menores valores das proporções nas áreas mais altas das variâncias, a relação entre melhores e piores desempenhos se mantém entre as diferentes técnicas. As variações máximas para os atributos selecionados estão apresentadas na Tabela 5-III, na qual os valores estão aproximados para o inteiro mais próximo.

Na Figura 5.18(a) estão os valores calculados segundo a Equação 5.14, utilizando  $p = 1$ , ou seja, permitindo apenas a contribuição das taxas de coincidência. Pode-se observar o desempenho superior das bordas produzidas com o gradiente de Sobel com limiar local nas variâncias mais baixas, até aproximadamente 4000, quando os desempenhos das outras técnicas se tornam semelhantes, com exceção do mapa auto-organizável com cromaticidade, classificação bayesiana e classificação por  $k$  vizinhos. A partir da variância 7000 as bordas obtidas por Sobel usando gradiente global produziram um comportamento médio levemente superior.

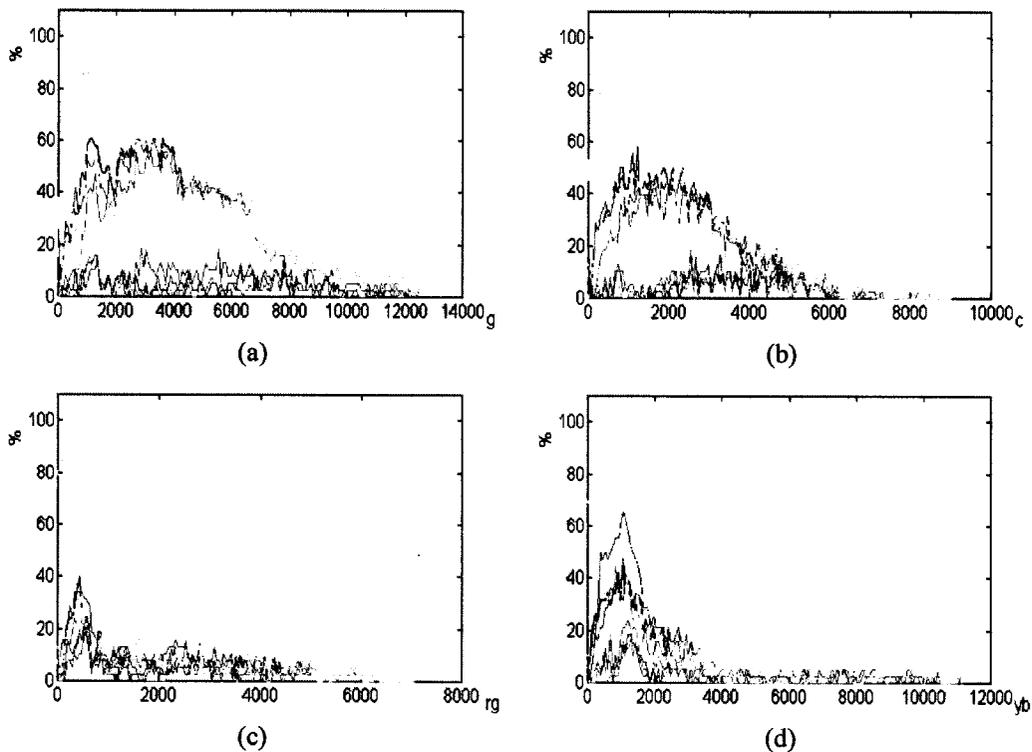


Figura 5.18. Porcentagem do número médio de vezes em que uma técnica apresentou o melhor desempenho relativo à taxa de coincidência para as variâncias (a)  $g$ , (b)  $c$ , (c)  $rg$  e (d)  $yb$ .

Para o atributo  $c$ , o comportamento médio mostrou-se semelhante, com variações locais, mas com a mesma tendência geral para as diversas técnicas. A Figura 5.18(b) mostra esta situação, possibilitando, ainda, observar que, para variâncias acima de 7000 não houve, praticamente, taxas de coincidência boas para método algum, com exceção do gradiente adaptável global, com porcentagem bastante pequenas.

Os atributos  $rg$ , representado na Figura 5.18(c), e  $yb$ , na Figura 5.18(d), mostram distribuição dos dados ligeiramente diferente. Nota-se um desempenho bom para o gradiente vetorial, considerando-se o atributo  $yb$ , no intervalo de variâncias de, aproximadamente, 1000 até um pouco inferior a 2000.

Embora apresente resultados interessantes para uma técnica ou outra, os resultados da Figura 5.18 levam em conta unicamente as taxas de coincidência. Métodos que determinam bons pontos de borda podem, também, identificar como bordas vários outros pontos que não deveriam ser.

Tabela 5-III. Variações máximas dos atributos locais para a base de imagens.

Variância	Variação
Níveis de cinza ( $g$ )	[0, 12488]
Vetores de cores ( $c$ )	[0, 9036]
Oposição Vermelho e Verde ( $rg$ )	[0, 7073]
Oposição Amarelo e Azul ( $yb$ )	[0, 11181]

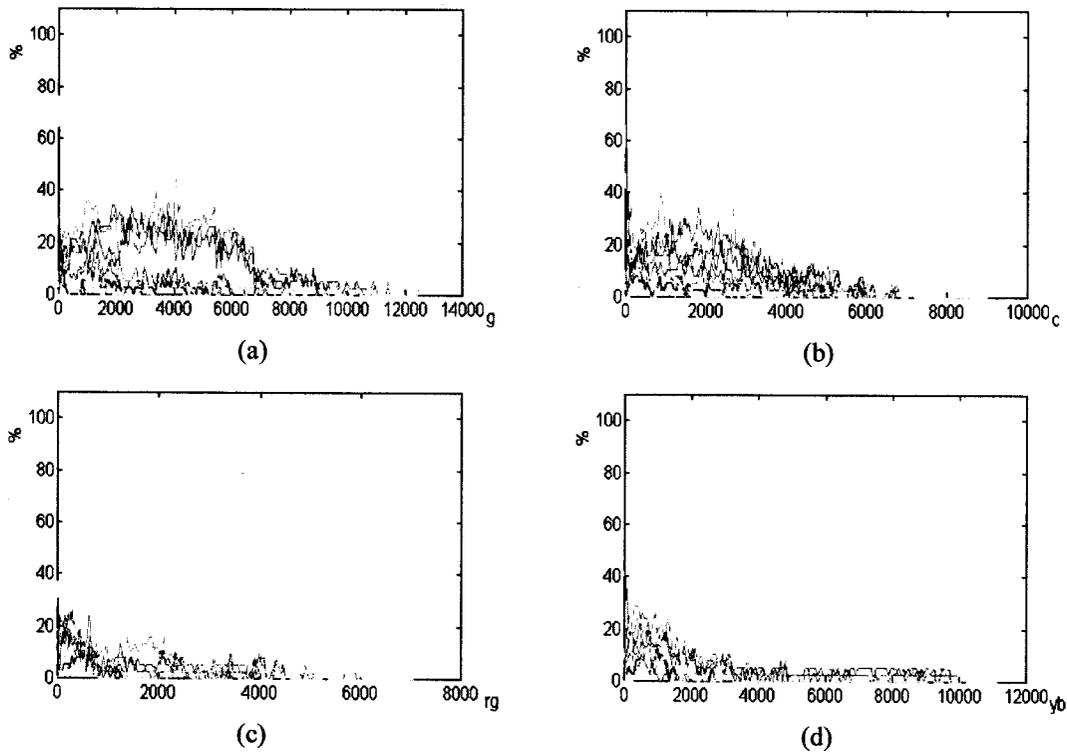


Figura 5.19. Porcentagem do número médio de vezes em que uma técnica apresentou o melhor desempenho relativo à taxa de excesso (menor taxa) para as variâncias (a)  $g$ , (b)  $c$ , (c)  $rg$  e (d)  $y_b$ .

A Figura 5.19 ilustra o comportamento dos métodos em relação às taxas de excesso. Os gráficos apresentam desempenhos (isto é, número relativo de menores taxas de excesso) em relação aos pontos assinalados como bordas que não encontram contrapartida na imagem de bordas de referência. Pode-se notar uma certa inversão nos resultados quando comparados às melhores taxas de coincidência. As técnicas que produzem menos bordas incorretas são  $k$  vizinhos, classificação bayesiana e mapas auto-organizáveis usando cromaticidade, para a maior parte dos intervalos considerados. Estas condições podem ser observadas para todos os atributos considerados, destacando-se, porém, a ausência destes métodos para variâncias  $rg$  acima de um 5000, observado na Figura 5.19(c), e o bom desempenho para o mapa auto-organizável utilizando iluminação nas variâncias  $y_b$  mais baixas, apresentado na Figura 5.19(d).

Porém, considerar apenas as coincidências ou apenas os menores excessos correspondem a situações extremas. O balanceamento destas características permite controlar a importância dada a cada uma das taxas envolvidas. Variar o valor de  $p$  na Equação 5.14 viabiliza esta ponderação, desde  $p = 0$ , com consideração exclusiva das taxas de excesso, até  $p = 1$ , considerando apenas as taxas de coincidência. Na Figura 5.20 podem ser observados os desempenhos das técnicas quando combinadas com proporções diferentes, considerando-se a variância  $c$ , notando-se a transição gradual entre os gráficos à medida que se altera a proporção. A Figura 5.21 ilustra o comportamento variando-se a proporção da combinação para o atributo  $y_b$ . Deve-se ob-

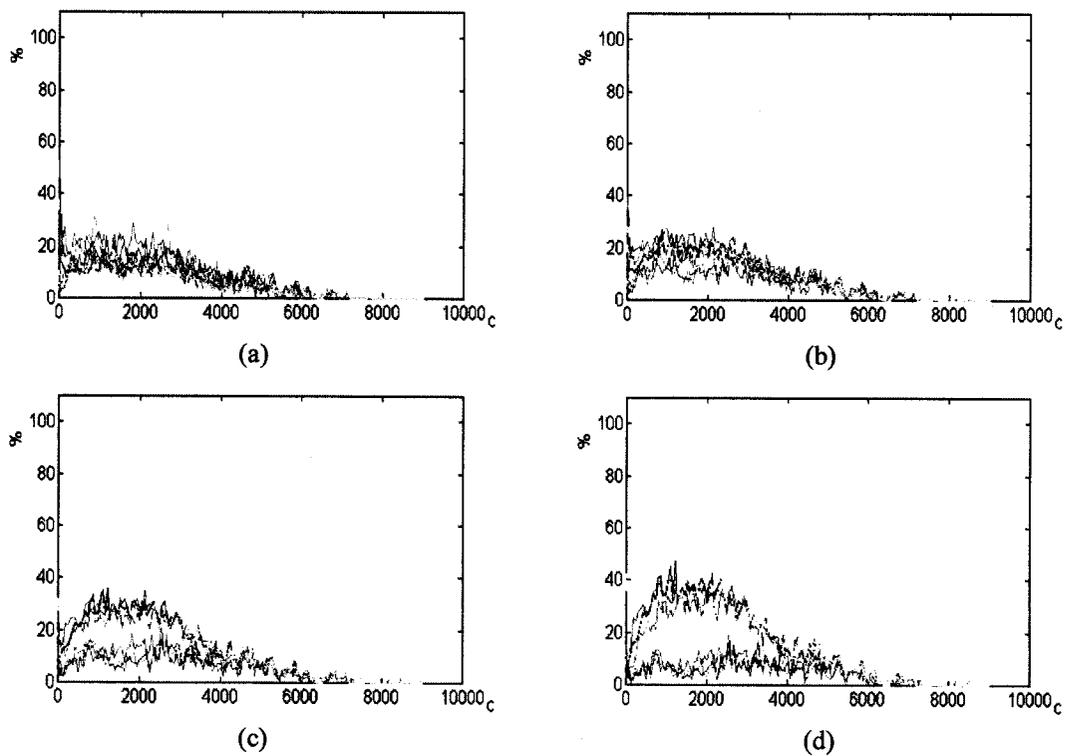


Figura 5.20. Combinação do número relativo de melhores taxas de coincidência e menores taxas de excesso, utilizando o atributo  $c$ . (a)  $p = 0,2$ , (b)  $p = 0,4$ , (c)  $p = 0,6$  e (d)  $p = 0,8$ .

servar, para ambas as figuras, que os casos extremos ( $p$  com valor zero ou um) estão apresentados nas Figuras 5.19(b) e (d), respectivamente, para a variância dos vetores de cor e nas Figuras 5.18(b) e (d), respectivamente, para a oposição YB.

Os resultados apresentados nas Figuras 5.18 a 5.21 são gerais, pois foram obtidos para todas as imagens da base. Considerar os comportamentos das várias formas de segmentação para classes de imagens, ou seja, para imagens que compartilhem alguns atributos gerais comuns é interessante, viabilizando a verificação dos desempenhos para imagens “semelhantes”<sup>\*</sup>.

Por considerações de espaço e organização da informação, os gráficos que ilustram os desempenhos para as classes de imagens estão agrupados no Apêndice E. As Figuras E.1 a E.4 mostram os comportamentos para a classe *casas*, cujos coeficiente foram calculados considerando-se as cinco imagens referenciadas na Tabela 5-II. Mostram, para valores distintos da proporção de combinação, os desempenhos para os atributos  $r$ ,  $c$ ,  $rg$  e  $yb$ , respectivamente.

A Figura E.1(a) permite observar os coeficientes calculados utilizando-se somente a minimização das taxas de excesso. Nota-se o melhor desempenho para a classificação por  $k$  vizinhos nas variâncias em cinza mais baixas e mais altas (em cor ciano), com destaques pontuais para a classificação bayesiana (verde). Técnicas volta-

<sup>\*</sup> O termo *semelhante*, no contexto, é utilizado em conformidade com as classes de imagens descritas na seção 5.1.5.

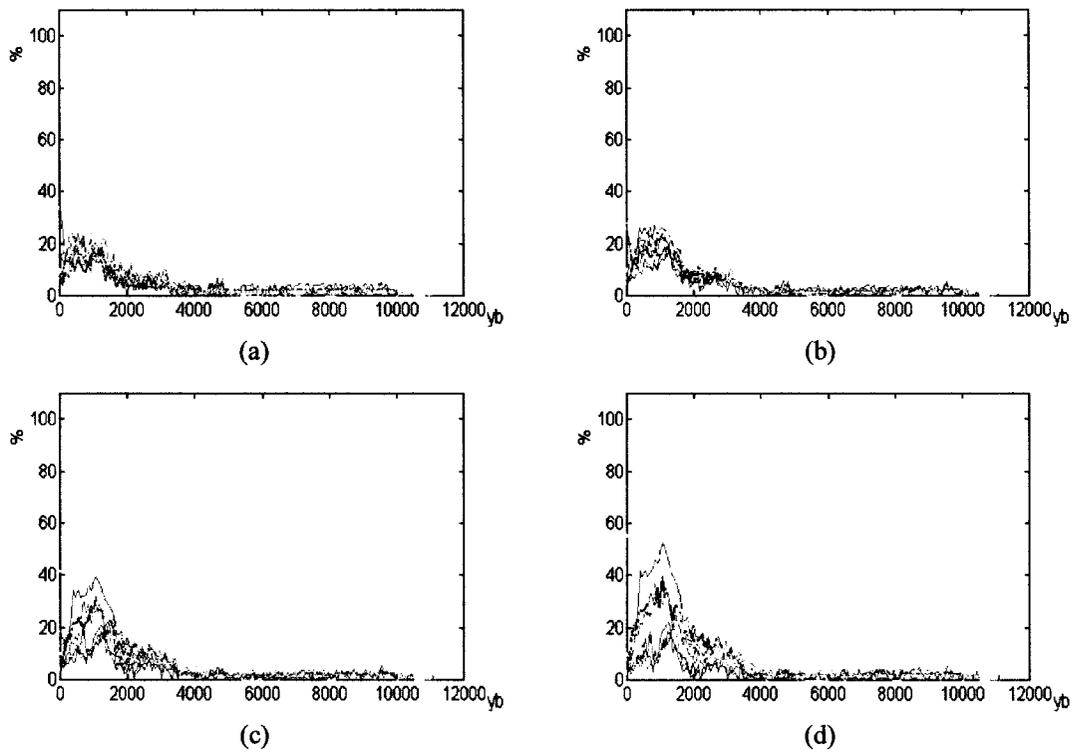


Figura 5.21. Combinação do número relativo de melhores taxas de coincidência e menores taxas de excesso, utilizando o atributo  $y_b$ . (a)  $p = 0,2$ , (b)  $p = 0,4$ , (c)  $p = 0,6$  e (d)  $p = 0,8$ .

das a bordas, como os gradientes por Sobel, apresentam-se sistematicamente com desempenho inferior, mas geralmente não nulos. No outro extremo, a Figura E.1(f), que ilustra o comportamento somente para a taxa de coincidência, mostra a nítida superioridade do Sobel adaptável local sobre as outras técnicas para as variâncias mais baixas, sendo substituído pelo Sobel adaptável global nas variâncias intermediárias e mais altas. Os mapas auto-organizáveis utilizando a cromaticidade e a combinação de iluminação e cromaticidade, embora não obtenham o melhor desempenho, apresentam coeficientes relativamente altos para as variâncias entre 2000 e 3000. Nos intervalos entre 9000 e 11000, aproximadamente, a classificação bayesiana também se sobressai com um bom desempenho. As combinações intermediárias, nas Figuras E.2(b) a (e), permitem observar, ainda, a situação de transição, nas quais o Sobel com limiar manual e adaptável local, além do mapa auto-organizável com cromaticidade e do gradiente vetorial apresentam os melhores desempenhos, nos intervalos de 2000 a 7000.

O atributo  $c$  é considerado na Figura E.2. Novamente quando a consideração é maior para a taxa de coincidência, o limiar adaptável local se sobressai. Para combinações intermediárias, com  $p$  assumindo valores 0,2 a 0,6, respectivamente representados nas Figuras E.2(b) a (d), o gradiente vetorial apresenta os melhores resultados. Para a maioria das técnicas, suas posições relativas tendem a ser mantidas, quase que de forma independente do valor de  $p$ , consideradas as exceções destacadas.

As variâncias *rg* têm pouca representatividade para esta classe, com variâncias limitadas abaixo da variância 1000, como observado nos gráficos da Figura E.3. Tendências similares às dos atributos comentados se mantêm, o que se repete para o atributo *yb*, ilustrado na Figura E.4.

Para a classe *cubos*, cujos gráficos são apresentados nas Figuras E.5 a E.8, nota-se, quando opta-se por um valor de *p* que favoreça o menor número de excessos de bordas, que a classificação bayesiana tem um bom desempenho para variâncias de 2000 a 4000, enquanto o Sobel com limiar adaptável local apresenta o pior comportamento, evidenciando seu excesso de bordas em relação à imagem de bordas de referência. Para valores de *p* mais próximo de 1, passam a se destacar os métodos que utilizam a intensidade luminosa como atributo, como o gradiente vetorial, o Sobel com limiares manual e as duas versões do limiar adaptável.

Todos os atributos locais, para a classe *externas*, apresentam maior distribuição entre os limites mínimo e máximo (Tabela 5-III). Notam-se, nas Figuras E.9 a E.12, características gerais similares às das outras classes, com predominâncias, para *p* próximo de 1, das técnicas voltadas a bordas e das que utilizam de alguma forma a intensidade luminosa. Novamente, o coeficiente calculado com favorecimento a minimizar o excesso de bordas identifica melhor comportamento para as técnicas baseadas em regiões, que utilizam a cromaticidade (por exemplo, mapa auto-organizável utilizando cromaticidade e classificações por *k* vizinhos e bayesiana). Em particular, o mapa auto-organizável utilizando cromaticidade apresenta um bom desempenho para as variâncias *c* mais baixas, até em torno de 3000 e *p* igual a 0 e 0,2, como mostra Figura E.10(a) e (b).

As Figuras E.13 a E.16 mostram, finalmente, os comportamentos para valores variados de *p* considerando-se a classe *laboratório*. O comportamento se aproxima da classe *cubos* (que constitui um subconjunto da classe *laboratório*), excetuando-se o número de vezes que o gradiente vetorial apresentou melhor desempenho, principalmente para a variância *yb*. Este desempenho pode ser notado na Figura E.16(d), (e) e (f).

Permite-se notar, de um ponto de vista mais geral sobre as classes, incluindo a *geral*, que a distinção para os desempenhos estabelecidos pelo coeficiente acumulativo não é grande, salvo algumas variantes já comentadas. Sob este prisma de comparação, é interessante perceber que, embora numericamente distintos, os gráficos sugerem um comportamento relativamente independente do conteúdo das imagens (classes), de forma que, com exceção de algumas imagens em particular, as técnicas apresentam desempenhos similares.

### 5.2.3 Coeficiente proporcional

Definido pela Equação 5.16, o coeficiente proporcional avalia os desempenhos considerando não o número de vezes em que uma técnica apresentou resultados me-

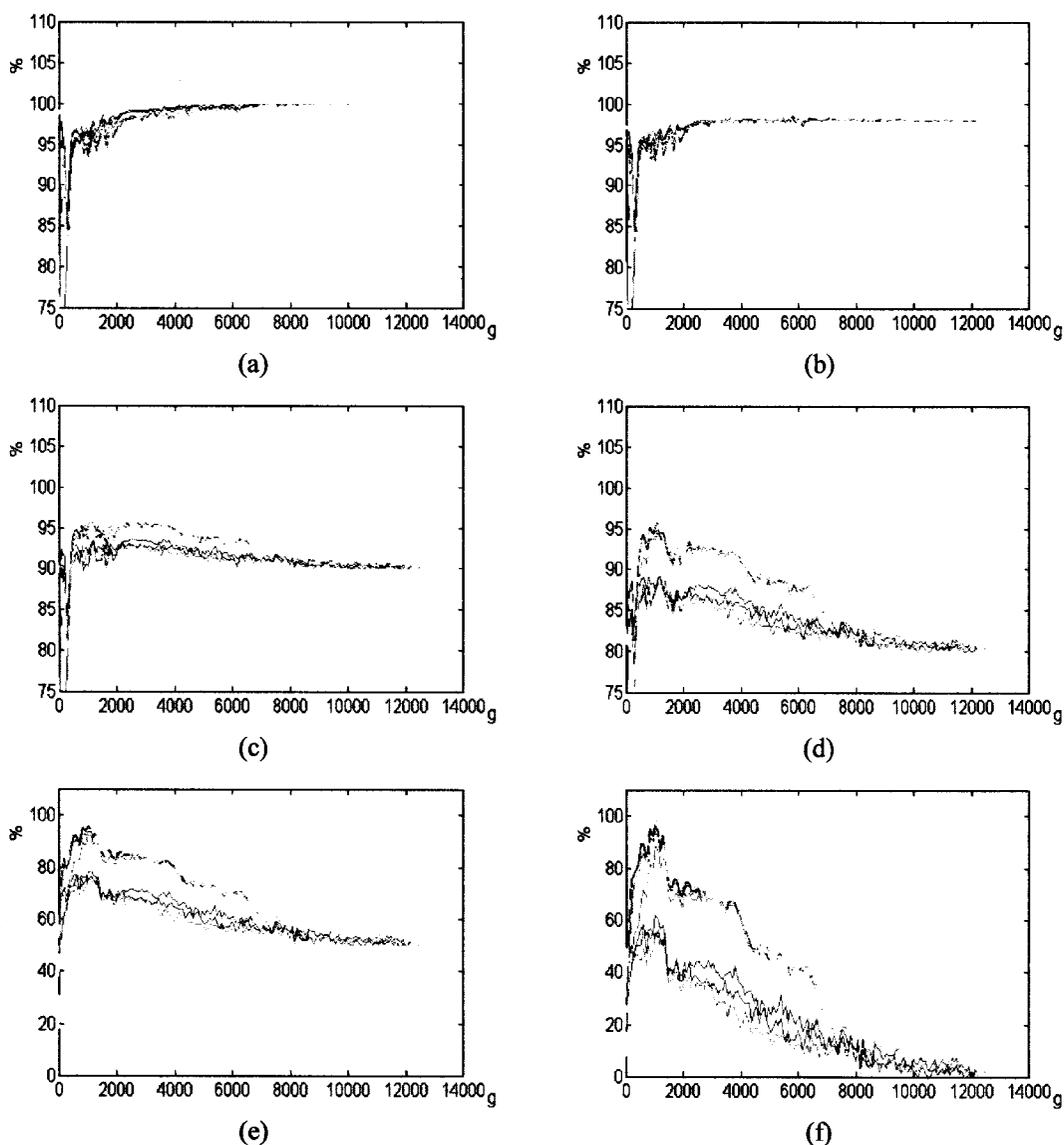


Figura 5.22. Variação dos coeficientes proporcionais segundo diferentes proporções de combinação. As distribuições referem-se ao atributo  $g$ . (a)-(f)  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5, e 1, respectivamente.

lhores que as outras\*, mas as taxas médias de coincidência e excesso. Os gráficos das Figuras 5.22 a 5.25 ilustram os coeficiente proporcionais para vários intervalos de variância, combinando as taxas de coincidência e excesso através de proporções distintas (parâmetro  $p$ ). Quando os valores de  $p$  são menores (abaixo de 0,2), a escala dos gráficos é modificada; perde-se na visualização de alguns pontos com coeficiente abaixo de 0,75, mas torna-se possível observar as curvas de cada técnica mais distintamente.

Os gráficos para o atributo local  $g$  são apresentados na Figura 5.22. Devido às diferenças de proporção entre as taxas de coincidência, que variam de 0 a 100%, e as

\* Confrontar com o coeficiente acumulativo.

taxas de excesso, que freqüentemente ultrapassam os 100%, o parâmetro  $p$  não é linear na combinação para o cálculo do coeficiente. Para a ilustração das várias combinações, as proporções utilizadas foram  $p$  com valores 0, 0,02, 0,1, 0,2, 0,5, e 1, escolhidas de forma a visualizar melhor a evolução das curvas de desempenho. Na Figura 5.22(a), embora com a visualização limitada pelas características do gráfico, é possível observar que, para  $p = 0$ , o Sobel com limiar adaptável local e a classificação por  $k$  vizinhos, além do mapa auto-organizável utilizando cromaticidade, apresentam as maiores taxas de excesso. Esta taxa é caracterizada nas variâncias baixas pela forte descendência das curvas, lembrando que a contribuição da taxa de excesso é feita com sinal negativo no cálculo do coeficiente proporcional.

Para variâncias acima de 10000 há uma tendência para que todas as técnicas tenham comportamento assintótico em 1, indicando um menor número de pontos de bordas em excesso para estas variâncias. Valores muito próximos a 1 também são causados pelo menor número de pontos de bordas para determinados intervalos de variância, ou até a ausência destas bordas. Quando há ausência de bordas, o sentido em se calcular taxas de coincidência ou excesso inexistente.

Nota-se, entretanto, uma inclinação para que métodos voltados a regiões, em especial utilizando a cromaticidade, apresentem um desempenho ligeiramente superior para as variâncias entre 500 e 8000, aproximadamente. Esta disposição relativa dos métodos tende a se inverter à medida que o valor de  $p$  favorece mais as coincidências, o que já é visível na Figura 5.22(c), na qual  $p = 0,1$ . As técnicas voltadas a bordas e que utilizam a iluminação apresentam desempenhos superiores em termos do coeficiente proporcional, havendo um menor número de inversões entre os dois grupos de técnicas. Esta tendência se mantém até o valor de  $p = 1$ , o qual considera somente a taxa de coincidência.

As Figuras 5.23 a 5.25 mostram os coeficientes proporcionais calculados para, respectivamente, as variâncias  $c$ ,  $rg$  e  $y_b$ . Podem ser observados comportamentos similares, ou seja, para valores  $p$  mais próximos de zero há um desempenho melhor, para diversos intervalos, das técnicas de segmentação por classificação bayesiana, por  $k$  vizinhos e mapas auto-organizáveis utilizando cromaticidade. Esta última, porém, tem desempenho notadamente ruim nos intervalos de variâncias menores.

A principal razão desta disparidade é a influência de uma ou outra imagem que produziu uma taxa de excesso muito alta (bastante distante da média), o que altera a média calculada. Uma verificação das variâncias das taxas de excesso revelam valores altos, indicando uma dispersão grande em torno da média apresentada nos gráficos.

Os comportamentos de todos os métodos são relativamente coerentes uns com os outros, descrevendo curvas com formato similar, embora numericamente diferentes. Quando os coeficientes proporcionais decrescem, por exemplo, todas as técnicas são afetadas de forma semelhante, mesmo que em maior ou menor grau.

Da mesma forma, é também interessante ressaltar que não há nenhuma técnica que possa ser considerada ruim, dado que todas apresentam desempenhos melhores ou piores em relação umas às outras, mas não um comportamento ruim absoluto. Os

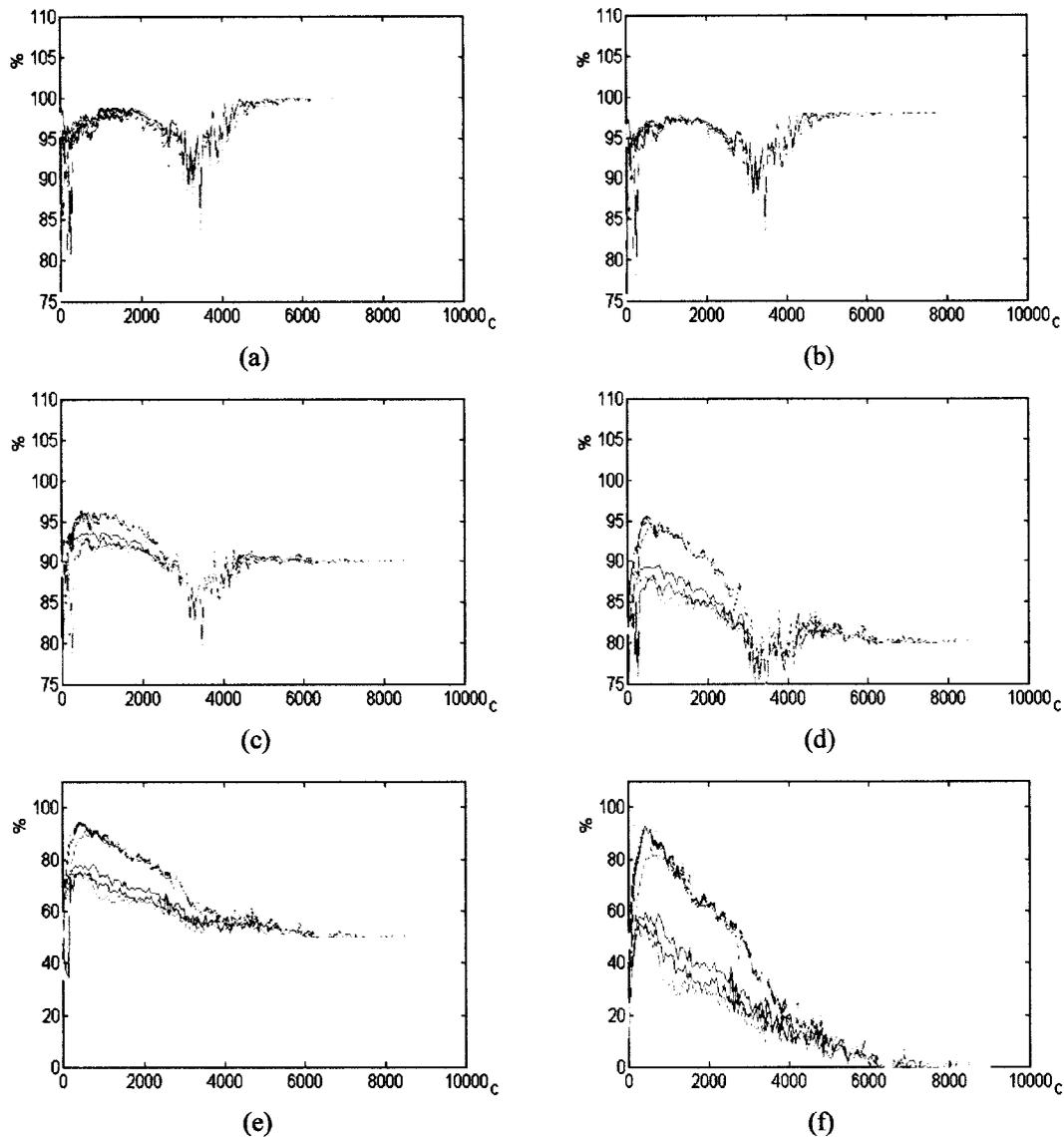


Figura 5.23. Variação dos coeficientes proporcionais segundo diferentes proporções de combinação. As distribuições referem-se ao atributo  $c$ . (a)-(f)  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5, e 1, respectivamente.

gráficos permitem observar que raramente as taxas induzem a coeficientes nulos, salvo para os intervalos nos quais não há contribuição de nenhuma técnica.

A análise dos coeficientes calculados com as restrições das imagens divididas em grupos, apresentadas no Apêndice F. Não podem ser distinguidas modificações significativas, o que leva a considerar um certo grau de independência das técnicas em relação ao conteúdo das imagens de uma forma geral. É claro que, para uma imagem em particular, pode haver variações.

Um destaque no tratamento por classes é para a variância  $c$ , considerando-se a classe *externas*. Pode-se observar, na Figura F.10, que entre as variâncias 3000 e 6000, aproximadamente, não há distinções explícitas entre as técnicas, todas elas re-

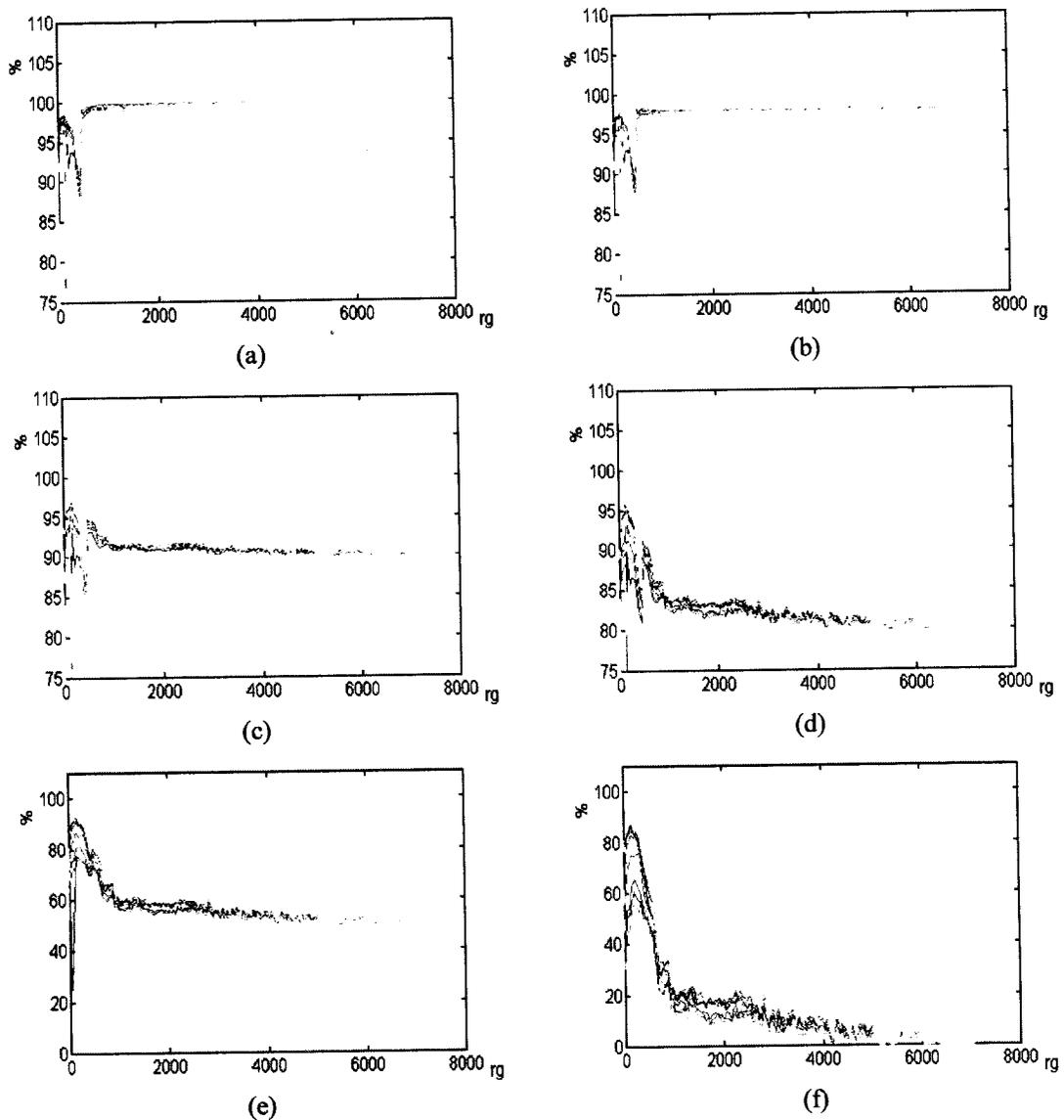


Figura 5.24. Variação dos coeficientes proporcionais segundo diferentes proporções de combinação. As distribuições referem-se ao atributo  $rg$ . (a)-(f)  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5, e 1, respectivamente.

sultando em um coeficiente proporcional bastante similar, de forma que em alguns poucos intervalos um técnica se destaca das outras.

### 5.3 Considerações gerais sobre os coeficientes acumulativo e proporcional

Dos resultados apresentados na seção 5.2, algumas considerações podem ser feitas, relativamente à utilização das classes e também relacionadas às fases que o gráfico assume para os diferentes valores de  $p$ .

Quando são usadas as classes, as tendências mais fortes das técnicas se mantêm, pelo menos sob um ponto de vista mais geral. As técnicas que utilizam a intensidade luminosa da imagem como parâmetro para segmentação produzem bordas mais

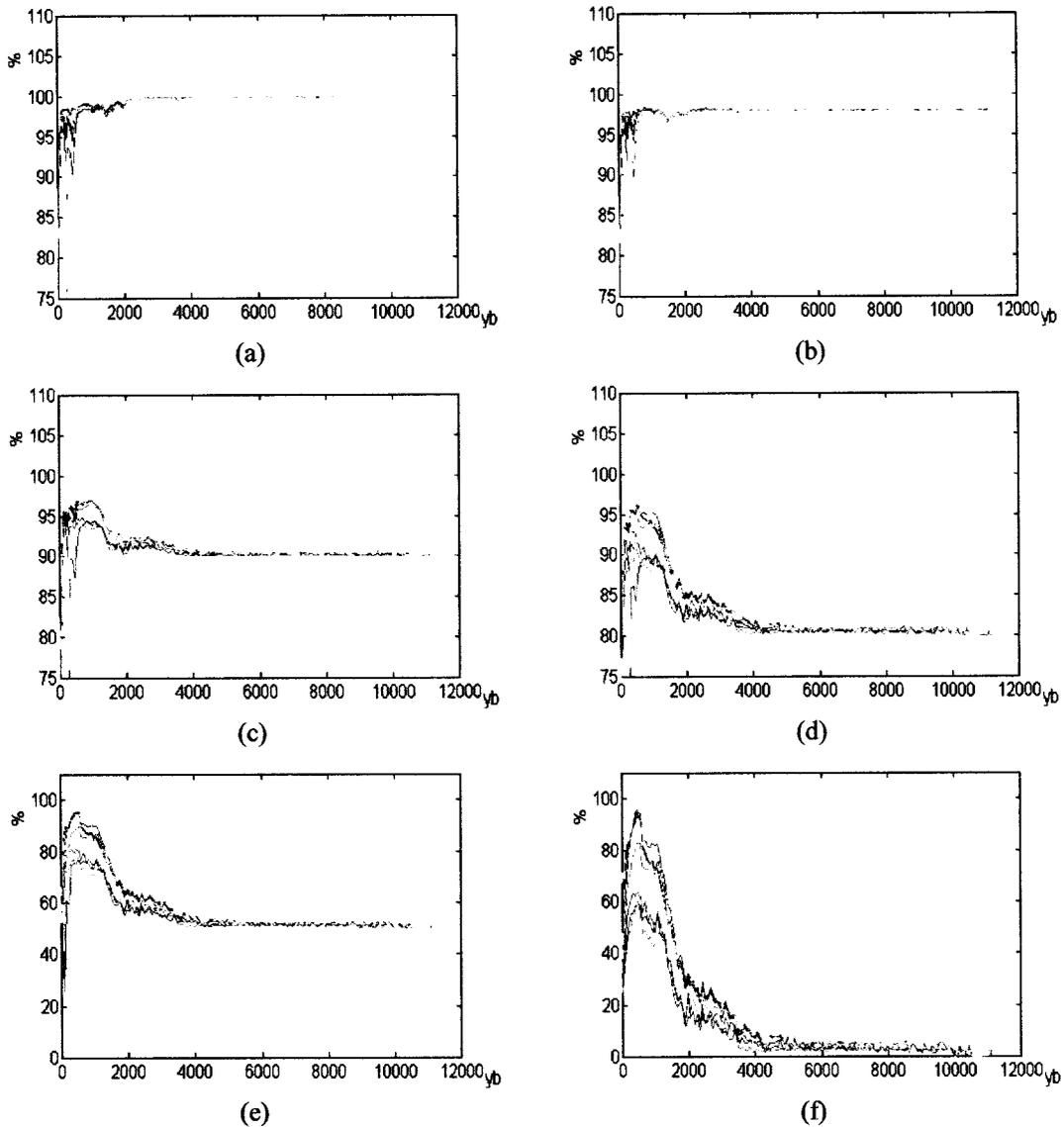


Figura 5.25. Variação dos coeficientes proporcionais segundo diferentes proporções de combinação. As distribuições referem-se ao atributo  $y_b$ . (a)-(f)  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5, e 1, respectivamente.

precisas, refletidas nas taxas de coincidência mais elevadas. Em contrapartida, estas técnicas tendem a ter uma taxa de excesso maior. Esta característica é patente, por exemplo, para a segmentação por Sobel com utilização do limiar adaptável local, que produz um excesso grande de bordas, apesar de também detectar corretamente as esperadas. As imagens *Cube015*, *Frogs01* e *Sign04* ilustram bem este excesso.

Apesar de uma observação mais generalizada não indicar diferenças qualitativas entre os resultados para as várias classes, mesmo considerando-se os atributos locais distintos, é fácil observar que, numericamente, há diferenças e que, utilizando-se um valor adequado para  $p$  é possível indicar uma proporção que maximize as coincidências e minimize os excessos. Esta proporção é, sem dúvidas, dependente das classes e dos atributos.

Tabela 5-IV. Sumário do desempenho relativo entre as técnicas de segmentação utilizadas.

Técnicas	Prioridade para maximização da taxa de coincidência	Prioridade para minimização da taxa de excesso
(i) Sobel (limiar manual)	Desempenho em variâncias intermediárias baixas apenas inferior a (ii); desempenho similar às outras, exceto (v), (vii) e (ix) nas variâncias restantes.	Desempenho geral muito baixo em todas as variâncias; valores mais altos de desempenho em variâncias mais baixas.
(ii) Sobel (limiar adaptável local)	Melhor desempenho global para variâncias baixas; tendência mantida até entre variâncias intermediárias e altas, quando se assemelha às outras técnicas.	Desempenho geral baixo em todas as variâncias, alternando-se com (i) e (iii).
(iii) Sobel (limiar adaptável global)	Similar a (i), embora ligeiramente inferior nas variâncias baixas.	Similar a (i) e (ii).
(iv) Gradiente vetorial	Similar a (iii), com exceção para a variância $y_b$ , para a qual é similar a (ii) no intervalo em torno de 2000.	Ligeiramente superior a (i), (ii) e (iii), exceto para variâncias médias e altas.
(v) Mapa auto-organizável (cromaticidade)	Desempenho baixo em todo o intervalo.	Similar a (viii), alternando melhores e piores desempenhos nas variâncias entre médias e baixas, principalmente.
(vi) Mapa auto-organizável (intensidade em cores)	Ligeiramente inferior a (iv) para variâncias baixas e similar a (i) acima das variâncias médias.	Similar a (v), (viii) e (ix) nas variâncias muito baixas; nas variâncias restantes similar a (i) até (iv).
(vii) Mapa auto-organizável (intensidade e cromaticidade)	Similar a (vi) nas variâncias mais baixas; inferior a (iv) para variâncias médias e entre médias e baixas; similar a (ii) nas variâncias restantes.	Similar a (iii).
(viii) Classificação bayesiana	Um pouco superior a (vii) até variâncias intermediárias; similar a (vii) nas outras variâncias.	Desempenho similar a (ix) para todas as variâncias, com exceção para valores entre baixos e médios, mais notadamente para o atributo $c$ . Ligeiramente superior às demais técnicas para o atributo $y_b$ em variâncias mais altas.
(ix) Classificação por $k$ vizinhos mais próximos	Similar a (viii).	Melhor desempenho até variâncias médias, porém somente ligeiramente superior aos demais; para as variâncias mais altas, comportamento similar às demais técnicas.

No que concerne aos gráficos em si, é possível determinar três fases distintas, independentes das classes de imagens e do atributo em consideração. A primeira é a que tem em conta a minimização das taxas de excesso, caracterizando a cromaticidade e as técnicas voltadas a regiões como melhores em desempenho. A segunda, em contraposição, considera a maximização da taxa de coincidência, descartando as taxas de excesso. Destacam-se, neste caso, as técnicas voltada a bordas e as que utilizam os níveis de cinza (intensidade luminosa da imagem) ou os valores não normalizados das componentes RGB (que também incorporam a iluminação). A terceira fase refere-se à transição entre estes dois extremos, caracterizada pela ponderação das duas taxas. O valor de  $p$  que indica esta transição, entretanto, depende de como as taxas foram geradas, o que está vinculado às classes e aos atributos locais selecionados.

A Tabela 5-IV apresenta um sumário dos desempenhos das técnicas estudadas, elaborado com base na classe *geral* (todas as imagens) e considerando-se, principalmente, o desempenho medido pelo coeficiente acumulativo. Quando se considera o coeficiente proporcional, as tendências gerais descritas na tabela se mantêm, embora alguns comentários mais específicos não possam mais ser aplicados. Para as classes consideradas várias exceções podem também ser notadas, conforme se pode observar nos gráficos dos Apêndices E e F.

Estas informações de transição podem ser usadas para identificar os melhores desempenhos das técnicas em regiões da imagem (determinadas pelos intervalos das variâncias), o que permitiria otimizar, até certo grau, as taxas de coincidência e de excesso. A seleção das bordas de uma técnica em particular pode, assim, estar vinculada aos atributos locais que a imagem apresenta, como apresentado e discutido nas próximas seções.

#### 5.4 Combinação de bordas baseada no contexto local

A avaliação comparativa de técnicas de segmentação, como proposta na seção 5.1, abre a possibilidade de se promover uma seleção dos métodos que proporcionam os melhores resultados, seja considerando a abordagem global, seja restringindo a abordagem a classes preestabelecidas. Considerando-se, assim, que se dispõe do desempenho de cada técnica para intervalos de um atributo local, propõe-se um modo de combinar os vários resultados também por intervalos. Técnicas que apresentam comportamento bom para determinados intervalos, assim, podem contribuir condicionalmente para um mapa de bordas geral.

Introduzidos na seção 5.1.4, os coeficientes acumulativo e proporcional representam uma quantificação da qualidade de cada método para um dado atributo local.

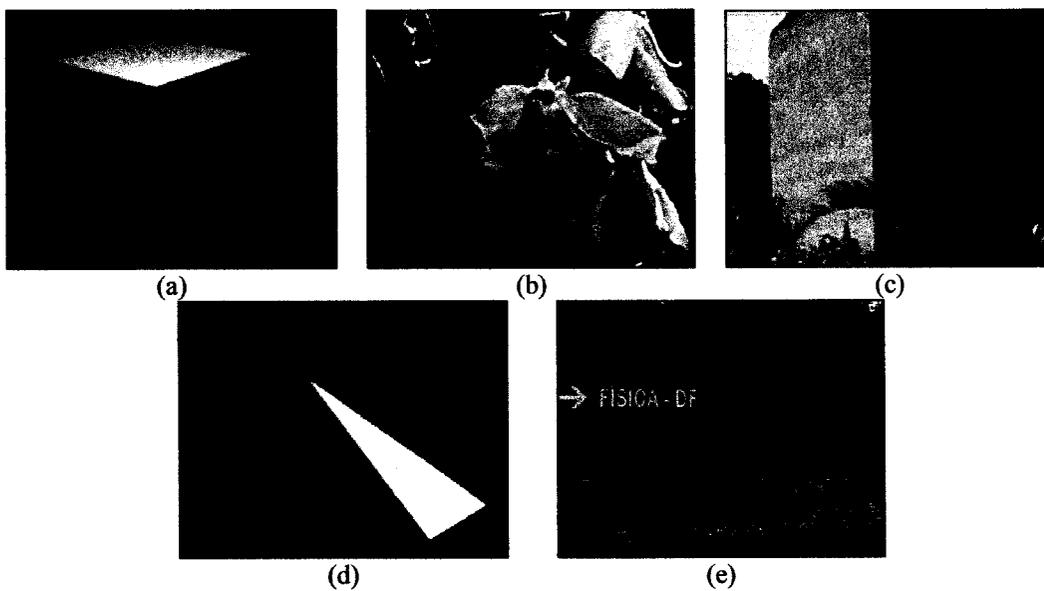


Figura 5.26. Imagens utilizadas para os resultados apresentados nas Figuras 5.27 a 5.34. (a) Cube011; (b) Flower02; (c) House03; (d) Pyram04 e (e) Sign04.

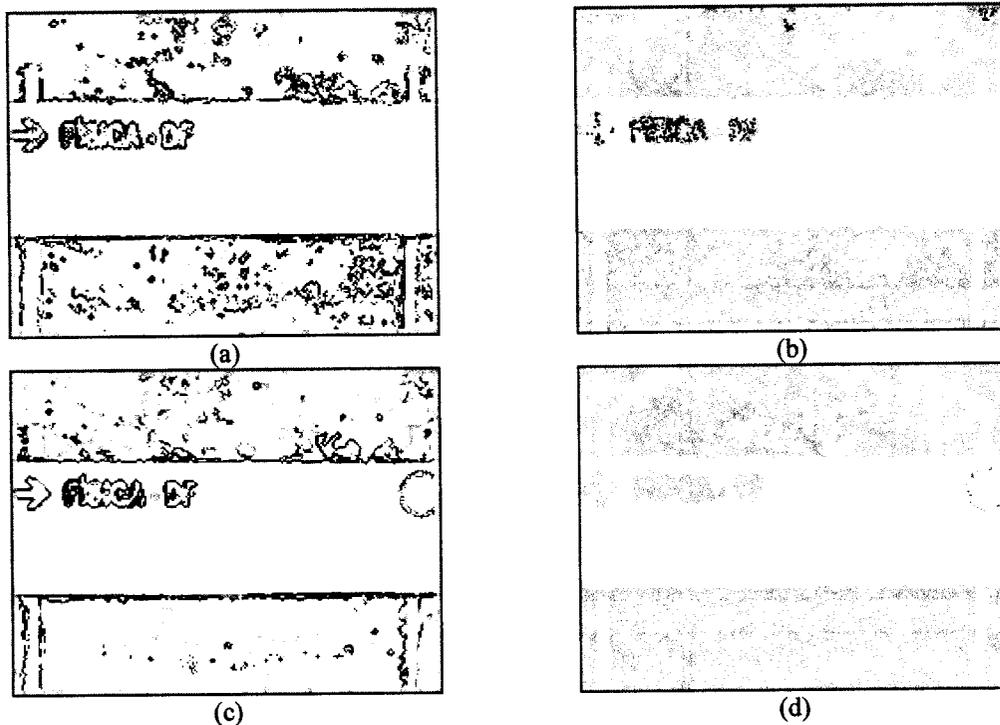


Figura 5.27. Representação da combinação de bordas para a imagem *Sign04*, utilizando o coeficiente acumulativo. (a) Atributo *g*, com  $p = 0,2$ ; (b) Atributo *g*, com  $p = 0,5$ ; (c) Atributo *rg*, com  $p = 0,1$ ; e (d) Atributo *rg*, com  $p = 0,5$ . O código de cores é o definido na Figura 5.17.

Pela seleção, intervalo a intervalo, do melhor método, pode-se tentar otimizar o mapa de bordas segundo o critério estabelecido para a comparação dos desempenhos.

A combinação dos mapas de bordas requer a divisão da imagem segundo os intervalos utilizados para as variâncias. Para cada intervalo, que representa um conjunto de pontos da imagem, opta-se pelas bordas produzidas pelo método de melhor desempenho para aqueles valores da variância. Seja para o coeficiente acumulativo ou para o coeficiente proporcional, opta-se, assim, pela técnica que proporciona o maior valor no intervalo. A “qualidade” da combinação, em termos de melhores taxas de coincidência ou menores taxas de excesso, é controlada também pela proporção da combinação, estabelecida pelo valor de  $p$  nas Equações 5.14 e 5.16.

### 5.5 Resultados da combinação de bordas e discussão

Na Figura 5.26 estão apresentadas as imagens utilizadas nos resultados ilustrados nas Figuras 5.27 a 5.34. A Figura 5.27 mostra a combinação obtida usando os coeficientes acumulativos calculados para os atributos *g* e *rg*, considerando-se a imagem *Sign04*. Na Figura 5.27(a) estão as bordas combinadas pelo atributo *g*, usando-se  $p = 0,2$ . As cores utilizadas identificam cada técnica em particular, usando a convenção estabelecida na Figura 5.17. Nota-se que há contribuições para as bordas de várias técnicas, algumas mais precisas que outras em relação à imagem de bordas de referência. A modificação da proporção entre coincidências e excessos para  $p = 0,5$  modifica a composição das bordas, como ilustra a Figura 5.27(b). Para esta combina-

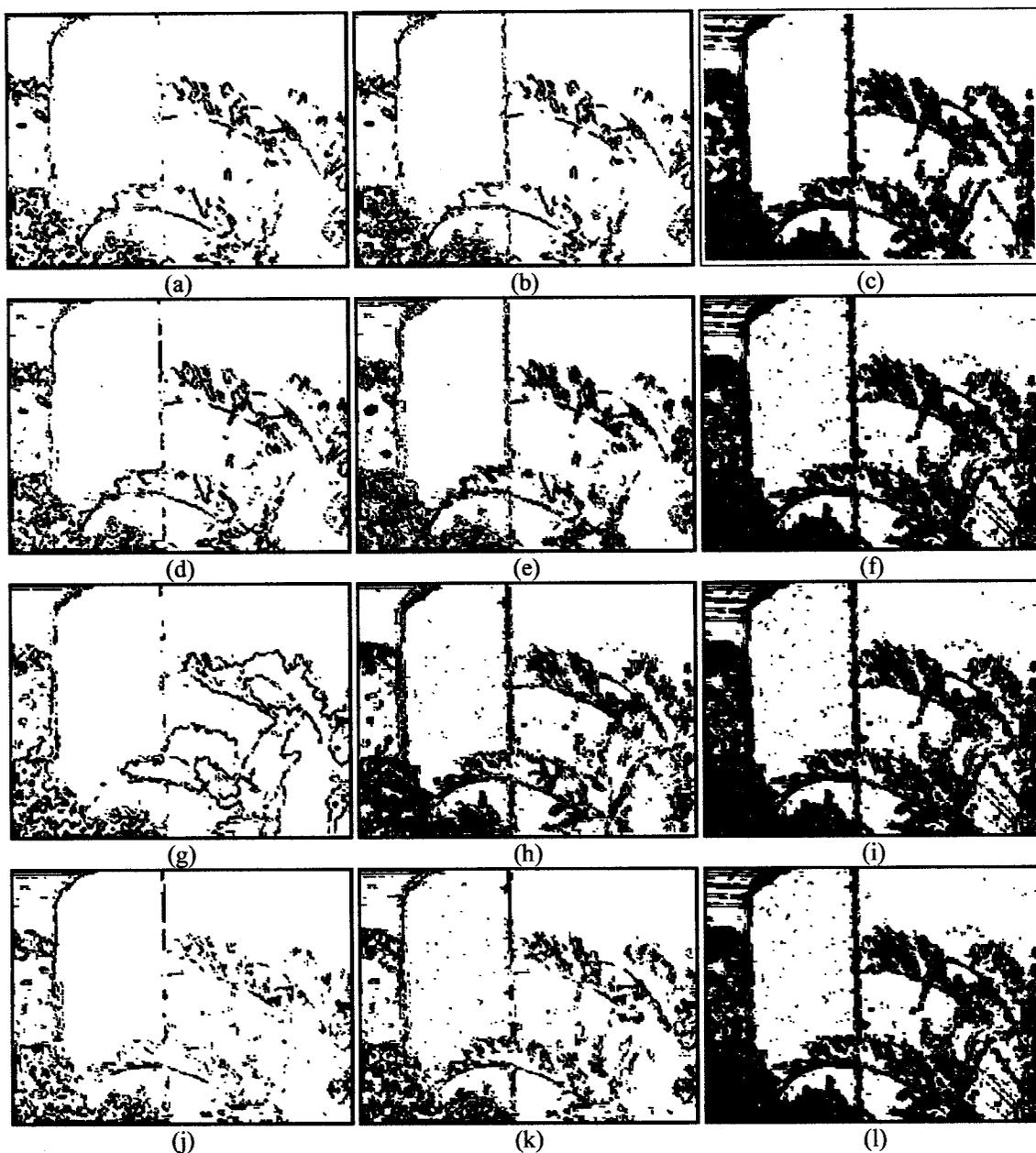


Figura 5.28. Combinações de bordas para a imagem House03. (a)-(c): atributo  $g$ ; (d)-(f): atributo  $c$ ; (g)-(i): atributo  $rg$ ; e (j)-(l): atributo  $y_b$ . As proporções de combinação para cada tríade de imagens é, respectivamente, 0, 0,2 e 1.

ção específica, notam-se as poucas contribuições da maioria das técnicas, praticamente levando-se em consideração apenas as bordas produzidas pelo Sobel com limiar adaptável local, codificado em ocre.

As Figuras 5.27(c) e (d) mostram combinações similares com a utilização das variâncias  $rg$ . Para a proporção 0,1, as contribuições principais de  $k$  vizinhos (em azul claro) e alguns trechos de contorno pela classificação bayesiana (em verde) podem ser observadas. A proporção de 0,5 mostra a grande preponderância do Sobel com limiar adaptável local, bem como o excesso de bordas que esta técnica apresenta. Nota-se,

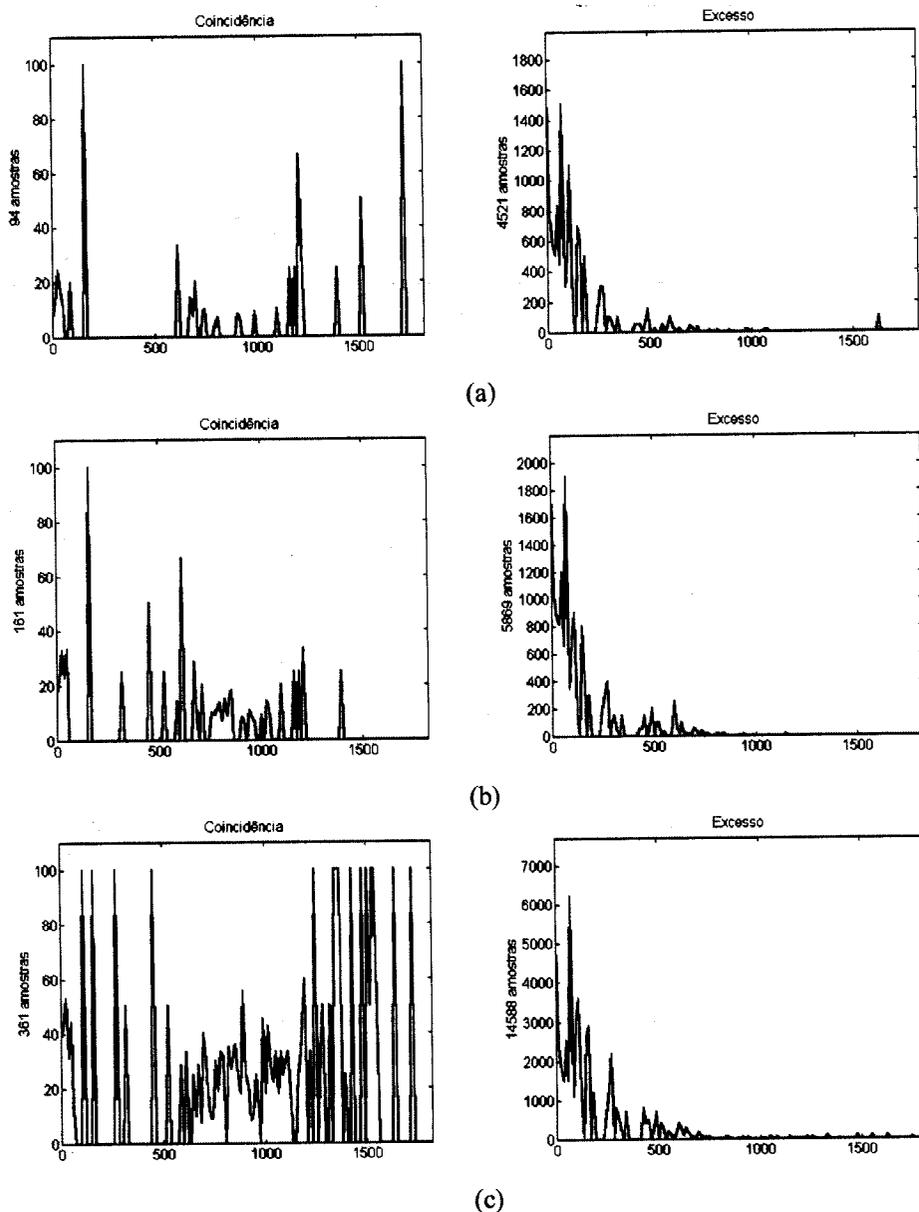


Figura 5.29. Taxas de coincidência (à direita) e de excesso para as combinações de bordas da Figura 5.28(a) a (c), respectivamente. (a)  $p = 0$ ; (b)  $p = 0,2$ ; e (c)  $p = 1$ .

ainda, que boa parte do círculo na placa é caracterizada por bordas obtidas pela classificação por  $k$  vizinhos.

Dada a dificuldade de visualização das bordas em cores, opta-se, nas figuras seqüentes, pela utilização de mapas com as bordas em preto, sem especificação explícita da combinação. Excetuam-se, porém, algumas das imagens, nas quais a codificação das cores nas bordas seja importante. Embora certas informações deixem de ser explicitadas, ainda se permite um boa avaliação visual das diferenças entre as combinações.

A combinação de bordas usando os resultados produzidos para a imagem House03 estão apresentados na Figura 5.28, para valores distintos da proporção em que são consideradas as coincidências e os excessos das bordas. Os mapas são obti-

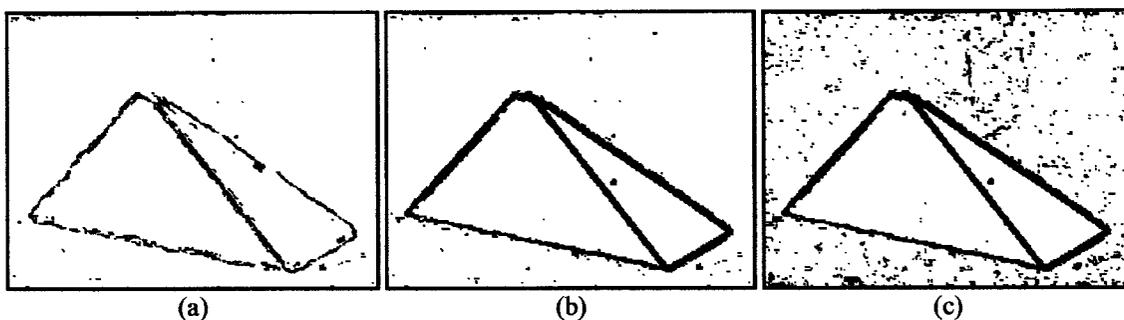


Figura 5.30. Combinações de bordas para a imagem *Pyram04*. Proporções: (a) 0; (b) 0,3 e (c) 1.

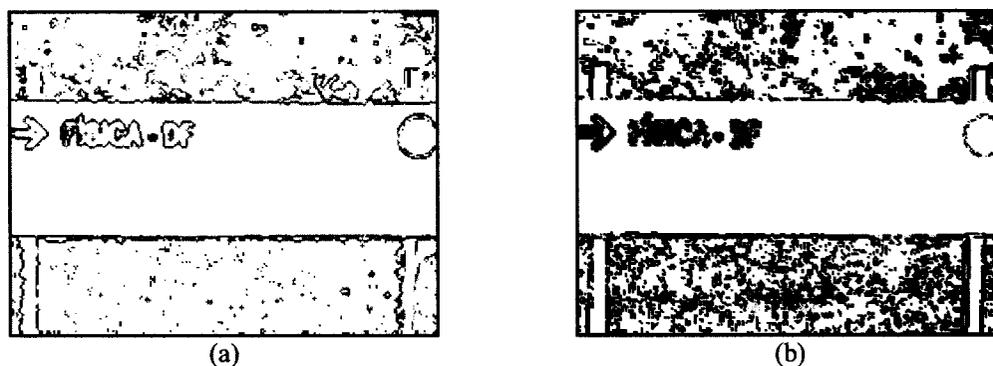


Figura 5.31. Representação da combinação de bordas para a imagem *Sign04*, utilizando o coeficiente proporcional calculado sobre o atributo *rg*: (a)  $p = 0$ ; (b)  $p = 0,3$ . O código de cores é o definido na Figura 5.17.

dos através da comparação dos desempenhos levando-se em consideração todas as imagens da base (classe *geral*) e os atributos definidos.

A Figura 5.28(a) mostra a combinação com consideração exclusiva das técnicas que minimizam a taxa de excesso, caracterizada por  $p = 0$ , calculada para a variância  $g$ . Para um valor de  $p$  intermediário, igual a 0,2, a combinação das bordas resulta suavemente diferente da anterior, notadamente acrescentando nitidez à aresta formada pelas faces das paredes, como ilustrado na Figura 5.28(b). A combinação que maximiza a taxa de coincidência, desconsiderando as taxas de excesso, é apresentada na Figura 5.28(c). Pode-se notar o progressivo aumento do número de pontos de bordas nos mapas, até o excesso de pontos do último mapa. A Figura 5.28(d) a (f) ilustram, para os mesmos valores de  $p$ , respectivamente, os resultados da combinação relativos ao atributo  $c$ .

Da mesma forma, as Figuras 5.28(g) a (i) e as Figuras 5.28(j) a (l) apresentam, respectivamente, combinações similares para os atributos  $rg$  e  $yb$ . De modo geral o comportamento é similar, desde um mapa de bordas menos denso, calculados com  $p$  igual a zero, passando por um mapa intermediário (com  $p$  arbitrário), até um mapa mais denso, para o qual são consideradas apenas as coincidências das borda. Algumas particularidades, porém, podem ser destacadas nos mapas, como o mapa um pouco menos denso produzido para o atributo  $g$ , com  $p = 1$ , apresentado na Figura 5.28(c), quando comparado aos outros atributos. Além deste, a variância  $rg$ , na Figura

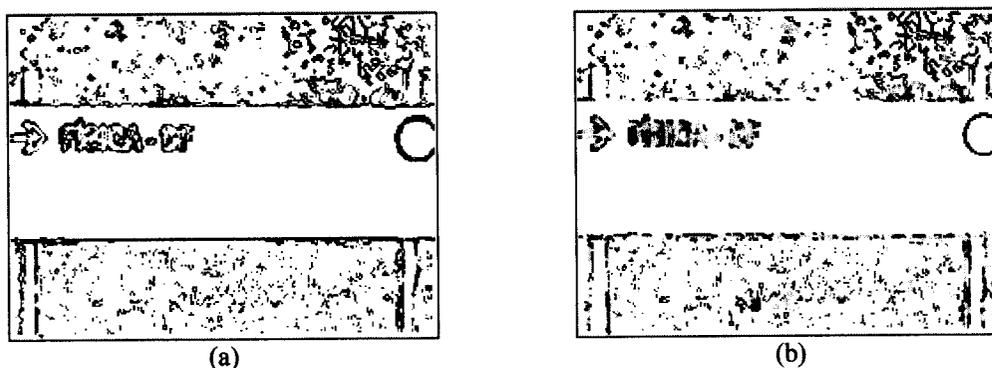


Figura 5.32. Representação da combinação de bordas para a imagem *Sign04*, utilizando o coeficiente proporcional calculado sobre o atributo *rg*: (a)  $p = 0$ ; (b)  $p = 0,2$ . O código de cores é o definido na Figura 5.17.

5.28(g), produz um mapa bastante completo mesmo considerando-se somente as taxas de excesso.

A Figura 5.29 apresenta as taxas de coincidência e de excesso para os intervalos da variância  $g$ , calculados relativamente aos mapas de bordas das Figuras 5.28(a) a (c). Podem ser observados os aumentos progressivos nas taxas de coincidência e os decréscimos nas taxas de excesso, à medida em que varia o fator da combinação  $p$ .

As combinações apresentadas na Figura 5.30 mostram, para a imagem *Pyram04*, os mapas de bordas produzidos por proporções 0, 0,3 e 1, respectivamente. Novamente a classe *geral* é utilizada para produzir os histogramas dos coeficientes acumulativos e a variância local dos níveis de cinza é considerada. O mapa da Figura 5.30(a) considera somente as taxas de excesso, com  $p$  igual a 0. Valores de  $p$  iguais a 0,3 e 1 resultam nos mapas de bordas apresentados nas Figuras 5.30(b) e (c), respectivamente. Nota-se, novamente, o progressivo aumento da densidade das bordas que compõem os mapas, em conformidade com o aumento do fator de combinação que pondera entre as duas taxas utilizadas.

A utilização do coeficiente proporcional para as combinações é ilustrada na Figura 5.31. Combinações utilizando  $p$  igual a 0 e  $p$  igual a 0,3 resultam nos mapas apresentados nas Figuras 5.31(a) e (b), respectivamente. Notam-se as composições específicas das diferentes técnicas, caracterizadas pelas cores específicas (referenciadas na Figura 5.17). Notam-se as influências maiores das bordas da classificação por  $k$  vizinhos (em ciano) e as contribuições da classificação bayesiana (verde) e do gradiente vetorial (azul). Na Figura 5.31(b) a combinação utiliza  $p = 0,3$ , para o mesmo atributo *rg*. Nota-se a nítida predominância do Sobel com limiar manual (preto), a complementação das bordas da placa com Sobel com limiar adaptável local (ocre) e do círculo na plana com o gradiente vetorial (azul). É possível observar, na segunda imagem, uma quantidade maior de pontos de borda detectados além das expectativas, característicos na área da grama e na folhagem das árvores na parte superior.

A Figura 5.32 apresenta as combinações para a variância  $yb$ , com variantes nos resultados. A maior contribuição do gradiente vetorial, em azul, é patente, mas as

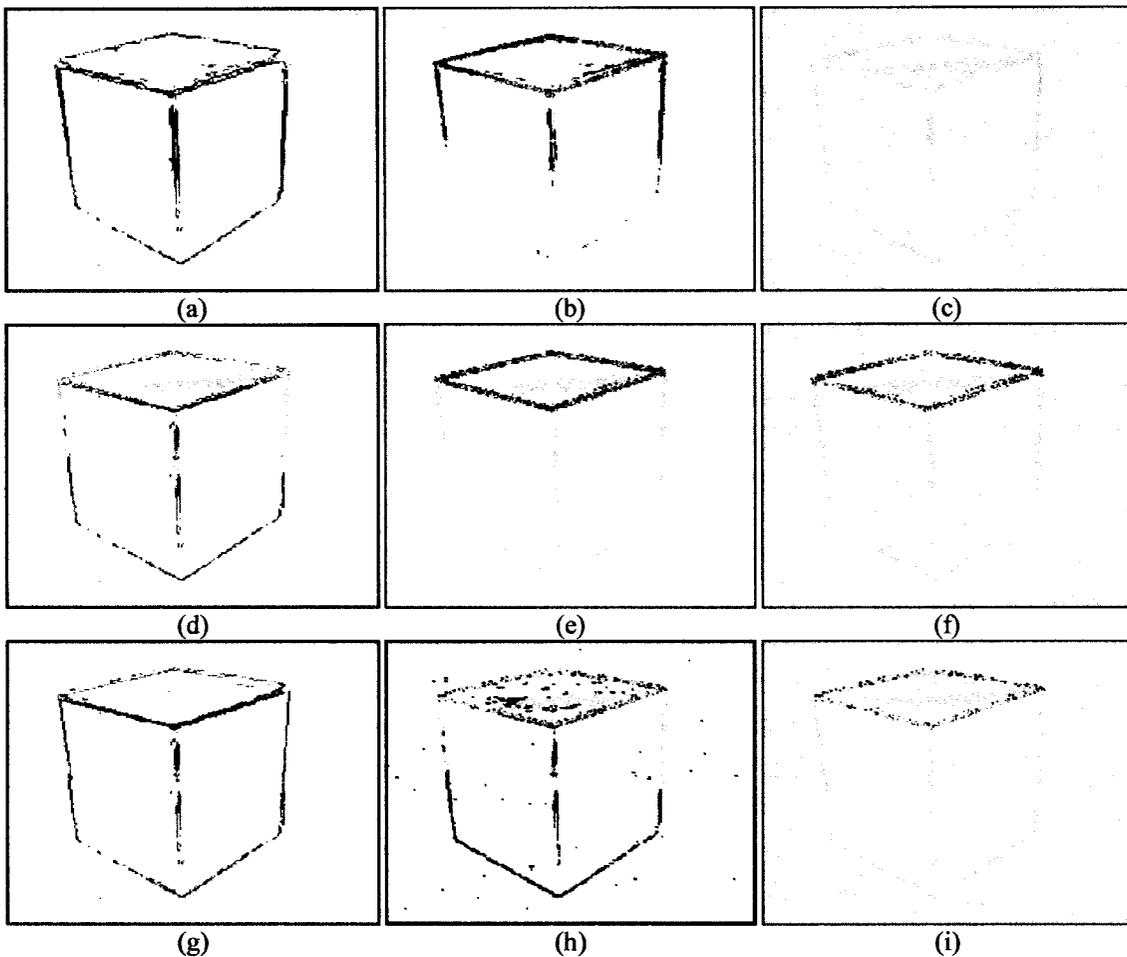


Figura 5.33. Combinações de bordas utilizando o coeficiente proporcional. (a)-(c): Combinação baseada na classe *geral*; (d)-(f): Combinação baseada na classe *laboratório*; (g)-(i): Combinação baseada na classe *cubos*. Para cada tríade foram utilizadas as proporções  $p$  iguais a 0, 0,2 e 1. Todas as bordas consideram o atributo  $c$ .

bordas principais da placa são definidas pela classificação por  $k$  vizinhos (ciano) e bayesiana (verde). Algumas outras técnicas podem ter sua bordas observadas em pontos dispersos, principalmente na Figura 5.32(b), na região das letras.

Todas combinações apresentadas até agora foram geradas através dos coeficientes vinculados a todas as imagens da base, ou seja, considerando-se a classe *geral*. A utilização das classes para o cálculo dos coeficientes está ilustrada, para a imagem *Cube011*, na Figura 5.33. As Figuras 5.33(a) a (c) utilizam a classe *geral*, enquanto as Figuras 5.33(d) a (f) consideram a classe *laboratório* e as Figuras 5.33(g) a (i) são relativas à classe *cubos*. Cada grupo de mapas de bordas utiliza, para as três combinações, proporções de 0, 0,2 e 1, respectivamente. Muitas diferenças na composição das bordas podem ser observadas, bem como as características gerais dos mapas formados. Novamente, mapas gerados com proporções intermediárias tendem a possibilitar uma manipulação mais fina das bordas e proporções que favorecem muito as coincidências geram, usualmente, excesso de bordas. Entre as três tríades de mapas apresentadas, pode-se notar, para a primeira coluna da Figura 5.33, que a precisão das

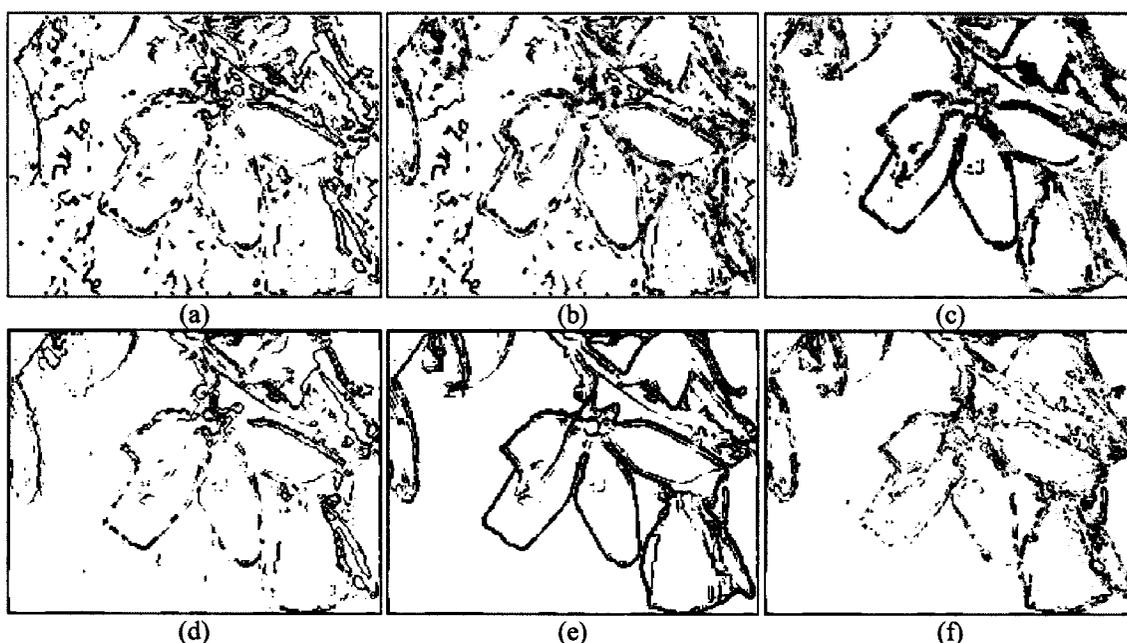


Figura 5.34. Combinações de bordas utilizando o coeficiente proporcional. (a)-(c): Combinação baseada na classe *geral*; (d)-(f): Combinação baseada na classe *externas*. Para cada tríade foram utilizadas as proporções  $p$  iguais a 0, 0,2 e 0,4. Todas as bordas consideram o atributo  $y_b$ .

bordas é maior quando as classes são mais restritas (*laboratório* e *cubos*), enquanto, para a classe *geral*, as bordas da face superior do cubo apresentam certa ondulação (causada pelas técnicas voltadas a bordas, particularmente por  $k$  vizinhos).

Outro exemplo de combinação para a imagem `Flower02` é apresentado na Figura 5.34. Utilizando o atributo  $y_b$ , as Figuras 5.34(a) a (c) consideram a classe *geral* e as Figuras 5.34(d) a (f) utilizam os coeficientes proporcionais calculados para a classe *externas*. Para estes mapas, observam-se as várias diferenças nas composições das bordas não somente para as proporções distintas, mas também entre ambas as classes.

Através da experimentação para várias imagens e vários atributos, a escolha de um valor ótimo para a proporção que maximize as coincidências ao mesmo tempo que minimiza os excessos, entretanto, não é uma tarefa simples. Há uma tendência para que este valor de  $p$  fique entre 0,2 e 0,5 para o coeficiente acumulativo e entre 0,1 e 0,2 para o coeficiente proporcional, para a maioria da imagens da base.

## 5.6 Comentários finais

O estudo comparativo proposto e analisado neste capítulo consiste, de um primeiro ponto de vista, em um modelo de integração e comparação do desempenho de técnicas de segmentação através dos mapas de bordas gerados. Sob um segundo ângulo, estes desempenhos relativos podem ser utilizados para que o mapa de bordas seja produzido aproveitando-se resultados de técnicas diferentes de forma seletiva. Estes dois enfoques são discutidos nas seções seqüentes.

### 5.6.1 Comentários sobre a comparação das técnicas de segmentação

Algumas suposições iniciais, elaboradas intuitivamente, foram comprovadas pelos experimentos realizados. Assim, puderam ser confirmadas as expectativas que a intensidade luminosa, enquanto componente da informação da imagem, é um atributo valioso na detecção de bordas. Além disto, técnicas diretamente voltadas à detecção de discontinuidades (isto é, bordas) obtiveram desempenho superior, em relação às coincidências, às técnicas voltadas a regiões, das quais o mapa de bordas não era o objetivo final.

A importância da luminosidade é evidenciada pelos desempenhos das segmentações por Sobel (com limiares manual, adaptável global e adaptável local), e do mapa auto-organizável utilizando a intensidade luminosa, associando-a ou não à cromaticidade. Estas técnicas, embora não obtivessem sistematicamente os melhores desempenhos para todos os casos, apresentam-se sempre entre as curvas de maior coeficiente para quando o coeficiente beneficia as taxas de coincidência. Em contraposição, estas mesmas técnicas tendem a garantir uma boa taxa de coincidência pela detecção “exagerada” de bordas. Em outras palavras, as taxas de coincidência mais altas são obtidas determinando-se vários pontos da imagem como sendo pontos de bordas, além dos que a segmentação manual (referência) qualifica como mais corretos. Esta situação é mais evidente para o caso do Sobel com limiar adaptável local, técnica que, em geral, apresenta os melhores desempenhos para  $p$  acima de 0,5, seja para o coeficiente acumulativo ou para o proporcional. O comprometimento do mapa pela detecção de bordas em demasia é facilmente observado pela análise visual do resultado da desta forma de segmentação. As imagens *Aerial02*, *Cube011*, *Cube015*, *Flowers* e *Lenna* demonstram claramente esta afirmação, como pode ser observado no Apêndice D.

Além desta técnica em particular, outras, como mapas auto-organizáveis com cromaticidade, iluminação ou com os dois atributos simultaneamente, além da classificação bayesiana e por  $k$  vizinhos, apresentam, esporadicamente, excessos nas detecções das bordas. Estes excessos de bordas, entretanto, acabam diluídos entre os vários mapas de bordas, o que acabou por permitir qualificar algumas destas técnicas como as que apresentam menor taxas de excesso. Quanto maior o número de imagens analisadas, maior a probabilidade, portanto, de que resultados eventualmente discrepantes afetem o comportamento geral das técnicas.

Considerando-se as técnicas individualmente para um determinado valor fixo de  $p$ , é também interessante notar que algumas delas prevalecem sobre as outras em determinada regiões dos intervalos das variâncias. Estas características se mantêm para pequenas variações de  $p$ . Para variações maiores do parâmetro de combinação, como já comentado, há a inversão dos desempenhos das técnicas. Entretanto, grupos de técnicas apresentam comportamento semelhante, tendendo a ter suas curvas dos coeficiente variando juntamente.

Em relação aos dois coeficientes definidos, o acumulativo e o proporcional, ambos revelam um comportamento similar, de forma que uma análise mais geral pode

optar por um ou outro indistintamente. Para uma consideração mais específica, entretanto, o coeficiente proporcional evita a característica “ou tudo ou nada” que o acumulativo utiliza a cada imagem. Por não utilizar somente o melhor método a cada etapa, métodos com desempenho muito próximos ao melhor, mas sistematicamente abaixo dele, não são sumariamente desconsiderados. Ao invés disto, suas taxas (coincidência ou excesso), são consideradas em relação à sua média. Estas diferenças refletem-se para valores de  $p$  na situação de transição entre os extremos da coincidência e do excesso. Embora a análise visual dos gráficos apresentados não permita identificar estas variações com facilidade em um nível mais detalhado das curvas, esta situação reflete-se na maior suavidade das curvas do coeficiente proporcional em relação ao acumulativo. Esta influência reflete-se, também, na combinação das bordas proposta na seção 5.4.

Apesar de ter se tentado manter um número mínimo de imagens para caracterizar uma classe de imagens, uma delas acabou por contar com apenas cinco imagens (classe *cubos*). A expectativa de que, para classes com pouca representação em termos numéricos, pudesse haver alguma tendenciosidade que favorecesse os resultados não pôde, porém, ser comprovada ou refutada.

Finalmente, relativamente ao conjunto de atributos locais selecionado, notou-se que a tendência geral dos métodos acabou por independer do tipo de variância. Contudo, variações locais (nos intervalos das variâncias) podem ser observados.

### 5.6.2 Comentários sobre a composição seletiva de bordas

A utilização da análise das bordas pelo modelo apresentado proporcionou uma ferramenta bastante interessante para a composição de um mapa de bordas que selecionasse, com base nos atributos locais, qual a técnica mais apropriada para determinadas regiões da imagem.

Os resultados apresentados, embora não completamente extensivos, puderam comprovar a eficiência da seleção na formação de mapas de bordas segundo um critério que oscila entre o máximo de acerto (derivado a partir da taxa de coincidência) e o mínimo de erro (taxa de excesso). Como observado nos exemplos e comentado na seção anterior, determinar um ponto de equilíbrio entre estes extremos é necessário, apesar de ainda não se dispor de um mecanismo mais objetivo que indique o melhor valor para  $p$ , o qual pode assumir valores dependentes de uma aplicação específica. Experimentalmente é viável, entretanto, escolher o valor da proporção da combinação não somente pela visualização do mapa de bordas, mas também pela observação das taxas de coincidência e excesso que este mapa combinado produz em relação à imagem de referência, quando disponível.

Os atributos locais mostraram-se mais importantes na composição das bordas do que na análise geral dos comportamentos, influenciando de forma diferenciada os mapas compostos (em particular as Figuras 5.28(g) a (i), por exemplo). Os atributos cromáticos RG e YB se destacaram em relação às variâncias  $g$  e  $c$  na diferenciação das composições de bordas.

## 6 Integração entre módulos: cor e estéreo

A visão, como abordada no Capítulo 2, não é somente um mecanismo de interpretação da informação visual, mas um sistema integrado no qual muitos aspectos da informação que chega aos olhos são considerados. Muitos são os atributos visuais, que incorporam forma, cor, movimento, padrões, texturas e intensidades, entre outros [Goldstein (1989)]. O processamento, por outro lado, não é somente combinado, mas tratado separadamente e integrado para uma percepção global daquilo que se vê [Zeki & Shipp (1988); Tovée (1996)]. Esta separação de processamento, sua modularização e a cooperação e integração dos resultados é uma das bases para o sistema de visão *Cyvis-1* [Costa *et al.* (1994)].

A proposta de como atributos distintos podem cooperar para atingir um objetivo determinado é o assunto exposto neste capítulo. Especificamente, dois subsistemas são considerados: cor e estéreo. Cada subsistema, ou módulo, estabelece metas diferentes de resultados. O atributo da cor é estipulado, neste contexto, para a segmentação das imagens adquiridas, proporcionando uma representação mais primária das informações da cena, o que inclui regiões e suas bordas. O módulo de estéreo, por sua vez, utiliza estas informações de nível mais baixo para determinar uma representação mais abstrata da imagem. Não apenas bordas são representadas, mas também informação de profundidade, isto é, incorpora-se a tridimensionalidade à representação.

Baseado nesta representação espacial dos dados, um mecanismo para o reconhecimento de objetos poliédricos é proposto, o qual passa a representar, desta forma, um conhecimento de nível mais alto em relação à cena.

A organização do sistema de visão que abrange tais características é descrita na seção 6.1. A seção 6.1.1 descreve a comunicação entre os módulos, os quais são estruturados como um sistema distribuído, e a seção 6.1.2 caracteriza e descreve cada módulo envolvido, apresentando suas principais particularidades de implementação. A seção 6.2 discorre sobre a forma de validação dos resultados, enquanto a seção 6.3

apresenta os principais resultados obtidos. As conclusões da abordagem proposta são apresentadas na seção 6.4.

## 6.1 Integração de modalidades

Em um sistema de visão como o *Cyvis-1*, a funcionalidade de seus componentes deve estar bem definida, determinando-se, de um ponto de vista mais geral para um mais específico, quais os módulos constituintes, suas interconexões e quais atribuições cada módulo deve assumir. É importante ressaltar que, como se trata de uma implementação inicial do sistema, alguns dos módulos foram estruturados para acumular funções de módulos ainda inexistentes.

### 6.1.1 Caracterização e estruturação básica da comunicação

O compartilhamento de informações dentro do sistema de visão *Cyvis-1* está definido para permitir estabelecer as formas de comunicação e os elementos envolvidos. Os principais aspectos envolvendo a comunicação entre módulos incluem:

- a) Definição da topologia do sistema, em termos de computadores e distribuição dos processos (módulos) e meios físicos de comunicação (rede local);
- b) Estruturação dos tipos de dados a serem compartilhados através dos mecanismos de comunicação;
- c) Definição dos protocolos que garantam a confiabilidade nas trocas de mensagens e dados;
- d) Efetivação da comunicação através da criação de um canal binário entre módulos.

A topologia adotada configura o esquema básico de um sistema distribuído, o qual incorpora características como processos (aproximadamente) independentes, executando possivelmente em computadores diferentes, cuja interconexão se estabelece através de uma rede local (esquema misto com segmentos usando cabo coaxial e segmentos usando par trançado). Os computadores em questão são microcomputadores da linha PC com processadores Pentium ou Pentium II e sistema operacional Microsoft Windows 95/98. Os módulos são implementados usando o compilador Borland Delphi, o qual proporciona rápido desenvolvimento para o ambiente Windows, com facilidades de composição de janelas, formulários etc., e incorpora orientação a objetos na linguagem Object Pascal.

A necessidade da estruturação adequada dos dados e dos protocolos é evidente para que a troca de informações se torne confiável. As informações a serem trocadas compõem dois grupos distintos: mensagens e dados. As mensagens compreendem cadeias de caracteres que estabelecem a forma básica da comunicação, enviadas e recebidas através de ligações baseadas no protocolo TCP/IP. São definidos códigos que indicam ações e estados, os quais os módulos compartilham uns com os outros. Os dados, por sua vez, incluem tanto informações internas, como coordenadas ou parâmetros que devam ser compartilhados, quanto imagens, as quais constituem o elemento principal do compartilhamento. Estes dados, reunindo em si usualmente grande volume de informação, são transmitidos através do protocolo IPX. Esta esco-

lha, embora envolva a utilização de arquivos para a troca de dados, foi determinada pelo desempenho do protocolo para a transmissão de arquivos de granulação alta [Bruno & Costa (1997)].

Mensagens envolvem estruturas simples que incluem requisição de serviços ou informações, abertura e fechamento de uma negociação de transferência e os processos internos da comunicação estabelecida. Os serviços compreendem o processamento de determinada informação por um módulo específico (extração de linhas retas ou a captura de uma nova imagem, por exemplo) e a devolução do resultado. As negociações são realizadas através de um canal permanente aberto entre módulos.

### 6.1.2 Estruturação geral do sistema

A Figura 6.1 apresenta um esquema simplificado do *Cyvis-1*, indicando seus módulos e principais elementos de controle, já apresentados na seção 1.1. Para esta proposta específica de implementação (cor e estéreo), alguns módulos, como o de textura, por exemplo, são desconsiderados, bem como alguns sistemas de controle são ignorados ou incorporados aos módulos existentes. Mantém-se, entretanto, a filosofia na qual o sistema como um todo se baseia.

A Figura 6.2 apresenta o diagrama de blocos do sistema estruturado para esta proposta. Destacam-se os cinco módulos implementados: aquisição de imagens, segmentação por cor, segmentação por detecção de bordas (níveis de cinza), análise de formas e reconhecimento de objetos.

Como se trata de um projeto que envolve cooperação com outras pessoas, é importante ressaltar que o desenvolvimento e a implementação do módulo de estéreo foi realizado por Alan Salvany Felinto, tendo sua participação sido intensa (e essencial) na validação do módulo de reconhecimento e no estabelecimento dos critérios de erro. A ferramenta básica de comunicação foi implementada por Odemir Bruno, que também participou das discussões até que se chegasse ao protocolo atualmente utilizado.

#### 6.1.2.1 Aquisição

O módulo de aquisição consiste em um programa independente e autônomo, o qual mantém controle sobre a aquisição de imagens através de câmera digital (CCD). As imagens são adquiridas aos pares, mantendo-se uma distância conhecida entre as posições que a câmera ocupa na ocasião da aquisição das imagens esquerda e direita, informação que é utilizada para o cálculo da profundidade no módulo de estéreo. As imagens são capturadas sem convergência, ou seja, os eixos para a captura são paralelos.

Uma vez iniciado, o módulo de aquisição se conecta à placa de aquisição de vídeo e entra em modo de espera, que consiste em aguardar uma conexão externa e solicitações de serviço. Neste caso específico, o serviço disponível caracteriza-se pela imagem adquirida via câmera.

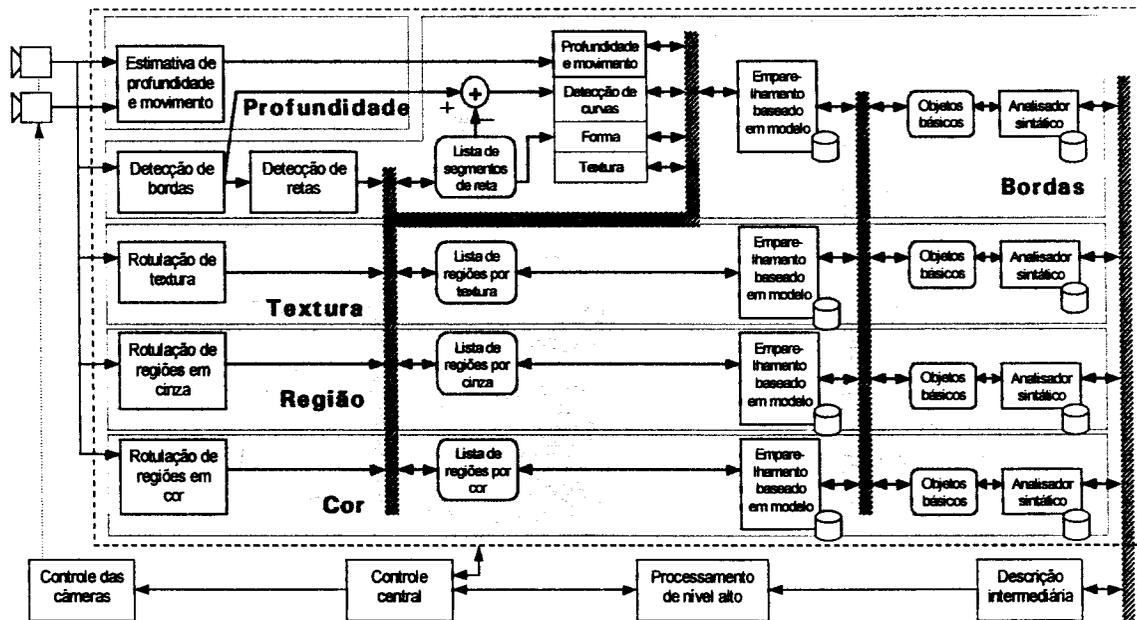


Figura 6.1. Esquema geral do sistema de visão *Cyvis-1*. (Adaptado de Costa *et al.* [Costa *et al.* (1994)].)

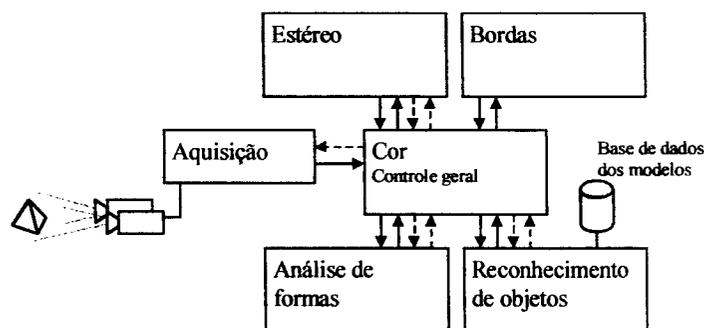


Figura 6.2. Esquema da estruturação entre os módulos para a integração entre estéreo e cor. Setas sólidas indicam troca de dados, enquanto as tracejadas indicam trocas de sinais de controle.

### 6.1.2.2 Cor

A segmentação de uma das imagens do par estéreo é realizada pelo módulo de cor, o qual emprega a técnica proposta no Capítulo 4. O mapa auto-organizável é treinado a partir dos dados de ambas as imagens e, após convergir, tem suas classes identificadas e a imagem esquerda do par tem seus pontos classificados para a segmentação por regiões. Em seguida as fronteiras entre as regiões são encontradas, gerando o mapa de bordas da imagem esquerda. A imagem direita não é segmentada devido ser desnecessário ao módulo de estéreo.

### 6.1.2.3 Análise de formas

A partir do mapa de bordas, que representa parte da informação contida na imagem, o módulo de análise de formas detecta as principais retas existentes, dado que estas informações são relevantes para o reconhecimento de objetos poliédricos.

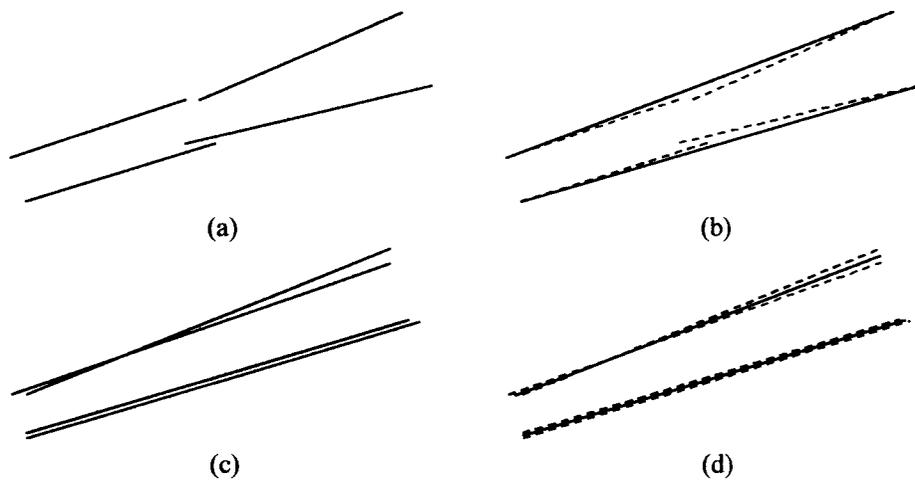


Figura 6.3. Representação da simplificação de segmentos de reta. (a) Dois pares de segmentos consecutivos e (b) sua simplificação para dois segmentos. (c) Dois pares de segmentos coincidentes e (d) sua simplificação. As retas sólidas indicam a posição de segmentos reais e as tracejadas a posição original dos segmentos que foram desconsiderados ou modificados.

A detecção das retas é realizada por meio da transformada de Hough [Hough (1962)] parametrizada nas coordenadas polares  $\rho$  e  $\theta$  (Equação 6.1), brevemente descrita no Apêndice C. Além da transformada, que gera uma representação paramétrica das retas da imagem, emprega-se a técnica de rastreamento (*backmapping*) [Gerig & Klein (1986)] para realçar as mais significativas. Além destes aspectos, é também utilizada a restrição da variação do parâmetro  $\theta$  ao intervalo  $[-\pi/2, \pi/2]$ .

$$\rho = x \cos\theta + y \sin\theta \quad (6.1)$$

Por esta abordagem, então, são reconhecidas as principais retas, as quais são verificadas sobre o mapa para determinar seus limites e determinar uma lista dos segmentos de reta presentes. Cada segmento é identificado por um número de ordem e as coordenadas  $(i, j)$  da imagem de ambas suas extremidades. Pequenas variações locais são toleradas nesta verificação das retas sobre as bordas, como tolerância a pequenas descontinuidades e variações na posição das bordas (um ponto acima ou abaixo, à esquerda ou à direita).

Alguns parâmetros são importantes para a detecção de retas, destacando-se a quantização do espaço de parâmetros e o modo de detecção dos máximos neste espaço de acumuladores. As dimensões do espaço de parâmetros determinam os passos em que os parâmetros  $\rho$  e  $\theta$  variam. Por exemplo, a utilização de 100 intervalos para  $\theta$  indicam discernimento do ângulo não inferior  $1,8^\circ$ . A quantização estabelecida para  $\rho$  pode ser considerada da mesma forma, lembrando que o máximo valor para esta dimensão do mapa corresponde ao comprimento da diagonal da imagem.

Um processamento adicional dos segmentos de reta ainda é utilizado para identificar se há segmentos consecutivos ou coincidentes [Costa & Sandler (1993)]. Um limite de distância entre as extremidades é utilizado para estabelecer uma tolerân-

cia para esta verificação. Segmentos consecutivos são os que apresentam inclinação e distância da origem bastante próximas, além de manterem uma distância pequena entre suas extremidades, como representado na Figura 6.3(a). Para estes segmentos são desconsideradas as extremidades mais próximas e um único segmento é criado unindo-se as extremidades restantes, como ilustrado na Figura 6.3(b). Segmentos considerados coincidentes, como os esquematizados Figura 6.3(c), são substituídos por um único segmento com extremidades nos pontos médios das extremidades dos segmentos originais, o que é ilustrado na Figura 6.3(d). Reduzem-se, assim, tanto a quantidade de segmentos quanto o número de segmentos truncados (no caso de segmentos consecutivos), uma característica interessante para o reconhecimento dos objetos.

#### 6.1.2.4 Estéreo

O cálculo das profundidades é realizado para cada extremidade dos segmentos de reta detectados na imagem, através das disparidades localizadas entre o par de imagens adquirido. Para que seja possível avaliar a disparidade, é necessário que sejam determinadas as posições em que cada extremidade de segmento está representada nas imagens. Estes pontos que identificam a mesma posição física nas duas imagens são chamados pontos conjugados ou de correspondência [Shirai (1987)].

Existem várias abordagens para resolver o problema de determinar os pontos de correspondência, as quais incluem correlação, diferenças quadráticas mínimas, transformadas de Gabor e programação dinâmica, além de outras [Theimer & Marlot (1994); Faugeras (1996); Qian (1997)]. Para esta implementação do sistema, optou-se pela utilização de janelas quadradas centradas nas posições dos pontos, pela qual é feita a busca do conjugado na outra imagem. A similaridade é medida através da correlação entre os conteúdos das janelas. Eventuais ambigüidades são contornadas através da mudança sistemática do tamanho da janela, como proposto por Shirai [Shirai (1987)]. As restrições às regiões cobertas para esta busca são determinadas pela linha epipolar\* e pela determinação da disparidade máxima entre os pontos. A correlação é importante por determinar, na maioria dos casos, a posição correta e mais precisa dos pontos correspondentes [Shirai (1987)], visto que os segmentos detectados tendem a ter as extremidades próximas aos vértices, o que as caracteriza apropriadamente.

A calibração das câmeras utiliza o procedimento proposto por Tsai [Tsai (1987)], o qual tem se mostrado apropriado à avaliação dos dados de profundidade.

---

\* A linha epipolar, calculada através da geometria das câmeras, restringe as posições na segunda imagem do par estéreo nas quais um ponto conjugado pode ser localizado. Dada a posição de um ponto, necessariamente seu conjugado deve estar ao longo da linha epipolar na segunda imagem.

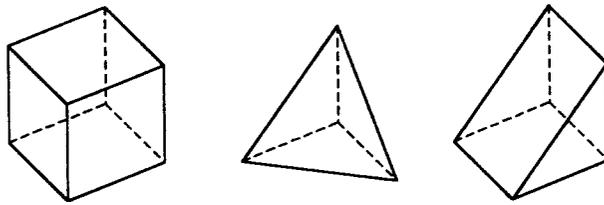


Figura 6.4. Segmentos que descrevem os modelos da base de dados.

#### 6.1.2.5 *Bordas*

Para a comparação dos resultados, um segundo critério para a extração do mapa de bordas é utilizado. Optou-se pela tradicional detecção de bordas em níveis de cinza, através dos operadores de Sobel e escolha manual do limiar na seleção das maiores magnitudes, uma alternativa que já vinha sendo utilizada nos cálculos de profundidade.

Este subsistema ainda não opera em modo contínuo com os outros já implementados. A segmentação é processada e os mapas de bordas gerados armazenados em arquivos. A recuperação destes arquivos é feita posteriormente pelo módulo de cor.

#### 6.1.2.6 *Reconhecimento de objetos*

O reconhecimento de objetos faz uso de uma base de dados, ainda bastante limitada, de modelos de objetos poliédricos. São considerados objetos como cubo e pirâmides, como os apresentados na Figura 6.4. Cada objeto é descrito por um conjunto de segmentos de reta com extremidades especificadas por coordenadas tridimensionais. A escolha dos objetos, embora introduza uma simplificação, foi motivada pela necessidade da precisão nas medidas comparativas, uma vez que as distâncias calculadas podem, deste modo, ser confrontadas com as medidas reais dos objetos. Aferindo-se os parâmetros para objetos conhecidos, outras medidas de distância podem ser, então, consideradas. Cada objeto real é construído em alumínio e apresenta dimensões conhecidas, que compreendem aspectos importantes na avaliação dos resultados dos cálculos das profundidades, embora estas considerações estejam fora do escopo deste texto.

O conjunto de segmentos extraídos das imagens é comparado a cada modelo da base, visando identificar o objeto que mais se aproxima do conteúdo da cena. A verificação é feita através de uma comparação usando “força bruta”, ou seja, uma verificação exaustiva é feita procurando-se determinar o melhor emparelhamento dos dados. A escala das descrições dos objetos (do modelo e da imagem) não é relevante, dado que uma normalização dos segmentos é realizada antes do início da comparação. O emparelhamento inicial é feito entre dois segmentos (ambos normalizados para

comprimento  $l$ ) e, através de rotações e translações rígidas\*, os demais segmentos são posicionados e um erro geral do emparelhamento é calculado. É estabelecido um limite, expresso por uma porcentagem do comprimento do primeiro segmento emparelhado, que fornece o grau de tolerância para que uma extremidade de um segmento seja considerada próxima suficiente de outra. Caso ambas as extremidades de um segmento estejam dentro desta tolerância, o emparelhamento dos segmentos pode ser considerado. Caso contrário, há a rejeição e combinações com outros segmentos são avaliadas. O símbolo  $\tau$  é utilizado para especificar o parâmetro da tolerância.

Além destas considerações, o número de segmentos provenientes da imagem é, usualmente, superior ao que representa os objetos da base de dados. Muitos destes segmentos são oriundos de outros objetos, de retas detectadas no fundo, ou mesmo são segmentos incorretamente detectados. Quando segmentos encontrados na imagem não podem ser associados a nenhuma das arestas dos modelos, são apenas desconsiderados, assim como o erro que introduziriam na avaliação do desempenho. Arestas do modelo, entretanto, não podem ser desconsideradas e, portanto, são incluídas nas considerações sobre o reconhecimento relativas ao cálculo do erro. O erro considerado é descrito em detalhes na seção 6.2.

#### 6.1.2.7 *Controle central*

O controle central do sistema é responsável pela coordenação dos módulos e controle do fluxo de informações. Por questões de implementação, o controle geral foi incorporado ao mesmo programa que realiza a segmentação por cor, embora continue, do ponto de vista da organização, sendo uma unidade independente do sistema. Assim, o controle central opera o módulo de cor para solicitar imagens à câmera, as repassa para os outros módulos e intermedia, também, as outras informações geradas, como listas de segmentos de reta bidimensionais e tridimensionais. Controla, desta forma, o tempo e ordem de ativação dos demais módulos, além de verificar a interdependência dos dados. Promove, assim, hierarquia e integração dos módulos.

## 6.2 **Descrição do erro do emparelhamento**

O critério de erro para a avaliação dos resultados do reconhecimento foi estabelecido como o erro médio quadrático entre as distâncias das extremidades correspondentes de todos os segmentos emparelhados do objeto da imagem e do modelo da base de dados. A Equação 6.2 mostra o cálculo do erro, sendo que  $N$  indica o número de extremidades de segmentos,  $o_i$  e  $m_i$  representam os pontos conjugados no espaço do objeto da imagem e do modelo da base, respectivamente, e  $D(\cdot)$  é a função de distância euclidiana.

---

\* Transformações rígidas não permitem modificação da escala de forma diferenciada nos três eixos, ou seja, não introduzem distorções nos objetos.

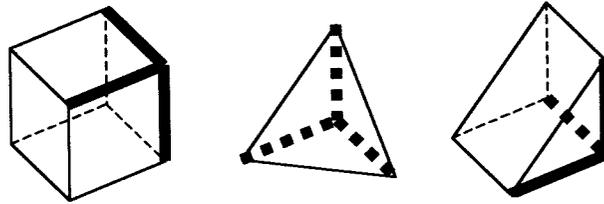


Figura 6.5. Três segmentos, em destaque, emparelhados com as arestas dos modelos.

$$\varepsilon_q = \frac{1}{N} \sum_{i=1}^N [D(o_i, m_i)]^2 \quad (6.2)$$

Permite-se, assim, estabelecer um critério objetivo de comparação dos resultados de emparelhamentos distintos.

Há, entretanto, um outro problema em relação ao reconhecimento que é relativo à similaridade entre os objetos da base. Como pode ser observado nos objetos da Figura 6.4, o prisma (à direita) é formado por um corte do cubo por um plano que passa por arestas opostas; da mesma forma o tetraedro (no centro) é formado por um corte do cubo por um plano que passa por três vértices não compartilhados pela mesma face e que não contenha uma diagonal. Desta forma, quando são detectados na imagem somente três segmentos ortogonais e com uma extremidade em comum, o emparelhamento pode ser feito tanto com o cubo quanto com o prisma ou o tetraedro, como indicado pelos segmentos destacados na Figura 6.5. Para estes casos, o erro médio quadrático  $\varepsilon_q$  produziria o mesmo resultado. Então, com a finalidade de ajustar o reconhecimento para um dos objetos, introduziu-se um erro que penaliza o número de arestas visíveis que não foram consideradas no emparelhamento.

Para o exemplo dos três segmentos, torna-se bastante razoável considerar que é mais provável que alguns segmentos foram perdidos e que se trata de um tetraedro o objeto da imagem. Seria, neste enfoque, menos provável que fosse um prisma ou um cubo, pois é menos provável, também, que tenham sido perdidas muitas bordas (embora não necessariamente). Opta-se, segundo estas considerações, por reconhecer o objeto mais simples. Considerar o número máximo de arestas visíveis ao invés do número total de arestas remove o peso excessivo que seria atribuído ao cubo, por exemplo, do qual três arestas estão permanentemente ocultas e, portanto, não podem ser detectadas. Para os objetos da Figura 6.4, desta forma, são considerados 9, 6 e 8 o número máximo de arestas visíveis, respectivamente.

A Equação 6.3 expressa o erro relativo ao emparelhamento de bordas, com  $N_v$  e  $N_o$  indicando, respectivamente, o número de arestas visíveis do modelo e o número de arestas (segmentos) da imagem que foram emparelhados. O fator  $k$  é um fator de correção da magnitude do erro, para que tenha a mesma variação máxima que o erro quadrático médio  $\varepsilon_q$ . O erro  $\varepsilon_e$  estabelece a razão entre a falha no emparelhamento esperado (número de arestas visíveis não relacionadas a arestas da imagem) e a expectativa inicial (número de arestas visíveis).

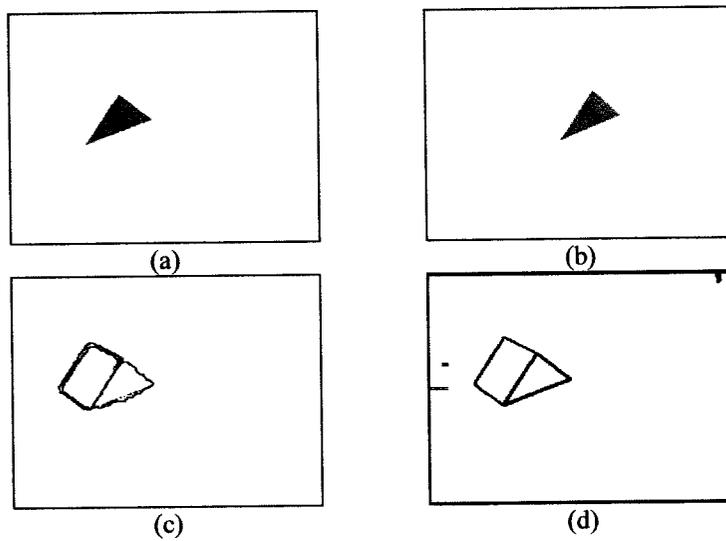


Figura 6.6. Exemplo de um prisma. (a) e (b): Par de imagens em estéreo; (c) Imagem em (a) segmentada por mapas auto-organizáveis utilizando cromaticidade; (d) imagem em (a) segmentada por Sobel.

$$\varepsilon_e = k \frac{N_v - N_o}{N_v} \quad (6.3)$$

Ressalta-se, ainda que o número de arestas visíveis é tomado em relação a um ponto de vista arbitrário, pelo qual é possível observar o maior número possível. Situações circunstanciais de pontos de vistas diferentes não são incluídas no erro.

O erro total do emparelhamento, por sua vez, consiste na ponderação entre os dois erros definidos, como apresenta a Equação 6.4. O fator  $h$  representa a ponderação entre os erros que permite ao erro total discernir corretamente entre os modelos da base.

$$\varepsilon = h\varepsilon_q + (1-h)\varepsilon_e \quad (6.4)$$

O erro  $\varepsilon$ , portanto, é o erro que permite distinguir, para um dado conjunto de segmentos obtidos da imagem, a qual modelo da base que mais se assemelha. Considera, assim, o número de arestas emparelhadas, que reflete o grau de similaridade com o modelo, e o erro quadrático médio, que representa a “qualidade” deste emparelhamento.

Para a análise comparativa entre Sobel e o mapa auto-organizável, adota-se somente a qualidade do emparelhamento, representada por  $\varepsilon_q$ , levando-se em conta que o reconhecimento do objeto já foi efetuado.

### 6.3 Resultados e discussão

As Figuras 6.6(a) e (b) ilustram um par de imagens em estéreo obtida em laboratório. Na Figura 6.6(c) está apresentado o mapa de bordas da imagem direita\* do par, obtido através da segmentação por mapas auto-organizáveis. O mapa de bordas obtido por Sobel está apresentado na Figura 6.6(d). Notam-se dois mapas de bordas com bons resultados, destacando-se a melhor precisão das bordas por Sobel. Estas imagens representam os objetos e imagens típicas utilizadas nos experimentos.

Uma característica das imagens é as dimensões que os objetos assumem, que acabam não ocupando grande parte da imagem. Este fato é consequência das dimensões dos objetos e da sua distância da câmera, o que é exigido por duas razões principais: calibração e enquadramento. A calibração é feita considerando-se o foco da câmera, o qual é dependente da distância dos objetos [Shirai (1987)]. Relativamente ao enquadramento, é necessário que se tenha uma certa margem de segurança para que os pontos importantes (para os quais há uma expectativa que sejam considerados pontos conjugados) permaneçam em áreas da imagem na qual é feita a comparação por correlação. A primeira razão, entretanto, é a mais restritiva, visto que a incorreção da calibração induz distorções às vezes grosseiras na profundidade calculada.

A Figura 6.7 mostra várias imagens dos objetos, que correspondem à imagem direita do par estéreo, além de suas segmentações por mapa auto-organizáveis e por Sobel. Na primeira coluna estão as imagens originais, na segunda estão as segmentações por mapas auto-organizáveis e, na terceira, por Sobel.

Notam-se as diferenças básicas nos resultados da segmentação, como bordas produzidas por Sobel mais definidas, do ponto de vista de apresentarem menos ondulações que as produzidas pelo mapa auto-organizável. Por outro lado, os mapas de bordas que utilizam a cromaticidade tendem a identificar uma quantidade muito menor de pontos de bordas não significativos, produzidos usualmente por diferenças de iluminação ou na transição entre partes do fundo da imagem com cores similares (por exemplo, entre a mesa e a parede). As Figuras 6.7(h) e (i) e as Figuras 6.7(k) e (l) ilustram bem estas situações.

Os resultados produzidos para a segmentação por mapas auto-organizáveis foram estabelecidos para o conjunto de parâmetros discutidos no Capítulo 4. Para estes parâmetros, a segmentação apresenta pouca variação entre as aplicações e identifica, de forma geral, as classes da forma esperada, gerando mapas de bordas satisfatórios do ponto de vista de uma análise visual. São utilizados, sem alterações para o conjunto de experimentos realizados, mapas de dimensões  $32 \times 32$ , vizinhança circular inicial com raio 25,6 (equivalente a 80% da dimensão do mapa), taxa de aprendizado inicial de 60% e taxas de decaimento de 0,05% tanto para a taxa de aprendizado

---

\* As referências direita e esquerda para o par estéreo são relativas às posições da câmera, e não às posições em que são apresentadas no texto.

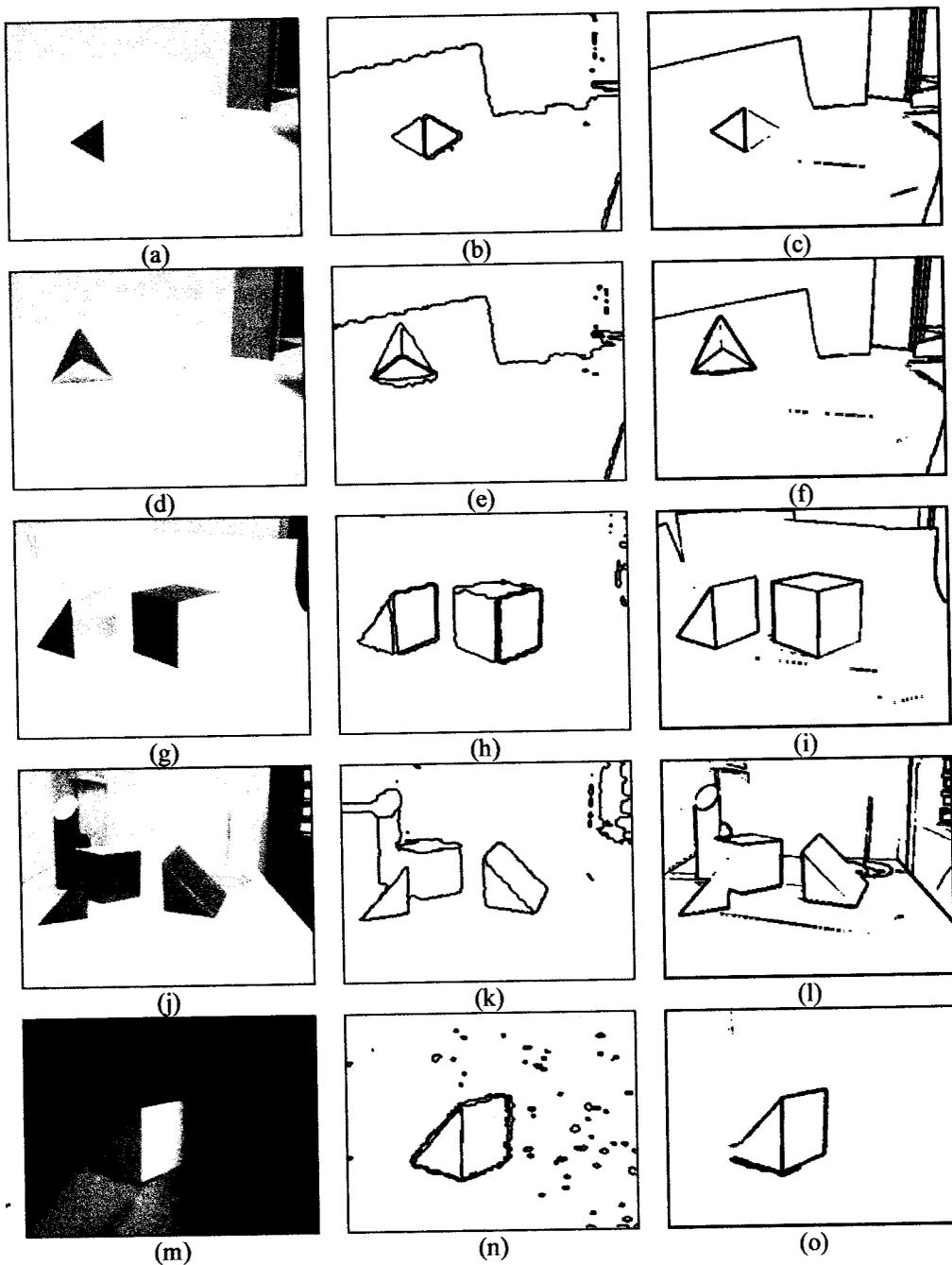


Figura 6.7. Imagens dos objetos e suas segmentações por mapas auto-organizáveis e por Sobel. (a), (d), (g), (j) e (m): Imagens direitas do par estéreo. (b), (e), (h), (k) e (n): Mapas de bordas por mapas auto-organizáveis. (c), (f), (i), (l) e (o): Mapas de bordas por Sobel.

como para o raio da vizinhança. O erro adotado para a verificação da convergência foi estipulado em  $10^{-4}$ .

Uma vez obtidos os mapas de bordas por ambas as técnicas, cada um pode ser submetido ao módulo de análise de formas, que implementa a transformada de Hough para a detecção dos segmentos de reta. Na Figura 6.8 estão representados os segmentos detectados em alguns dos mapas de bordas apresentados na Figura 6.7. Pode ser observada, em algumas áreas, a detecção de vários segmentos (áreas com mais

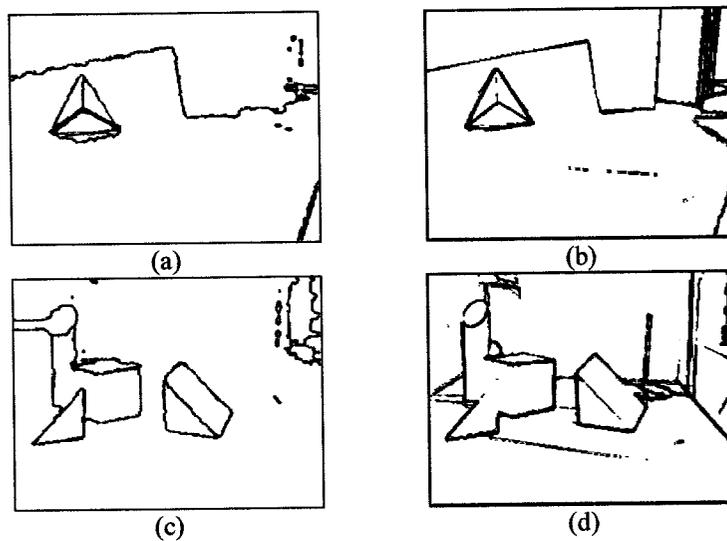


Figura 6.8. Segmentos de reta detectados nos mapas de bordas, apresentados. São considerados os seguintes mapas: (a) Figura 6.7(e); (b) Figura 6.7(f); (c) Figura 6.7(k); e (d) Figura 6.7(l).

retas vermelhas) sobre uma mesma borda, o que é devido em geral, a sua espessura. Vários destes segmentos já foram, ainda, resultados da fusão de outros segmentos através busca de segmentos consecutivos ou coincidentes.

Nota-se também que o excesso de pontos de borda (sejam os causados por falhas ou características da técnica ou ainda pelo grande número de bordas efetivamente presentes na imagem), através do processo de rastreamento\*, acaba por desconsiderar alguns dos segmentos, como pode ser notado na Figura 6.8(d), para as bordas determinadas por Sobel. Uma das arestas do prisma e uma aresta do cubo não foram detectadas, embora o mapa de bordas as representasse.

Outra dificuldade se apresenta quando há bordas mais grossas, para as quais, próximos aos vértices dos objetos há um tendência a inclinações não desejadas. A Figura 6.9 apresenta um esquema que ilustra esta ocorrência, na qual os segmentos mais finos indicam a aresta na qual a borda foi detectada. A reta tracejada seria uma reta excelente para representar a posição do segmento da aresta, que tende a ser inclinada se as bordas tiverem espessura grande. A reta indicada com a linha sólida cobre um número maior de pontos na imagem e representa, sob esta perspectiva, uma reta mais significativa. Há, assim, em posições da imagem que apresentam mais ruídos ou excesso de bordas, uma chance maior de haver certa inclinação das bordas. A influência deste desvio se reflete no erro do reconhecimento, mas caso a inclinação não seja muito significativa, o desvio afeta em menor intensidade o reconhecimento do objeto,

\* O processo de rastreamento incrementa os acumuladores no espaço de parâmetros somente quando correspondem a "boas" retas da transformada. Como a detecção das retas é feita localizando-se pontos de máximo local no espaço de parâmetros, ocorrem casos em que os acumuladores não caracterizam adequadamente máximos locais. Há esta situação, por exemplo, quando pontos vizinhos no espaço de parâmetros compartilham um mesmo valor máximo.

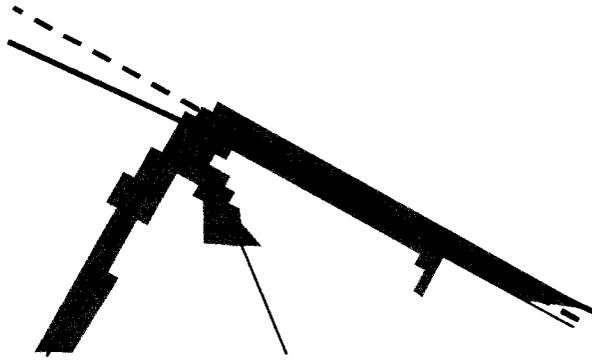


Figura 6.9. Esquema de uma borda com certa espessura em um vértice da imagem. Destaca-se, tracejada, a reta desejada, que corresponde à melhor orientação para determinar a aresta. A reta sólida indica a orientação da reta obtida por Hough.

dado que as extremidades do segmento não tendem a ficar muito distantes da posição esperada.

O resultado do módulo de análise de formas, que é retornado ao módulo de controle (cor), é repassado para o módulo de estéreo, juntamente com o par de imagens. Este último determina, para cada extremidade de segmento de reta encontrado, o valor da profundidade correspondente, verificando os pontos conjugados nas duas imagens. Em desenvolvimento por Alan Salvany Felinto, o módulo utiliza as imagens em cinza para o cálculo da correlação na determinação das correspondências entre pontos. Estão sendo discutidas, entretanto, alternativas para que seja possível, também, utilizar informações de cor neste processo (inserindo cromaticidade ou as componentes RGB no cálculo da correlação, por exemplo).

A determinação da profundidade pela disparidade dos pontos não é sempre possível de ser avaliada, sendo os principais fatores as diferenças de perspectiva, que causam oclusões parciais ou introduzem reflexos diferenciados, e a imprecisão dos segmentos. Este segundo fator, mais importante no contexto deste trabalho, ocorre em função da presença de sombras (no caso do Sobel, principalmente) ou reflexos das cores (no caso do mapa auto-organizável). Estas situações fazem com que as bordas encontradas não fiquem precisamente sobre os limites reais dos objetos, o que faz com que alguns segmentos tenham suas extremidades fora do objeto. Em outras palavras, o segmento pode começar ou terminar sobre a mesa ou sobre a parede do fundo. O valor calculado para a profundidade se torna, então, relativo à mesa ou à parede, o que pode destoar bastante das dimensões e posição do objeto de interesse.

De uma forma geral, pontos para os quais não são encontrados seus conjugados são desconsiderados, causando a exclusão de todo o segmento da lista. Os demais segmentos são mantidos para verificação, mas como usualmente não é possível o emparelhamento de todos com os segmentos dos modelos da base de dados, os “incorretos” também acabam desconsiderados. Isto acontece com segmentos encontrados em outras posições da imagem, como fundo, mesa ou eventuais sombras que não se confundem as bordas do poliedro, além de outros segmentos detectados nas circunvizinhanças do objeto.

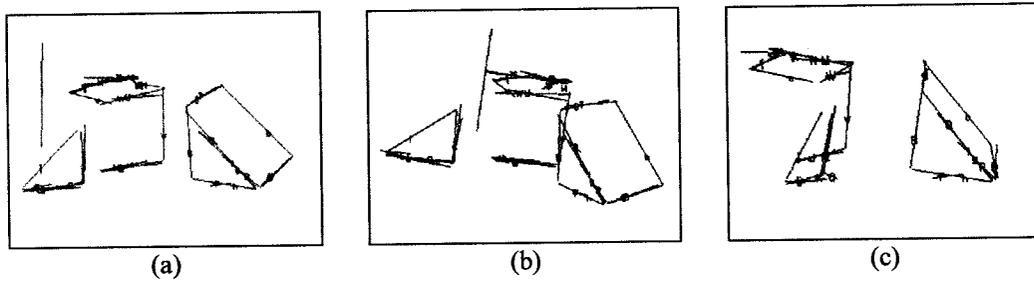


Figura 6.10. Segmentos detectados na imagem da Figura 6.7(j), utilizando a segmentação por mapas auto-organizáveis da Figura 6.8(c), sob três pontos de vista diferentes.

A última etapa do processamento é a confrontação da descrição do conteúdo da imagem especificada pelos segmentos de reta no espaço com as descrições da base de dados. Para esta verificação, o parâmetro  $k$  da Equação 6.3 deve ser definido e, baseado nos valores típicos do número de arestas visíveis na base de dados, adotou-se o valor 0,02. Para este valor específico de  $k$ , o valor de  $\varepsilon_e$  assume aproximadamente a mesma magnitude máxima do erro quadrático médio  $\varepsilon_q$ , tornando possível balancear com maior facilidade o erro geral de emparelhamento introduzido pela Equação 6.4. Considerando-se os erros  $\varepsilon_q$  produzidos experimentalmente, foi adotado o valor 0,3 para o balanceamento estabelecido por  $h$ , o qual favorece mais o número de arestas que encontraram correspondências na base. Considerando-se que há um erro máximo tolerável para que se tome a decisão que um emparelhamento entre dois segmentos é aceitável, o critério mais importante para a decisão de rotular o objeto da cena torna-se, necessariamente, o ajuste das arestas.

O parâmetro  $\tau$ , introduzido na seção 6.1.2.6, é usado para indicar esta tolerância no emparelhamento. Considerando-se que a proposta é a avaliação dos erros do reconhecimento dos objetos e que um certo grau de precisão é exigido, a utilização de um valor de  $\tau$  igual a 2% é adotada. Na prática, os segmentos iniciais são normalizados antes dos emparelhamentos das demais arestas, o que leva pontos que distem mais que 0,02 unidade de distância a não serem considerados próximos para o emparelhamento. É esta restrição que fornece, em última instância, a magnitude máxima para  $\varepsilon_q$ .

A Figura 6.11 apresenta alguns resultados do reconhecimento. Na primeira coluna são apresentadas as imagens direitas do par estéreo, na segunda estão as superposições dos segmentos de reta detectados com o modelo reconhecido, considerando-se a utilização do mapa de bordas gerado pela segmentação por mapas auto-organizáveis e, na terceira, as superposições entre segmentos reconhecidos e modelo relativos à segmentação por Sobel. Considerando-se os resultados obtidos pela segmentação pelo mapa auto-organizável, o cubo da Figura 6.11(b) foi reconhecido através da identificação de 6 arestas, as quais produziram um erro de emparelhamento  $\varepsilon = 7,0 \times 10^{-3}$  e um erro de qualidade  $\varepsilon_q = 1,29 \times 10^{-2}$ . O cubo reconhecido com o mapa de bordas por Sobel, apresentado na Figura 6.11(c), foi identificado por 7 arestas, com erros  $\varepsilon = 4,6 \times 10^{-3}$  e  $\varepsilon_q = 5,1 \times 10^{-3}$ . Visualmente o melhor desempenho da abor-

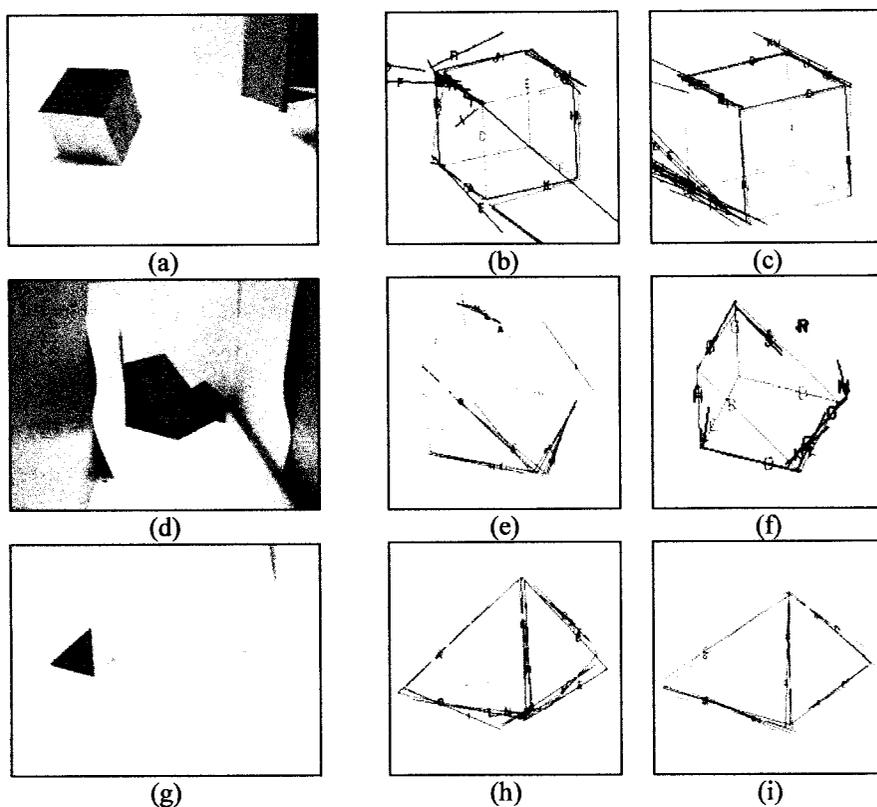


Figura 6.11. Exemplos de reconhecimento de objetos com superposição dos segmentos detectados ao modelo reconhecido. (a), (d) e (g): Imagens direitas do par estéreo; (b), (e) e (h): superposição para mapas auto-organizáveis e (c), (f) e (i): superposição para Sobel. Para os segmentos do objeto, a cor magenta indica os que foram emparelhados com o modelo e a cor preta os restantes. Para o modelo, a cor azul representa as arestas emparelhadas e a cor ocre aquelas que não foram utilizadas.

dagem por Sobel também pode ser notado. Nas Figuras 6.11(e) e (f) são apresentados os reconhecimentos do prisma da Figura 6.11 pelas duas técnicas, ambas reconhecendo quatro arestas, porém em posições diferentes. Com a utilização do mapa auto-organizável, os erros obtidos foram  $\varepsilon = 7,4 \times 10^{-3}$  e  $\varepsilon_q = 1,4 \times 10^{-3}$ , enquanto para Sobel os valores resultaram em  $\varepsilon = 8,5 \times 10^{-3}$  e  $\varepsilon_q = 5,1 \times 10^{-3}$ . Para este objeto, o desempenho conseguido por Sobel foi inferior, induzido principalmente pelas sombras e pelo outro objeto atrás do prisma, o qual causou a interrupção da borda entre eles. O segmento preto que ocupa apenas metade da aresta posterior do prisma, na Figura 6.11(f), indica esta falha na borda. Para o tetraedro, apresentado nas Figuras 6.11(g) a (i), os erros para o mapa auto-organizável foram  $\varepsilon = 1,2 \times 10^{-2}$  e  $\varepsilon_q = 5,9 \times 10^{-3}$  e, para Sobel,  $\varepsilon = 3,1 \times 10^{-3}$  e  $\varepsilon_q = 2,6 \times 10^{-3}$ , revelando uma precisão no reconhecimento melhor para o caso da utilização do gradiente calculado por Sobel.

Experimentos similares foram conduzidos para cerca de 20 pares de imagens, tomadas em condições de iluminação, fundo e quantidade de objetos diferentes. O número de reconhecimentos incorretos, entretanto, ainda é relativamente alto, independente do método de segmentação utilizado. Para os reconhecimentos corretos, os erros, em geral, oscilam para valores similares aos apresentados para a Figura 6.11. Há, entretanto, um menor erro geral de qualidade para o caso de se obter o mapa de

bordas por Sobel. A maior afinidade entre a intensidade e as posições das bordas é, então, novamente confirmada. As vezes em que a utilização de Sobel não leva ao reconhecimento dos objetos estão, na maior parte dos casos, ligadas ao excesso de pontos de borda, que levam a uma detecção também incorreta dos segmentos. As incorreções existem tanto no módulo de análise de formas, com segmentos bidimensionais por bordas grossas (ou muitas bordas próximas), que levam a orientações incorretas para os segmentos, quanto no módulo de estéreo, no qual há a propagação do erro em duas dimensões, visto que as profundidades são calculadas para pontos que não se encontram nos vértices ou arestas dos objetos. A utilização do mapa auto-organizável, por outro lado, tende a reduzir a quantidade de pontos de borda indesejáveis, sacrificando, porém, a precisão. As falhas no reconhecimento, para este método, estão em geral vinculados à imprecisão das bordas. Ao invés de, por exemplo, serem localizadas bordas formadas pela convergência de duas arestas, as oscilações podem arredondar o vértice, fazendo com que as retas acabem por não ser reconhecidas ou sejam detectadas em orientações divergentes da correta. Como discutido no Capítulo 4, há também casos de sobre-segmentação para casos em que o número de classes foi incorretamente determinado.

## **6.4 Comentários finais**

Há uma grande variabilidade nos segmentos de reta produzidos pela abordagem descrita, que é introduzida pela qualidade de iluminação das imagens, parâmetros do mapa auto-organizável, limiar para o gradiente calculado por Sobel e dimensões do espaço de parâmetros na transformada de Hough, além dos parâmetros e calibrações para a avaliação da profundidade, por exemplo. Este extenso conjunto de fatores não permitiu que se sistematizasse uma avaliação de erro para condições específicas (por exemplo, variações somente na transformada de Hough ou somente nos parâmetros do estéreo), de forma que são apresentadas considerações sobre as principais características dos reconhecimentos obtidos. Pode-se, através destas considerações, verificar as principais características e tendências deste sistema integrado.

### **6.4.1 Detecção de segmentos de retas**

A geração de uma representação intermediária dos dados da imagem através de segmentos de reta se mostrou bastante adequada para o reconhecimento proposto, embora a aplicabilidade ainda seja restrita. A grande maioria dos segmentos de reta identificados correspondem aos presentes na imagem e as falhas por falta de detecção podem ser contornadas por modificações nos parâmetros, adequando-os ao tipo de imagem e bordas considerada. As principais deficiências recaem sobre bordas mais grossas, dentro das quais várias retas podem ser ajustadas, o que acontece com certa frequência. Outro problema introduzido por bordas mais grossas é o desvio da posição ideal do segmento nas proximidades dos vértices.

Há muitos aspectos na detecção de retas a serem corrigidos para aumentar a qualidade dos segmentos, embora tenham tido, neste trabalho, uma prioridade menor

em função dos outros módulos. Dentre as possibilidades de modificações na transformada de Hough, duas podem ser destacadas. A primeira é a roboração de um segmento detectado (e pós-processado para simplificações), confrontando-o novamente com o mapa de bordas e com a imagem original, em cinza ou colorida, procurando dar consistência e corrigir distorções. Incorporar um esquema mais eficiente para a redução do número de segmentos sobre uma aresta, evitando que o processo de fusão de segmentos consecutivos ou coincidentes introduzam uma incorreção ainda maior a sua orientação e posição.

#### 6.4.2 Mapas auto-organizáveis

Os mapas de bordas obtidos através da nova abordagem para a segmentação apresentada no Capítulo 4 podem ser considerados sob dois enfoques distintos em função de sua aplicação ao problema de reconhecimento de objetos poliédricos. O primeiro é relativo à qualidade das bordas e o segundo, relativo ao tipo de bordas.

Considerando-se a precisão das bordas como sua proximidade às bordas esperadas para os objetos poliédricos, a qualidade das bordas obtidas não oferece uma precisão muito grande, visto que o principal fator que identifica o limite dos objetos ainda é a intensidade. As bordas, de uma perspectiva mais local, tendem a apresentar ondulações, as quais são causadas pela variabilidade das cores nas transições entre regiões da imagem. Esta variabilidade é também influenciada pelo sistema de captura, relativamente a sua qualidade (câmera, lentes, CCD e placa de aquisição). As tolerâncias da análise de formas e do reconhecimento de objetos, entretanto, são suficientes para contornar estas limitações e permitir, com certo grau de confiabilidade, reconhecimentos satisfatórios cujos erros não se apresentam substancialmente altos.

O tipo de segmentação, que considera as cores através dos atributos de cromaticidade da imagem, por outro lado, mostrou-se bastante adequado ao tipo de problema. São gerados mapas de bordas em função das cores em detrimento das variações da intensidade da iluminação, o permite identificar não somente bordas contínuas, mas também desconsiderar o excesso de bordas que a iluminação, em geral, promove. Esta característica é interessante, pois as bordas extras, devidas ao fundo, sombras ou outros elementos na cena, acabam sendo em grande parte eliminadas, resultando em um mapa de bordas mais limpo. Porém, não são perdidas, em geral, as bordas necessárias ao processamento estabelecido.

Deve-se, ainda, comentar que em situações nas quais a iluminação é deficiente (partes escuras ou pretas da cena), a cromaticidade não pode ser considerada adequadamente, propendendo a produzir mapas de bordas excessivamente carregados (sobre-segmentados).

#### 6.4.3 Sobel

A segmentação pelas máscaras de Sobel apresentam resultados muito interessantes na detecção de bordas, geralmente as identificando com precisão. Havendo um bom contraste nas imagens, todas as bordas de interesse podem ser conseguidas, ca-

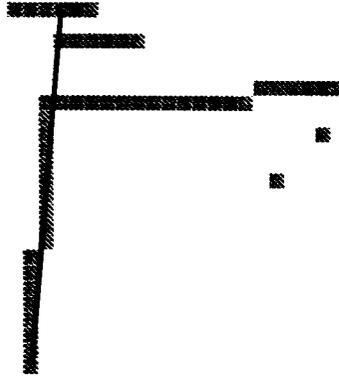


Figura 6.12. Ilustração da extensão de um segmento de reta (em preto) para o caso de serem toleradas descontinuidades de até três pontos. Os pontos hachurados representam as posições nas quais existem bordas.

racterizando mapas de bordas com boa representatividade da cena. Para aplicações nas quais as bordas devem ser distinguidas umas das outras, entretanto, o método acaba introduzindo uma grande quantidade de bordas pelas quais o interesse é muito limitado ou até nulo. Entre estas se incluem as induzidas por sombras ou outras variações de iluminação que não representam bordas dos objetos sob análise.

As principais conseqüências destas propriedades são o excesso de segmentos, o que penaliza com uma sobrecarga o processo de reconhecimento, e algumas incorreções introduzidas na definição dos segmentos de reta. Nesta última situação, deve-se considerar que os segmentos são definidos sobre as retas identificadas pela transformada de Hough e que há a tolerância de algumas descontinuidades nos segmentos. Dependendo das bordas localizadas, não são raras as ocorrências de que um segmento é estendido além de seus limites reais, unindo bordas disjuntas. A Figura 6.12 mostra um esquema desta ocorrência, sendo que, havendo uma tolerância de intervalos nas bordas de até três pontos, o segmento de reta acaba sendo estendido além do vértice.

Além disto, são freqüentes as ocorrências de bordas com espessura superior a um ponto. Nestes pontos de borda existe a tendência de que várias retas possam ser ajustadas, com deslocamentos mais ou menos sutis entre elas. Algum excesso de segmentos pode ocorrer e, mesmo após a filtragem de segmentos próximos, pode ainda haver certa multiplicidade.

#### 6.4.4 Considerações gerais sobre os resultados

Superando algumas expectativas prévias em relação às medidas de erro para os reconhecimentos, foi possível constatar erros relativamente pequenos entre os emparelhamentos do modelo da base de dados com os segmentos extraídos da imagem. Considerando que os erros  $\varepsilon_q$  são, em geral, inferiores a  $5 \times 10^{-2}$  para comprimentos dos segmentos iniciais iguais a 1, é possível considerá-los reduzidos, o que indica ajustes de boa qualidade, mesmo em situação nas quais o reconhecimento foi mais crítico. Estas ponderações aplicam-se ao caso em que o objeto foi corretamente reco-

nhecido, havendo casos em que o número de arestas reconhecidas é reduzido, apesar do reconhecimento correto.

É importante, ainda, ressaltar que, em vários casos, o reconhecimento é bem sucedido em imagens com múltiplos objetos, como a apresentada nas Figuras 6.7(g) e (j). Embora ainda não se tenha um mecanismo automático de remoção dos segmentos removidos para continuidade da análise da imagem, em busca de outros objetos, esta abordagem é também viável, introduzindo no sistema de visão um elemento adicional relativo à complexidade da cena.

#### **6.4.5 Comentários gerais**

A integração de atributos se apresenta com um importante e promissor destaque para os sistemas de visão por computador. Os estudos apresentados neste capítulo, embora requeiram outras investigações, ilustram o grande potencial latente da cooperação entre técnicas e atributos para a resolução de problemas. Há, porém, ainda vários caminhos a serem explorados, envolvendo outros atributos perceptivos e outras técnicas de análise.

## 7 Conclusões

**H**iabilizar a inclusão das cores como informação no processamento da visão não é somente muito interessante, mas representa uma área ainda a ser explorada mais profundamente. Este trabalho apresenta uma proposta para que esta inclusão seja efetivamente realizada, fazendo a incorporação da cor ao sistema de visão artificial *Cyvis-1*.

O sistema *Cyvis-1*, de um ponto de vista geral, é um sistema de visão cuja inspiração em sistemas biológicos de visão levou à adoção de modularidade, hierarquia, segregação de funções específicas, cooperação entre subsistemas, representação da informação visual em vários níveis de complexidade, entre outras características [Costa *et al.* (1994)]. A este contexto se agregam as propostas realizadas nesta tese, definindo um subsistema que manipula a cor para a segmentação de imagens em regiões ou bordas, apresentando um modelo para avaliação comparativa entre mapas de bordas e, não menos importante, definindo o embrião de um esquema para a integração do processamento de atributos distintos, no caso o objeto do estudo (cor) e informações espaciais (estéreo), promovendo cooperação, compartilhamento de informações e integração sensorial.

A seção seguinte descreve as principais conclusões de cada uma das três ênfases em que esta tese se apoiou, apresentando também suas principais contribuições. As seções 7.3 e 7.2 apresentam, respectivamente, as conclusões gerais do trabalho e as propostas de continuidade.

### 7.1 Conclusões do trabalho

Um dos objetivos propostos para o trabalho foi atingidos através da especificação um técnica de segmentação de imagens coloridas baseada na utilização de mapas auto-organizáveis e detecção de agrupamentos. Além da proposição a técnica, foi feito seu detalhamento segundo suas características de comportamento não somente

frente a imagens de características variadas, mas também relativamente à variação e influência de um conjunto de parâmetros que influenciam a segmentação.

Um segundo item consiste no estudo comparativo entre técnicas de segmentação, verificando o comportamento dos mapas de borda produzidos por cada uma relativamente a um conjunto de atributos locais, os quais estão diretamente associados às variações locais dos atributos da imagem e, em consequência, às discontinuidades presentes. O estudo considera as taxas em que as técnicas concordam ou divergem de uma segmentação manual, esta última realizada incorporando-se elementos da percepção humana na geração das regiões e do mapa de bordas. A ponderação entre as taxas em que há concordância ou discordância da segmentação manual foi também realizada, tendo seus resultados sido discutidos.

Por final, e importante já que incorpora grande parte da motivação de todo o trabalho realizado, foi apresentada uma proposta de um esquema de implementação para os módulos do *Cyvis-1*, incluindo aquisição de imagens, processamento por cor, processamento de bordas em níveis de cinza, um módulo de análise de formas para a detecção de segmentos de reta e a inclusão do módulo para o cálculo de profundidade (estéreo) desenvolvido pelo aluno de doutorado Alan Salvany Felinto, além do reconhecimento de objetos poliédricos da imagem através da confrontação com representações em uma base de dados [Moreira *et al.* (1999)]. Foi possível, nesta etapa, chegar a conceitos importantes do sistema de visão, caracterizando os módulos envolvidos, definindo as estruturas de controle internas e externas, incorporando um mecanismo de troca de mensagens e dados baseado na proposta de comunicação do aluno de doutorado Odemir M. Bruno [Bruno & Costa (1997)], identificar e implementar estruturas de representação das informações em vários níveis de sofisticação, promover a hierarquia e a integração dos módulos constituintes, os quais processam funções distintas de atributos visuais.

Apresentadas as características gerais dos três principais objetivos do trabalho, encontram-se nas seções seqüentes as conclusões de cada um individualmente.

### **7.1.1 Método de classificação com mapa auto-organizável**

No Capítulo 4, uma proposta para um método para a segmentação de imagens coloridas foi descrito e avaliado. Corresponde a uma nova abordagem para a segmentação, a qual incorpora a inspiração biológica através da utilização de mapas auto-organizáveis, aliados à detecção automática das classes presentes, por um critério definido entre o espalhamento dos dados dentro das classes e entre as classes.

A abordagem permitiu, correspondendo às expectativas prévias, produzir resultados satisfatórios para uma grande variedade de imagens, as quais possuem variabilidade em termos de iluminação, condições de aquisição, presença de reflexos e quantidade e tipo dos objetos presentes. A utilização da cromaticidade como a informação interna das unidades da rede neural que implementa o mapa auto-organizável também provou ser uma escolha adequada para o reconhecimento das classes, reduzindo a influência da iluminação na imagem. Obteve-se, assim, um certo

grau de independência em relação a sombras e a outros contrastes de iluminação. Cada elemento cromático pôde, desta forma, ser detectado apropriadamente, o que ocorreu para a maioria das imagens utilizadas para os experimentos. Esta robustez indica, dentro de limitações, uma aplicabilidade relativamente geral da técnica proposta.

Os experimentos também puderam evidenciar, obedecendo-se a alguns limites, que a metodologia se apresenta com razoável independência dos parâmetros de configuração, como tamanho do mapa e taxas de aprendizado e decaimento, entre outros. Isto torna a utilização do método mais simples para que seja ajustado a uma aplicação específica. Situações extremas, como mapas com poucos nós ou taxas de decaimento que aceleram demasiadamente o processo de convergência, devem, obviamente, ser evitadas, sob o custo da convergência não ser atingida de modo apropriado e a segmentação ser incorreta.

A validação da técnica foi realizada através da comparação com uma segmentação por regiões feita manualmente. A segmentação manual é feita considerando-se, além dos dados mais objetivos da imagem (cor dos pontos e suas relações de vizinhança, por exemplo), informações também subjetivas como inferência do conhecimento do objeto na determinação dos limites entre regiões. Nota-se uma grande correspondência entre as regiões esperadas (com uma interpretação humana da cena) e as efetivamente detectadas pelo algoritmo. Esta relação é importante, pois a metodologia de segmentação proposta acaba refletindo uma proximidade com aspectos subjetivos da percepção humana.

A comparação do desempenho do método em relação a duas outras técnicas de segmentação por classificação foi também apresentada, indicando que o comportamento, embora não consistentemente superior, é comparável, na maioria dos experimentos, à classificação por  $k$  vizinhos mais próximos e um pouco inferior, em geral, à classificação bayesiana. Ressalta-se, ainda, que ambas as classificações usadas na comparação necessitam de uma fase supervisionada, na qual tanto o número de classes quanto seus ajustes para classificação dependem de um operador que determine um conjunto de amostras. Para a classificação por  $k$  vizinhos, um conjunto de amostras deve ser previamente rotulado para, posteriormente, ser confrontado com os demais dados para as rotulações restantes. As funções de probabilidade da classificação bayesiana também dependem de uma rotulação prévia de um conjunto de amostras, bem como é necessário estabelecer as probabilidades *a priori* para as classes. A ausência da supervisão corresponde, então, a uma vantagem em relação às outras técnicas, produzindo resultados similares em qualidade sem necessidade de intervenção.

Há, entretanto, uma certa instabilidade nos limites entre regiões nas segmentações geradas pela técnica proposta. Devido tanto a problemas relativos à resolução na aquisição das imagens quanto à utilização da cromaticidade como informação para a separação das classes, nas áreas da imagem em que há a transição de uma classe para outra não é incomum haver oscilações nas cores, o que é refletido em uma precisão menor na fronteira entre duas regiões. Como conseqüência, mesmo em partes da ima-

gem em que a fronteira aparenta ser reta surgem algumas ondulações, que aumentam de intensidade à medida que outros fatores como baixa iluminação ou reflexos se tornam mais presentes. O acréscimo da informação da intensidade luminosa como dado para a segmentação tende a minimizar este problema, fazendo com que as oscilações nas bordas sejam bastante reduzidas. Em contrapartida, novas bordas são introduzidas em função da iluminação, bem como observa-se uma tendência para que ocorresse sobre-segmentação nestes casos. Em casos específicos, entretanto, o uso da iluminação para a formação do mapa auto-organizável pode ser considerado como uma alternativa conveniente.

As deficiências da segmentação correspondem, principalmente, a duas condições: texturas e iluminação. Nas regiões em que há texturas a classificação não é apropriada, observação que também é válida para outras formas de segmentação por classificação; não apenas há um excesso de bordas como o número de pequenas regiões é grande. Relativamente à iluminação, a cromaticidade, enquanto normalização das componentes RGB, tem valores mais estáveis e coerentes quando pelo menos uma das três componentes não tem valores muito baixos. Quando as componentes são todas de valores pequenos, o que corresponde a regiões escuras, uma instabilidade é introduzida na cromaticidade, a qual não tem significado para regiões de pouca iluminação. Estas áreas muito escuras tendem, sob condições gerais, a serem excessivamente segmentadas.

### 7.1.2 Comparação de desempenho de técnicas de detecção de bordas e combinação baseada em contexto local

Uma proposta de um novo modelo de comparação entre técnicas de segmentação foi apresentada, introduzindo atributos locais como parâmetros de separação para as comparações. A relação entre bordas e atributos como variâncias locais, que foram utilizadas no modelo, mostrou-se um elemento bastante adequado na comparação, permitindo agrupar técnicas de desempenho similar e distinguir outras com comportamento diferenciado.

A comparação é vinculada, em um primeiro plano, ao número de *coincidências* entre bordas produzidas por técnicas de segmentação e mapas de bordas de referência. Em um segundo plano, as divergências entre as bordas também são consideradas, através da definição de uma *taxa de excesso* que representa o quanto uma técnica “discorda” do mapa de referência. Através da ponderação entre as taxas de coincidência e de excesso é possível analisar a qualidade relativa dos mapas de bordas para situações específicas. Esta ponderação foi estabelecida nas definições dos coeficientes acumulativo e proporcional. Torna-se viável, então, priorizar as taxas de coincidência e identificar as técnicas que determinam melhor as bordas e, por outro lado, priorizar as menores taxas de excesso, o que reflete encontrar uma minimização dos erros na detecção das bordas.

As análises apresentadas permitiram ilustrar o comportamento genérico das técnicas, revelando melhores ou piores desempenhos através do coeficiente acumula-

tivo. Este coeficiente, em particular, através da contagem do número de melhores desempenhos, tende a separar de uma maneira mais forte umas técnicas das outras. Isto é obtido quando técnicas que não tiveram o melhor desempenho para um dado intervalo da variância são desconsideradas em função daquela com melhor desempenho. Desta forma, os melhores desempenhos são considerados de modo mais absoluto. Os coeficientes proporcionais, por sua vez, produziram resultados gerais similares, corroborando os desempenhos relativos das técnicas. Entretanto, como utilizam a média das taxas de coincidência e excesso, tendem a apresentar uma comparação mais precisa entre as várias técnicas, mesmo que estas não sejam as melhores ou as piores.

Enquadradas no modelo de comparação, foi possível observar que técnicas voltadas a bordas introduzem uma maior precisão nos resultados, quando comparadas às técnicas que utilizam classificação. Da mesma forma foi comprovada a importância das diferenças de iluminação na localização das bordas, as quais são mais nítidas e precisas quando a segmentação considera a componente de intensidade. Por outro lado, esta inclusão tende a aumentar o número de bordas em excesso. Ponderar entre estes limites é, portanto, necessário.

O estudo feito para classes de imagens comprovam os resultados gerais obtidos para todas as imagens da base de dados. A utilização de classes, entretanto, permitem uma análise mais detalhada e precisa dos desempenhos das técnicas quando a variabilidade de conteúdo das imagens é reduzida.

Derivada da comparação das técnicas, a proposta de uma combinação seletiva baseada nos atributos locais foi apresentada. Considerar os resultados individuais das técnicas para a seleção condicional dos pontos de bordas em função das regiões da imagem mostrou-se importante para reduzir as taxas de excesso, ao mesmo tempo em que se procura maximizar as taxas de coincidência. Os experimentos realizados comprovam esta posição, gerando um modelo que, após calibrado segundo as taxas e coeficientes definidos, pode ser empregado para “otimizar” mapas de bordas em relação a um conjunto de técnicas para uma dada aplicação.

### **7.1.3 Integração sensorial entre cor e estéreo**

Dada a importância da combinação de atributos visuais para o processamento e a análise de imagens, esta tese também apresenta um modelo de estruturação para o sistema de visão *Cyvis-1* que promove a integração em vários níveis entre seus subsistemas. A proposta apresentada envolve, em primeira instância, a cooperação e integração entre os módulos de segmentação por cor e estéreo. Além disto, são também considerados os módulos de aquisição, análise de formas e reconhecimento de objetos poliédricos.

A proposta define os mecanismos de comunicação e os protocolos envolvidos, permitindo a estruturação de módulos independentes e quase autônomos, que operam como um sistema de processamento distribuído. As características de modularidade e integração ficam, desta forma, incorporadas à estruturação, que também envolve a segregação de funções entre os diversos subsistemas. O estabelecimento de um meca-

nismo de coordenação do sistema introduz o controle central, que gerencia sua operação como um todo.

O reconhecimento de objetos poliédricos, utilizado para a avaliação do módulo de estéreo, foi descrito e implementado, estabelecendo-se como mecanismo para a validação desta abordagem para do sistema integrado. Através da verificação dos reconhecimentos (e de suas falhas) foi possível determinar as características básicas do comportamento do sistema, que se mostrou viável como solução para a classificação dos objetos. Um erro qualitativo para o reconhecimento e um erro geral de emparelhamento foram definidos, tendo sua utilização se mostrado adequada para a comparação dos resultados.

Permitiu-se, assim, comparar os mapas de bordas gerados pela técnica com mapas auto-organizáveis e por Sobel. Desta comparação, os principais aspectos são a precisão do reconhecimento e o tipo de mapa de bordas produzido. Concordando com as conclusões feitas para nova técnica de segmentação por mapas auto-organizáveis e para a comparação das técnicas de segmentação, uma tendência para que a precisão dos resultados produzidos por Sobel permanecesse maior foi observada. Entretanto, o excesso de bordas que Sobel introduz na imagem (outros objetos e sombras, por exemplo) tendem a alterar o desempenho da análise de formas, que prolonga segmentos de reta além do desejado e gera certa multiplicidade de segmentos para uma mesma borda. Embora esta última ocorrência também se manifeste em menor intensidade para a segmentação por mapas auto-organizáveis, há uma propensão deste último para um menor excesso de bordas. Sacrifica-se, porém, um certo grau de precisão.

Apesar do correto reconhecimento de objetos com erros considerados satisfatórios, ainda há um número grande de falhas no reconhecimento, o que não invalida a abordagem, mas salienta que ainda há ajustes a serem feitos.

## 7.2 Conclusões gerais

Esta tese descreve três contribuições principais, relacionadas à segmentação de imagens coloridas, ao estudo comparativo das bordas produzidas por técnicas de segmentação variadas e à estruturação do subsistema de cor do sistema de visão *Cyvis-1*, com integração ao subsistema de estéreo. Todas convergem para formar aspectos diferentes de um mesmo escopo, que é o sistema artificial de visão *Cyvis-1*. Cada objeto de estudo foi apresentado e avaliado, indicando suas características, limitações e deficiências, mas mostrando também que sua aplicação e qualidades sobrepõem, em grande parte, as insuficiências. Embora consideradas separadamente, as inter-relações entre os trabalhos realizados são muitas, destacando-se a utilização da segmentação por mapas auto-organizáveis, aliado a informações estereoscópicas, no reconhecimento de objetos e a viabilidade da integração do estudo comparativo das técnicas de segmentação no sistema de visão, para a separação de tarefas e integração de outros atributos visuais.

Tabela 7-I. Síntese das contribuições deste trabalho.

Abordagem	Principais contribuições
Segmentação por mapas auto-organizáveis	<ul style="list-style-type: none"> <li>- Estipulação de uma nova abordagem para segmentação de imagens coloridas;</li> <li>- Validação dos resultados;</li> <li>- Comparação com outras técnicas (classificação bayesiana e por <math>k</math> vizinhos mais próximos), mostrando, dentro de certos limites, sua proximidade de resultados;</li> <li>- Inclusão da cromaticidade para redução da influência de sombras e outras interferências de intensidade de luz;</li> <li>- Proximidade geral dos resultados da segmentação à segmentação manual</li> </ul>
Comparação de técnicas de segmentação por bordas	<ul style="list-style-type: none"> <li>- Determinação de uma nova abordagem comparativa para mapas de bordas;</li> <li>- Introdução de atributos locais como elemento de comparação entre técnicas;</li> <li>- Proposição de medidas de desempenho, como as taxas de coincidência e excesso e os coeficientes acumulativo e proporcional;</li> <li>- Proposição e avaliação da combinação seletiva de bordas;</li> <li>- Análise dos desempenhos por classes e geral.</li> </ul>
Integração entre cor e estéreo	<ul style="list-style-type: none"> <li>- Introdução da estruturação e implementação para a integração de atributos visuais no sistema de visão <i>Cyvis-1</i>;</li> <li>- Proposição um modelo de cooperação e integração entre subsistemas;</li> <li>- Avaliação do reconhecimento de objetos poliédricos através da integração;</li> <li>- Avaliação e validação do sistema através da definição de critérios de erro;</li> <li>- Caracterização de mapas de bordas de origens diferentes (bordas e por segmentação*), qualificando-os segundo o reconhecimento.</li> </ul>
Geral	<ul style="list-style-type: none"> <li>- Comprovação da viabilidade e potencial da integração de atributos visuais;</li> <li>- Validação da estrutura do sistema como mecanismo de avaliação para processos de visão.</li> </ul>

Demonstrou-se, também, que a estruturação do sistema de visão permite e estimula a validação dos resultados, avalizando os progressos realizados e viabilizando um trabalho integrado e multidisciplinar.

A Tabela 7-I apresenta as principais contribuições desta tese, sintetizando as conclusões e colocações deste capítulo.

Além disto, um último objetivo geral, que compreende não só os desenvolvimentos realizados separadamente, mas sua união como elementos de um sistema de visão mais abrangente também foi atingido, abrindo novas perspectivas de estudos e contribuindo para o conhecimento como um todo.

\* Considerando que um dos mapas de bordas é derivado da segmentação por regiões.

### 7.3 Continuidade

O trabalho desenvolvido, em seus três enfoques principais, abriu vários outros caminhos, para os quais a investigação se apresenta promissora. Apesar de apresentadas separadamente em relação à segmentação por mapas auto-organizáveis, à comparação das técnicas de segmentação e à integração de módulos do *Cyvis-1*, muitas das propostas apresentadas compartilham aspectos em comum.

Embora a segmentação por mapas auto-organizáveis tenha atingido um estado estável de desenvolvimento, muitas melhorias ainda podem ser introduzidas para corrigir falhas na segmentação. Trabalhos futuros podem ser direcionados, assim, para os seguintes pontos principais:

- a) *Inclusão de tratamento diferenciado nas regiões das bordas.* As bordas entre as regiões constituem uma deficiência da técnica de segmentação proposta, de forma que um tratamento diferenciado para estas regiões pode proporcionar uma melhor qualidade nas fronteiras entre as classes. Uma possibilidade é a associação da técnica a outra de detecção de bordas, a qual verificaria (através do gradiente colorido, por exemplo), as áreas em que há bordas, mesmo que esta detecção seja em excesso. Após a seleção das classes do mapa auto-organizável, estas regiões de bordas seriam revistas em termos de classificação. As áreas classificadas, menos as bordas, poderiam sofrer um processo de crescimento de regiões, até que as fronteiras de regiões vizinhas se unissem. Um critério de controle para o crescimento de uma classe poderia, por exemplo, adotar um esquema de “escalada” do gradiente. Valores mais altos do gradiente seriam classificados por último, tendendo a fazer coincidir o encontro entre regiões à bordas qualificada localmente pelo gradiente. Naturalmente, critérios de continuidade e coesão das classes precisariam ser estabelecidos.
- b) *Avaliar outras formas de agrupamento de dados.* Outras formas não supervisionadas para a determinação das classes poderiam ser adotadas, utilizando outros critérios para avaliar o número de classes e sua representatividade no mapa. Da mesma forma uma segunda rede neural (ou mapa auto-organizável) poderia ser incluída para processar os dados da primeira instância da rede, reduzindo mapas  $32 \times 32$  para mapas  $1 \times n$ , para os quais nós vizinhos pudessem quantificar melhor os limites entre classes distintas.

O estudo comparativo e a combinação de bordas podem ser enriquecidos através das seguintes considerações:

- a) *Avaliação concorrente dos atributos locais e aumento do número de atributos.* As análises apresentadas consideraram cada atributo de forma isolada. Na combinação de mapas de bordas, poder-se-ia também incluir um critério de seleção que comparasse simultaneamente atributos distintos, definindo mecanismos como “melhor atributo para uma região da imagem”, ou variações similares.

- b) *Incorporação da combinação seletiva de bordas ao sistema Cyvis-1.* Uma das principais dificuldades na combinação seletiva de bordas é a possibilidade de se ter que realizar várias segmentações para combinar resultados. Em sistemas que operassem de forma paralela, como se propõe o *Cyvis-1*, módulos variados fariam diversas segmentações, as quais seriam combinadas por outros módulos, cada qual segundo suas necessidades particulares.

Confundindo-se com as propostas para a comparação das técnicas de segmentação, a integração do *Cyvis-1* apresenta diversas possibilidades, dentre as quais podem ser destacadas as seguintes:

- a) *Incorporação de outros atributos visuais.* Novas investigações sobre cooperação entre módulos incluem diversas associações, dentre as quais a inclusão de textura se apresenta bastante interessante. Para este aspecto, novas estruturas de comunicação e representação deveriam ser desenvolvidas, bem como a inclusão de um controle geral mais sofisticado se tornaria necessária. [Moreira *et al.* (1999)]
- b) *Aumento do grau de compartilhamento de informações, com acréscimos de fluxo inverso nos dados.* Um dos pontos principais da integração é a reutilização da informação processada nos diversos módulos. A confirmação de informações pode ser feita entre módulos, de forma que cada um refinasse suas informações a partir da comparação com outros dados. Por exemplo, as informações dos dados em três dimensões poderiam ser incorporadas ao refinamento das segmentações por mapas auto-organizáveis, o que melhoraria, por sua vez, a qualidade da informação de profundidade.

Por fim, novas investigações para utilização das cores na segmentação de imagens podem incluir, por exemplo:

- a) *Caracterização estatística das distribuições das cores.* Há, ainda, uma falta de caracterização de como as diversas cores se distribuem nas imagens. Um melhor entendimento das cores, que dependem tanto da classe de imagens como das condições de aquisição, poderia ser obtido através de uma caracterização sua caracterização estatística. Não somente os atributos RGB poderiam ser avaliados, mas também outros sistemas de representação, em especial os relacionados a atributos perceptivos (por exemplo, HSI).
- b) *Desenvolvimento de novos mecanismos de aquisição e análise da composição espectral do ambiente.* As imagens adquiridas por câmeras digitais representam simplificações da composição espectral do ambiente. A compreensão mais pormenorizada de como determinados tipos de ambiente são formados em termos de composição de comprimentos de onda é importante para criar novos mecanismos de segmentação, processamento ou cooperação. Esta opção, naturalmente, envolveria também o desenvolvimento ou a adaptação de um *hardware* específico.

*"Nenhum livro é tão ruim a ponto de não se poder  
aprender nada com ele."  
Plínio, o Jovem*

*"Para que servirão os livros se não nos conduzem à vida, se não  
conseguem fazer com que bebamos dela com mais avidez?"  
Henry Miller*

## 8 Bibliografia

- Anderson, J. A. *An introduction to neural networks*. MIT Press, 1995.
- Auge, E. & Pitas, I. Segmentation of color images by using the learning vector quantizer. 1995 IEEE Workshop on Nonlinear Signal and Image Processing. 4p. Greece, June 20-22, 1995.
- Babu, G. P. Self organizing neural networks for spatial data. *Pattern Recognition Letters*, 18: 133-142, 1997.
- Ballard, D.H. & Brown, C. M. *Computer vision*, Prentice Hall, 1982.
- Barlow, H. B. & Mollon, J. D. Psychophysical measurements of visual performance. In: BARLOW, H. B. & MOLLON, J. D. *The senses*, Cambridge Univ. Press, Cambridge, U. K., 1982.
- Bruno, O. M. & Costa, L. F. Versatile Real-Time Vision Based on a Distributed System of Personal Computers. *Proceedings - Third IEEE International Conference on Engineering of Complex Computer Systems*. Como, Itália, 1997.
- Castleman, K, R. *Digital image processing*. Prentice Hall, Inc., New Jersey, 1996.
- Celenk, M. A color clustering technique for image segmentation. *Computer Vision, Graphics and Image Processing*, 52:145-170, 1990.
- Churchland, P. S. & Sejnowski, T, J. Perspectives in cognitive neuroscience. *Science*, 242: 741-745, 1988.
- Coleman, G. B. & Andrews, H. C. Image segmentation by clustering. *Proc. IEEE*, 67(5): 773-785, 1979.
- Costa, L. F. & Sandler, M. B. Effective detection of digital bar segments with Hough transform. *CVGIP: Gr. Models Image Process.*: 55(3): 180-191, 1993.

- Costa, L. da F.; Roda, V. O.; Köberle, R. A biologically-inspired system for visual pattern recognition. *IEEE International Symposium on Industrial Electronics*, Santiago-Chile, 1994.
- Cumani, A. Edge detection in multispectral images. *CVGIP Graphic Models and Image Processing*, 53:40-51, 1991.
- Demigny, D. & Kamlé, T. A discrete expression of Canny's criteria for step edge detection performances evaluation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(11): 1199-1211, 1997.
- DiZenno, S. A note on the gradient of a multi-image. *Computer Vision, Graphics and Image Processing*, 33:116-125, 1986.
- Duda, R. O. & Hart, P. E. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1): 11-15, 1972.
- Duda, R. O. & Hart, P. E. *Pattern classification and scene analysis*. John Wiley and Sons, Inc., New York, 1973.
- Fairchild, M. D. *Color appearance models*. Addison Wesley Longman, Inc., 1998.
- Faugeras, O. *Three-dimensional computer vision: a geometric viewpoint*. MIT Press, 2<sup>nd</sup> ed. 1996.
- Fausett, L. *Fundamentals of neural networks: architectures, algorithms, and applications*. Prentice-Hall, 1994.
- Frigui, H. & Krishnapuran, R. A robust algorithm for automatic extraction of an unknown number of clusters from noisy data. *Patt. Recogn. Letters*, 17: 1223-1232, 1996.
- Gerig, G. & Klein, F. Fast contour identification through efficient Hough transform and simplified interpretation strategy. In: Proc. 8th Int. Conference on Pattern Recognition, vol. 1, pp. 498-500, Paris, France, October 27-31, 1986.
- Godfrey, K. R. L. & Attikiouzel, Y. Applying neural networks to colour image data compression. IEEE Region 10 Conference, Tencon 92. Melbourne, Australia, 1992.
- Goethe, J. W. *Materialien zur Geschichte der Farbenlehre, Zur Farbenlehre*. 2 vol. Tübingen, 1810.
- Goethe, J. W. *Theory of colours*. MIT Press, 1970\*.
- Goldstein, E. B. *Sensation and perception*. Wadsworth Publ. Co., Belmont, U. S. A., 1989.
- Gonzales, R. C. & Woods, R. E. *Digital image processing*. Addison-Wesley Publ. Co., 1992.
- Gorte-Kroupnova, N. & Gorte, B. Method for multi-spectral images segmentation in case of partially available spectral characteristics of objects. *SPIE* -

---

\* Tradução da obra original em alemão de 1810 [Goethe (1810)] por C. L. Eastlake (em 1840); introdução por D. B. Judd (em 1970).

- International Society for Optical Engineering, Proceedings*. 2665: 210-218, 1996.
- Heath, M. D.; Sarkar, S.; Sanocki, T.; Bowyer, K. W. A robust visual method for assessing the relative performance of edge-detection algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(12): 1338-1359, 1997.
- Heath, M. D.; Sarkar, S.; Bowyer, K. W. Comparison of edge detectors. *Computer Vision and Image Understanding*, 69(1): 38-54, 1998.
- Hedley, M. & Yan, H. Segmentation of color images using spatial and color space information. *Journal of Electronic Imaging*, 4(1): 374-380, 1992.
- Hertz, J.; Krogh, A.; Palmer, R. G.. *Introduction to the theory of neural computation*. Addison-Wesley Publ. Co., 1991.
- Hough, P. V. C. *Methods and means for recognizing complex patterns*. U. S. Patent 3,069,654.
- Hueckel, M. An operator which locates edges in digitized pictures. *Journal of the ACM*, 20(4): 634-647, 1973.
- Huntsberger, T. L.; Jacobs, C. L.; Cannon, R. L. Iterative fuzzy image segmentation. *Pattern Recognition*, 18:131-138, 1985.
- Hurlbert, A. Color algorithms for image segmentation. In: Cotterill, R. M. J. *Models of brain function*, Cambridge Univ. Press, Cambridge, England, 1989.
- Jain, R.; Kasturi, R.; Schunck, B. G. *Machine vision*. McGraw Hill, 1995.
- Kamei-Harada, Y.; Miyakoshi, M.; Shimbo, M. A fuzzy clustering method for automatic segmentation of color scenes. *Japanese Journal of Fuzzy Theory and Systems*, 6(5):651-672, 1994.
- Kangas, J. A.; Kohonen, T. K.; Laaksonen, J. T. Variants of self-organizing maps. *IEEE Trans. Neural Networks*, 1(1):93-99, 1990.
- Khotanzad, A. & Bouarfa, A. Image segmentation by a parallel non-parametric histogram based clustering algorithm. *Pattern Recognition*, 23(9): 961-973, 1990.
- Klinker, G. J.; Shafer, S. A.; Kanade, T. A physical approach to color image understanding. *Intern. Journal of Computer Vision*, 4: 7-38, 1990.
- Kohonen, T. K. *Self-organization and associative memory*. Springer-Verlag, Berlin, 1984.
- Kohonen, T. K. Emergence of invariant-feature detectors in the adaptive-subspace self-organizing map. *Biological Cybernetics*, 75: 281-291, 1996.
- Lee, H.-C. & Cok, D. R. Detecting boundaries in a vector field. *IEEE Trans. Sign. Process.*, 39:1181-1194, 1991.
- Lee, C.-C. & Gyvez, J. P. Color image processing in a cellular neural-network environment. *IEEE Trans. Neural Networks*, 7(5): 1086-1098, 1996.

- Levine, M. D. *Vision in man and machine*. McGraw Hill Publ. Co., N. Y., 1985.
- Lim, Y. W. & Lee, S. U. On the color image segmentation algorithm based on the thresholding and the fuzzy *c*-means techniques. *Pattern Recognition*, 23:935–952, 1990.
- Livingstone, M. & Hubel, D. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, 240: 740-749, 1988.
- Martin, A. *et al.* Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, 270: 102-105, 1995.
- Masland, R. H. Unscrambling color vision. *Science*, 271: 616-617, 1996.
- Maxwell, B. A. & Shafer, S. A. A framework for segmentation using physical models of image formation. *Technical report* CMU-RI-TR-93-29, Carnegie Mellon University, 1993.
- Maxwell, B. A. & Shafer, S. A. Physics-based segmentation: looking beyond color. *Technical report* CMU-RI-TR-95-37, Carnegie Mellon University, 1995.
- Mollon, J. D. Colour vision and colour blindness. In: BARLOW, H. B. & MOLLON, J. D. *The senses*, Cambridge Univ. Press, Cambridge, U. K., 1982.
- Moreira, J. & Costa, L. F. Multispectral image segmentation by chromaticity classification. *VIII SIBGRAPI (Brazilian Symposium on Computer Graphics and Image Processing)*, pp. 119-125. São Carlos, 1995.
- Moreira, J. & Costa, L. F. Neural-based color image segmentation and classification using self-organizing maps. *IX SIBGRAPI (Brazilian Symposium on Computer Graphics and Image Processing)*, Caxambu, 1996.
- Moreira, J.; Felinto, A. S.; Costa, L. F. Cyvis-1: Integration of Color and Stereo. *International Workshop on Synthetic-Natural Hybrid Coding an 3D Imaging*. Santorini, Greece, September 15-17, 1999. (Aceito para publicação.)
- Nevatia, R. A color edge detector and its use in scene segmentation. *IEEE Trans. Systems, Man and Cybernetics*, 7(11): 820-826, 1977.
- Ohlander, R.; Price, K.; Reddy, D. R. Picture segmentation using a recursive region splitting method. *Computer Graphics and Image Processing*, 8:313–333, 1978.
- Ohta, Y.-I.; Kanade, T.; Sakai, T. Color information for region segmentation. *Computer Graphics and Image Processing*, 13:222–241, 1980.
- Pal, N. R. & Pal, S. K. A review on image segmentation techniques. *Pattern Recognition*, 26(9): 1227-1294, 1993.
- Pal, N. R.; Bezdec, J. C.; Tsao, E. C.-K. Generalized clustering networks and Kohonen's self-organizing scheme. *IEEE Trans. Neural Networks*, 4(4): 549- 557, 1993.
- Pratt, W. K. *Digital image processing*. John Wiley & Sons, Inc., 2<sup>nd</sup> ed., 1991.

- Pujas, P. & Aldon, M.-J. Estimation of the color image gradient with perceptual attributes. *Image Analysis and Processing, 9<sup>th</sup> International Conference, ICIAP'97*. Proceedings, vol. 1. pp. 103-110. Florence, Italy, September, 1997. In: *Lecture Notes in Computer Science*, 1310.
- Qian, N. Physiological computation of binocular disparity. *Vision Res.*, 37(13):1811-1827, 1997.
- Rachid, S. *et al.* Segmentation of sputum color image for lung cancer diagnosis based on neural networks. *Image Analysis and Processing, 9<sup>th</sup> International Conference, ICIAP'97*. Proceedings, vol. 2. pp. 461-468. Florence, Italy, September, 1997. In: *Lecture Notes in Computer Science*, 1311.
- Robinson, G. S. Color edge detection. *Optical Engineering*, 16(5): 479-484, 1977.
- Robinson, G. S. Edge detection by compass gradient masks. *Computer Graphics and Image Processing*, 6: 492-501, 1977.
- Schalkoff, R. J. *Digital image processing and computer vision*. John Wiley and Sons, Inc. 1989.
- Scharcanski, J. & Venetsanopoulos, A. N. Colour image edge detection using directional operators. *IEEE Trans. Circ. And Systems for Video Technology*, 7(2): 397-401, 1997.
- Schlünzen, E. T. M. *et al.* Classificação de dados multiespectrais utilizando redes neurais: a influência da amostragem no processo de treinamento. *Anais do Workshop sobre Visão Cibernética*. São Carlos, Agosto 1994.
- Shafer, S. A. Using color to separate reflection components. *Color Research and Application*, 10(4): 210-218, 1985.
- Shiozaki, A. Edge extraction using entropy operator. *Comp. Vision, Graphics and Image Processing*, 36: 1-9, 1986.
- Shirai, Y. *Three-dimensional computer vision*. Springer-Verlag, 1987.
- Sirosh, J. & Miikkulainen, R. Topographic receptive fields and patterned lateral interaction in a self-organizing model of the primary visual cortex. *Neural Computation*, 9(3): 577-594, 1997.
- Skarbek, W., Koschan, A. *Colour Image Segmentation - A Survey*, Technical Report 94-32, Technical University of Berlin, Department of Computer Science, October 1994.
- Sobel, I. E. *Camera models and machine perception*. PhD Thesis, Stanford University, 1970.
- Stenberg, R. J. & Davidson, J. E. *Nature of insight*. MIT Press, 1995.
- Theimer, W. M., Mallot, H. A. Phase-based binocular vergence control and depth reconstruction using active vision. *CVGIP: Image Understanding*. 60(3):343-358. November 1994.
- Tovée, M. J. *An introduction to the visual system*. Cambridge Univ. Press. Cambridge, 1996.

- Trahanias, P. E. & Venetsanopoulos, A. N. Color edge detection using vector order statistics. *IEEE Trans. Image Processing*, 2(2): 259-264, 1993.
- Trivedi, M. M. & Bezdek, J. C. Low-level segmentation of aerial images with fuzzy clustering. *IEEE Trans. Systems, Man and Cybernetics*, 16:589-598, 1986.
- Tsai, R. A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323-344, August 1987.
- Uchimura, K. Color image segmentation using tree representation. *T. IEE Japan*, 114-C(12): 1320-1321, 1994.
- Uchiyama, T. & Arbib, M. A. Color image segmentation using competitive learning. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 16(2): 1197-1206, 1994.
- Ultré, V. & Macaire, L. Multispectral edge detection by relaxation algorithm. *SPIE Intern. Soc. for Optical Engineering, Proceedings*, 2666: 252-258, 1996.
- Verikas, A.; Malmqvist, K.; Bergman, L. Colour image segmentation by modular neural network. *Pattern Recognition Letters*, 18: 173-185, 1997.
- Wyszecki, G. & Stiles, W. S. *Color science: concepts and methods, quantitative data and formulae*. John Wiley & Sons Ed. 1982.
- Yang, H. & Halme, A. An edge extracting method on color image. *The 3<sup>rd</sup> International Conference on Automation, Robotics, and Computer Vision (ICARCV'94)*. Singapore, Nov. 9-11, 1994.
- Yin, H. & Allinson, N. M. On the distribution and convergence of feature space in self-organizing maps. *Neural Computation*, 7:1178-1187, 1995.
- Zeki, S. & Shipp, S. The functional logic of cortical connections. *Nature*, 335: 311-317, 1988.
- Zheng, Y.-J. Feature extraction and image segmentation using self-organizing networks. *Machine Vision and Applications*, 8: 262-274, 1995.
- Zrenner, E. *Neurophysiological aspects of color vision in primates: Comparative studies on simian retinal ganglion cells and the human visual system*. Springer-Verlag, Berlin, Heidelberg, 1983.

## A. Glossário

A relação de termos apresentada neste apêndice corresponde às nomenclaturas com sentido restrito ou específico, apresentando a conotação utilizada no texto. Além destes, temos de uso geral em visão, reconhecimento de padrões, processamento de imagens e redes neurais, cujos significados foram considerados importantes para destaque, também se encontram incluídos para permitir um completo entendimento.

<b><i>Acromático</i></b>	Estímulo que não apresenta componentes de matiz, consistindo apenas de intensidade (do branco ao preto, através dos cinzas intermediários). Cf. <i>monocromático</i> .
<b><i>Agrupamento</i></b>	1. Ato de separar um conjunto de dados em grupos, segundo um critério de similaridade de atributos. 2. Um destes grupos.
<b><i>Borda degrau</i></b>	Corresponde a uma borda de ascensão rápida, cujo perfil delineia um degrau entre duas intensidades diferentes.
<b><i>Borda rampa</i></b>	Corresponde a uma borda de ascensão um tanto suave, formando, em seu perfil, uma rampa ou inclinação pouco abrupta.
<b><i>C.I.E.</i></b>	<i>Commission Internationale de l'Eclairage</i> , órgão que, em 1931, definiu o “observador padrão” para experimentos de colorimetria e fotometria.
<b><i>Campo receptivo</i></b>	Área da retina que, se estimulada, provoca alguma resposta em um neurônio que tenha seu sinal medido. Diz-se que é o campo receptivo daquele neurônio em particular.
<b><i>Cognição</i></b>	Aquisição do conhecimento ou identificação de um conhecimento prévio.
<b><i>Cor</i></b>	Característica da luz que permite a um observador distinguir diferenças no estímulo conforme varia sua composição espectral.

<b><i>Cromático</i></b>	Estímulo ou atributo correspondente a uma certa composição espectral com certa magnitude.
<b><i>Degrau</i></b>	Veja <i>borda de grau</i> .
<b><i>Estéreo (visão)</i></b>	Forma reduzida de estereoscópico. Visão binocular, isto é, relativa a ambos os olhos, permitindo a noção de profundidade. Visão estereoscópica.
<b><i>Foveal</i></b>	Relativo à fóvea e que possui suas características, dentre as quais destacam-se maior acuidade, maior número de fotorreceptores e grande sensibilidade a cores.
<b><i>Grupo</i></b>	Veja <i>agrupamento</i> (2).
<b><i>Iluminação</i></b>	Intensidade do estímulo luminoso, relativo à quantidade de luz que parece ser emitida.
<b><i>Insight</i></b>	Compreensão repentina, em geral intuitiva, das próprias atitudes e comportamentos, de um problema ou de uma situação.
<b><i>Irradiância</i></b>	Medida da potência de luz incidente sobre uma superfície.
<b><i>Luminância</i></b>	Componente de iluminação de um estímulo.
<b><i>Matiz</i></b>	Atributo da cor que permite classificá-la como azul, amarela, verde ou vermelha, ou como combinações delas.
<b><i>Monocromático</i></b>	Estímulo ou atributo que é especificado por um único comprimento de onda de certa magnitude.
<b><i>N.T.S.C.</i></b>	<i>National Television Systems Committee</i> , órgão estadunidense que regulamenta o padrão para transmissão e recepção de televisão. Aplica-se a sigla também ao próprio sistema de codificação do sinal (veja-se o sistema YIQ na seção 2.3.3.2).
<b><i>Neurônio</i></b>	1. Célula nervosa que transmite impulsos nervosos em organismos biológicos. 2. Cada unidade de redes neurais artificiais (veja <i>unidade</i> ).
<b><i>Nó</i></b>	Veja <i>unidade</i> .
<b><i>Percepção</i></b>	Experiência causada pelo estímulo a um dos sentidos.
<b><i>Radiância</i></b>	Medida da potência de luz proveniente de uma fonte de luz ou de uma superfície.
<b><i>Rampa</i></b>	Veja <i>borda rampa</i> .
<b><i>Refletância</i></b>	Veja <i>radiância</i> .
<b><i>Saturação</i></b>	Atributo da cor que equivale à quantidade de “branco”: o rosa é o vermelho menos saturado.
<b><i>Sensação</i></b>	1. Impressão causada numa formação receptora por um estímulo, e que, por via aferente, é conduzida ao sistema nervoso central. 2. Experiências elementares que, combinadas, formam a percepção.

<b><i>Topologia</i></b>	Relação de vizinhança entre os nós de uma rede neural. Através de conexões laterais, a ativação de uma unidade da rede não altera (fase de aprendizado) somente seus pesos internos, mas também os pesos de vizinhos, criando uma relação entre as unidades que mantém próximos nós com pesos semelhantes.
<b><i>Unidade, Unidade de processamento</i></b>	Cada “neurônio” de uma rede neural artificial, ao qual, geralmente, está associado um peso interno relativo a dados de entrada.

## B. Convenção de notação

A notação relacionada a seguir representa sua utilização mais comum, embora não exclusiva. Variações, quando utilizadas, são indicadas no texto.

$\lambda, \epsilon, \mu, \theta$	Indicam valores escalares relacionados a ângulos, comprimento e onda e erro, por exemplo.
RGB, HSI	Representam espaços de cores.
$R, G, B, r, g, b$ etc.	Indicam valores escalares das componentes cromáticas.
$M, N$	Correspondem a quantidades, como número de pontos ou de amostras.
$i, j, k$	Indicam valores escalares inteiros.
$a, b, x, y$	Representam valores escalares contínuos.
$\Sigma$	Além da indicação de somatórios, é a notação para a matriz de covariância.
$M, S$	Indicam matrizes.
$v, w, \bar{\mu}$	Representam vetores, utilizados tanto no sentido convencional quanto para a representação vetorial das cores.
$f, q$	Correspondem a funções ou espaços vetoriais.
$\bar{x}$	Indica o valor médio para um conjunto de dados. Exceção feita para as funções $\bar{r}, \bar{g}$ e $\bar{b}$ que representam os estímulos tricromáticos.

## C. A transformada de Hough

A transformada de Hough [Hough (1962)] é um procedimento usualmente utilizado para a localização de pontos colineares em imagens binárias. A efetivação da detecção das retas é obtida adotando-se uma representação (equação) que descreva uma reta genérica, como a apresentada na Equação C.1, proposta por Duda e Hart [Duda & Hart (1972)]. A distância da reta à origem é representada por  $\rho$ , enquanto  $\theta$  indica o ângulo que a reta normal forma com o eixo  $x$ .

$$\rho = x \cos\theta + y \sin\theta \quad (\text{C.1})$$

Por cada ponto  $(x, y)$  passam infinitas retas, cada uma delas descritas por valores distintos de  $\rho$  e  $\theta$ . Dois pontos definem, no espaço, uma reta e, através de suas coordenadas é possível encontrar os valores de  $\rho$  e  $\theta$  que caracterizam esta reta. Vários pontos colineares também compartilham os mesmos valores  $(\rho, \theta)$ .

O princípio da transformada de Hough é definir um espaço de parâmetros e fazer o mapeamento de cada ponto do espaço  $xy$  neste novo espaço. Um determinado ponto  $(x, y)$ , pelo qual passam infinitas retas, é descrito no espaço de parâmetros por uma curva senóide, para todos os valores de  $\rho$  e  $\theta$  que descrevem estas retas. Como  $\theta$  é o ângulo, restringir o espaço ao intervalo  $[-\pi/2, \pi/2]$  é suficiente para descrever corretamente as retas. Também ao parâmetro  $\rho$  é permitido assumir valores negativos, que indicam a direção na qual a distância é tomada.

Dentro destas diretrizes, pelos pontos  $(-4, 0)$ ,  $(-2, 5)$ ,  $(0, 10)$ ,  $(-4, 10)$  e  $(2, -5)$ , apresentados na Figura C.1(a), passam as retas determinadas pelos valores de  $\rho$  e  $\theta$  sobre as curvas da apresentadas na Figura C.1(b). As posições em que as curvas se interceptam identificam os parâmetros das retas que passam por dois pontos. Quanto mais curvas se interceptam em uma mesma posição no espaço de parâmetros, maior o número de pontos colineares. Notam-se, na Figura C.1(b), os cruzamentos de duas curvas (indicados como “duplos”), os quais de referem às retas definidas por cada par de pontos da Figura C.1(a). Além destes, cruzamentos de três

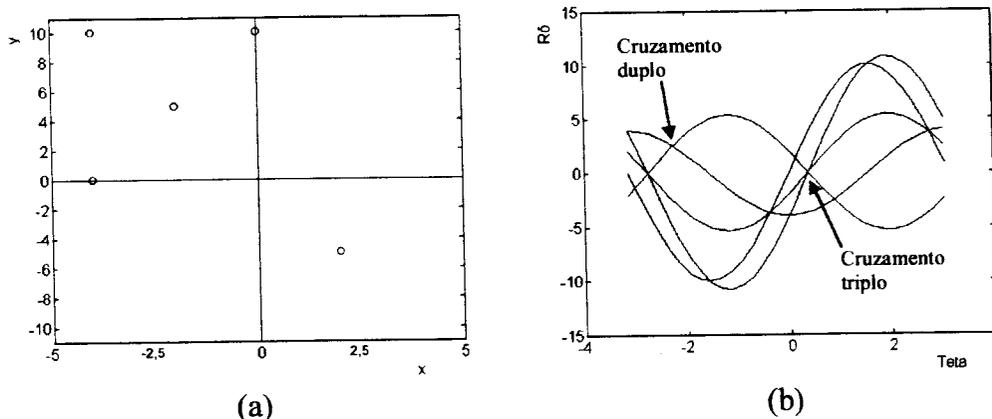


Figura C.1. (a) Conjunto de cinco pontos no espaço  $xy$ . (b) Curvas no espaço de parâmetros correspondentes a estes pontos.

curvas em um mesmo ponto (cruzamento “triplo”) indicam três pontos colineares. Deve-se observar que há quatro, e não apenas dois, cruzamentos triplos, referentes aos ângulos  $\theta$  que são equivalentes, como seriam os ângulos  $\pi/4$  e  $-3\pi/4$ , por exemplo.

Para aplicação em imagens digitais, dimensiona-se um espaço de acumuladores, equivalente a quantizações em  $\rho$  e em  $\theta$  [Schalkoff (1989); Gonzales & Woods (1992)]. Para um ponto da imagem, então, é calculada a senóide correspondente no espaço de parâmetros, considerando-se a origem no centro da imagem. Cada célula dos acumuladores interceptada pela curva tem seu conteúdo incrementado em uma unidade, sendo este processo repetido para os demais pontos. As células acumuladoras situadas nos cruzamentos das curvas identificam, pela contagem maior, que há um número maior de pontos na imagem que são colineares e compartilham aqueles parâmetros específicos (coordenadas  $\rho$  e  $\theta$ ). Localizando-se as células com maior valor também são identificadas as retas mais significativas presentes na imagem.

Obtidas as equações das retas é possível, então, traçá-las novamente sobre a imagem e verificar, assim, a extensão dos segmentos de reta que efetivamente provocaram a identificação da reta. Os segmentos de reta da imagem podem, deste modo, ser identificados. Além disso, uma análise de conectividade pode ser empregada para melhorar os segmentos [Costa & Sandler (1993)].

É importante, ainda, observar que outras formas de parametrização de retas podem ser utilizadas, como é o caso de se usar, por exemplo, a apresentada na Equação C.2. Esta forma, em particular, introduz uma dificuldade para retas verticais ou quase verticais, quando o valor de  $a$  deve ser muito alto e a representação computacional do espaço de parâmetros é mais complexa [Duda & Hart (1972); Schalkoff (1989)].

$$y = ax + b \quad (C.2)$$

### I. Rastreamento (*Backmapping*)

Em 1986, Gerig e Klein [Gerig & Klein (1986)] introduziram um passo adicional à transformada de Hough, incluindo um “rastreamento”, denominado *backmapping*, da transformada realizada. Um novo arranjo de acumuladores com as mesmas características do espaço de parâmetros é definido e, novamente, a transformada é calculada para cada ponto da imagem. Para este novo espaço, porém, somente são incrementadas as posições que correspondam ao acumulador de valor mais alto do espaço usual. Em outras palavras, para cada ponto da imagem é recalculada a senóide correspondente; sobre o espaço de acumuladores é traçada a curva, verificando-se, para as posições sobrepostas pela senóide, o acumulador que apresenta maior valor; somente a posição correspondente a este acumulador é incrementada no novo espaço.

Esta abordagem corresponde a um “reforço” aplicado à transformada mais usual, reduzindo drasticamente os picos causados por pontos colineares menos significativos (por exemplo, segmentos com poucos pontos ou interferência entre objetos distintos na imagem).

### II. Generalizações da transformada de Hough

Além da parametrização das retas, é possível parametrizar outras curvas, como círculos ou elipses, e aplicar um processo similar de incrementos de acumuladores no espaço de parâmetros. Extensões da transformada de Hough para a detecção de outros padrões são, portanto, viáveis. [Schalkoff (1989)]

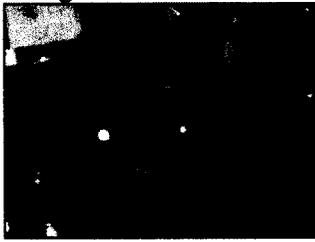
## D. Resultados para as diferentes técnicas de segmentação

Este apêndice apresenta, de forma mais extensiva, os resultados das segmentações das imagens. A Tabela D-I descreve as legendas utilizadas nas imagens.

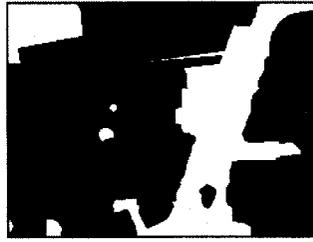
Tabela D-I. Descrição das legendas para a figuras do apêndice.

<b>Legenda</b>	<b>Descrição</b>
Original	Imagem colorida original, de tamanho 180×240 pontos.
Regiões de referência	Imagem de regiões de referência, obtida manualmente.
Mapa de referência	Imagem de bordas de referência, obtida a partir da imagem de regiões.
Regiões cromaticidade	Segmentação através de mapas auto-organizáveis utilizando cromaticidade como dados.
Mapa cromaticidade	Mapa de bordas derivado da segmentação por mapa auto-organizável usando cromaticidade.
Sobel (manual)	Mapa de bordas obtido pela aplicação de um limiar arbitrário à magnitude do gradiente calculado pelas máscaras de Sobel.
Regiões Bayes	Segmentação por classificação bayesiana.
Mapa Bayes	Mapa de bordas derivado da classificação bayesiana.
Regiões $k$ NN	Segmentação pelo método dos $k$ vizinhos mais próximos.
Mapa $k$ NN	Mapa de bordas derivado da classificação por $k$ vizinhos mais próximos.
Sobel (adaptável local)	Mapa de bordas através da escolha automática de um limiar local para a magnitude do gradiente calculado pelas máscaras de Sobel.
Sobel (adaptável global)	Mapa de bordas através da escolha automática de um limiar global para a magnitude do gradiente calculado pelas máscaras de Sobel.
Gradiente colorido	Mapa de bordas usando escolha arbitrária de um limiar para o gradiente calculado sobre os vetores de cor.
Regiões IC	Segmentação por mapas auto-organizáveis utilizando a intensidade e duas componentes de cromaticidade da imagem.
Mapa IC	Mapa de bordas derivado das regiões da segmentação por mapa auto-organizável utilizando intensidade e cromaticidade.
Regiões I	Segmentação por mapas auto-organizáveis utilizando as componentes de cor (RGB) da imagem.
Mapa I	Mapa de bordas derivado das regiões da segmentação por mapa auto-organizável utilizando as componentes de cor.

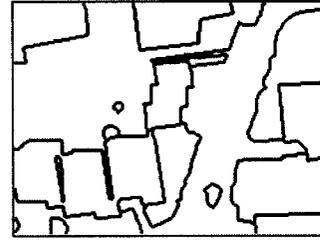
Imagem: Aerial01



Original



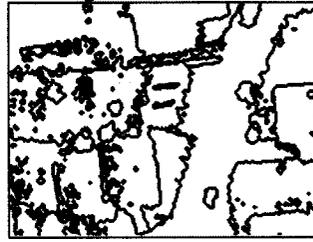
Regiões de referência



Mapa de referência



Regiões cromaticidade



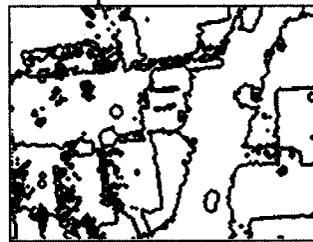
Mapa cromaticidade



Sobel (manual)



Regiões Bayes



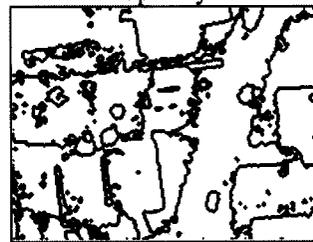
Mapa Bayes



Sobel (adaptável local)



Regiões kNN



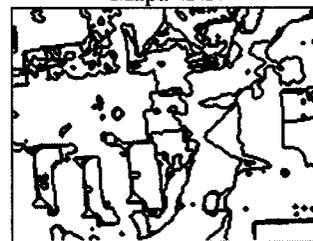
Mapa kNN



Sobel (adaptável global)



Regiões IC



Mapa IC



Gradiente colorido



Regiões I

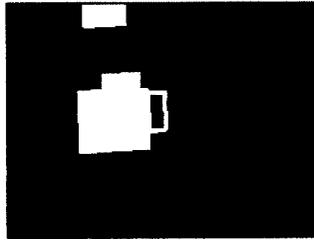


Mapa I

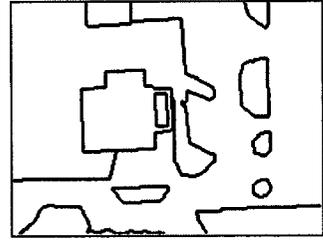
Imagem: Aerial102



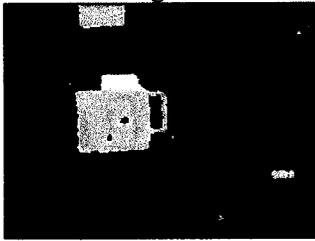
Original



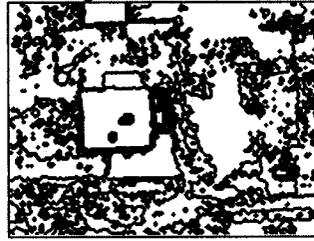
Regiões de referência



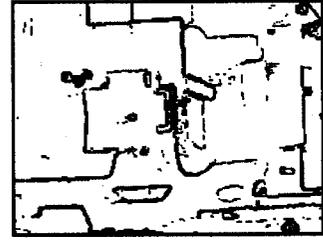
Mapa de referência



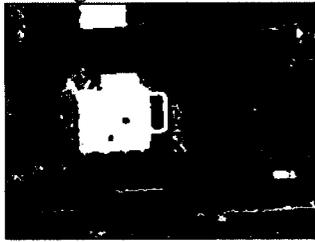
Regiões cromaticidade



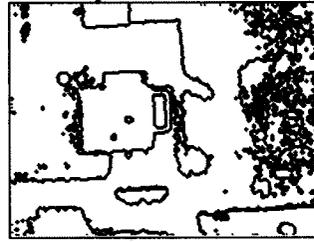
Mapa cromaticidade



Sobel (manual)



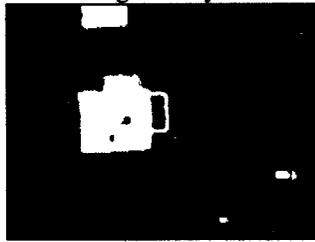
Regiões Bayes



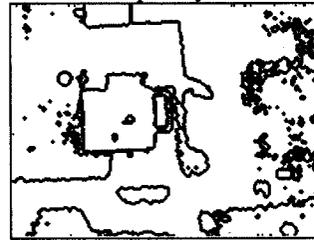
Mapa Bayes



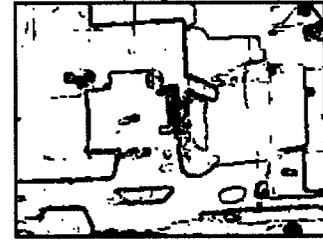
Sobel (adaptável local)



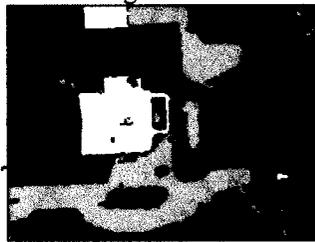
Regiões kNN



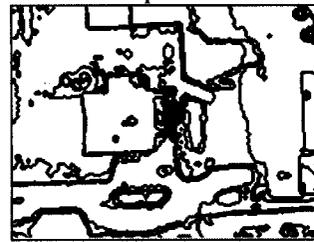
Mapa kNN



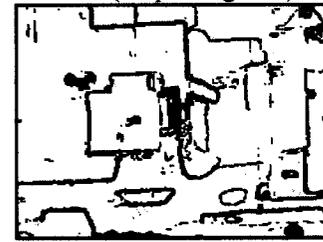
Sobel (adaptável global)



Regiões IC



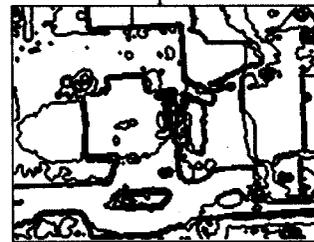
Mapa IC



Gradiente colorido



Regiões I



Mapa I

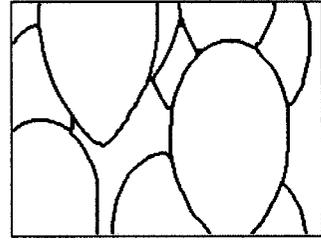
Imagem: Balloons



Original



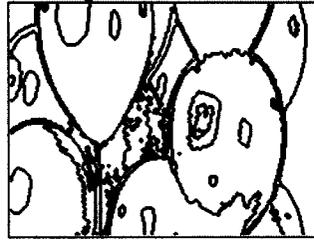
Regiões de referência



Mapa de referência



Regiões cromaticidade



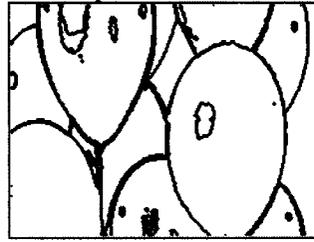
Mapa cromaticidade



Sobel (manual)



Regiões Bayes



Mapa Bayes



Sobel (adaptável local)



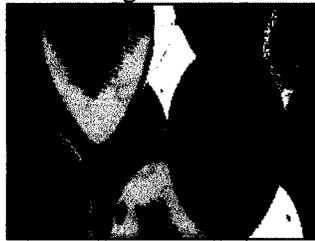
Regiões kNN



Mapa kNN



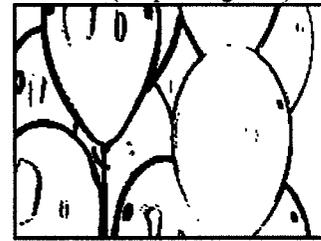
Sobel (adaptável global)



Regiões IC



Mapa IC



Gradiente colorido

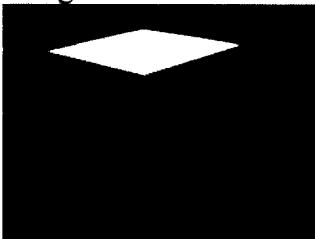


Regiões I

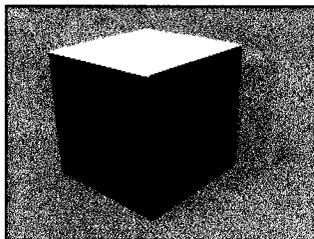


Mapa I

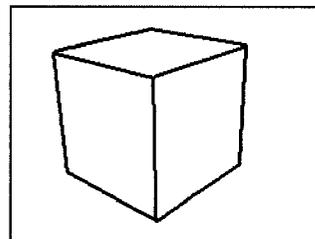
Imagem: Cube009



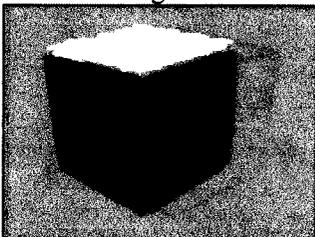
Original



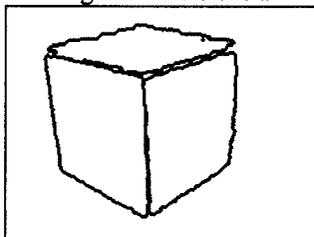
Regiões de referência



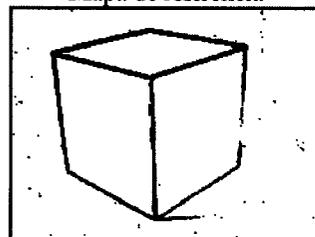
Mapa de referência



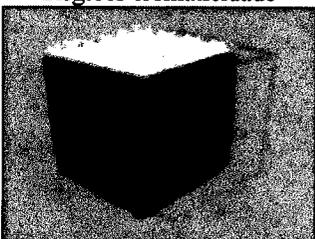
Regiões cromaticidade



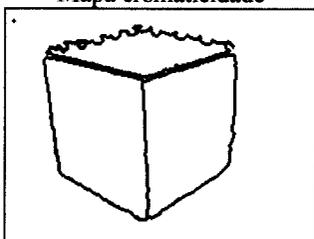
Mapa cromaticidade



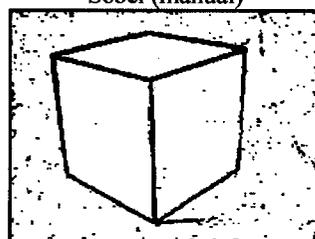
Sobel (manual)



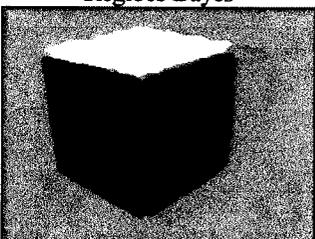
Regiões Bayes



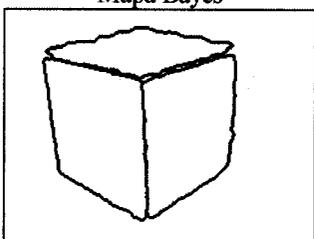
Mapa Bayes



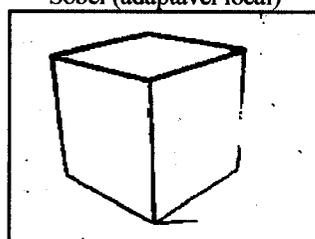
Sobel (adaptável local)



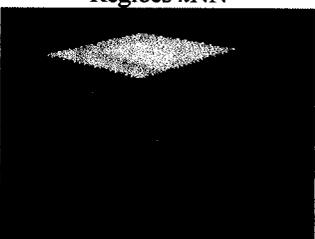
Regiões  $kNN$



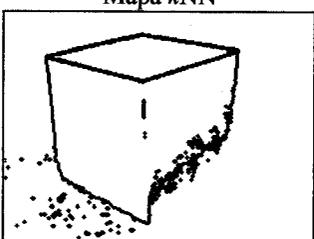
Mapa  $kNN$



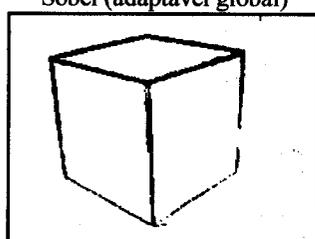
Sobel (adaptável global)



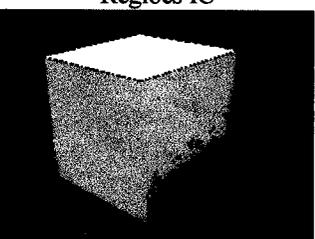
Regiões IC



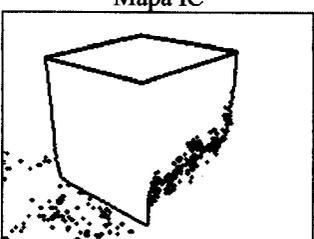
Mapa IC



Gradiente colorido

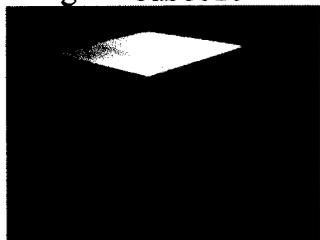


Regiões I

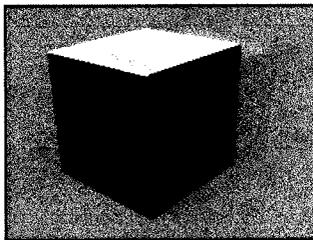


Mapa I

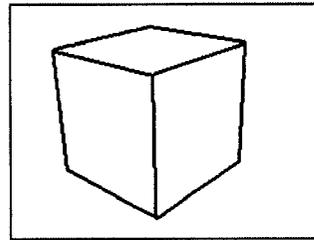
Imagem: Cube010



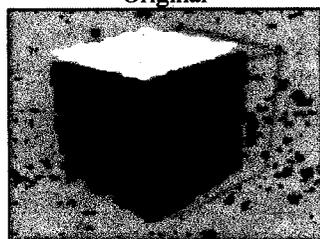
Original



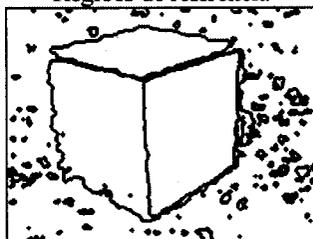
Regiões de referência



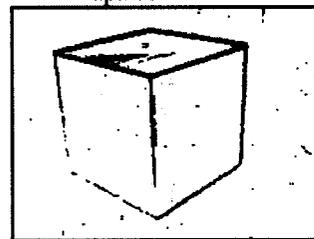
Mapa de referência



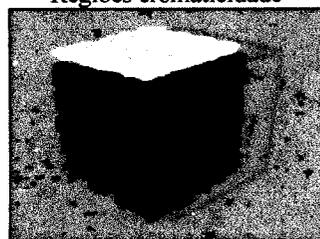
Regiões cromaticidade



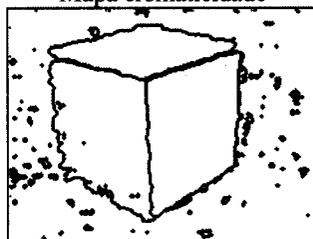
Mapa cromaticidade



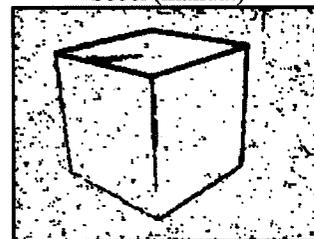
Sobel (manual)



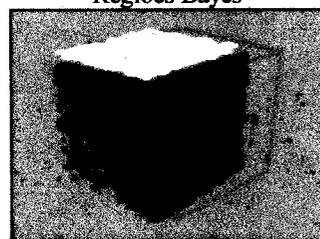
Regiões Bayes



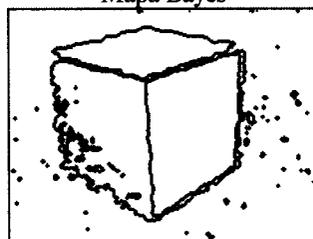
Mapa Bayes



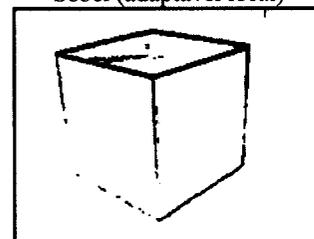
Sobel (adaptável local)



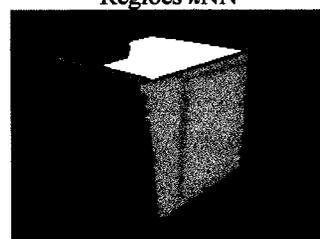
Regiões  $kNN$



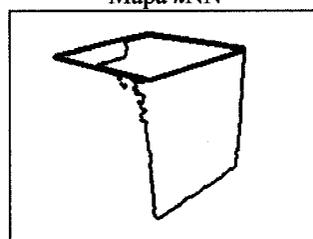
Mapa  $kNN$



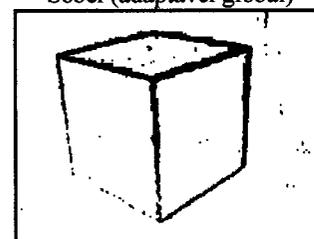
Sobel (adaptável global)



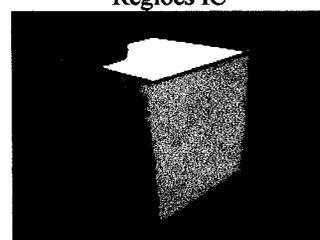
Regiões IC



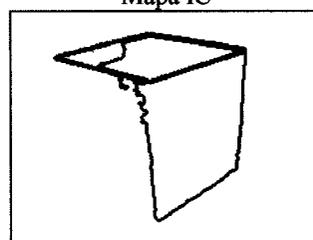
Mapa IC



Gradiente colorido

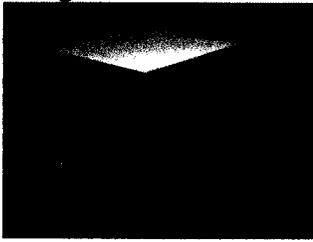


Regiões I

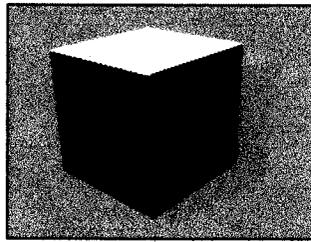


Mapa I

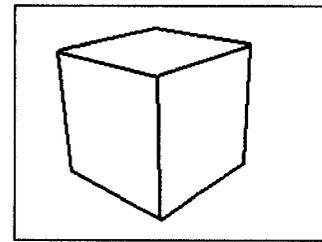
Imagem: Cube011



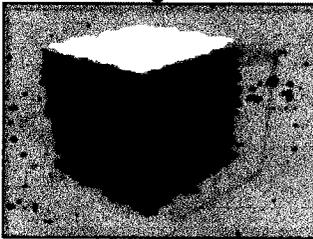
Original



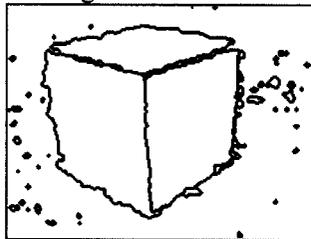
Regiões de referência



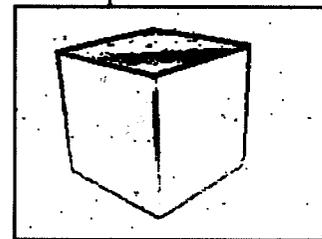
Mapa de referência



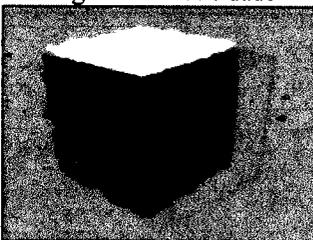
Regiões cromaticidade



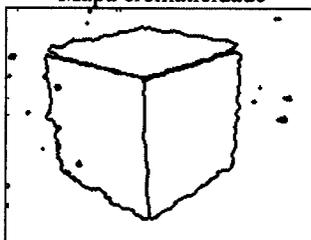
Mapa cromaticidade



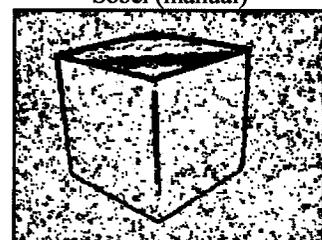
Sobel (manual)



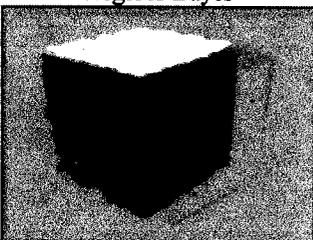
Regiões Bayes



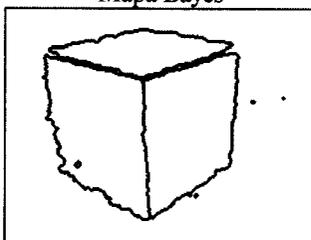
Mapa Bayes



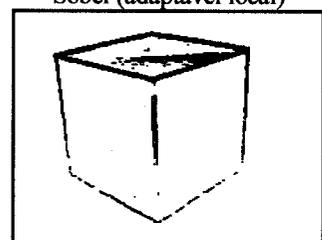
Sobel (adaptável local)



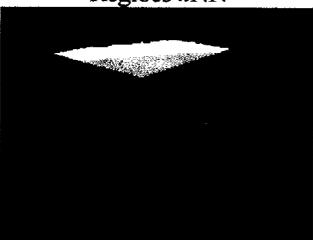
Regiões kNN



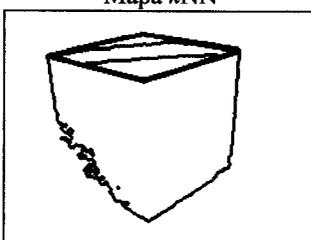
Mapa kNN



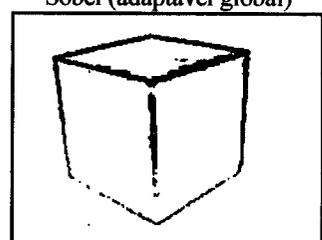
Sobel (adaptável global)



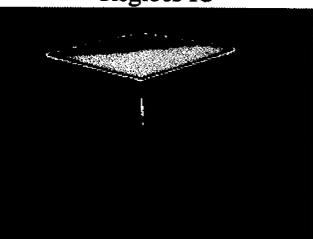
Regiões IC



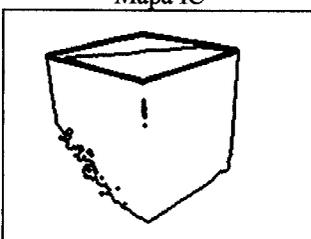
Mapa IC



Gradiente colorido

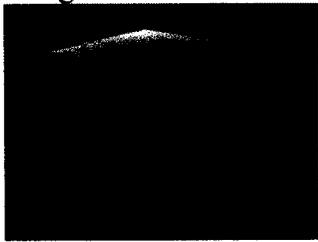


Regiões I

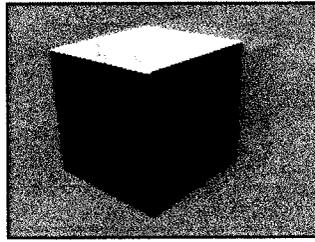


Mapa I

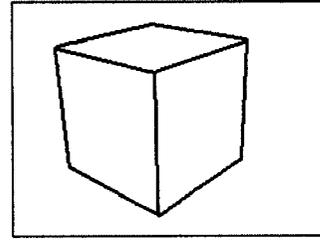
Imagem: Cube014



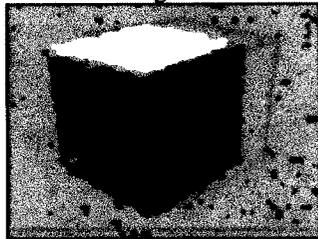
Original



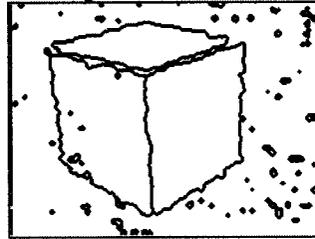
Regiões de referência



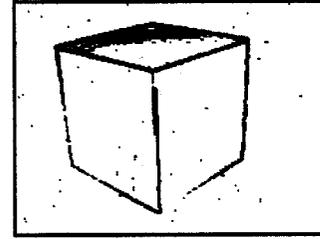
Mapa de referência



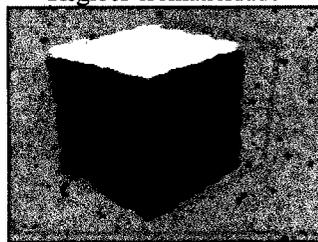
Regiões cromaticidade



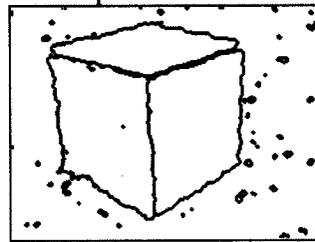
Mapa cromaticidade



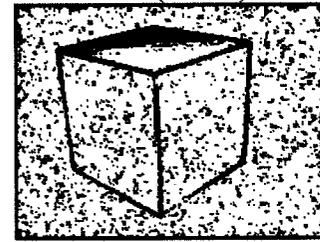
Sobel (manual)



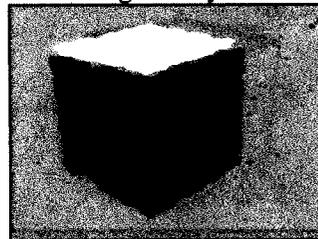
Regiões Bayes



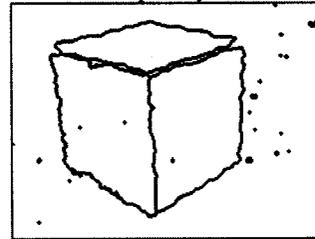
Mapa Bayes



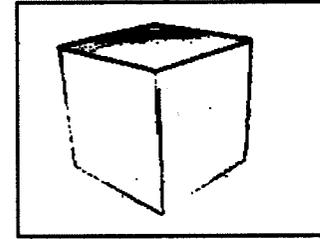
Sobel (adaptável local)



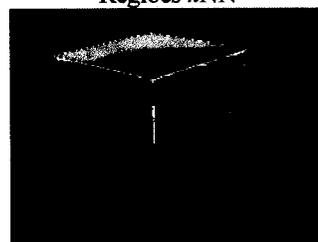
Regiões kNN



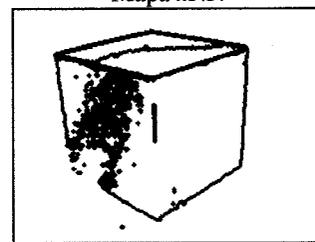
Mapa kNN



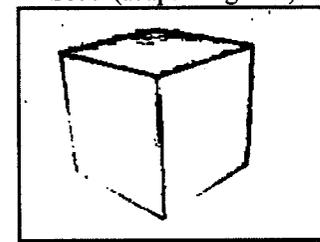
Sobel (adaptável global)



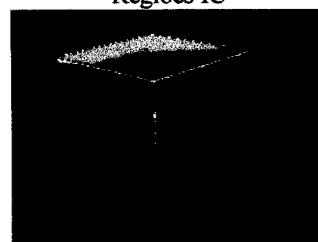
Regiões IC



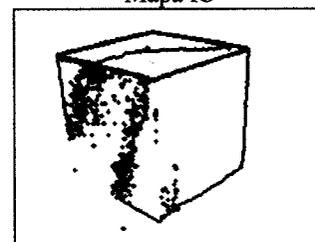
Mapa IC



Gradiente colorido



Regiões I

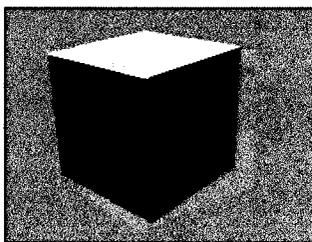


Mapa I

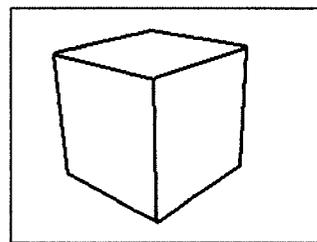
Imagem: Cube015



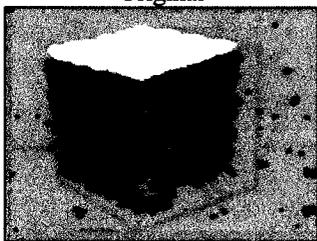
Original



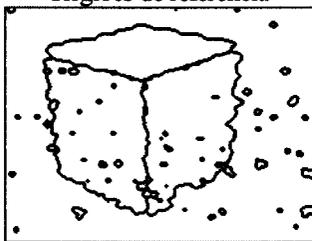
Regiões de referência



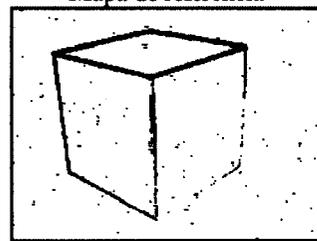
Mapa de referência



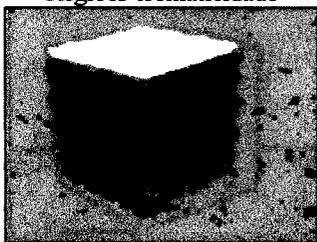
Regiões cromaticidade



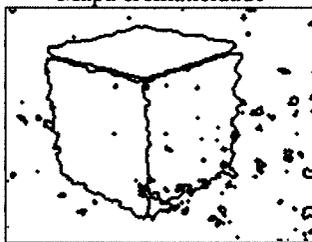
Mapa cromaticidade



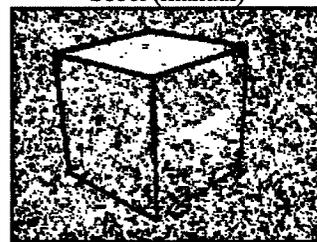
Sobel (manual)



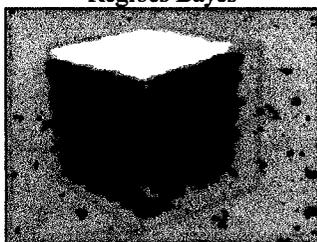
Regiões Bayes



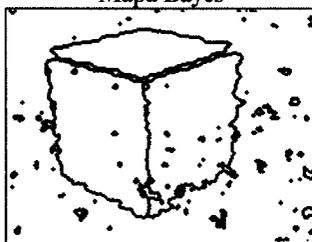
Mapa Bayes



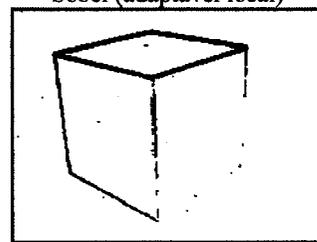
Sobel (adaptável local)



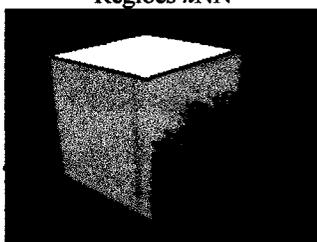
Regiões  $k$ NN



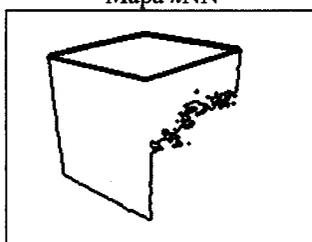
Mapa  $k$ NN



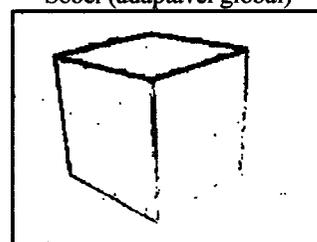
Sobel (adaptável global)



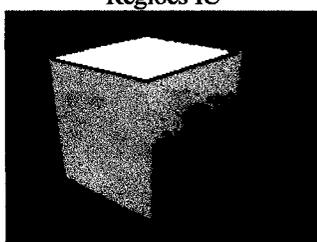
Regiões IC



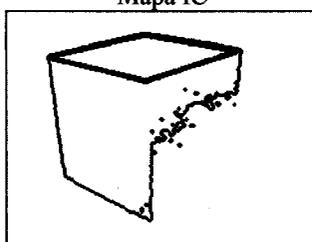
Mapa IC



Gradiente colorido

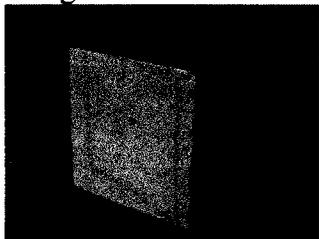


Regiões I

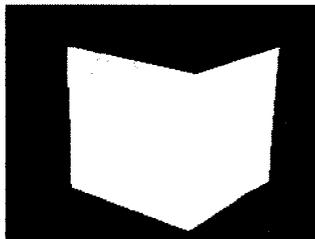


Mapa I

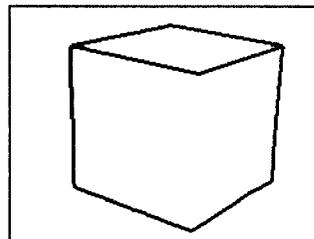
Imagem: Cube101



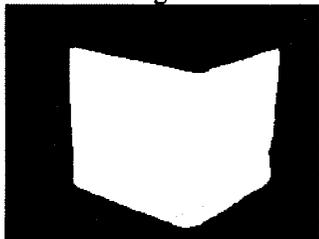
Original



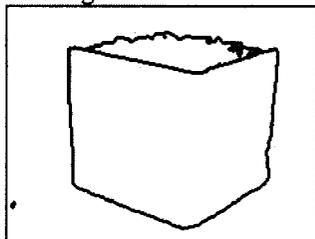
Regiões de referência



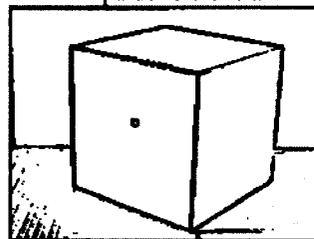
Mapa de referência



Regiões cromaticidade



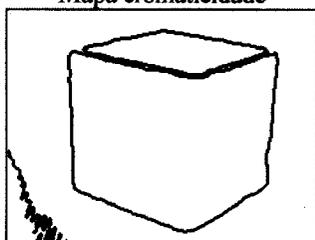
Mapa cromaticidade



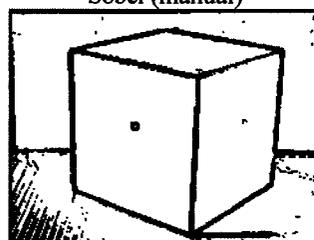
Sobel (manual)



Regiões Bayes



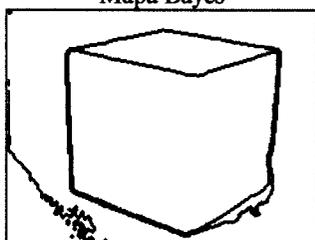
Mapa Bayes



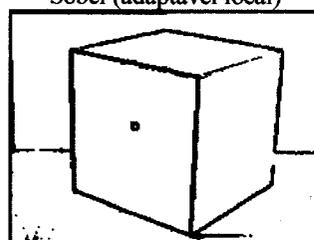
Sobel (adaptável local)



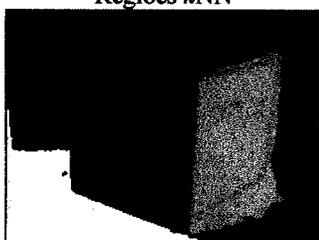
Regiões kNN



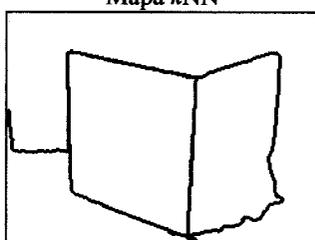
Mapa kNN



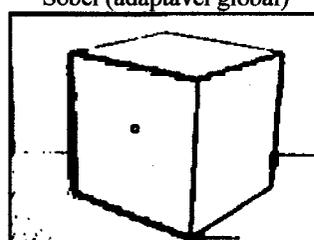
Sobel (adaptável global)



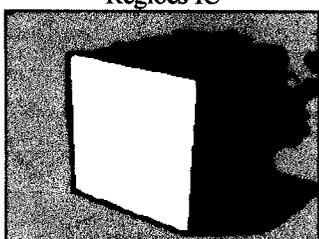
Regiões IC



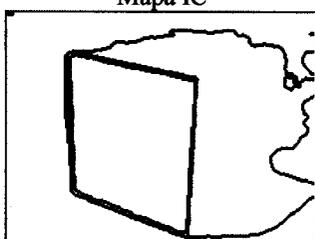
Mapa IC



Gradiente colorido

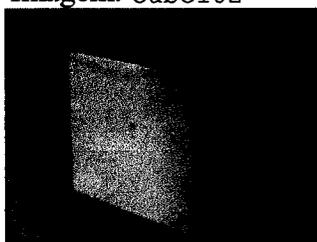


Regiões I

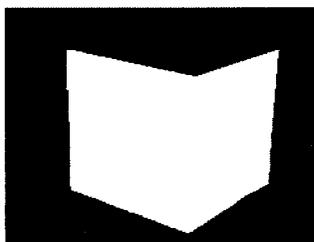


Mapa I

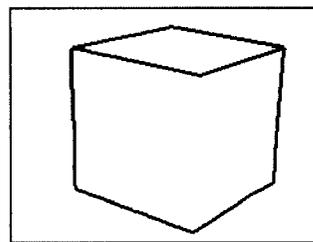
Imagem: Cube102



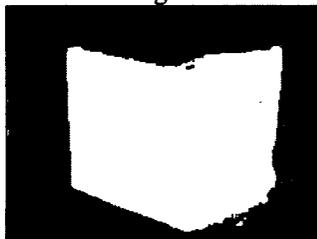
Original



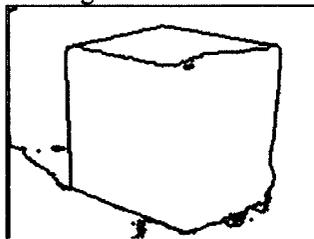
Regiões de referência



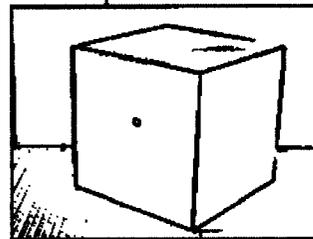
Mapa de referência



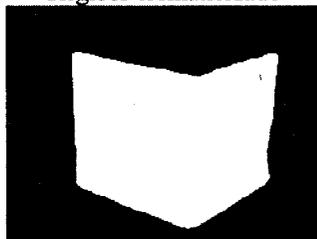
Regiões cromaticidade



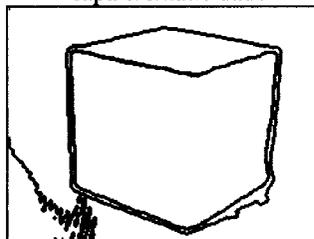
Mapa cromaticidade



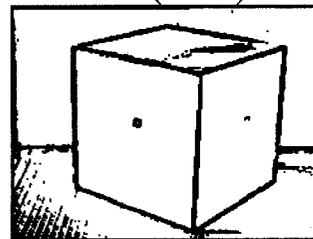
Sobel (manual)



Regiões Bayes



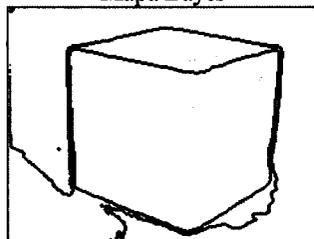
Mapa Bayes



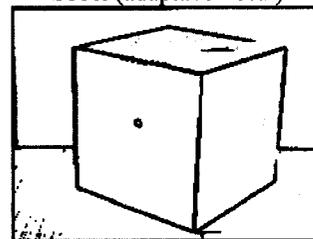
Sobel (adaptável local)



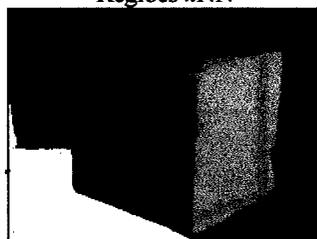
Regiões kNN



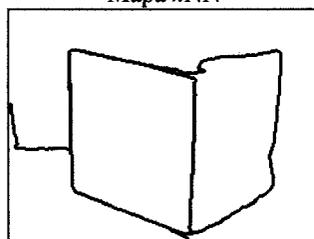
Mapa kNN



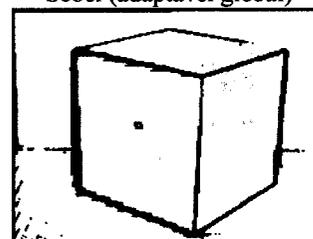
Sobel (adaptável global)



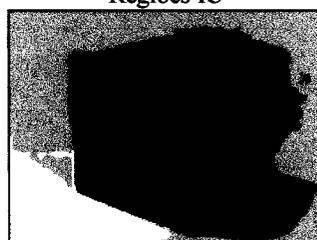
Regiões IC



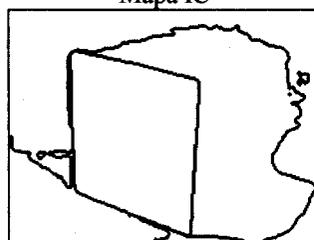
Mapa IC



Gradiente colorido



Regiões I

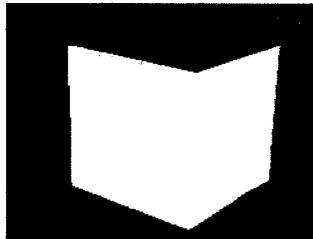


Mapa I

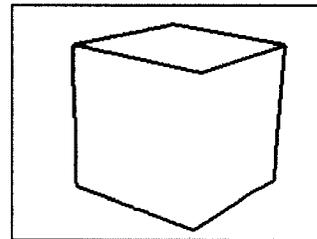
Imagem: Cube104



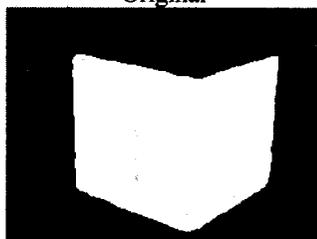
Original



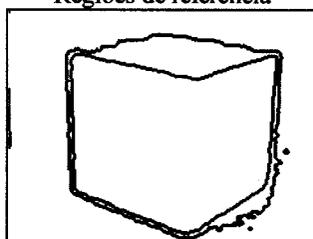
Regiões de referência



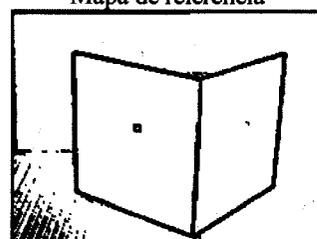
Mapa de referência



Regiões cromaticidade



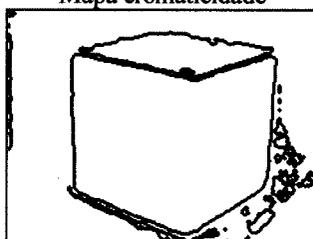
Mapa cromaticidade



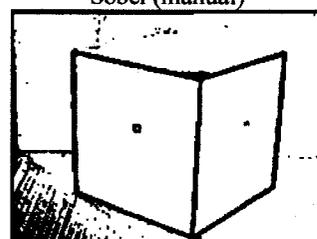
Sobel (manual)



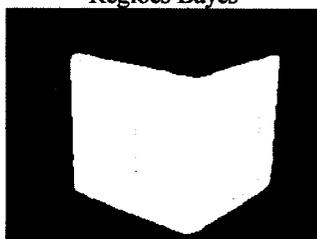
Regiões Bayes



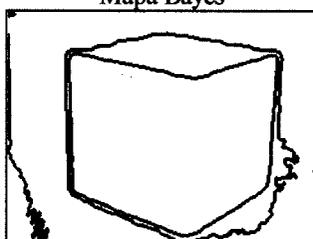
Mapa Bayes



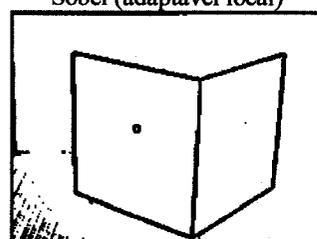
Sobel (adaptável local)



Regiões kNN



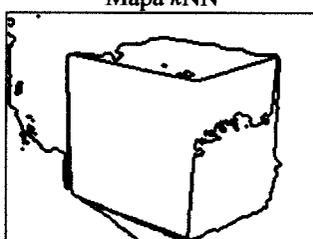
Mapa kNN



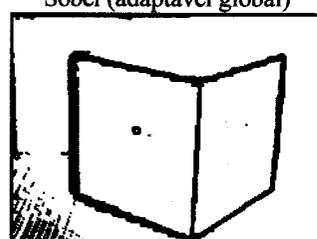
Sobel (adaptável global)



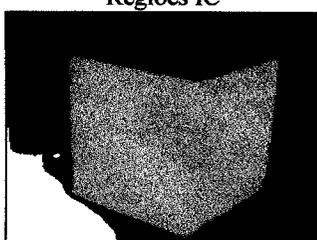
Regiões IC



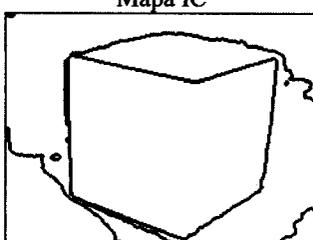
Mapa IC



Gradiente colorido

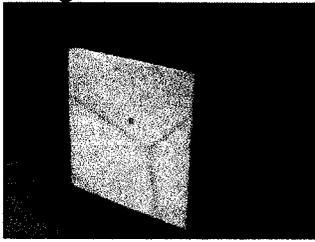


Regiões I

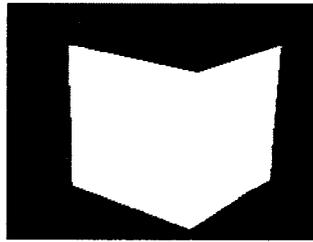


Mapa I

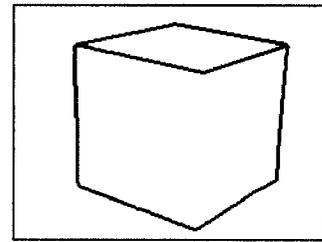
Imagem: Cube105



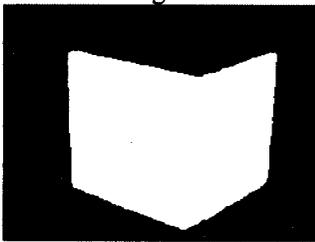
Original



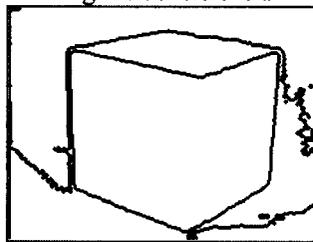
Regiões de referência



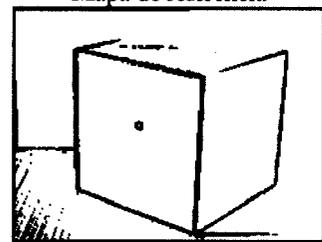
Mapa de referência



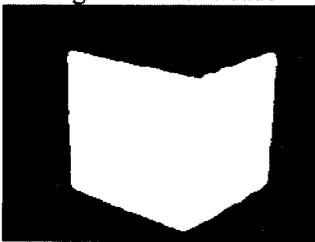
Regiões cromaticidade



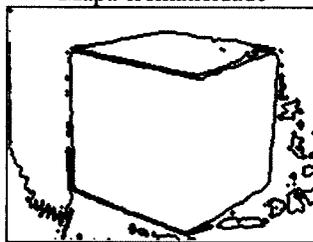
Mapa cromaticidade



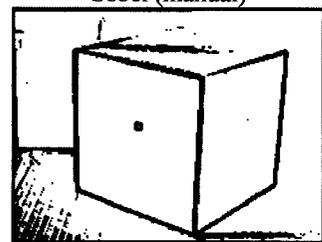
Sobel (manual)



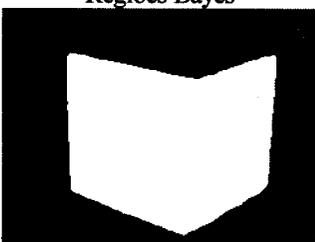
Regiões Bayes



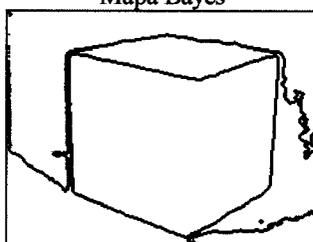
Mapa Bayes



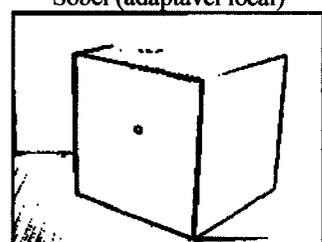
Sobel (adaptável local)



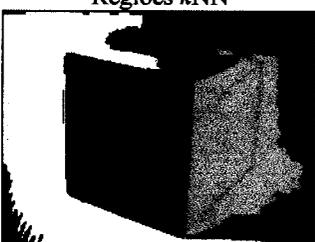
Regiões kNN



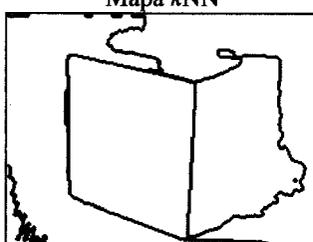
Mapa kNN



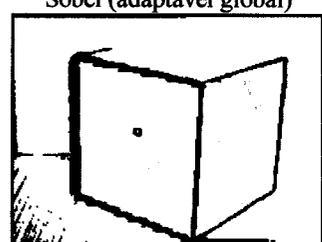
Sobel (adaptável global)



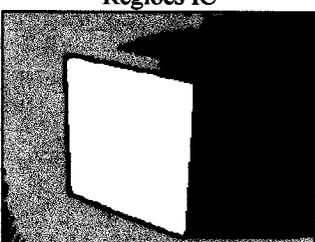
Regiões IC



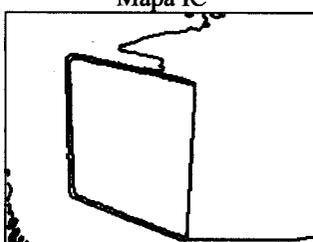
Mapa IC



Gradiente colorido

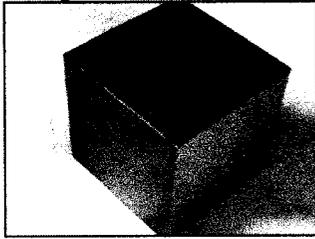


Regiões I

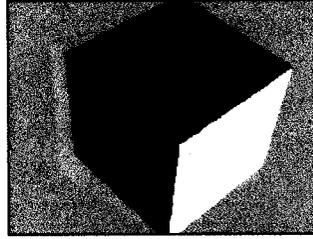


Mapa I

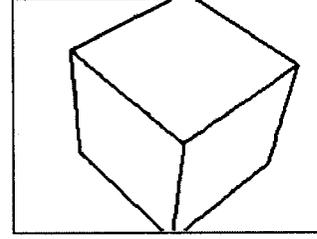
Imagem: Cube202



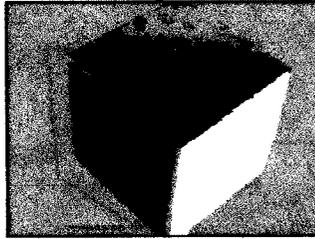
Original



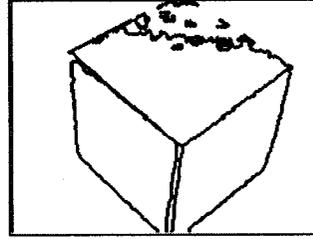
Regiões de referência



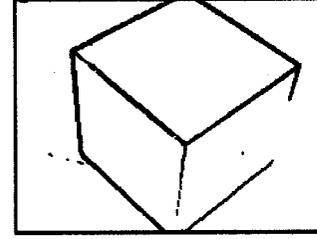
Mapa de referência



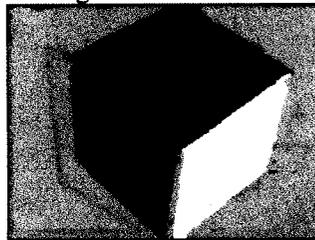
Regiões cromaticidade



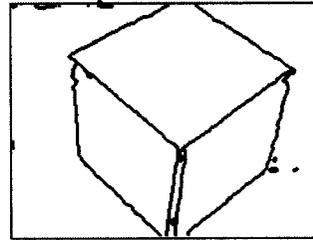
Mapa cromaticidade



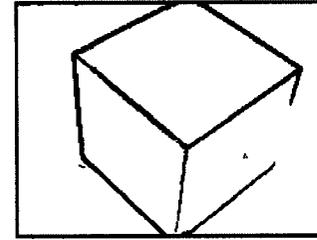
Sobel (manual)



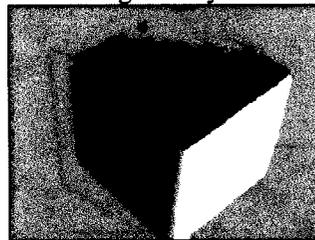
Regiões Bayes



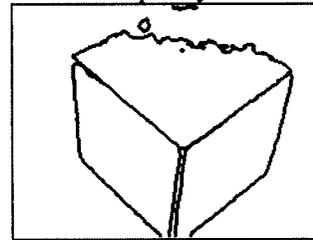
Mapa Bayes



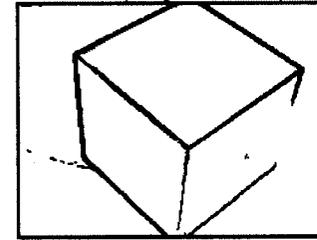
Sobel (adaptável local)



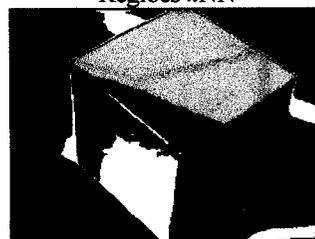
Regiões kNN



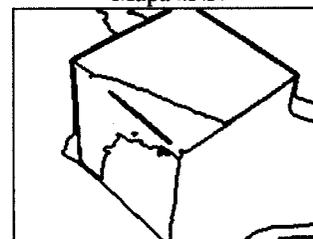
Mapa kNN



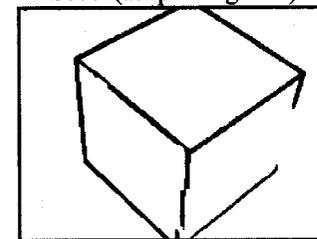
Sobel (adaptável global)



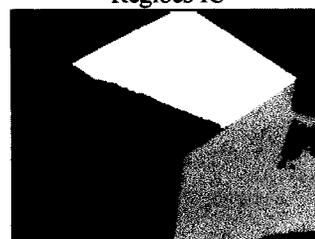
Regiões IC



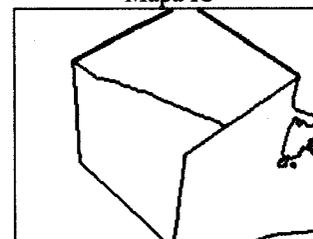
Mapa IC



Gradiente colorido

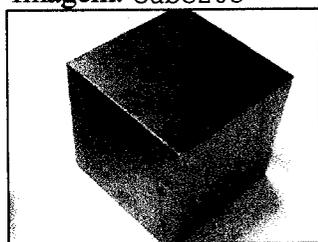


Regiões I

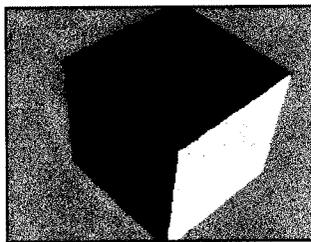


Mapa I

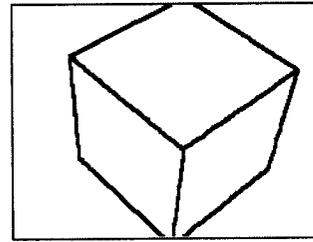
Imagem: Cube203



Original



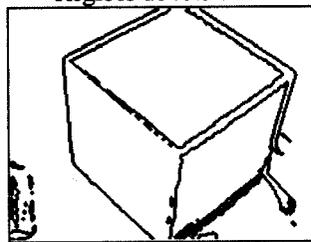
Regiões de referência



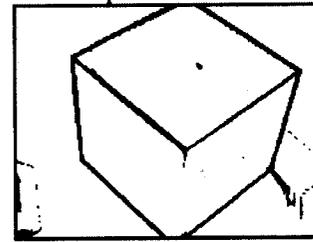
Mapa de referência



Regiões cromaticidade



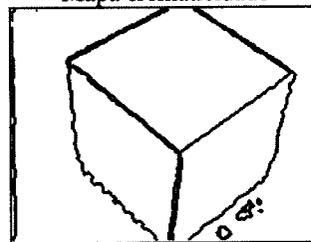
Mapa cromaticidade



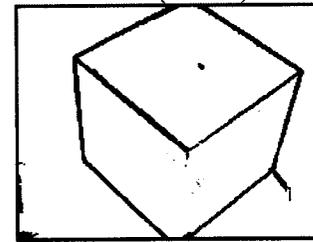
Sobel (manual)



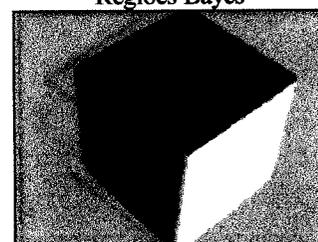
Regiões Bayes



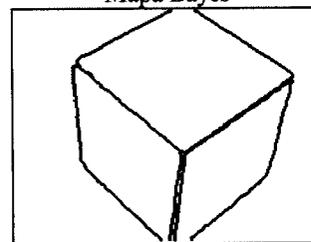
Mapa Bayes



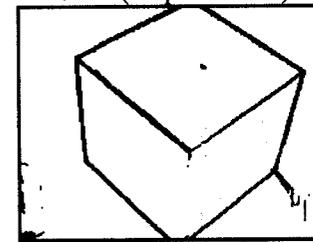
Sobel (adaptável local)



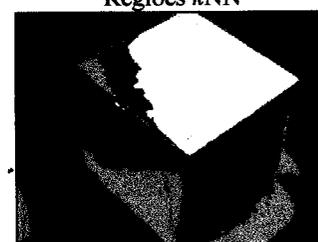
Regiões  $k$ NN



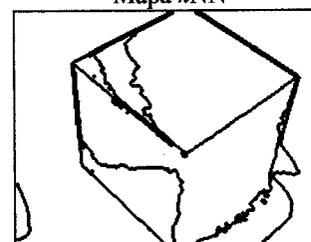
Mapa  $k$ NN



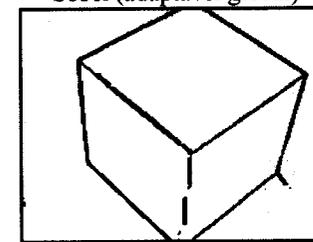
Sobel (adaptável global)



Regiões IC



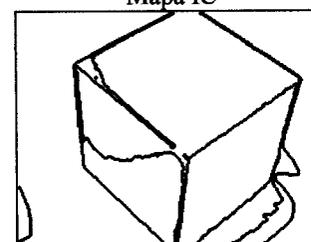
Mapa IC



Gradiente colorido



Regiões I

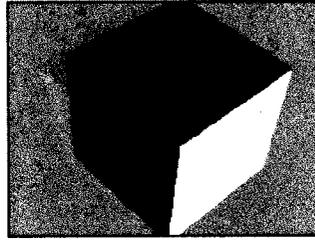


Mapa I

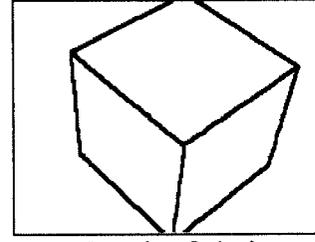
Imagem: Cube204



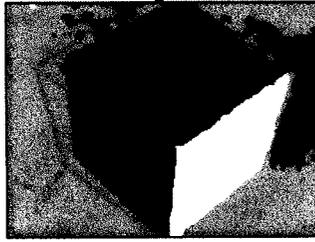
Original



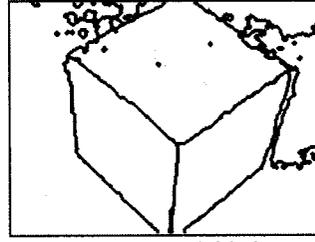
Regiões de referência



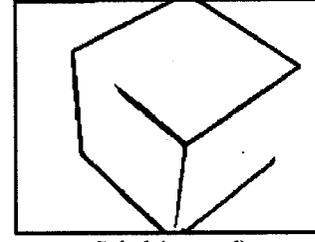
Mapa de referência



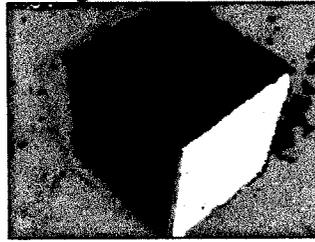
Regiões cromaticidade



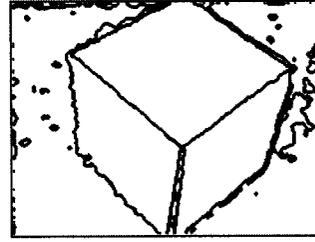
Mapa cromaticidade



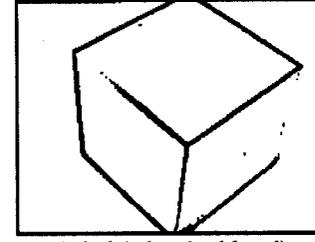
Sobel (manual)



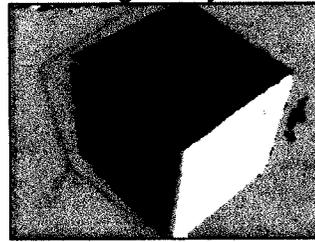
Regiões Bayes



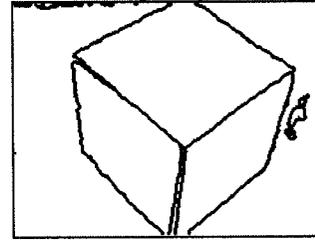
Mapa Bayes



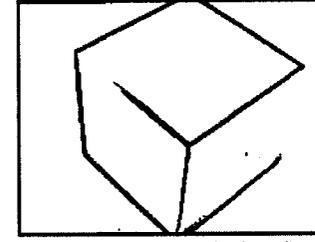
Sobel (adaptável local)



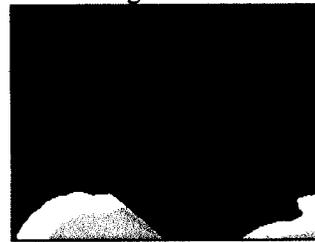
Regiões kNN



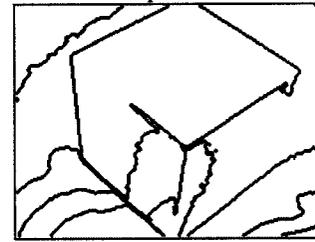
Mapa kNN



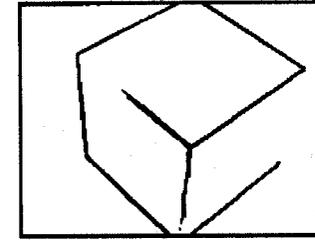
Sobel (adaptável global)



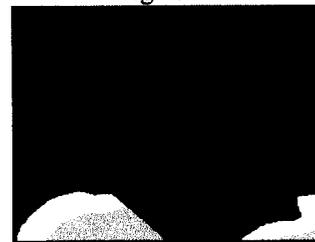
Regiões IC



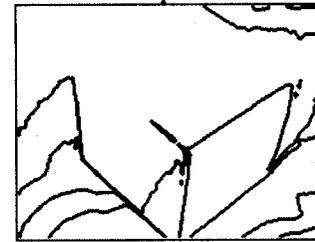
Mapa IC



Gradiente colorido

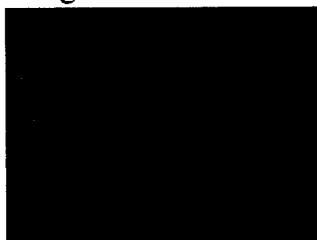


Regiões I



Mapa I

Imagem: Donna



Original



Regiões de referência



Mapa de referência



Regiões cromaticidade



Mapa cromaticidade



Sobel (manual)



Regiões Bayes



Mapa Bayes



Sobel (adaptável local)



Regiões  $k$ NN



Mapa  $k$ NN



Sobel (adaptável global)



Regiões IC



Mapa IC



Gradiente colorido

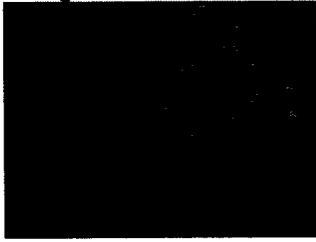


Regiões I

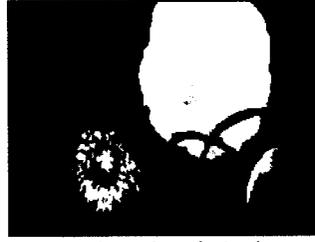


Mapa I

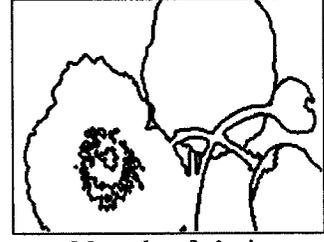
Imagem: Flowers



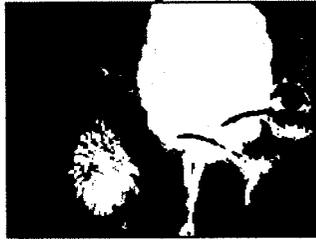
Original



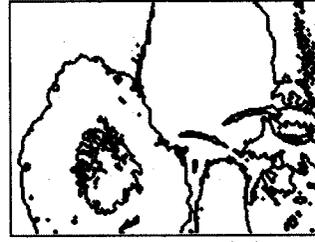
Regiões de referência



Mapa de referência



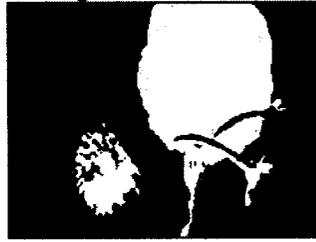
Regiões cromaticidade



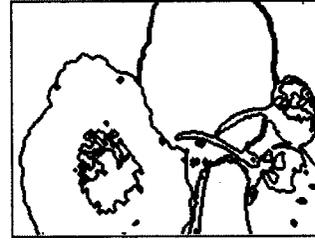
Mapa cromaticidade



Sobel (manual)



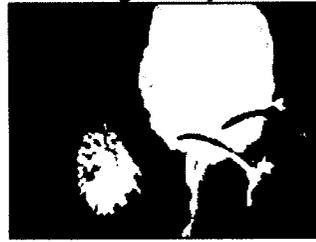
Regiões Bayes



Mapa Bayes



Sobel (adaptável local)



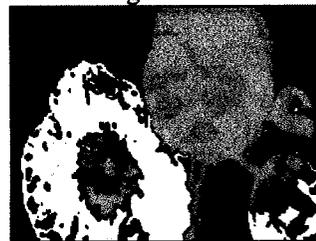
Regiões kNN



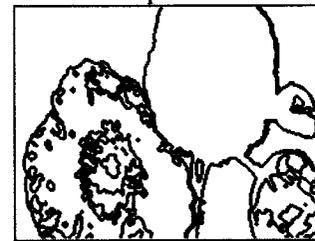
Mapa kNN



Sobel (adaptável global)



Regiões IC



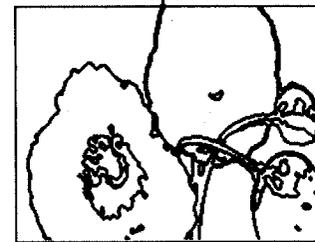
Mapa IC



Gradiente colorido



Regiões I



Mapa I

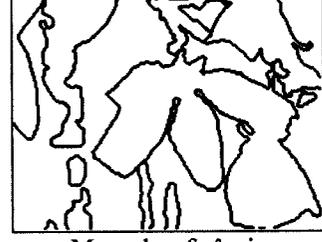
Imagem: Flower02



Original



Regiões de referência



Mapa de referência



Regiões cromaticidade



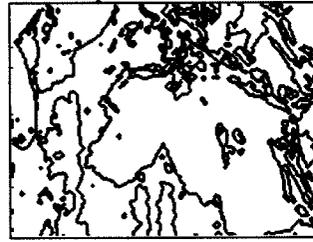
Mapa cromaticidade



Sobel (manual)



Regiões Bayes



Mapa Bayes



Sobel (adaptável local)



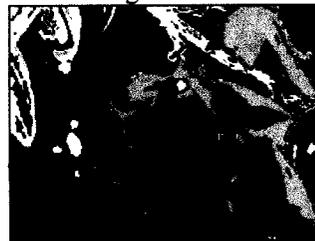
Regiões kNN



Mapa kNN



Sobel (adaptável global)



Regiões IC



Mapa IC



Gradiente colorido



Regiões I



Mapa I

Imagem: Frogs01



Original



Regiões de referência



Mapa de referência



Regiões cromaticidade



Mapa cromaticidade



Sobel (manual)



Regiões Bayes



Mapa Bayes



Sobel (adaptável local)



Regiões kNN



Mapa kNN



Sobel (adaptável global)



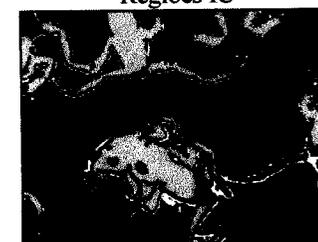
Regiões IC



Mapa IC



Gradiente colorido



Regiões I

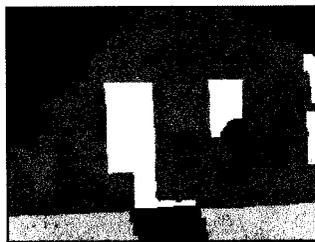


Mapa I

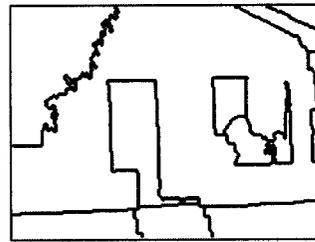
Imagem: House01



Original



Regiões de referência



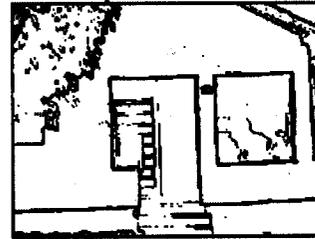
Mapa de referência



Regiões cromaticidade



Mapa cromaticidade



Sobel (manual)



Regiões Bayes



Mapa Bayes



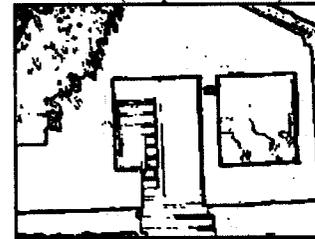
Sobel (adaptável local)



Regiões kNN



Mapa kNN



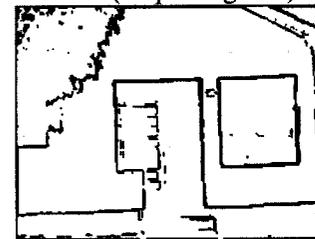
Sobel (adaptável global)



Regiões IC



Mapa IC



Gradiente colorido



Regiões I



Mapa I

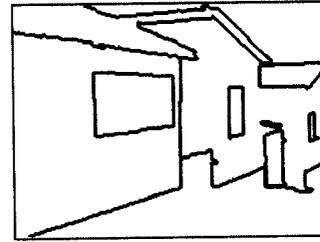
Imagem: House02



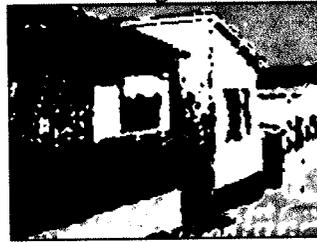
Original



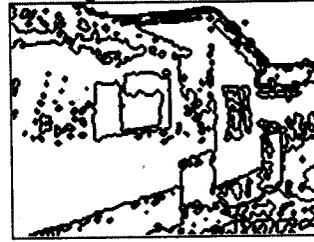
Regiões de referência



Mapa de referência



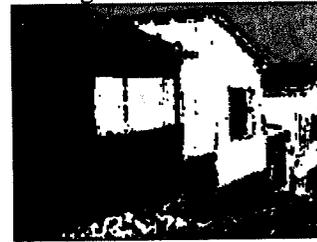
Regiões cromaticidade



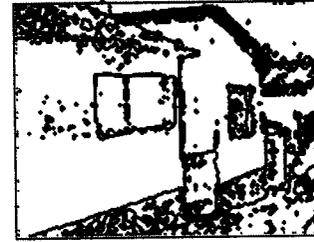
Mapa cromaticidade



Sobel (manual)



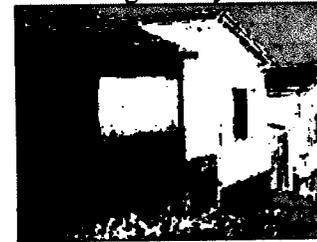
Regiões Bayes



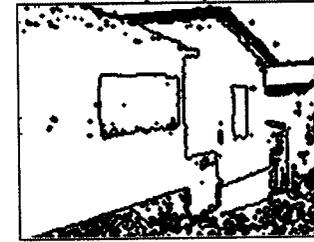
Mapa Bayes



Sobel (adaptável local)



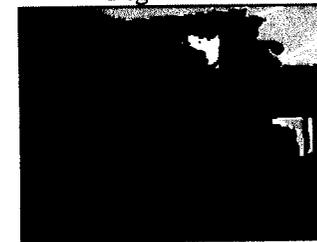
Regiões kNN



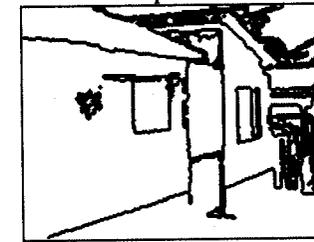
Mapa kNN



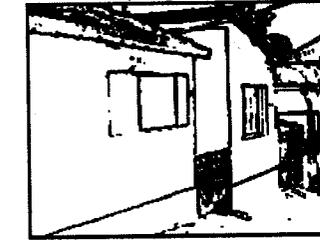
Sobel (adaptável global)



Regiões IC



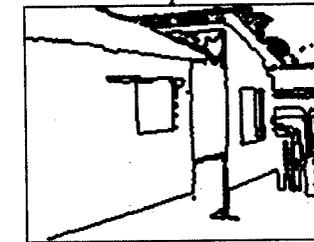
Mapa IC



Gradiente colorido



Regiões I



Mapa I

Imagem: House03



Original



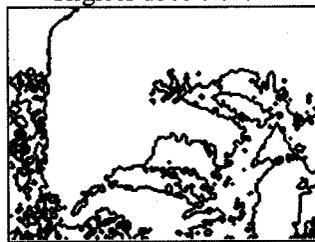
Regiões de referência



Mapa de referência



Regiões cromaticidade



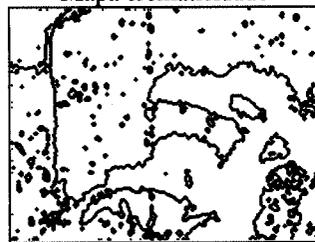
Mapa cromaticidade



Sobel (manual)



Regiões Bayes



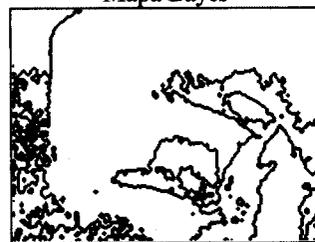
Mapa Bayes



Sobel (adaptável local)



Regiões kNN



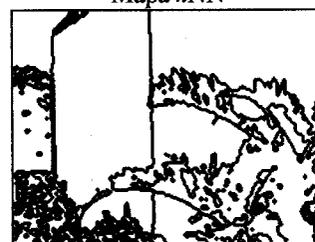
Mapa kNN



Sobel (adaptável global)



Regiões IC



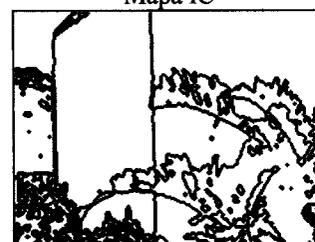
Mapa IC



Gradiente colorido

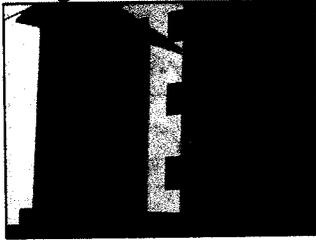


Regiões I

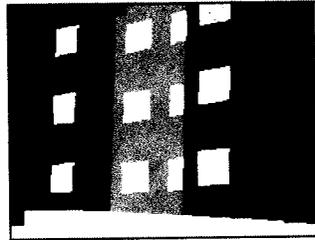


Mapa I

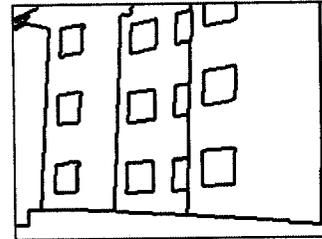
Imagem: House05



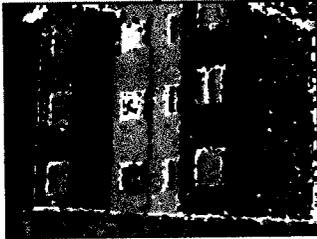
Original



Regiões de referência



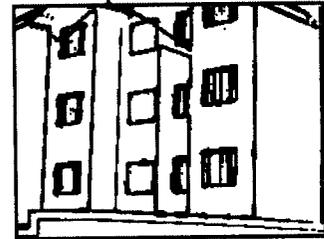
Mapa de referência



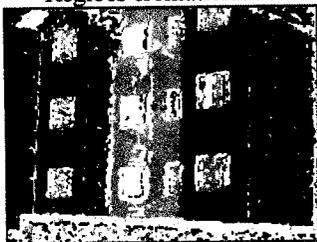
Regiões cromaticidade



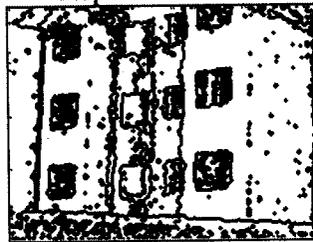
Mapa cromaticidade



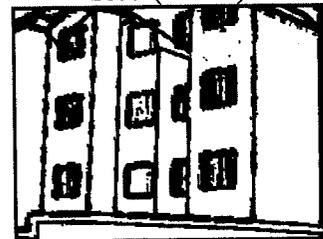
Sobel (manual)



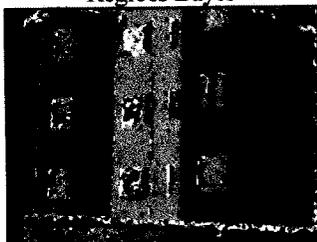
Regiões Bayes



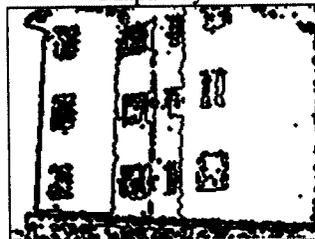
Mapa Bayes



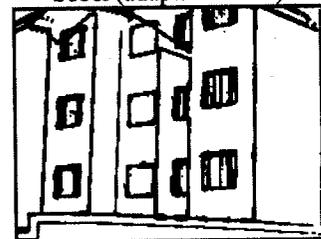
Sobel (adaptável local)



Regiões kNN



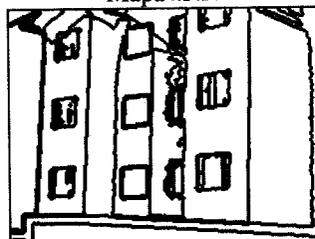
Mapa kNN



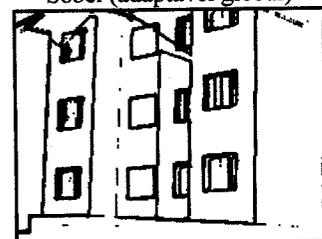
Sobel (adaptável global)



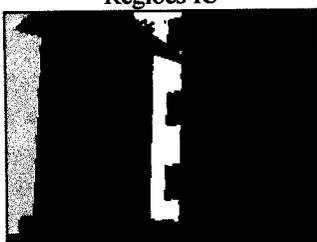
Regiões IC



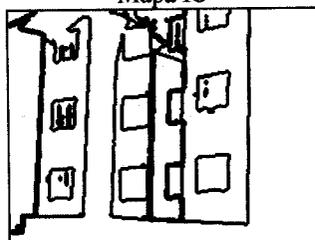
Mapa IC



Gradiente colorido



Regiões I

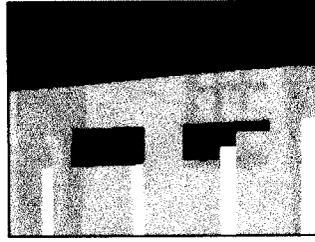


Mapa I

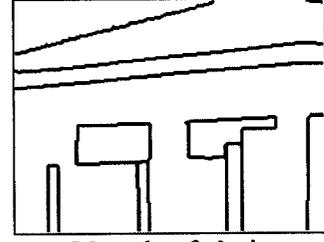
Imagem: House06



Original



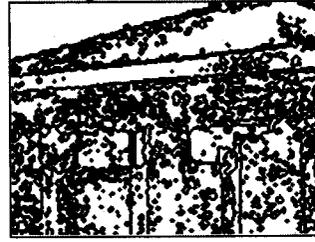
Regiões de referência



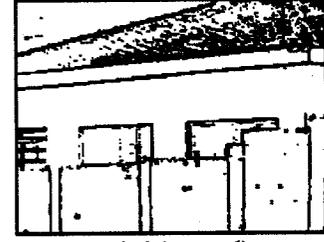
Mapa de referência



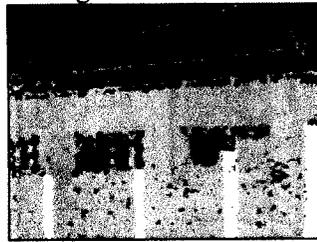
Regiões cromaticidade



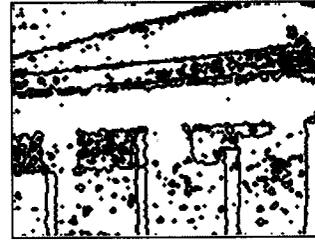
Mapa cromaticidade



Sobel (manual)



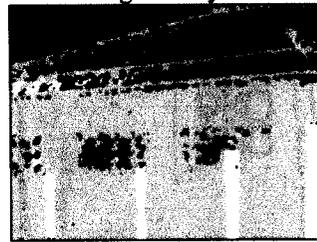
Regiões Bayes



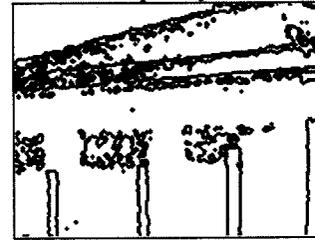
Mapa Bayes



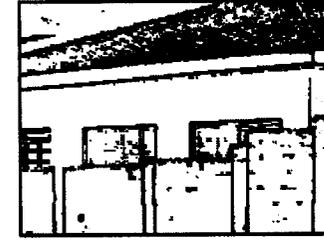
Sobel (adaptável local)



Regiões kNN



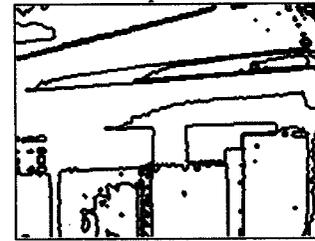
Mapa kNN



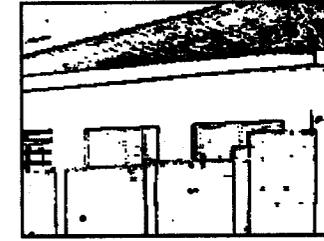
Sobel (adaptável global)



Regiões IC



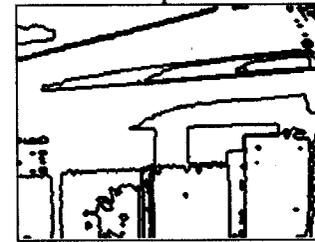
Mapa IC



Gradiente colorido

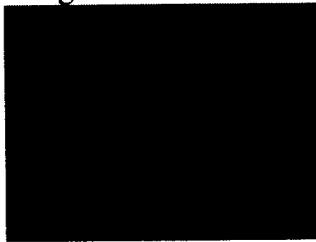


Regiões I

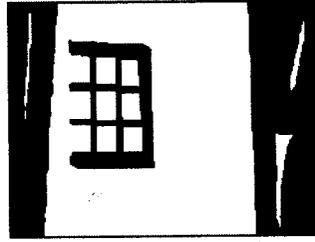


Mapa I

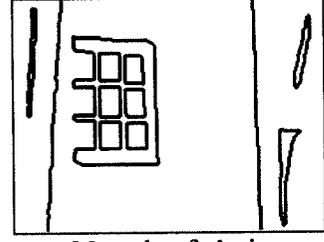
Imagem: House07



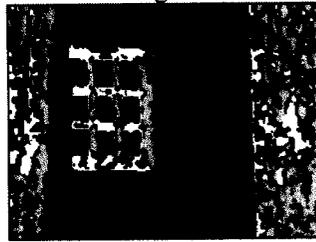
Original



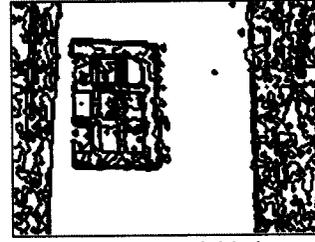
Regiões de referência



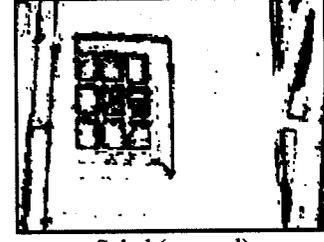
Mapa de referência



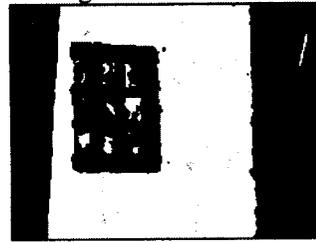
Regiões cromaticidade



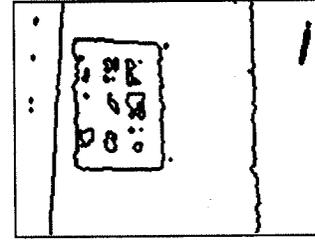
Mapa cromaticidade



Sobel (manual)



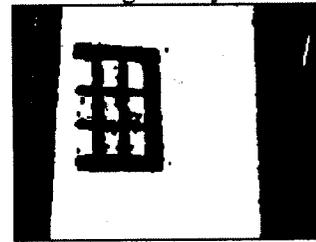
Regiões Bayes



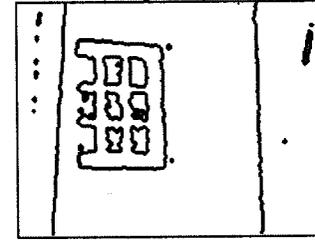
Mapa Bayes



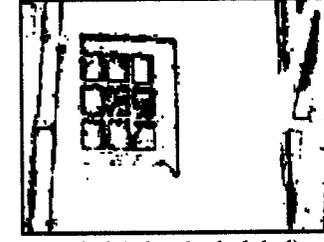
Sobel (adaptável local)



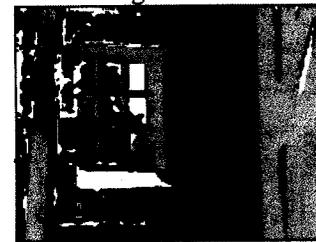
Regiões kNN



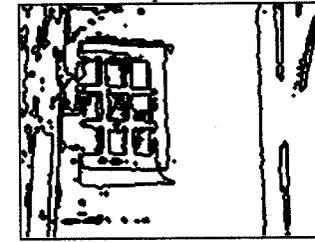
Mapa kNN



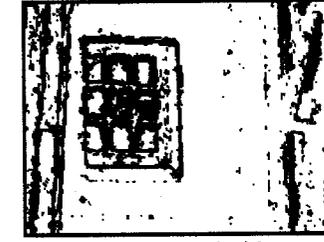
Sobel (adaptável global)



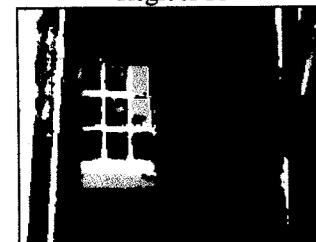
Regiões IC



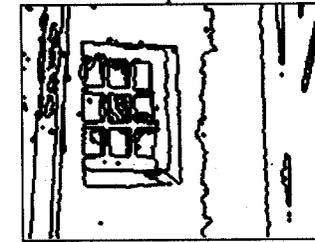
Mapa IC



Gradiente colorido



Regiões I



Mapa I

Imagem: Lenna



Original



Regiões de referência



Mapa de referência



Regiões cromaticidade



Mapa cromaticidade



Sobel (manual)



Regiões Bayes



Mapa Bayes



Sobel (adaptável local)



Regiões kNN



Mapa kNN



Sobel (adaptável global)



Regiões IC



Mapa IC



Gradiente colorido

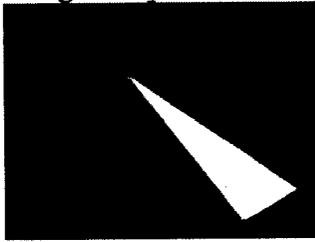


Regiões I

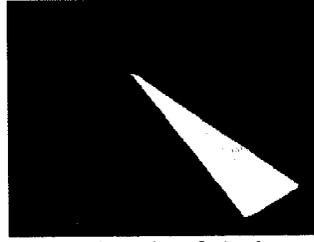


Mapa I

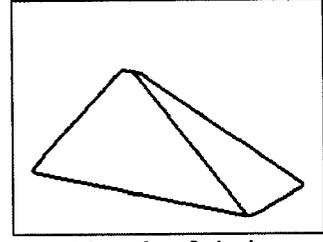
Imagem: Pyram01



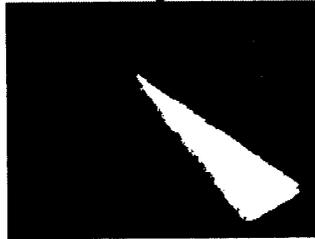
Original



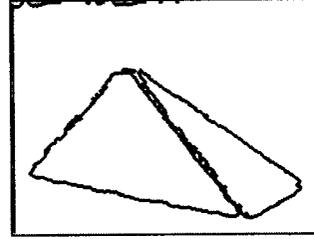
Regiões de referência



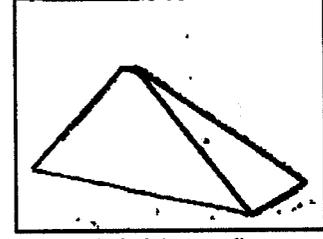
Mapa de referência



Regiões cromaticidade



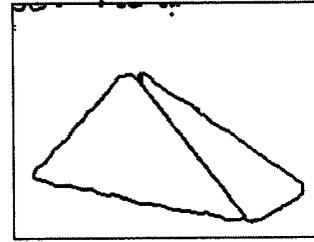
Mapa cromaticidade



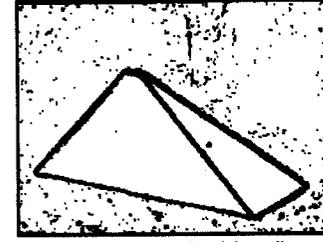
Sobel (manual)



Regiões Bayes



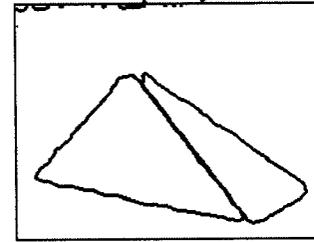
Mapa Bayes



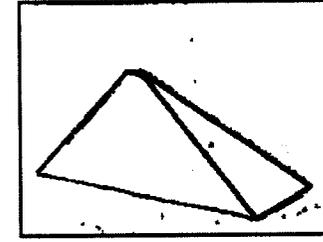
Sobel (adaptável local)



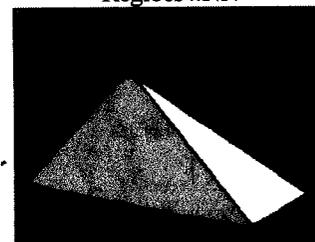
Regiões kNN



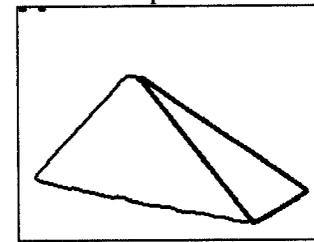
Mapa kNN



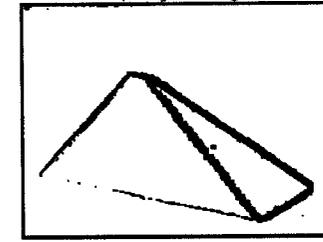
Sobel (adaptável global)



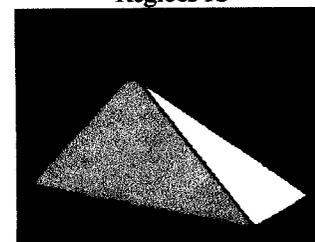
Regiões IC



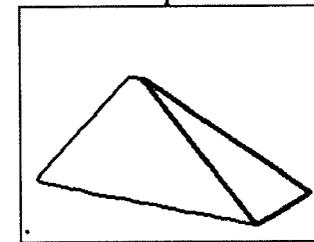
Mapa IC



Gradiente colorido

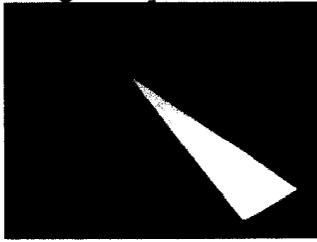


Regiões I

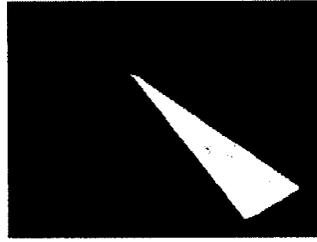


Mapa I

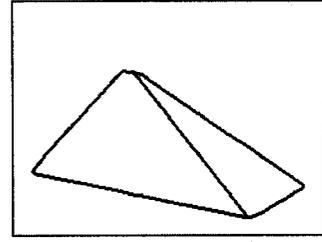
Imagem: Pyram03



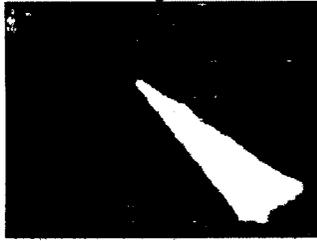
Original



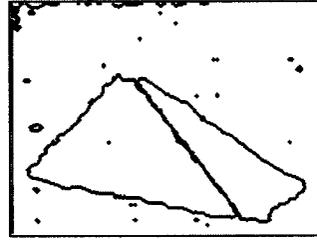
Regiões de referência



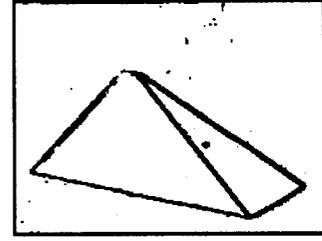
Mapa de referência



Regiões cromaticidade



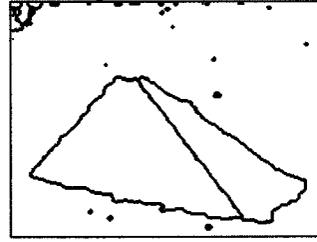
Mapa cromaticidade



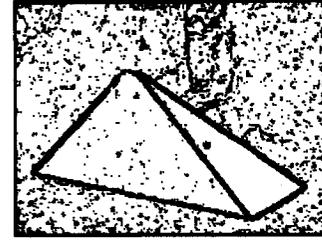
Sobel (manual)



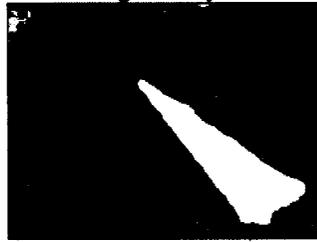
Regiões Bayes



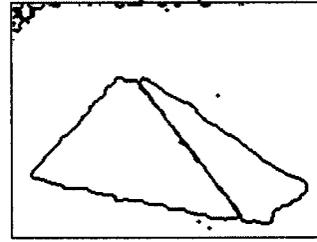
Mapa Bayes



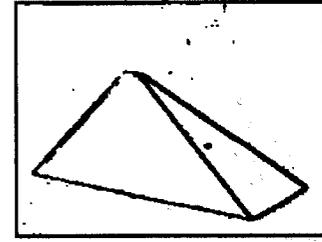
Sobel (adaptável local)



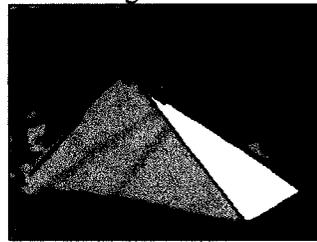
Regiões kNN



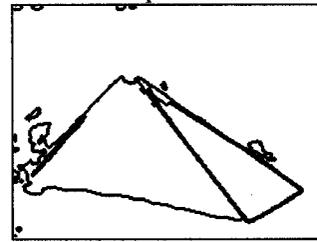
Mapa kNN



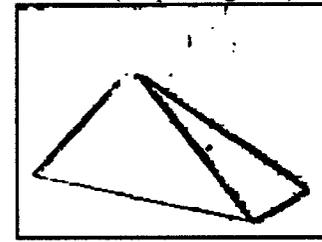
Sobel (adaptável global)



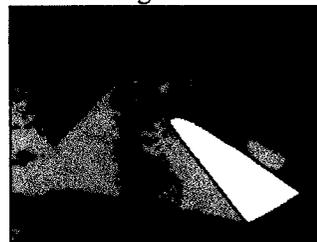
Regiões IC



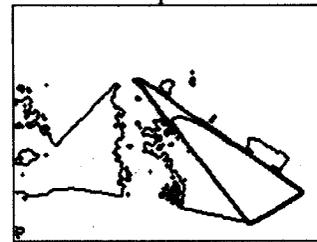
Mapa IC



Gradiente colorido

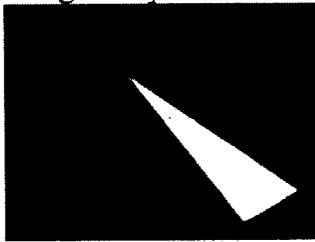


Regiões I

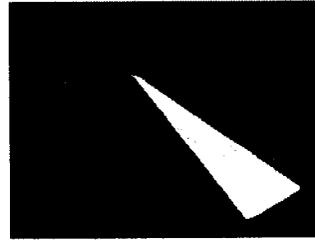


Mapa I

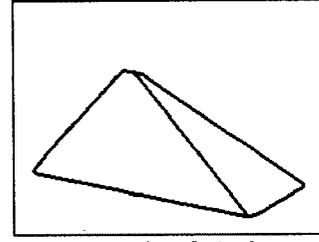
Imagem: Pyram04



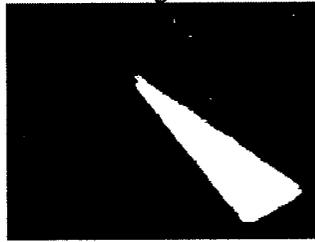
Original



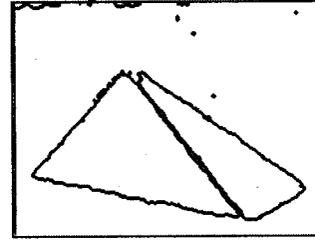
Regiões de referência



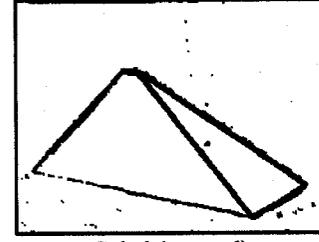
Mapa de referência



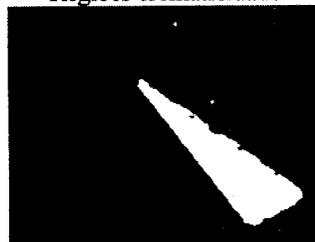
Regiões cromaticidade



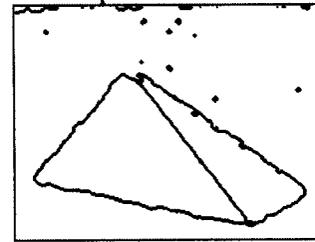
Mapa cromaticidade



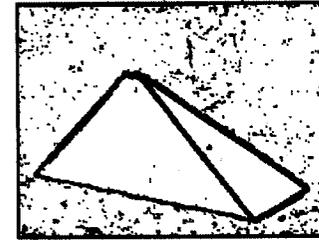
Sobel (manual)



Regiões Bayes



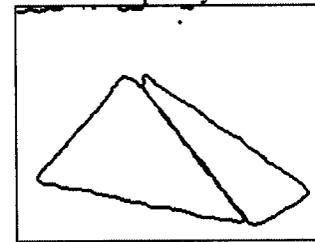
Mapa Bayes



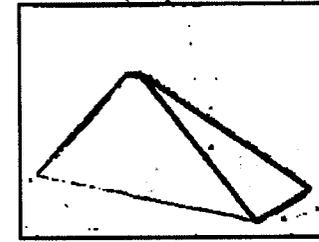
Sobel (adaptável local)



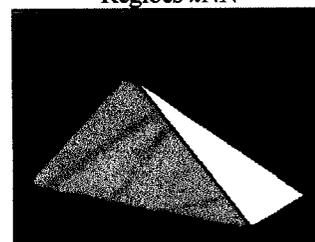
Regiões kNN



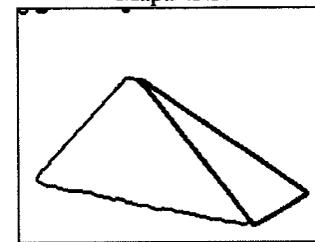
Mapa kNN



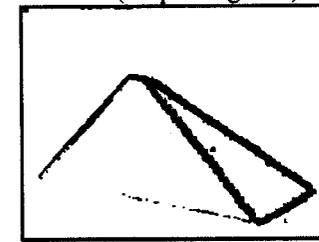
Sobel (adaptável global)



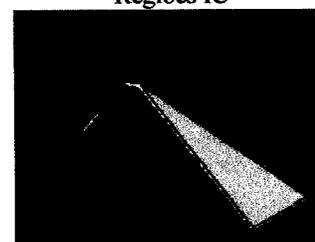
Regiões IC



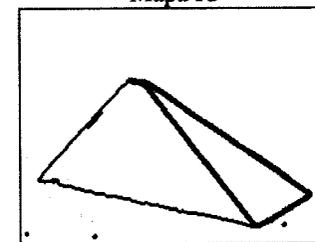
Mapa IC



Gradiente colorido



Regiões I

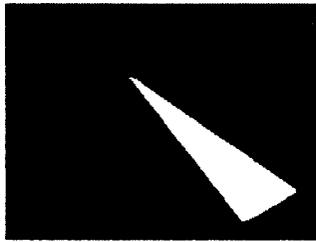


Mapa I

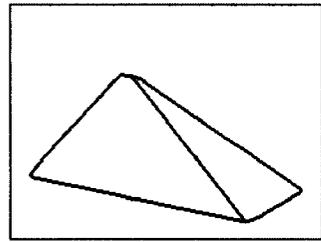
Imagem: Pyram06



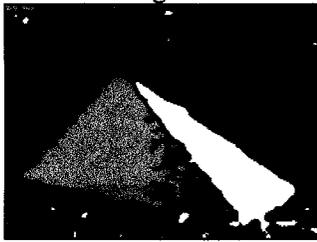
Original



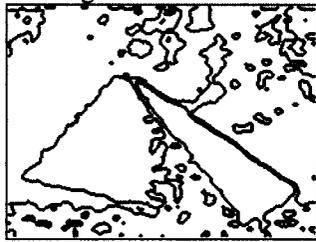
Regiões de referência



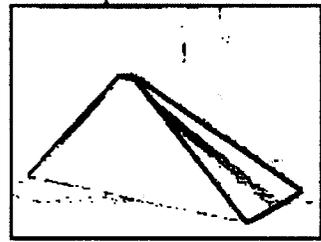
Mapa de referência



Regiões cromaticidade



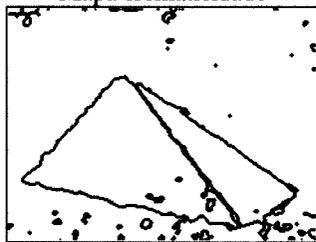
Mapa cromaticidade



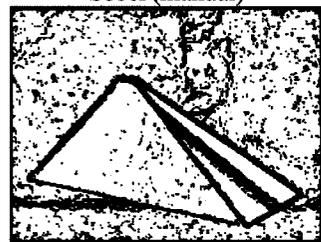
Sobel (manual)



Regiões Bayes



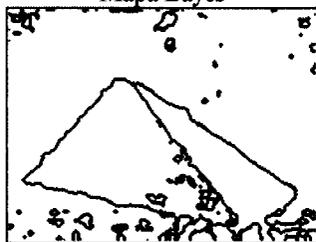
Mapa Bayes



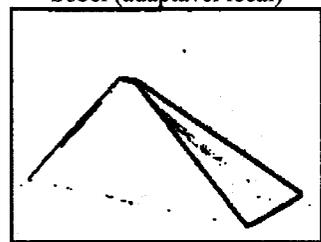
Sobel (adaptável local)



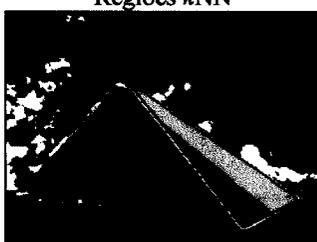
Regiões kNN



Mapa kNN



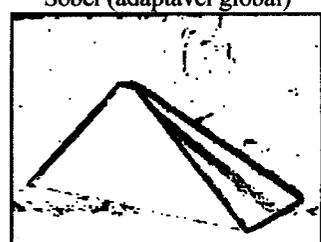
Sobel (adaptável global)



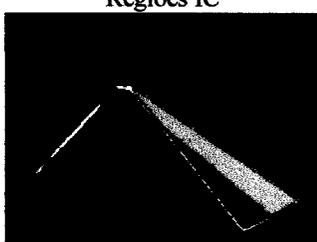
Regiões IC



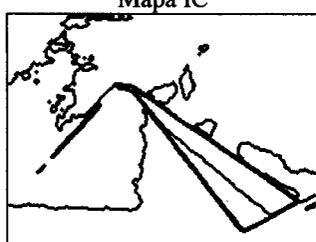
Mapa IC



Gradiente colorido



Regiões I



Mapa I

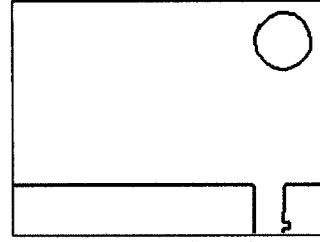
Imagem: Sign01



Original



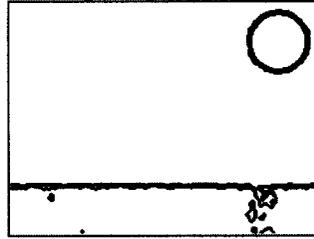
Regiões de referência



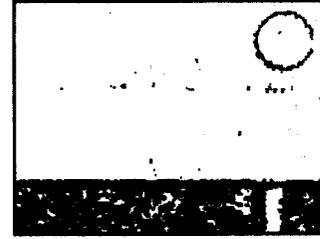
Mapa de referência



Regiões cromaticidade



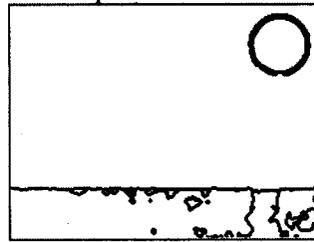
Mapa cromaticidade



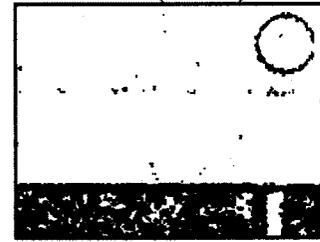
Sobel (manual)



Regiões Bayes



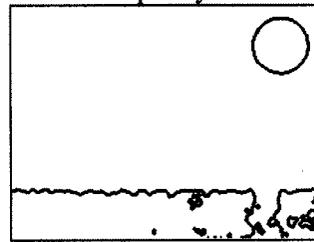
Mapa Bayes



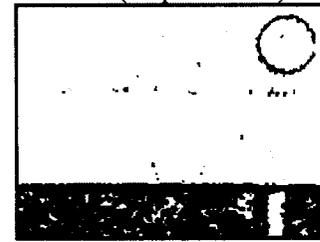
Sobel (adaptável local)



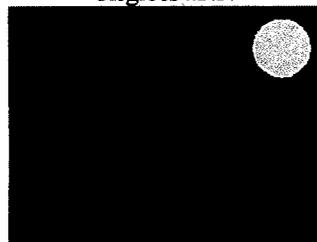
Regiões kNN



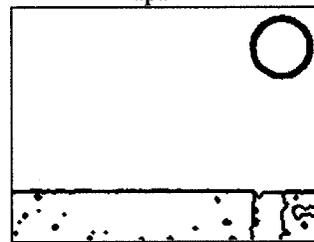
Mapa kNN



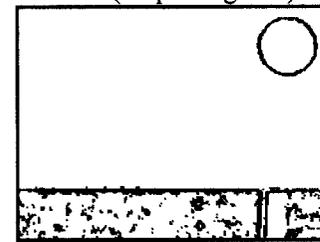
Sobel (adaptável global)



Regiões IC



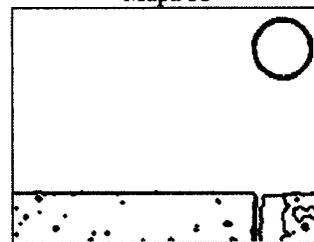
Mapa IC



Gradiente colorido



Regiões I



Mapa I

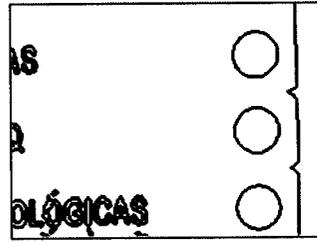
Imagem: Sign02



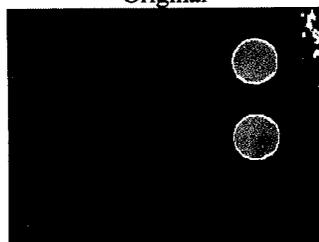
Original



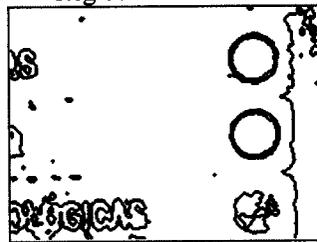
Regiões de referência



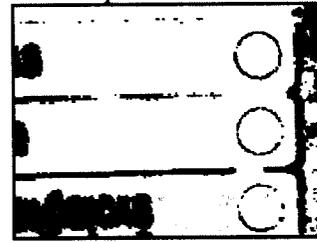
Mapa de referência



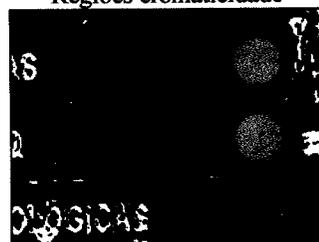
Regiões cromaticidade



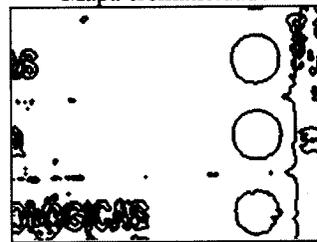
Mapa cromaticidade



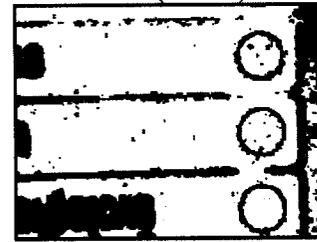
Sobel (manual)



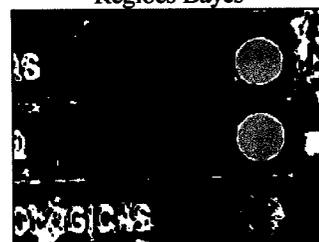
Regiões Bayes



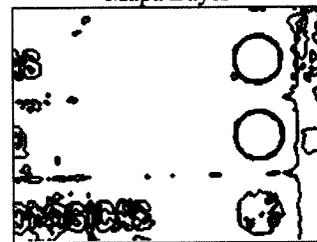
Mapa Bayes



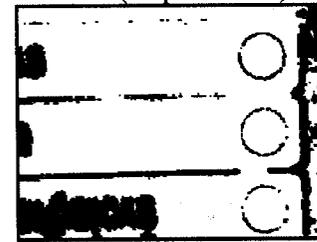
Sobel (adaptável local)



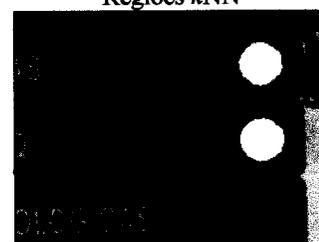
Regiões kNN



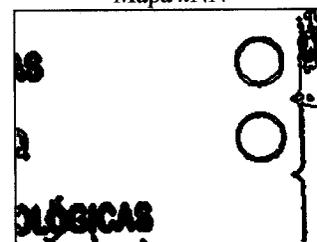
Mapa kNN



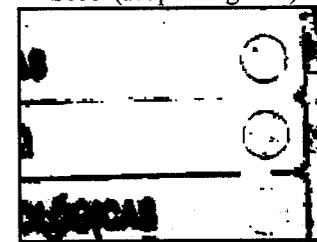
Sobel (adaptável global)



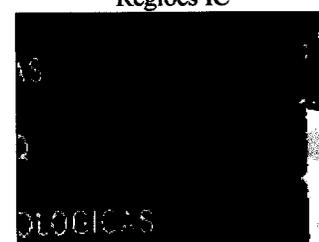
Regiões IC



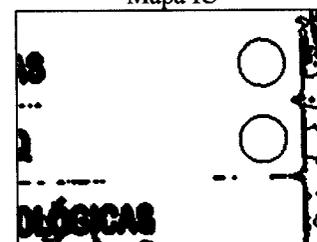
Mapa IC



Gradiente colorido

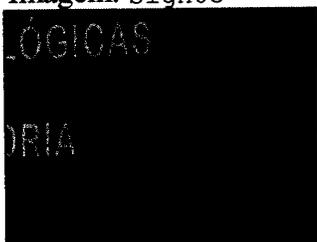


Regiões I



Mapa I

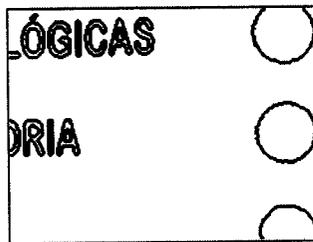
Imagem: sign03



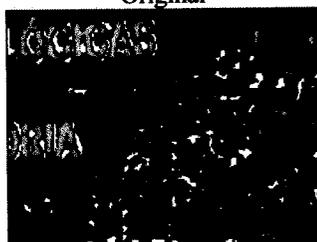
Original



Regiões de referência



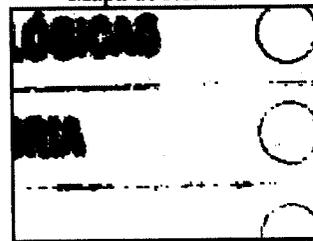
Mapa de referência



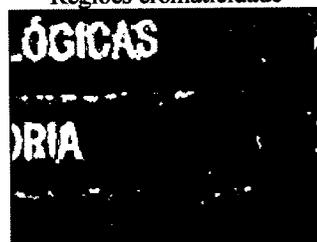
Regiões cromaticidade



Mapa cromaticidade



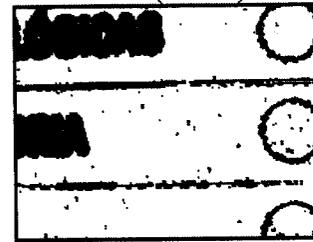
Sobel (manual)



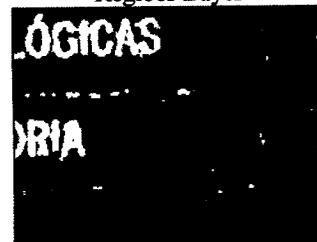
Regiões Bayes



Mapa Bayes



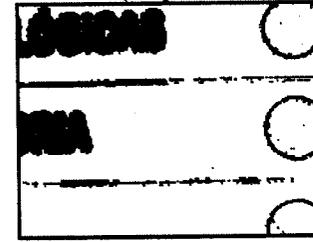
Sobel (adaptável local)



Regiões kNN



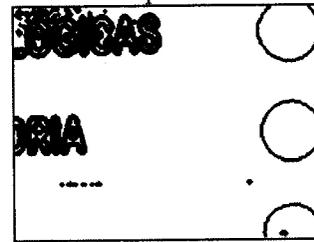
Mapa kNN



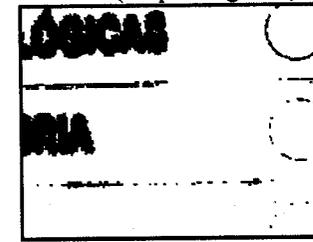
Sobel (adaptável global)



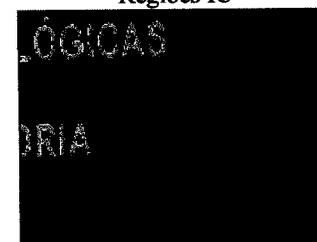
Regiões IC



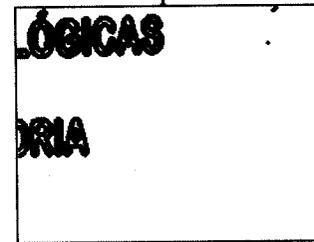
Mapa IC



Gradiente colorido

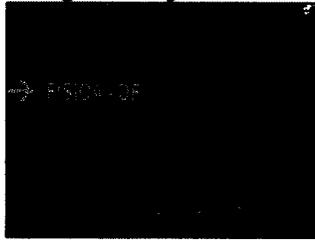


Regiões I

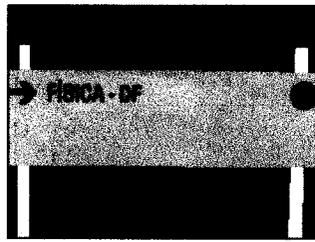


Mapa I

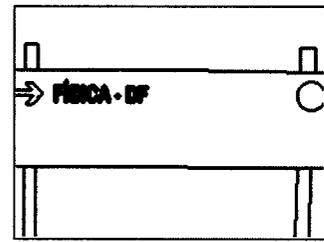
Imagem: sign04



Original



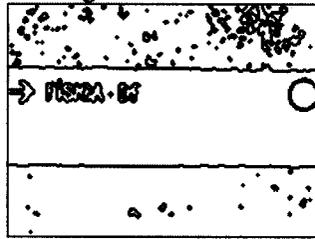
Regiões de referência



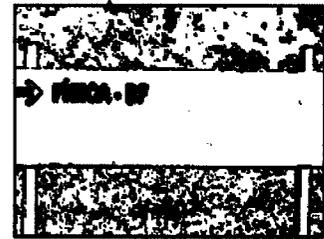
Mapa de referência



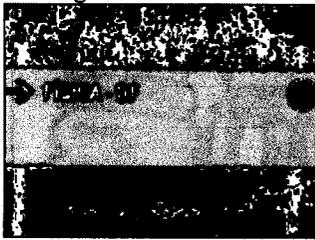
Regiões cromaticidade



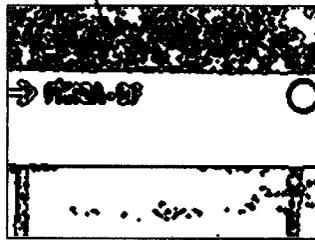
Mapa cromaticidade



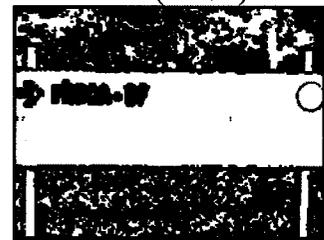
Sobel (manual)



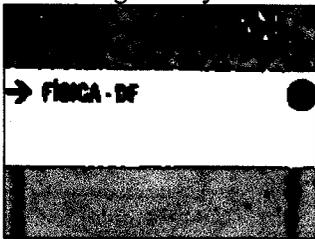
Regiões Bayes



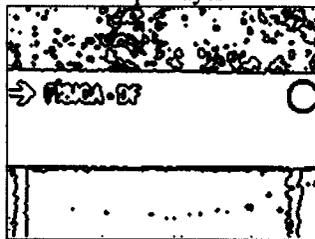
Mapa Bayes



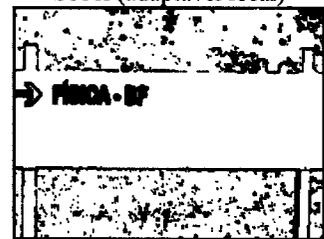
Sobel (adaptável local)



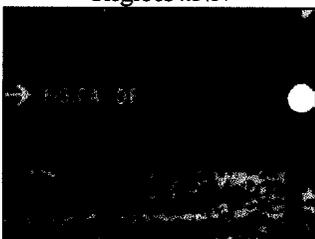
Regiões kNN



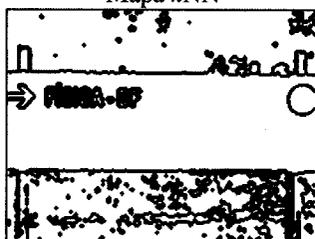
Mapa kNN



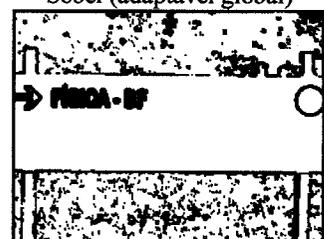
Sobel (adaptável global)



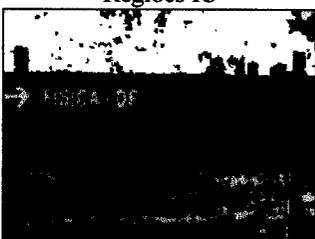
Regiões IC



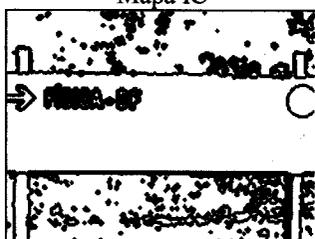
Mapa IC



Gradiente colorido



Regiões I

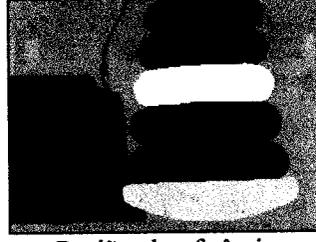


Mapa I

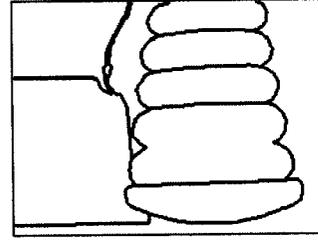
Imagem: Toy



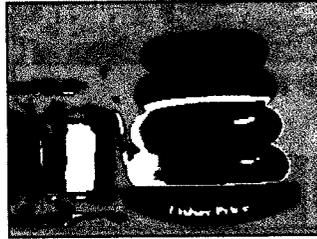
Original



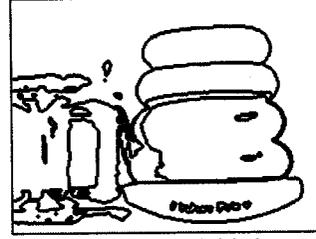
Regiões de referência



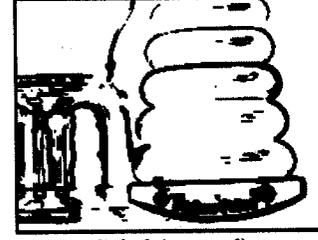
Mapa de referência



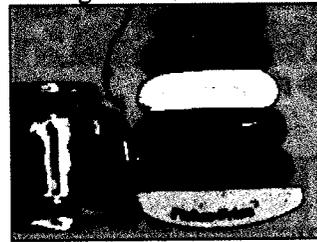
Regiões cromaticidade



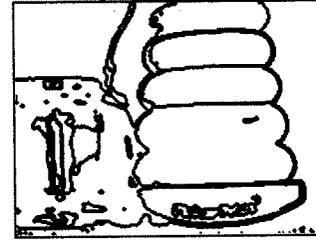
Mapa cromaticidade



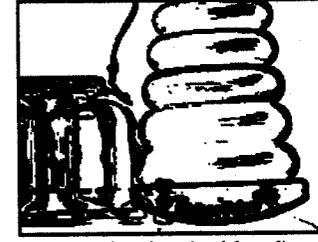
Sobel (manual)



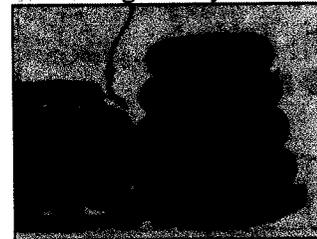
Regiões Bayes



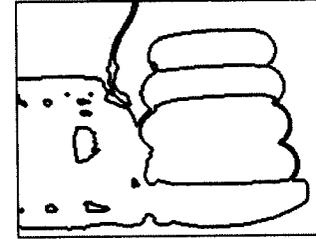
Mapa Bayes



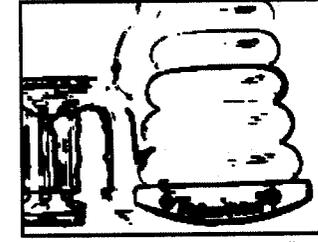
Sobel (adaptável local)



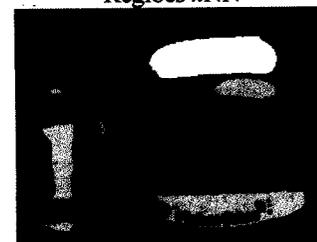
Regiões kNN



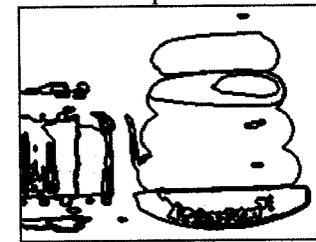
Mapa kNN



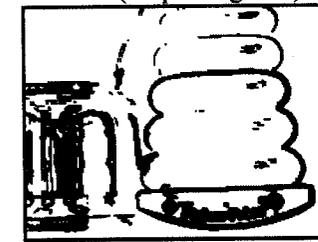
Sobel (adaptável global)



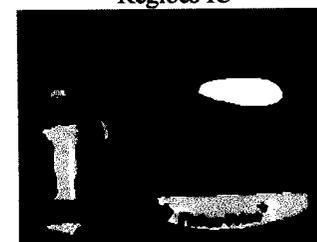
Regiões IC



Mapa IC



Gradiente colorido

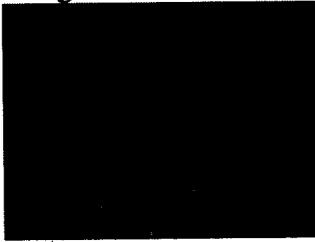


Regiões I

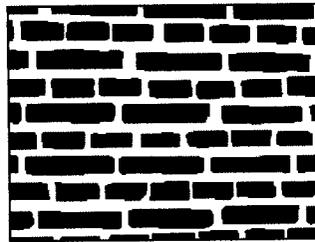


Mapa I

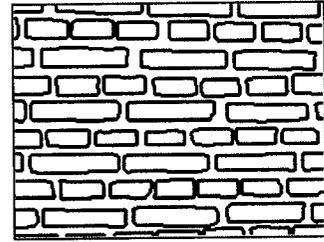
Imagem: wall01



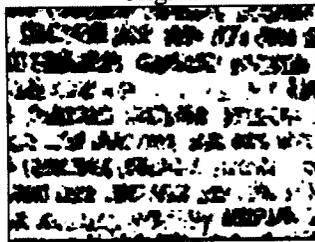
Original



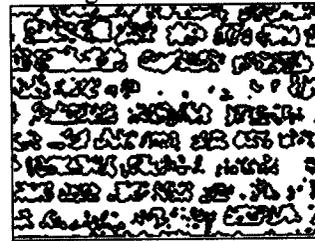
Regiões de referência



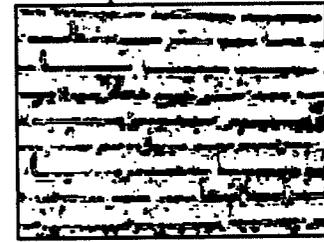
Mapa de referência



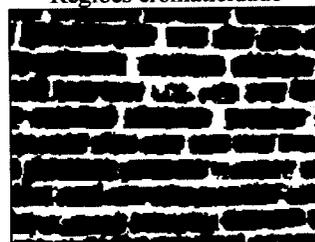
Regiões cromaticidade



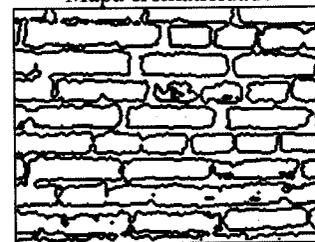
Mapa cromaticidade



Sobel (manual)



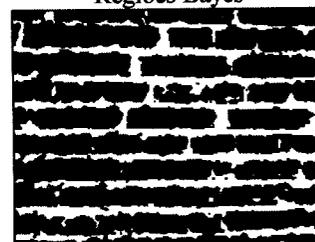
Regiões Bayes



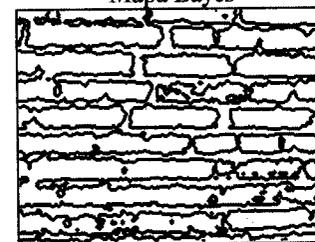
Mapa Bayes



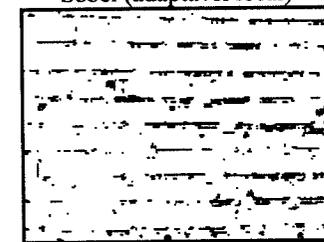
Sobel (adaptável local)



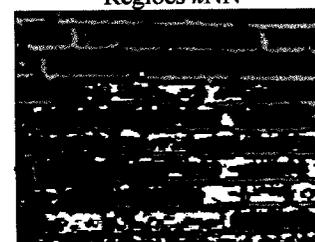
Regiões kNN



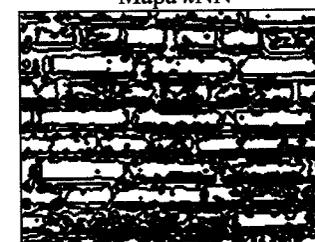
Mapa kNN



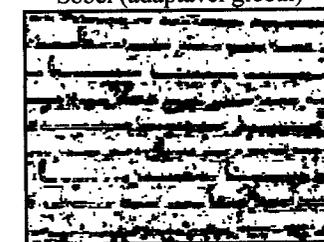
Sobel (adaptável global)



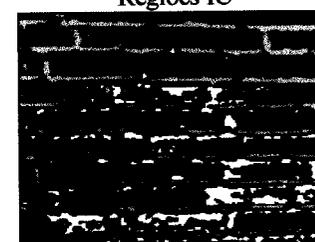
Regiões IC



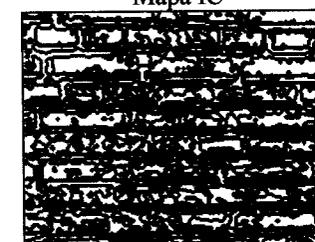
Mapa IC



Gradiente colorido

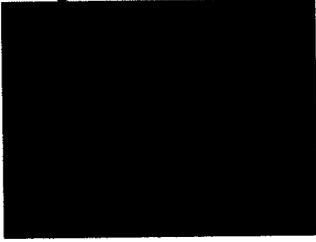


Regiões I

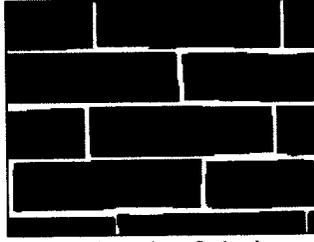


Mapa I

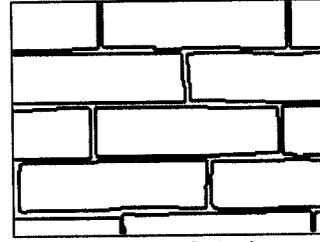
Imagem: wall102



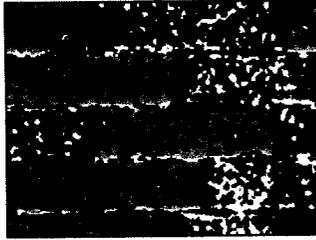
Original



Regiões de referência



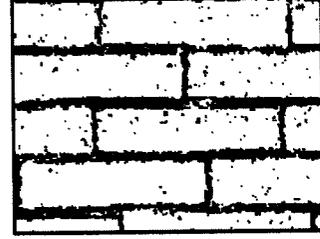
Mapa de referência



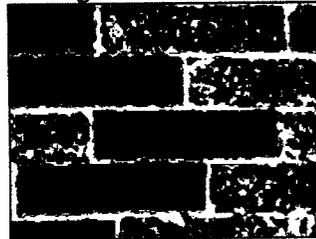
Regiões cromaticidade



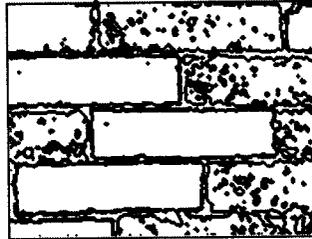
Mapa cromaticidade



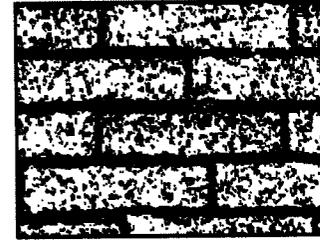
Sobel (manual)



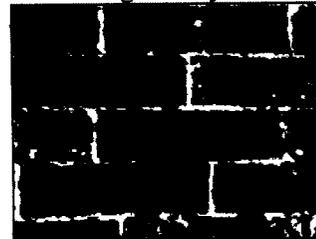
Regiões Bayes



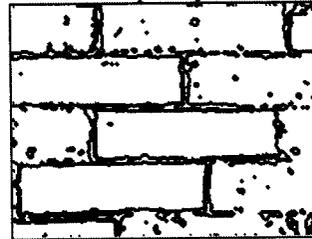
Mapa Bayes



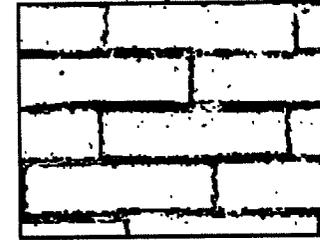
Sobel (adaptável local)



Regiões kNN



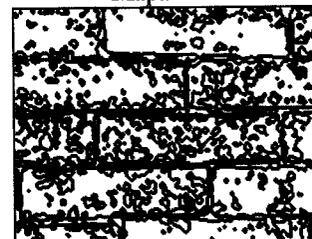
Mapa kNN



Sobel (adaptável global)



Regiões IC



Mapa IC



Gradiente colorido



Regiões I



Mapa I

**E. Resultados comparativos do coeficiente acumulativo para as classes de imagens**

Classe: *casas*

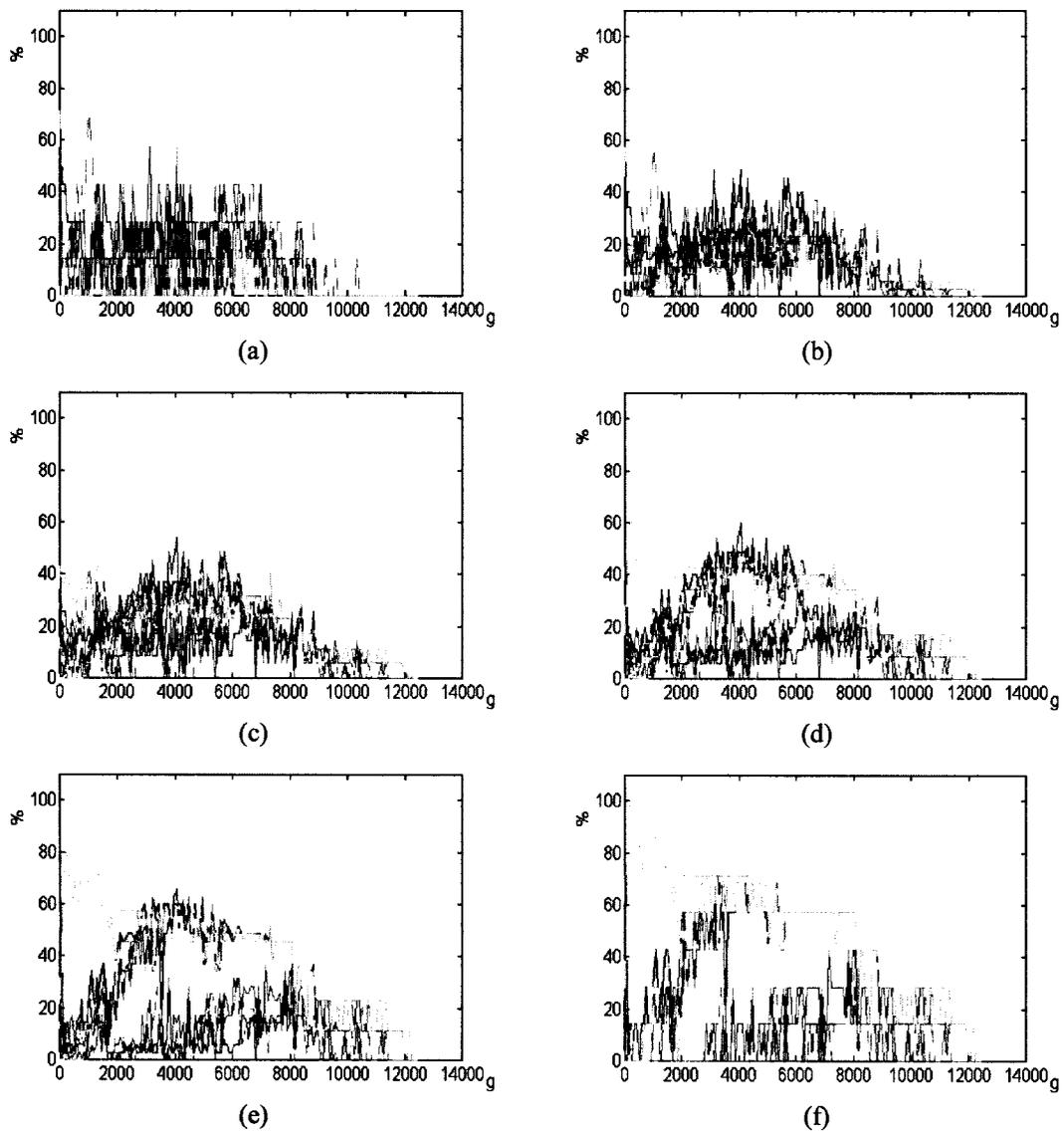


Figura E.1. Coeficiente acumulativo do atributo  $g$  para a classe *casas*. (a)-(f):  $p$  com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

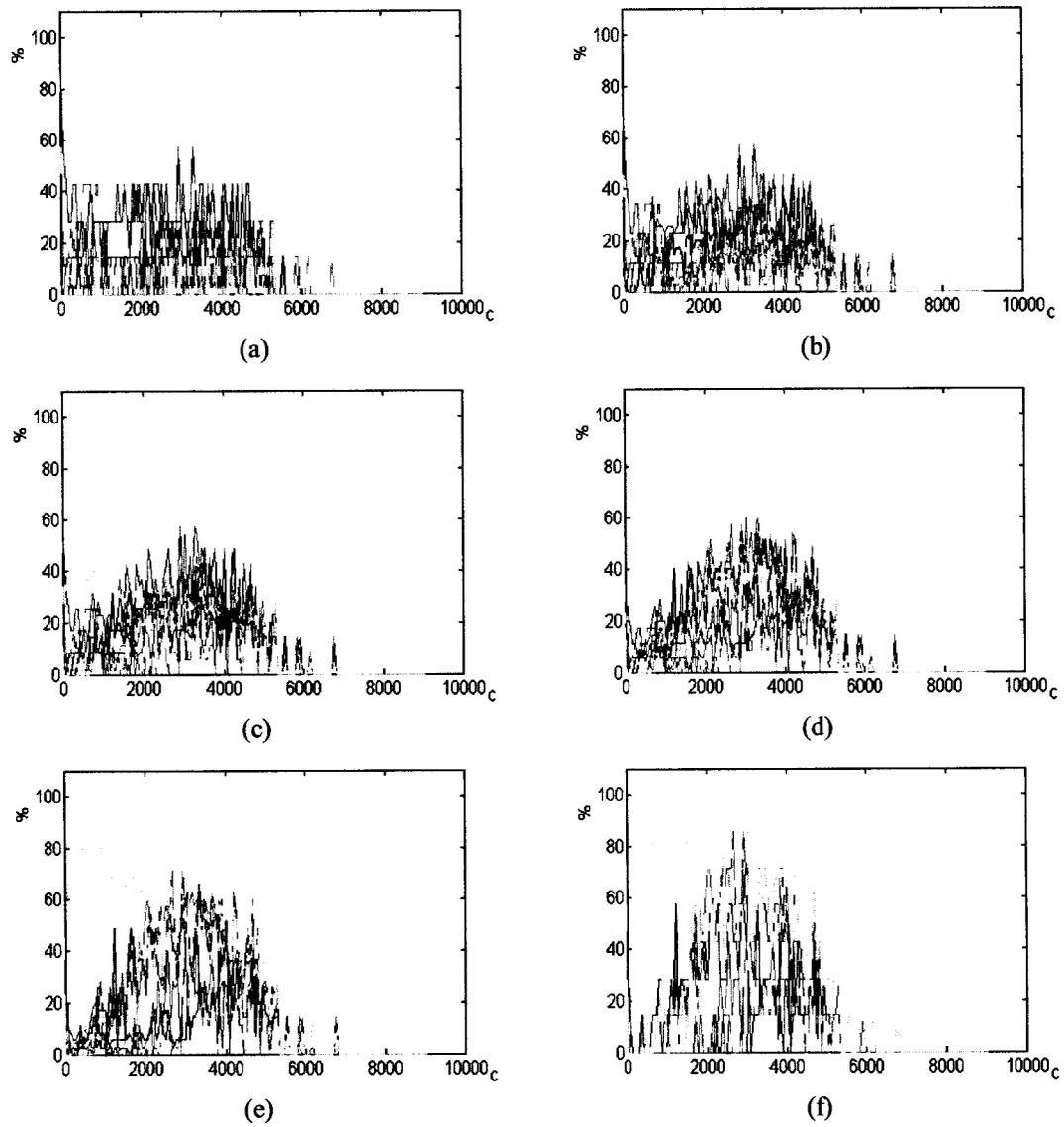


Figura E.2. Coeficiente acumulativo do atributo  $c$  para a classe *casas*. (a)-(f):  $p$  com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

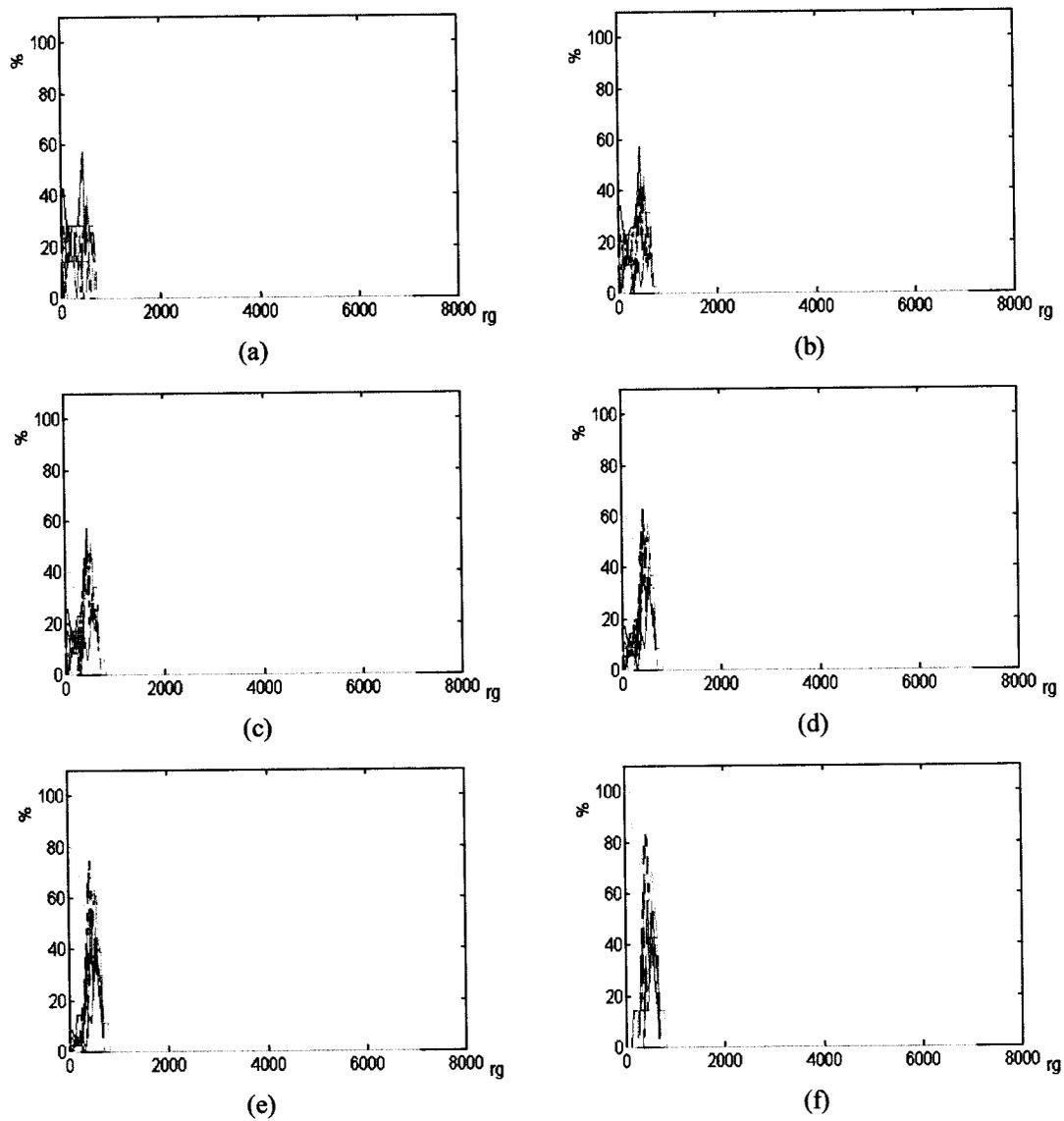


Figura E.3. Coeficiente acumulativo do atributo *rg* para a classe *casas*. (a)-(f):  $p$  com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

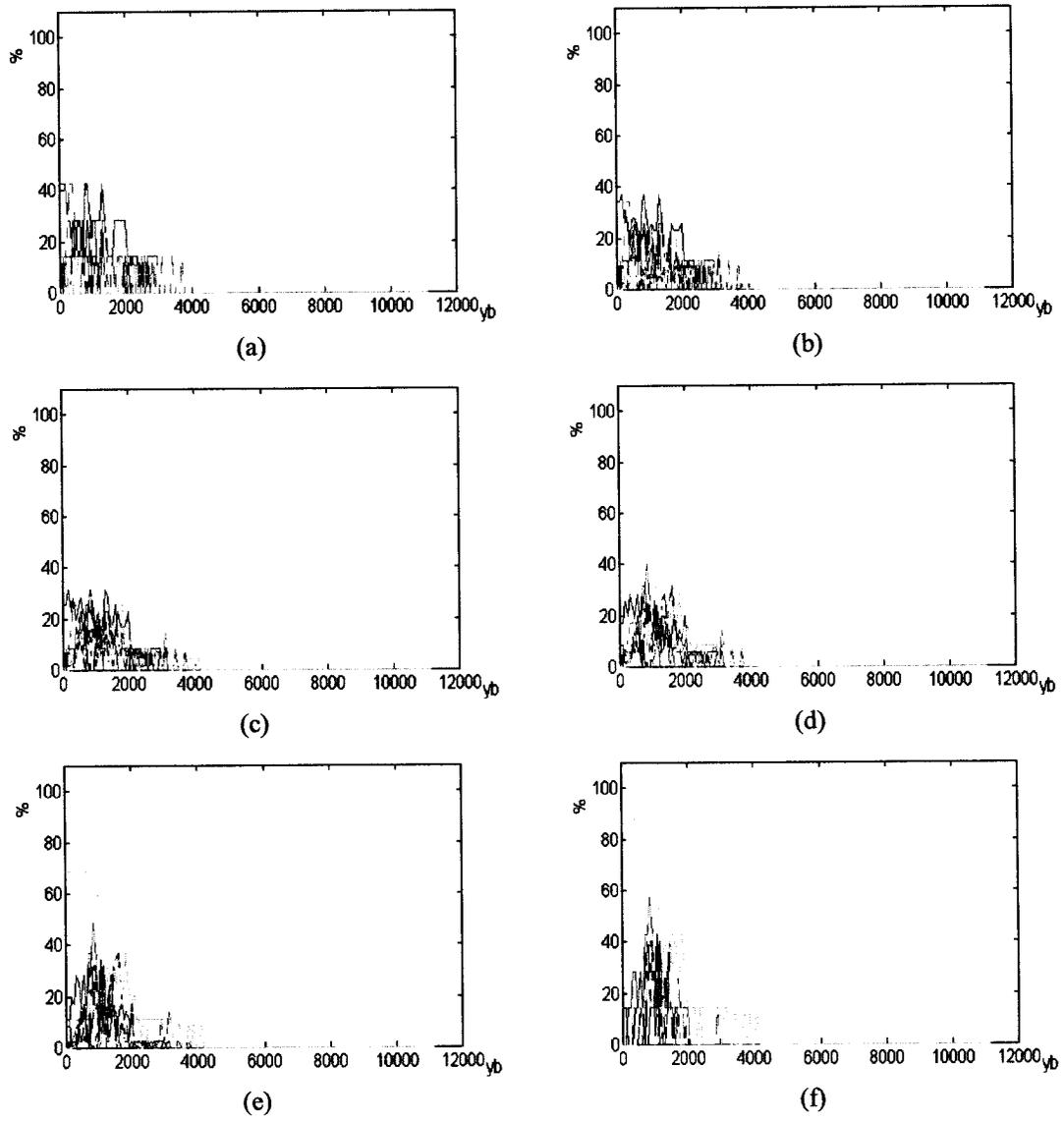


Figura E.4. Coeficiente acumulativo do atributo *yb* para a classe *casas*. (a)-(f): *p* com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

Classe: *cabos*

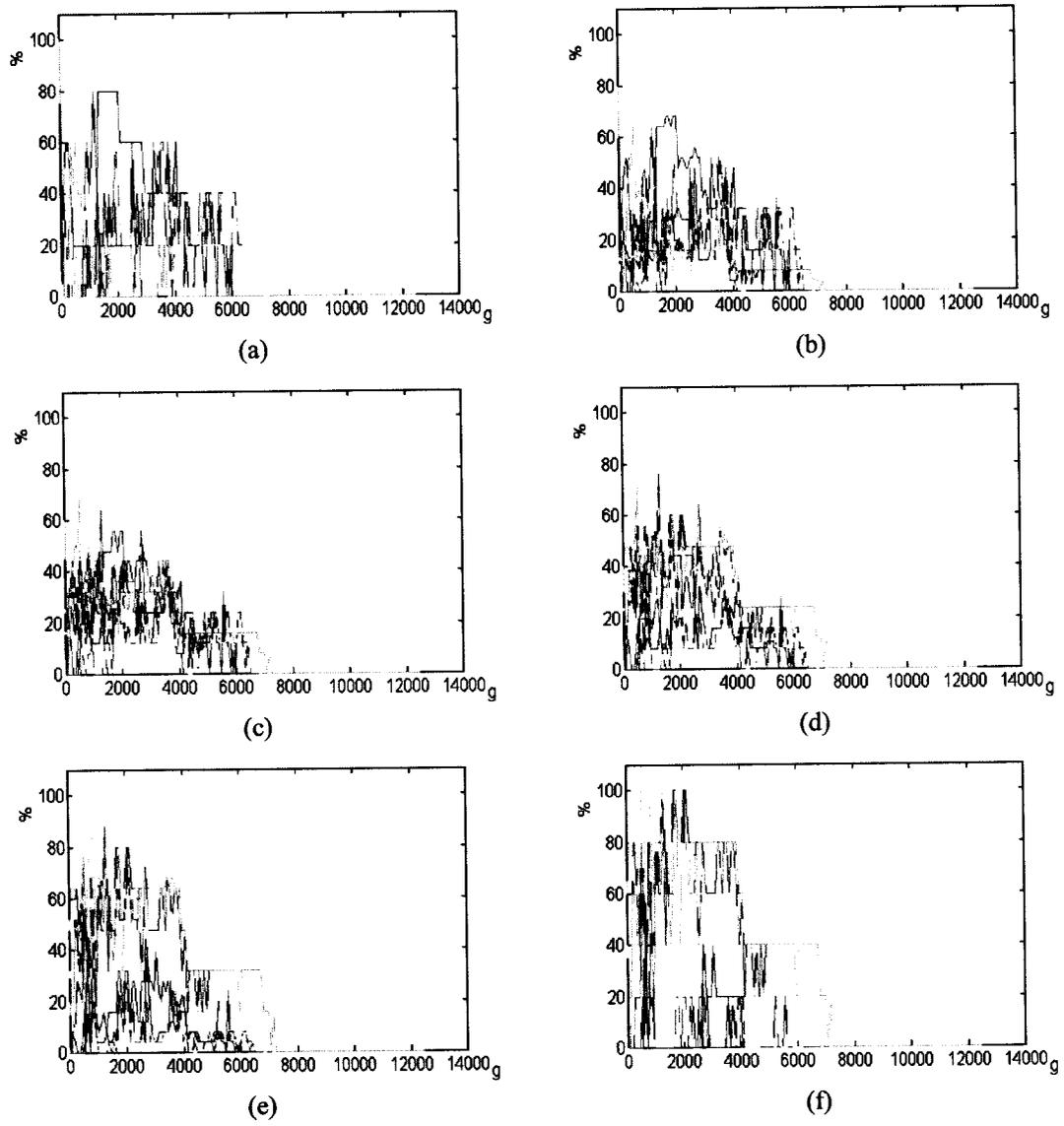


Figura E.5. Coeficiente acumulativo do atributo  $g$  para a classe *cabos*. (a)-(f):  $p$  com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

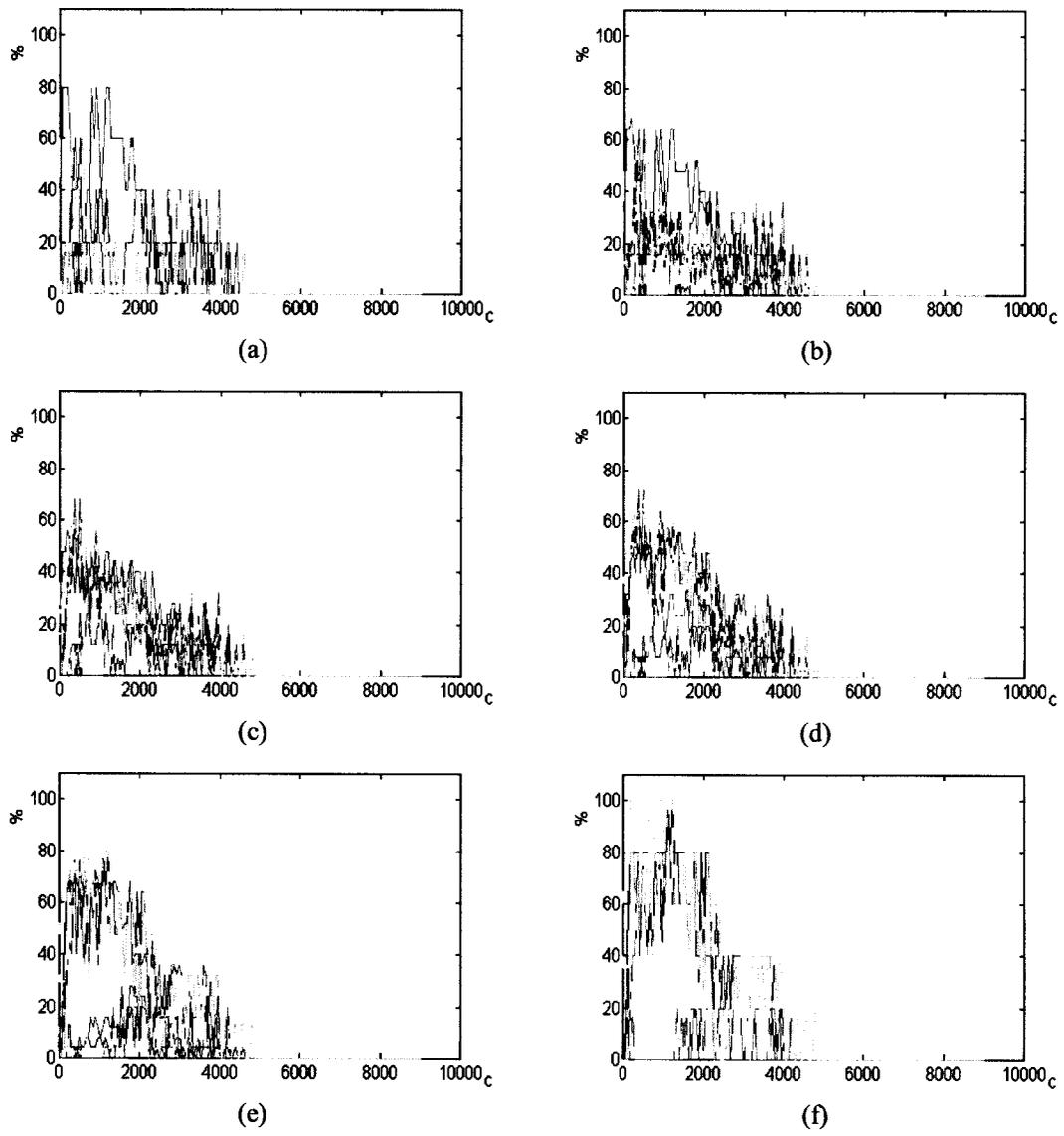


Figura E.6. Coeficiente acumulativo do atributo  $c$  para a classe *cubos*. (a)-(f):  $p$  com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

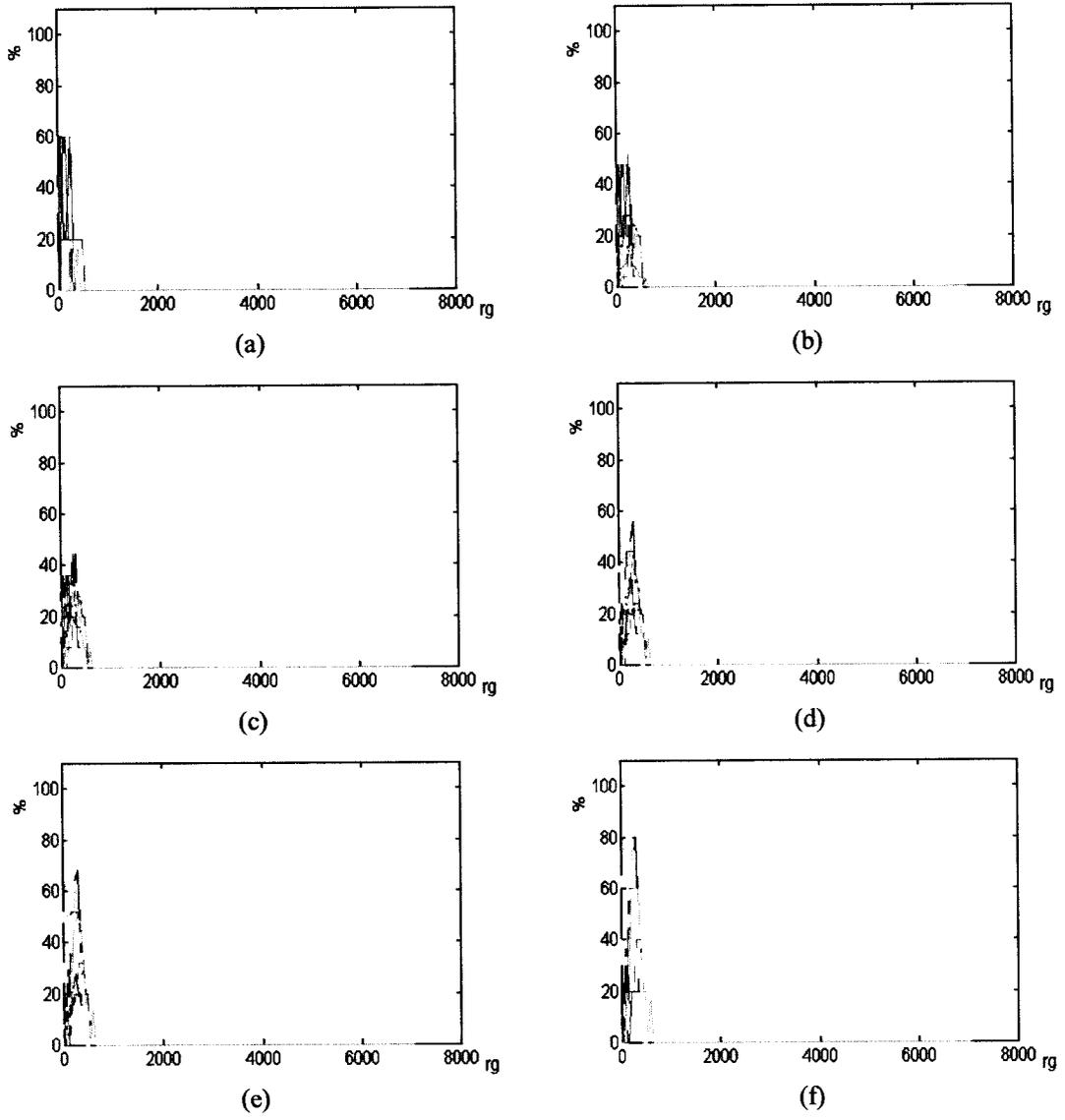


Figura E.7. Coeficiente acumulativo do atributo *rg* para a classe *cubos*. (a)-(f): *p* com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

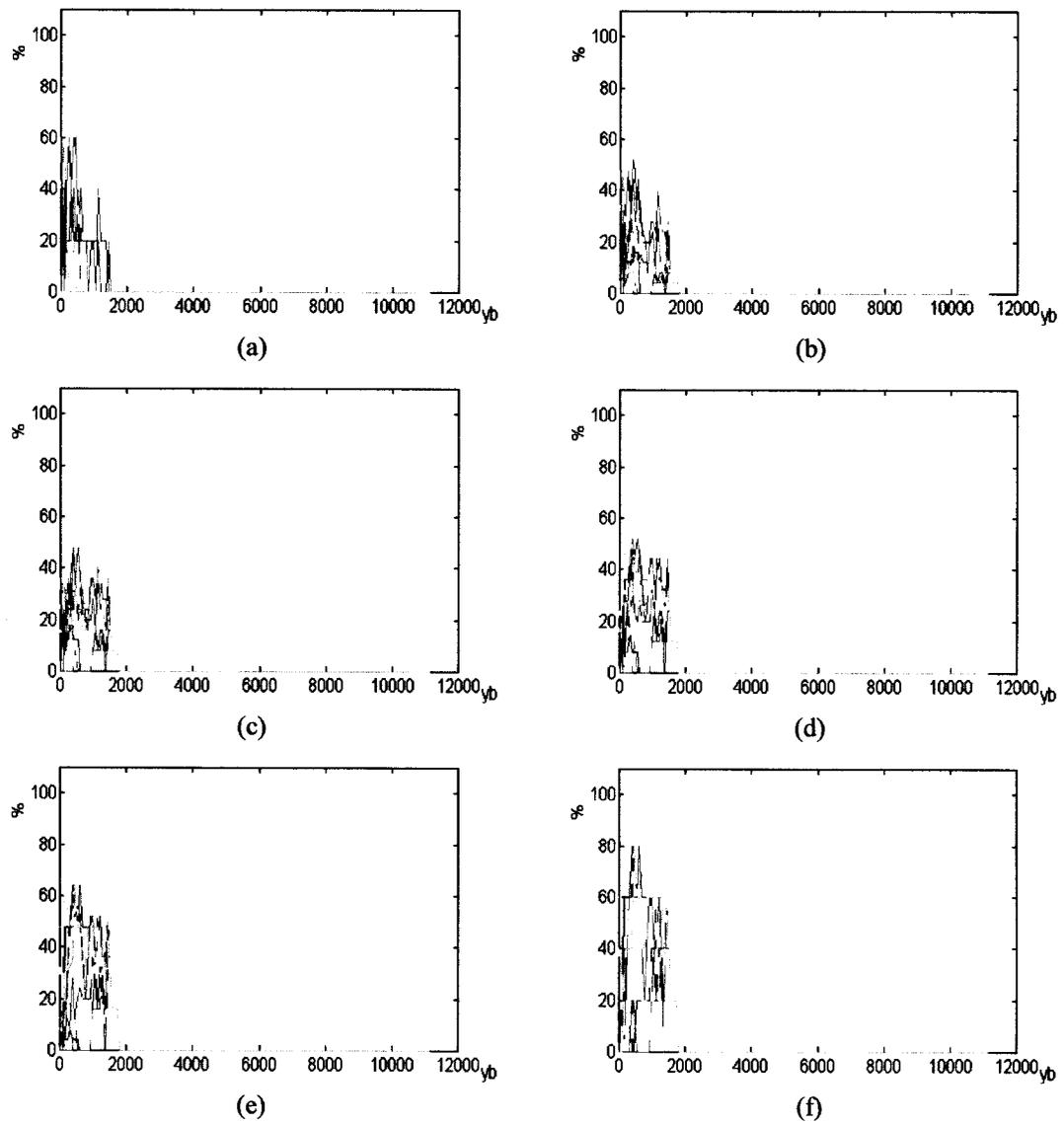


Figura E.8. Coeficiente acumulativo do atributo  $y_b$  para a classe *cubos*. (a)-(f):  $p$  com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

Classe: *externas*

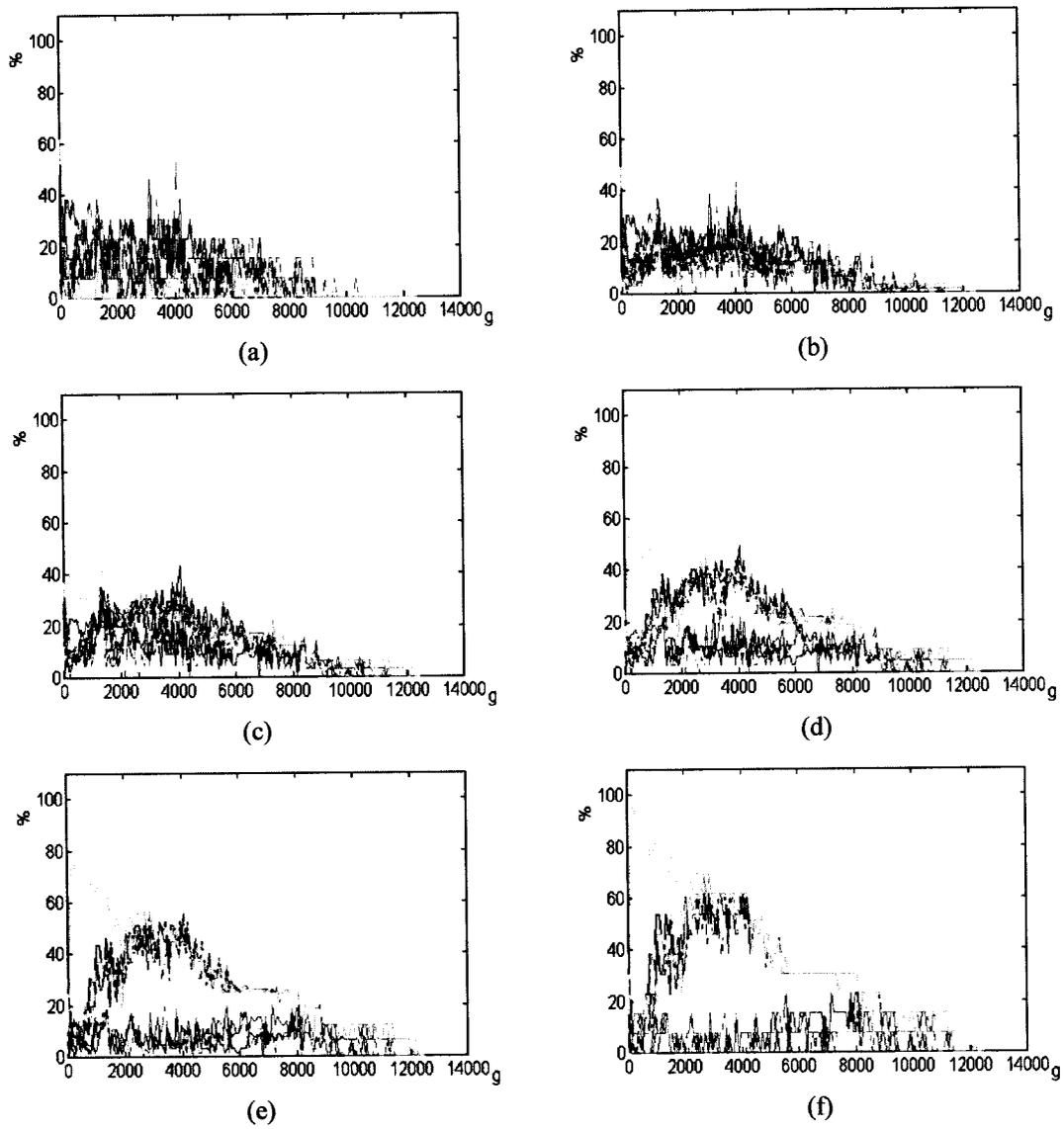


Figura E.9. Coeficiente acumulativo do atributo  $g$  para a classe *externas*. (a)-(f):  $p$  com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

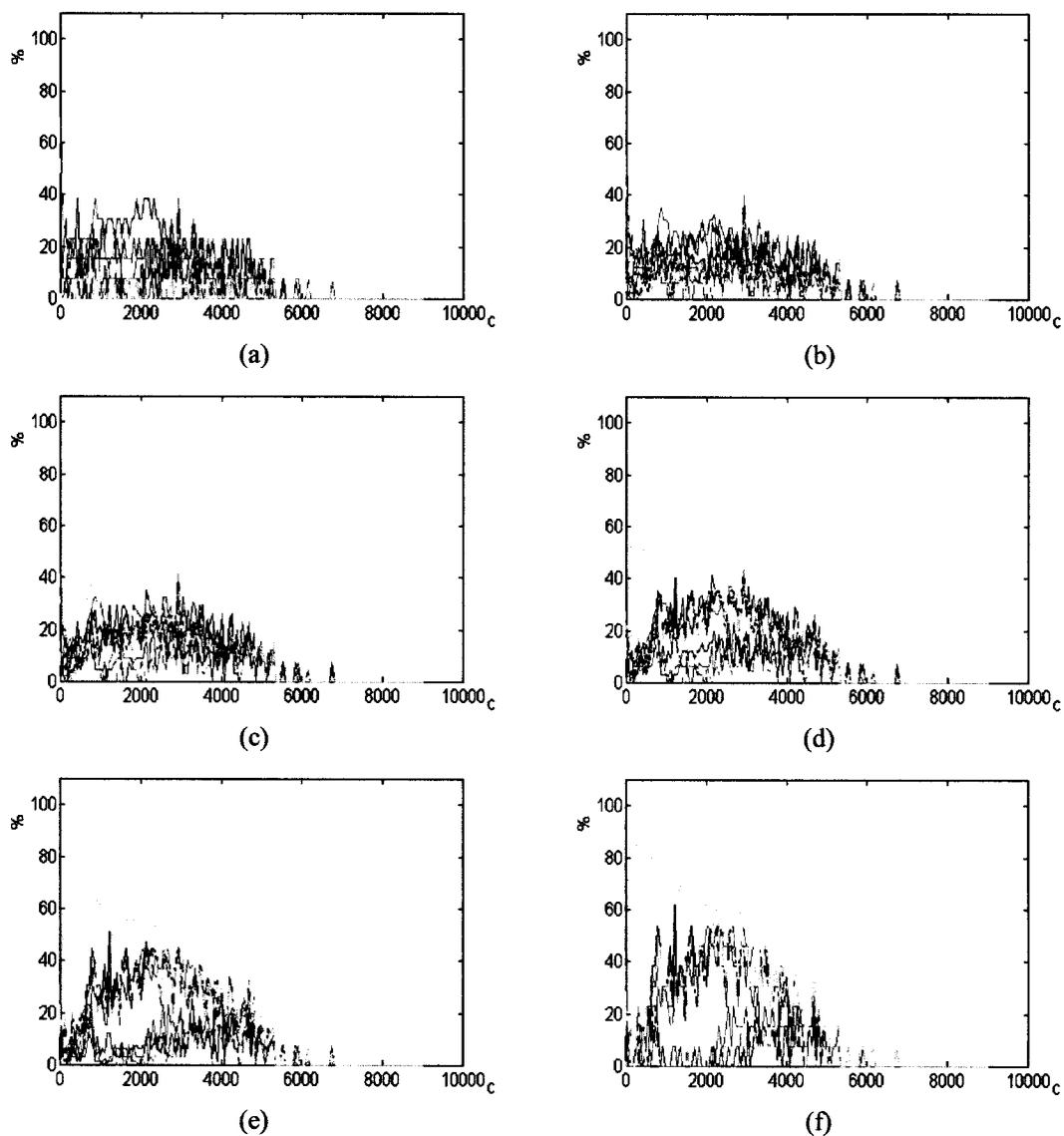


Figura E.10. Coeficiente acumulativo do atributo  $c$  para a classe *externas*. (a)-(f):  $p$  com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

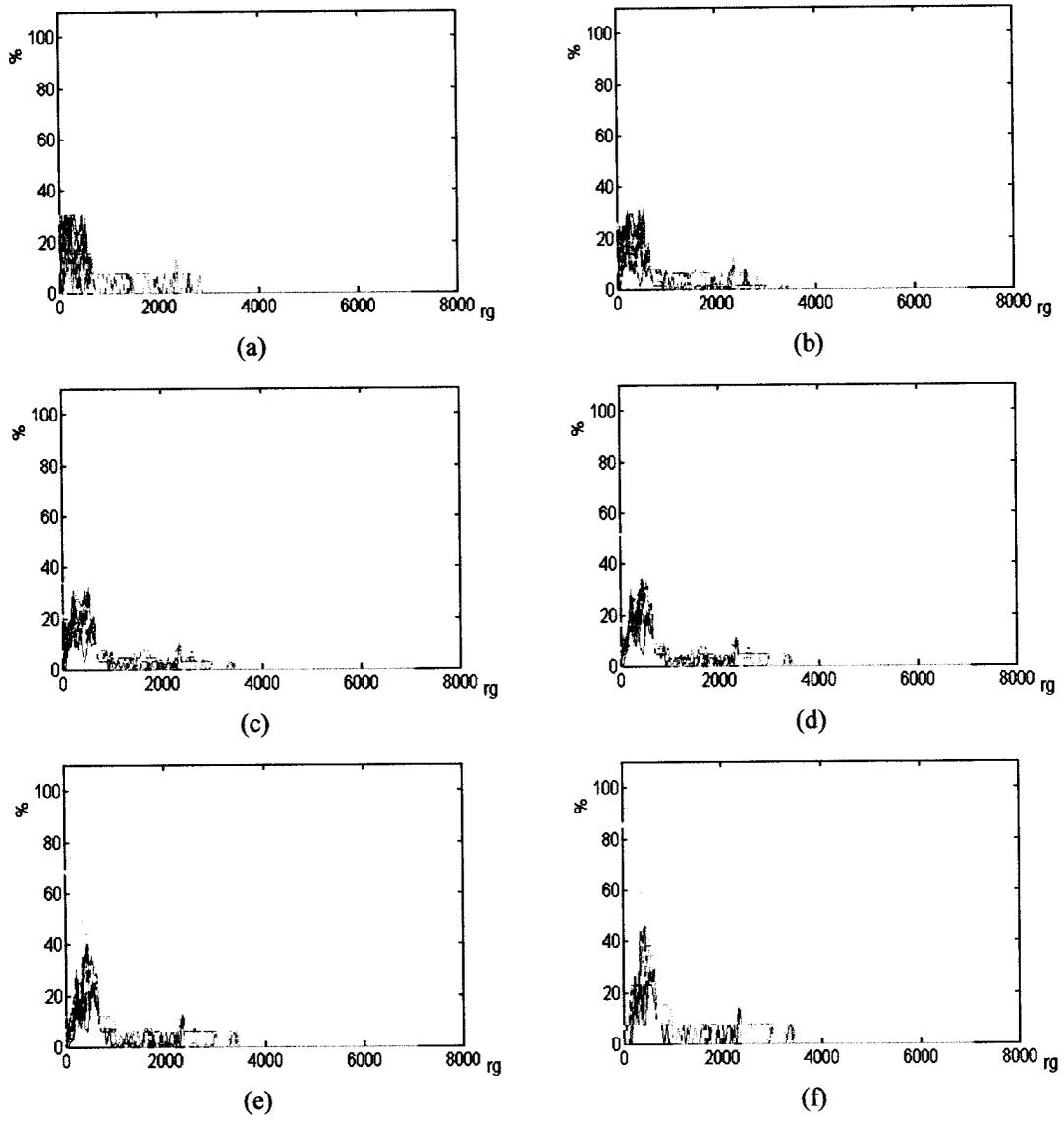


Figura E.11. Coeficiente acumulativo do atributo *rg* para a classe *externas*. (a)-(f): *p* com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

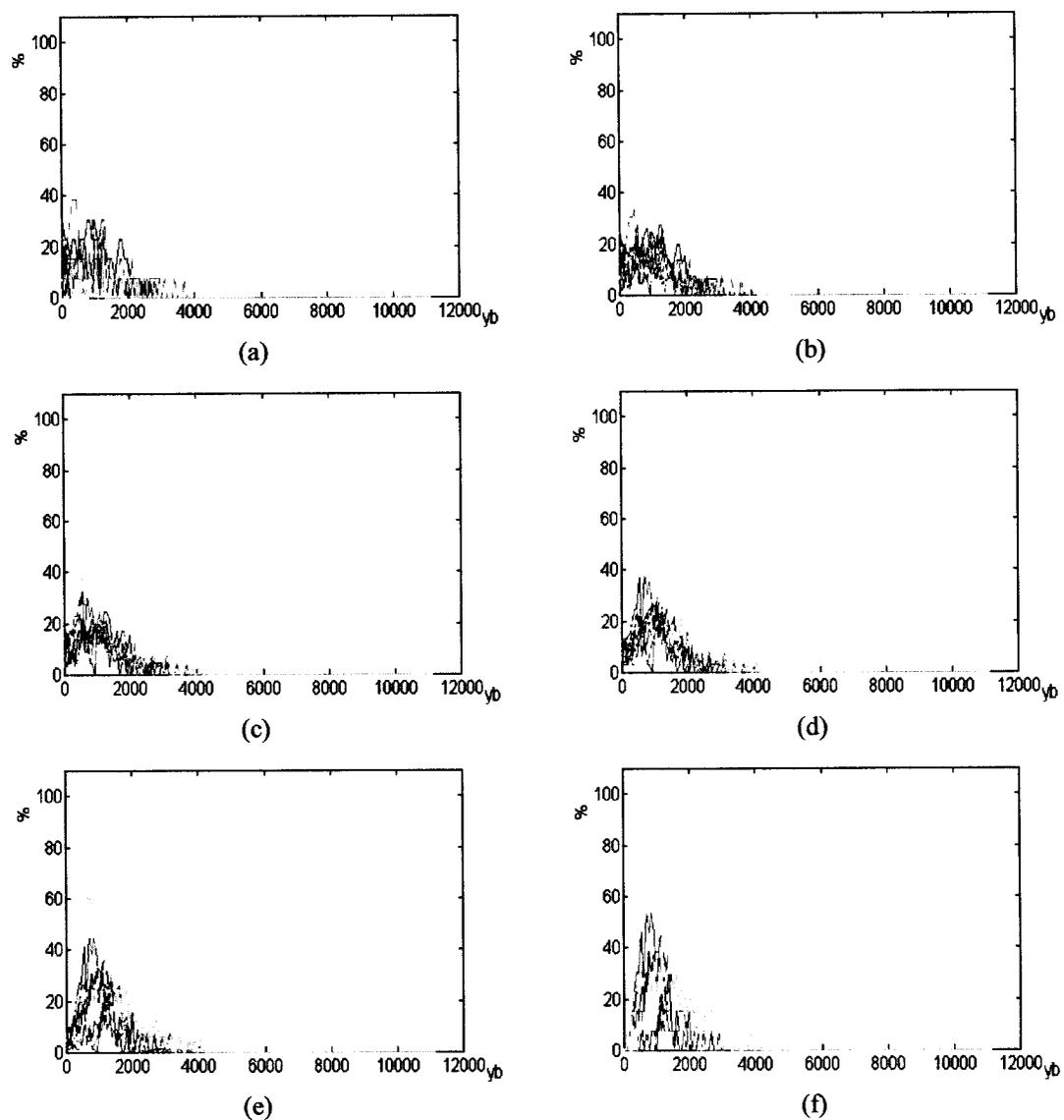


Figura E.12. Coeficiente acumulativo do atributo *yb* para a classe *externas*. (a)-(f): *p* com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

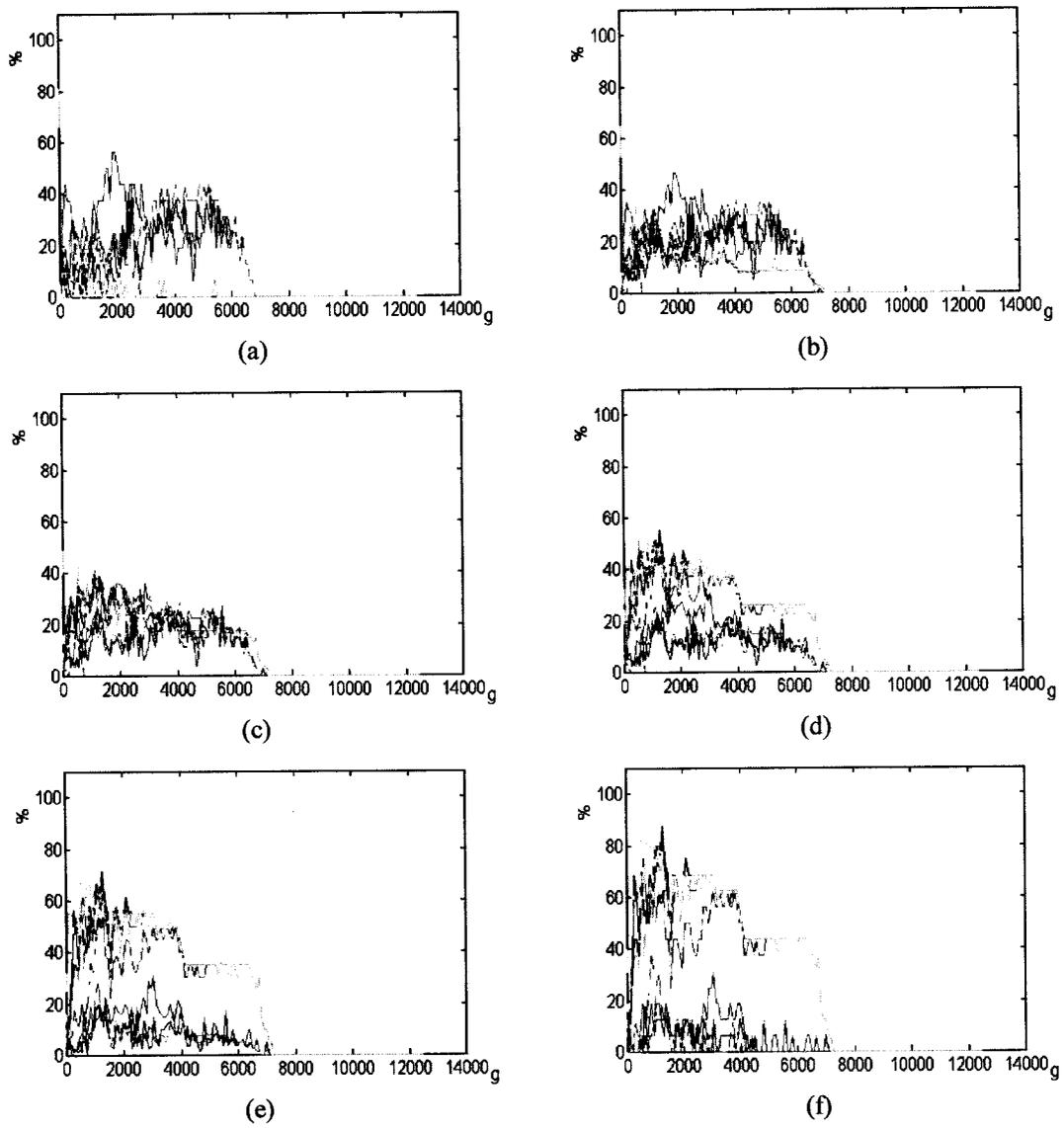
Classe: *laboratório*

Figura E.13. Coeficiente acumulativo do atributo  $g$  para a classe *laboratório*. (a)-(f):  $p$  com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

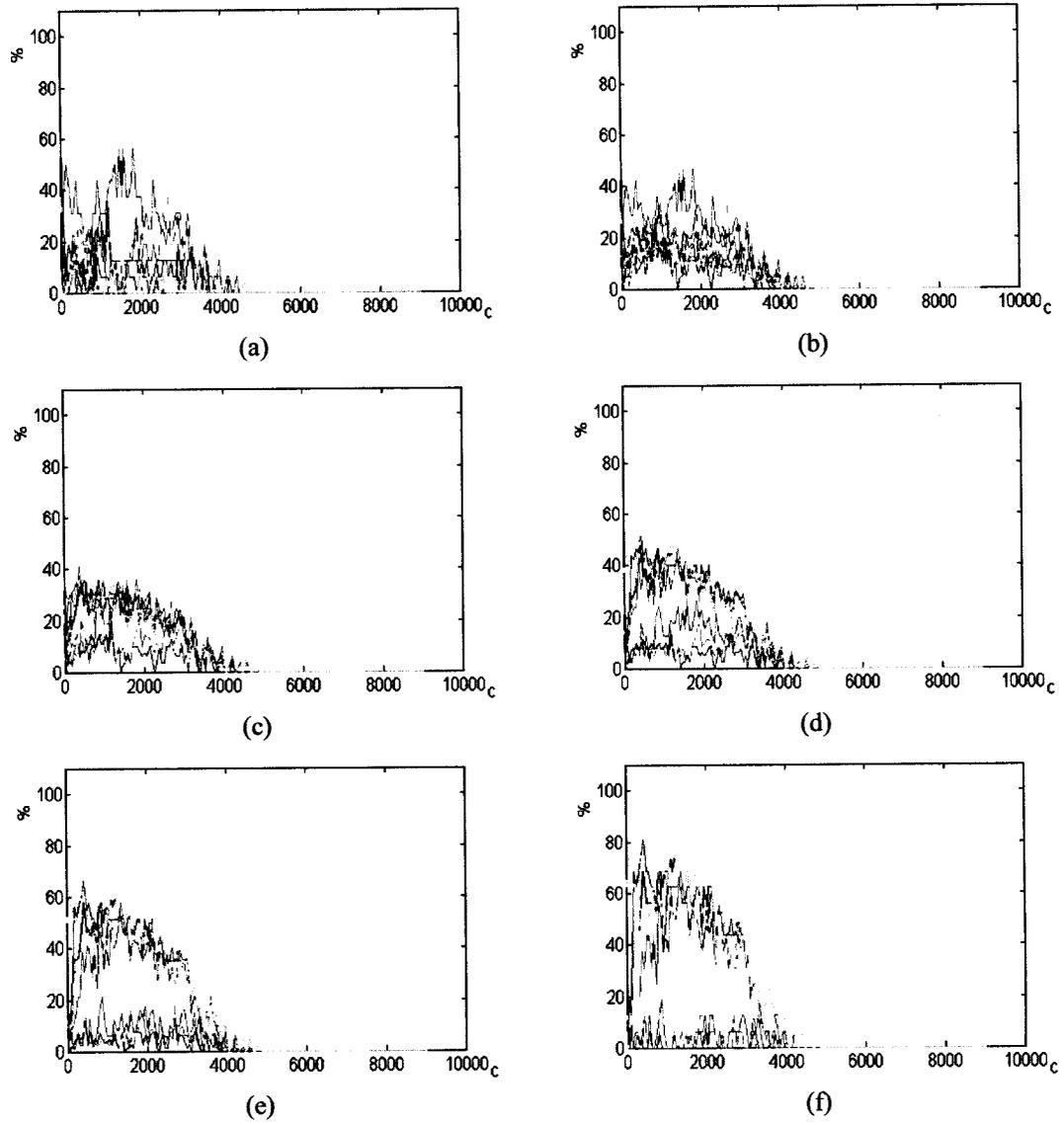


Figura E.14. Coeficiente acumulativo do atributo  $c$  para a classe *laboratório*. (a)-(f):  $p$  com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

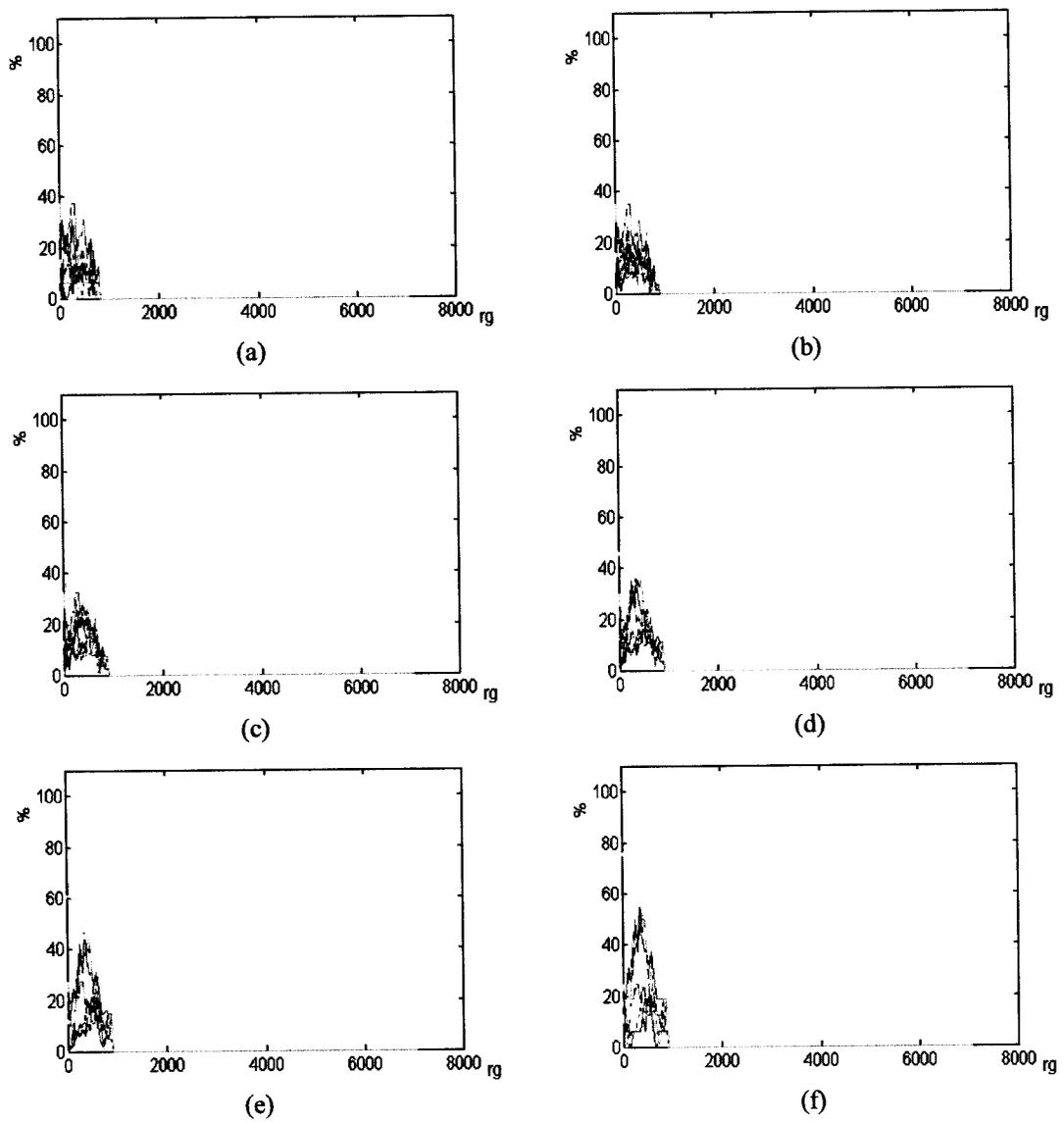


Figura E.15. Coeficiente acumulativo do atributo *rg* para a classe *laboratório*. (a)-(f):  $p$  com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

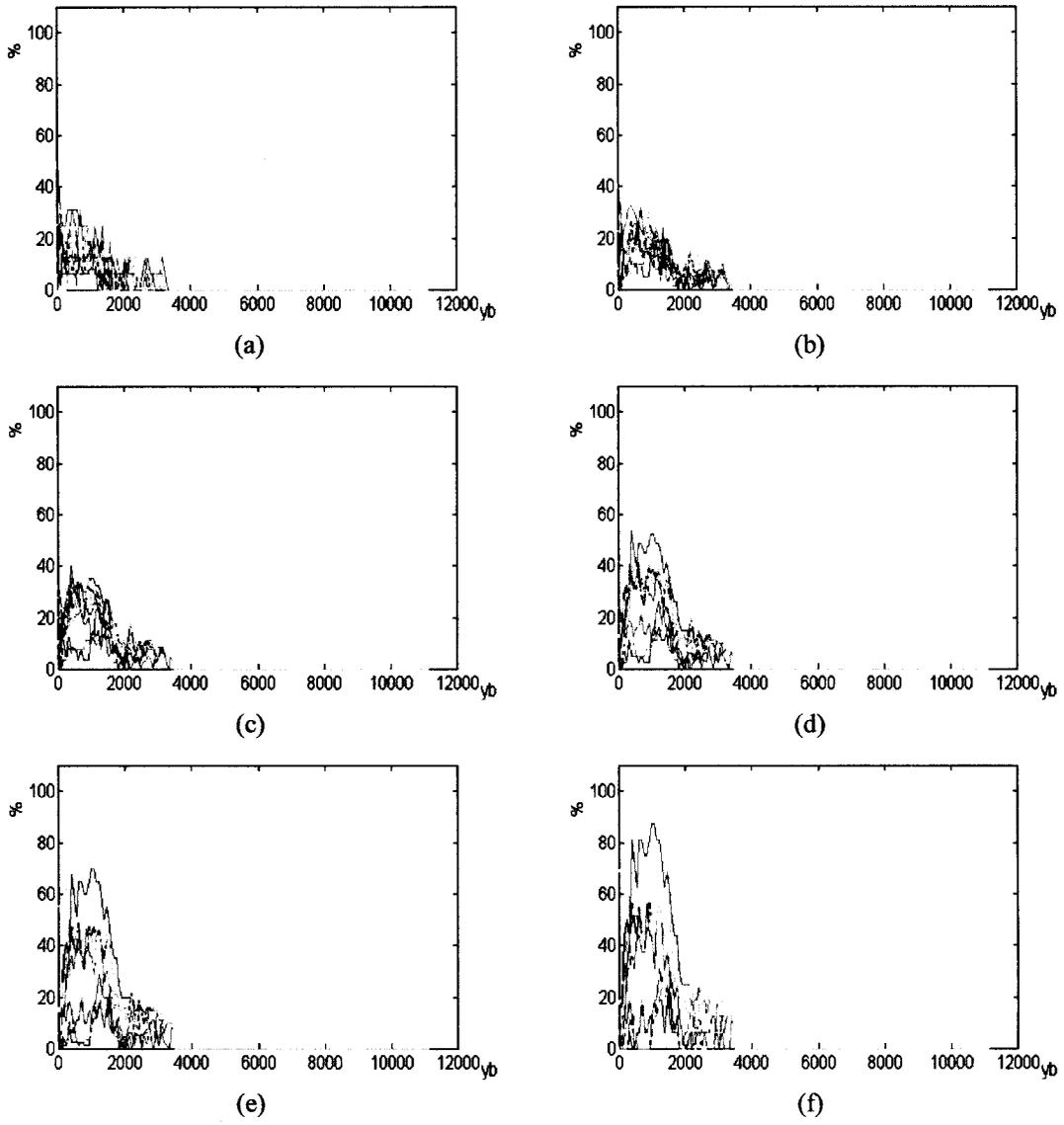


Figura E.16. Coeficiente acumulativo do atributo *yb* para a classe *laboratório*. (a)-(f): *p* com valores de 0, 0,2, 0,4, 0,6, 0,8 e 1, respectivamente.

## **F. Resultados comparativos do coeficiente proporcional para as classes de imagens**

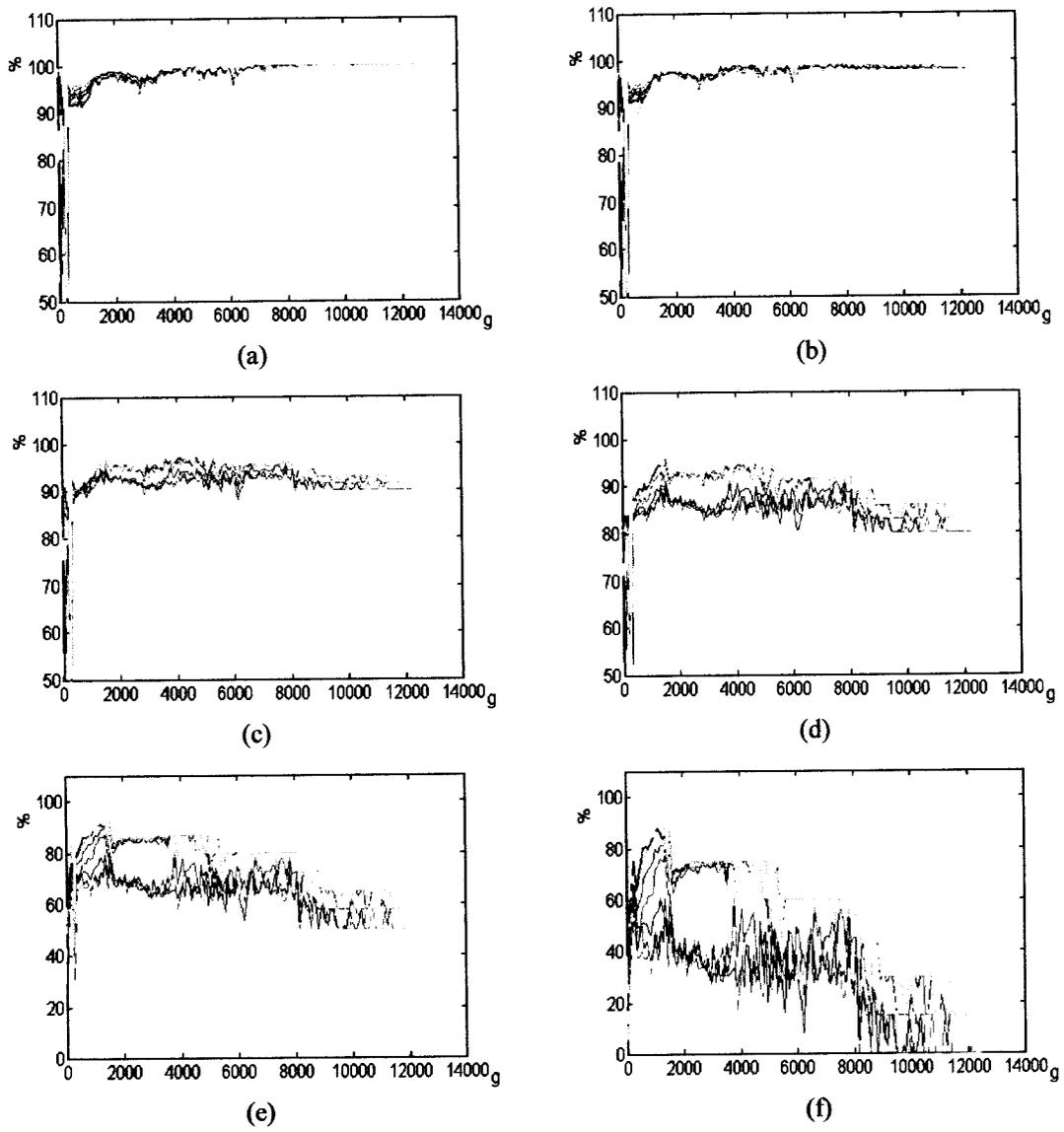
Classe: *casas*

Figura F.1. Coeficiente proporcional do atributo  $g$  para a classe *casas*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

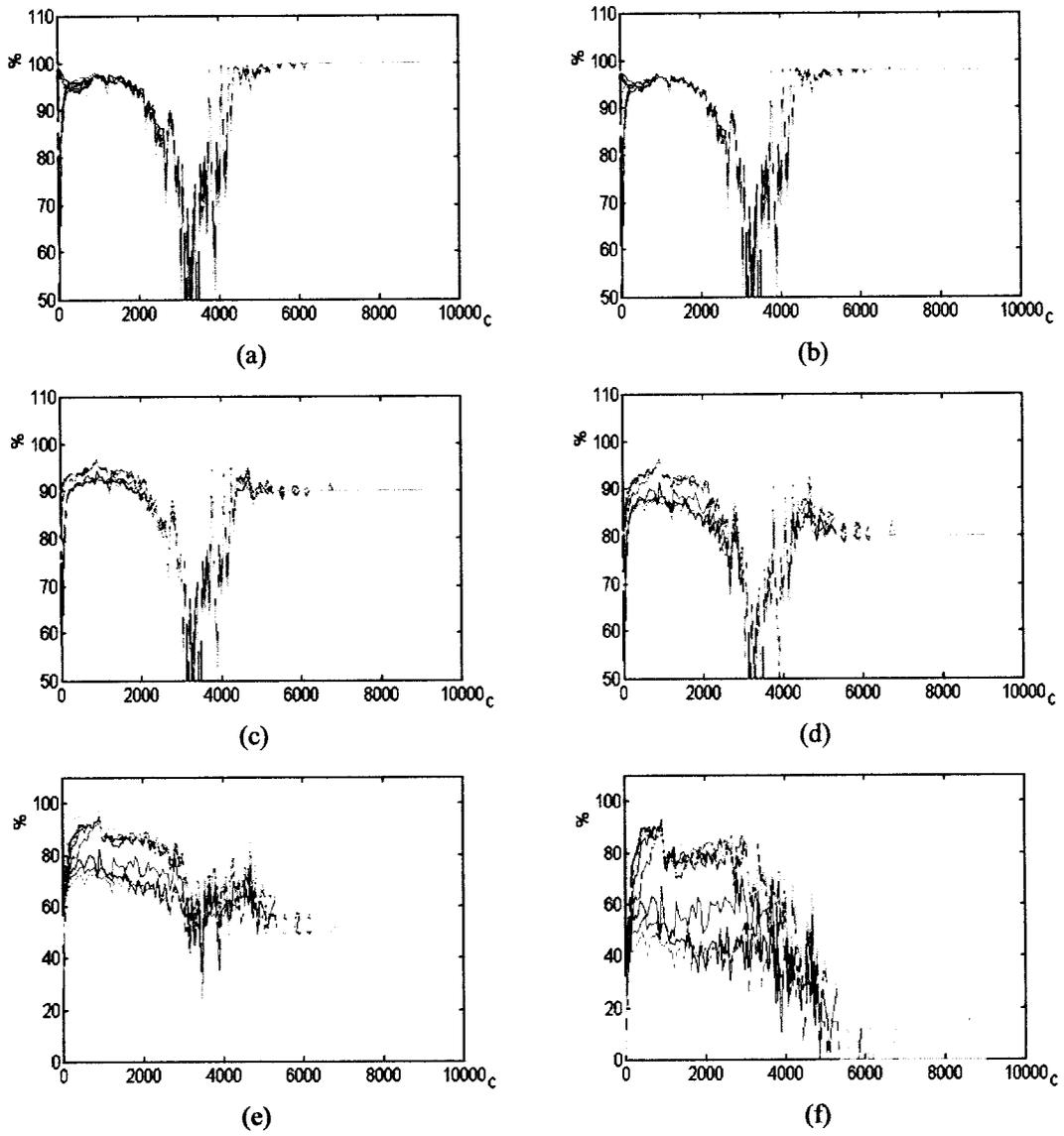


Figura F.2. Coeficiente proporcional do atributo  $c$  para a classe *casas*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

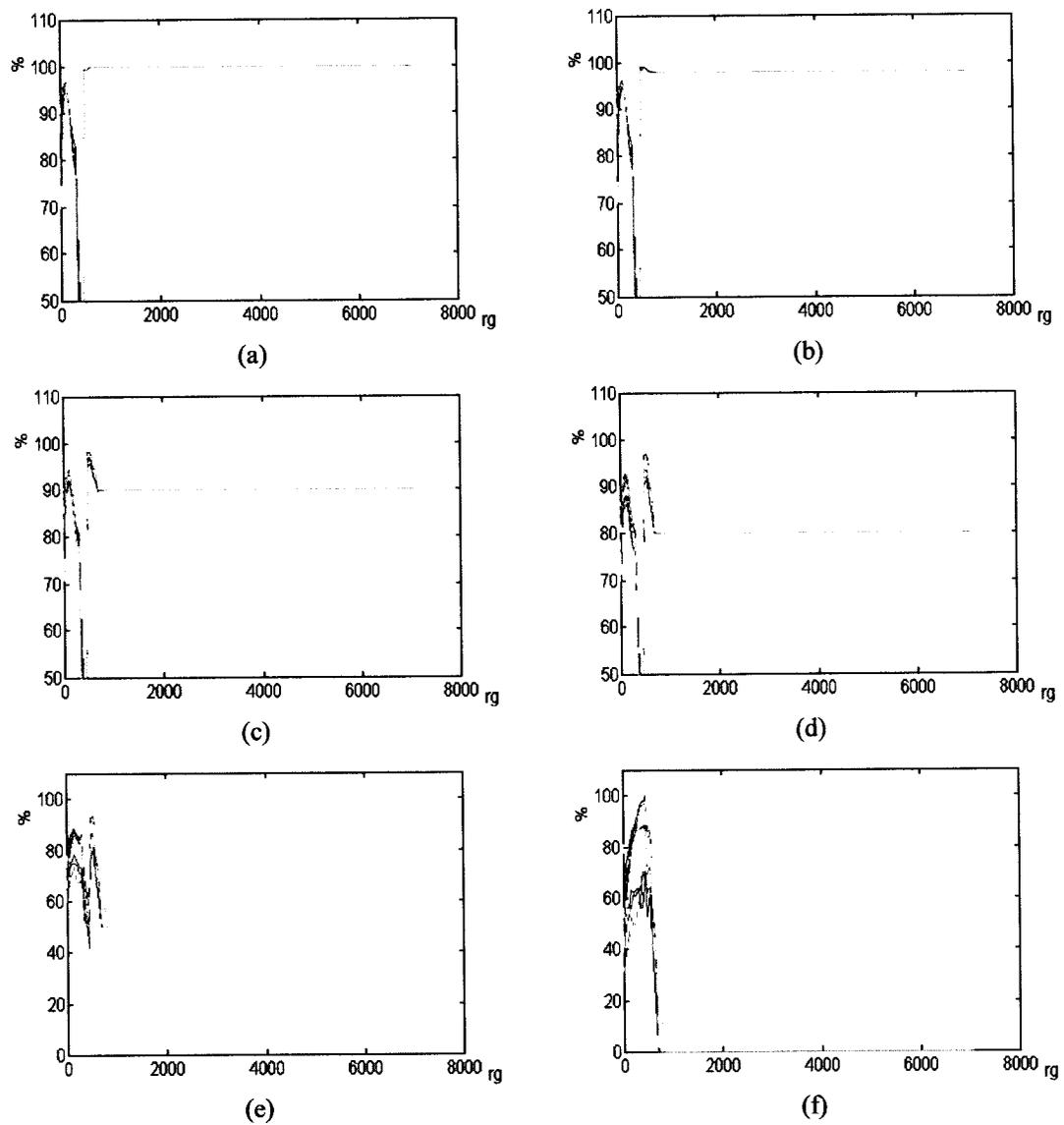


Figura F.3. Coeficiente proporcional do atributo  $rg$  para a classe *casas*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

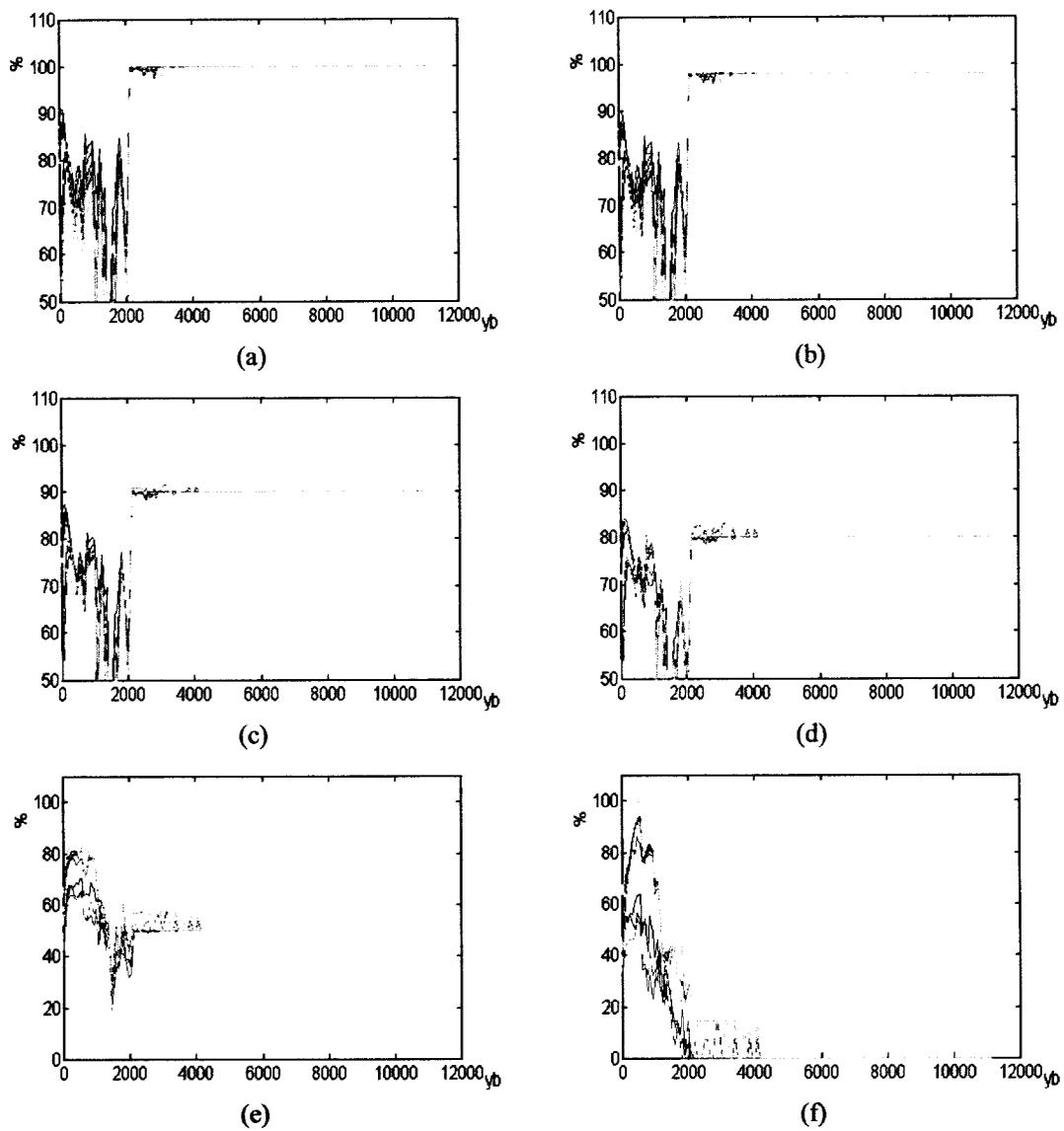


Figura F.4. Coeficiente proporcional do atributo  $y_b$  para a classe *casas*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

Classe:  *cubos*

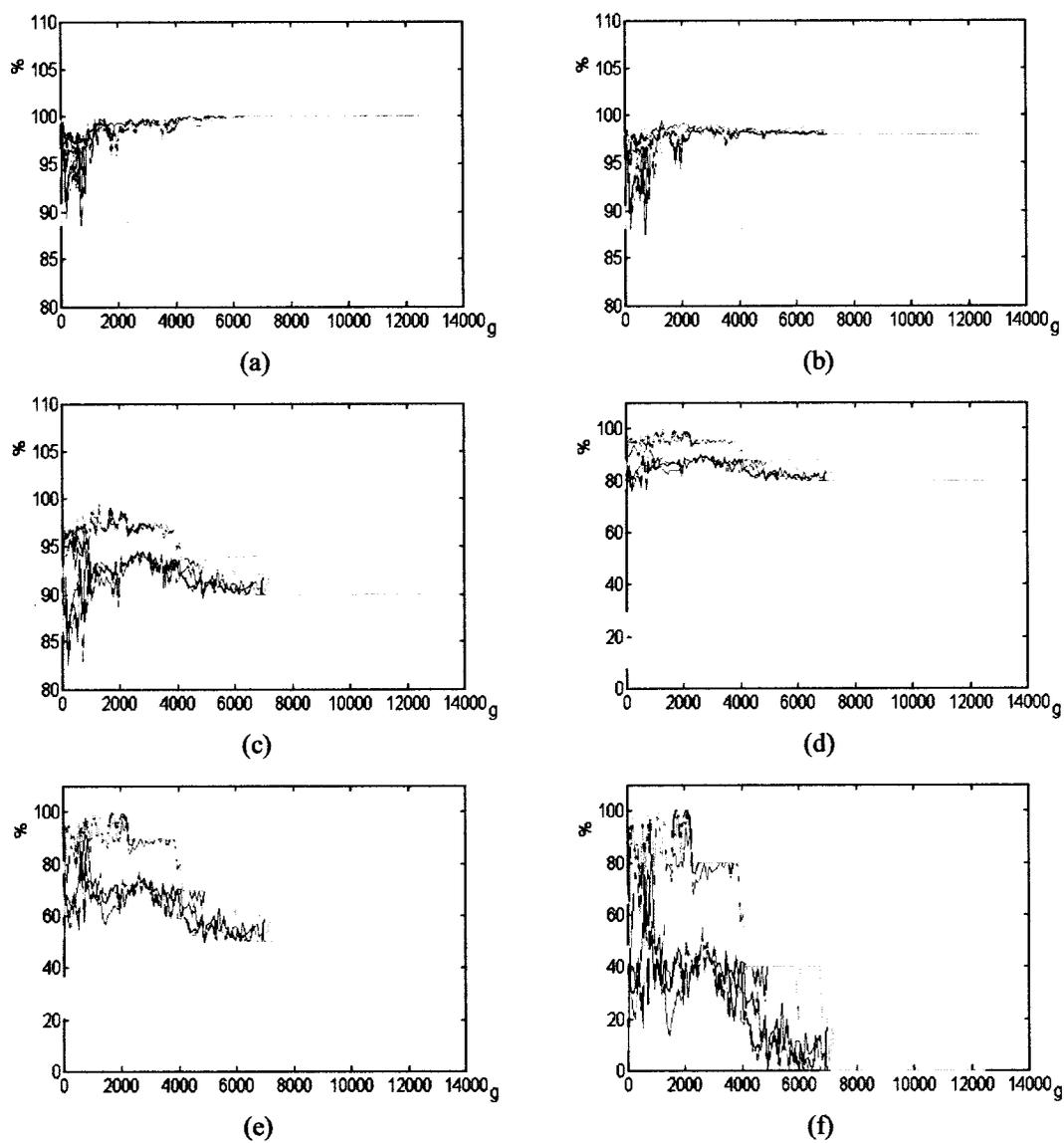


Figura F.5. Coeficiente proporcional do atributo  $g$  para a classe  *cubos*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

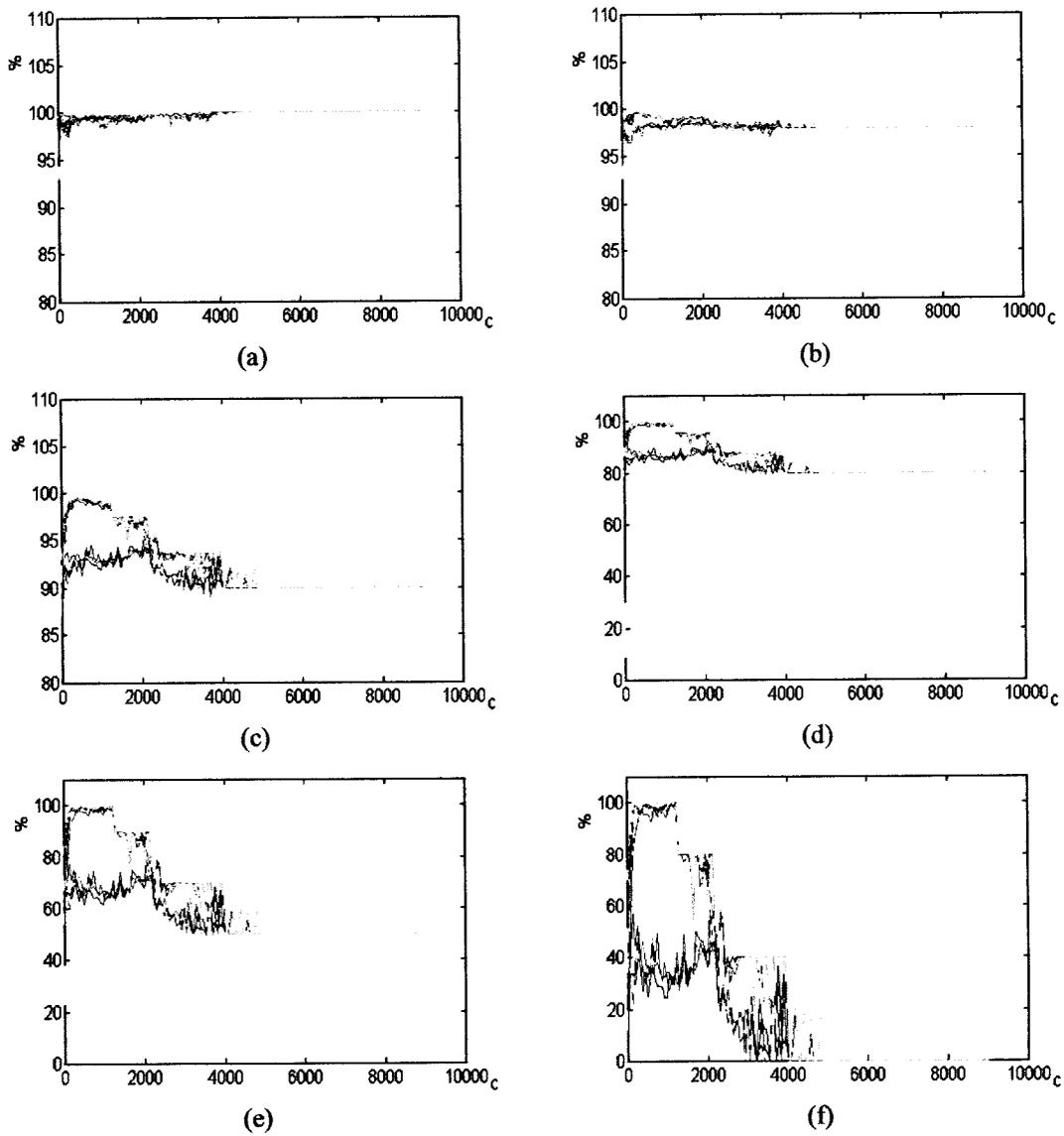


Figura F.6. Coeficiente proporcional do atributo  $c$  para a classe *cubos*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

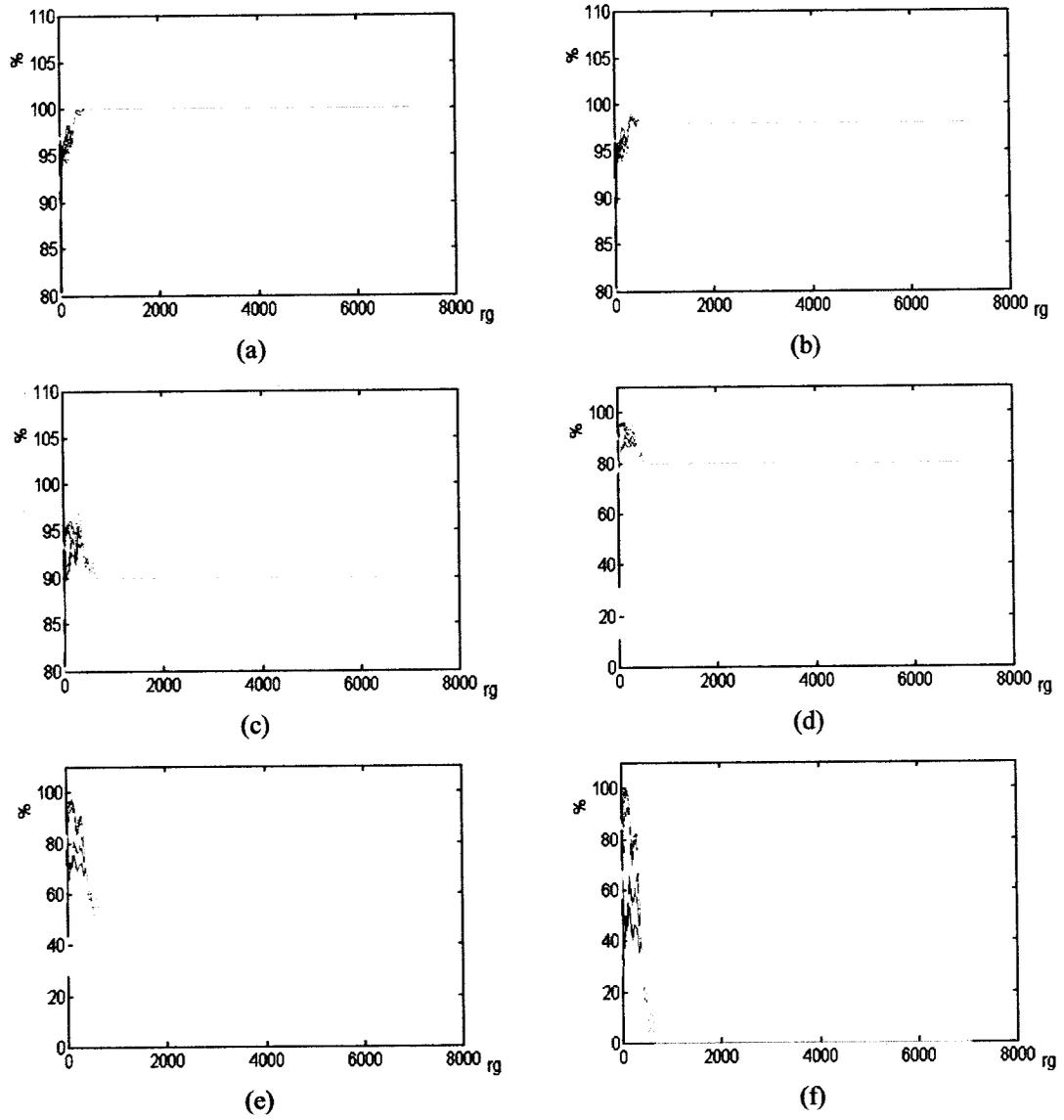


Figura F.7. Coeficiente proporcional do atributo *rg* para a classe *cubos*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

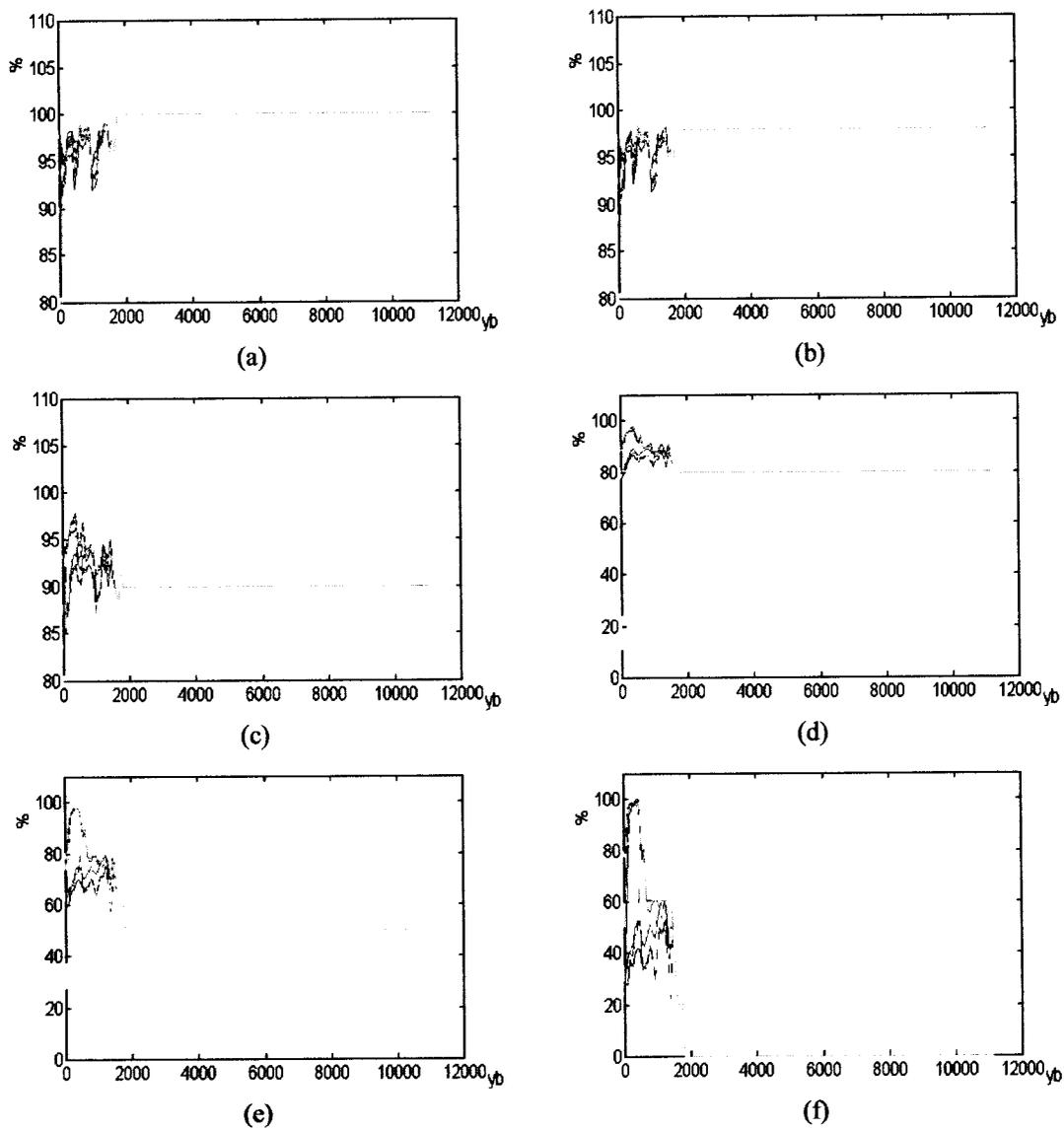


Figura F.8. Coeficiente proporcional do atributo *yb* para a classe *cubos*. (a)-(f): *p* com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

Classe: *externas*

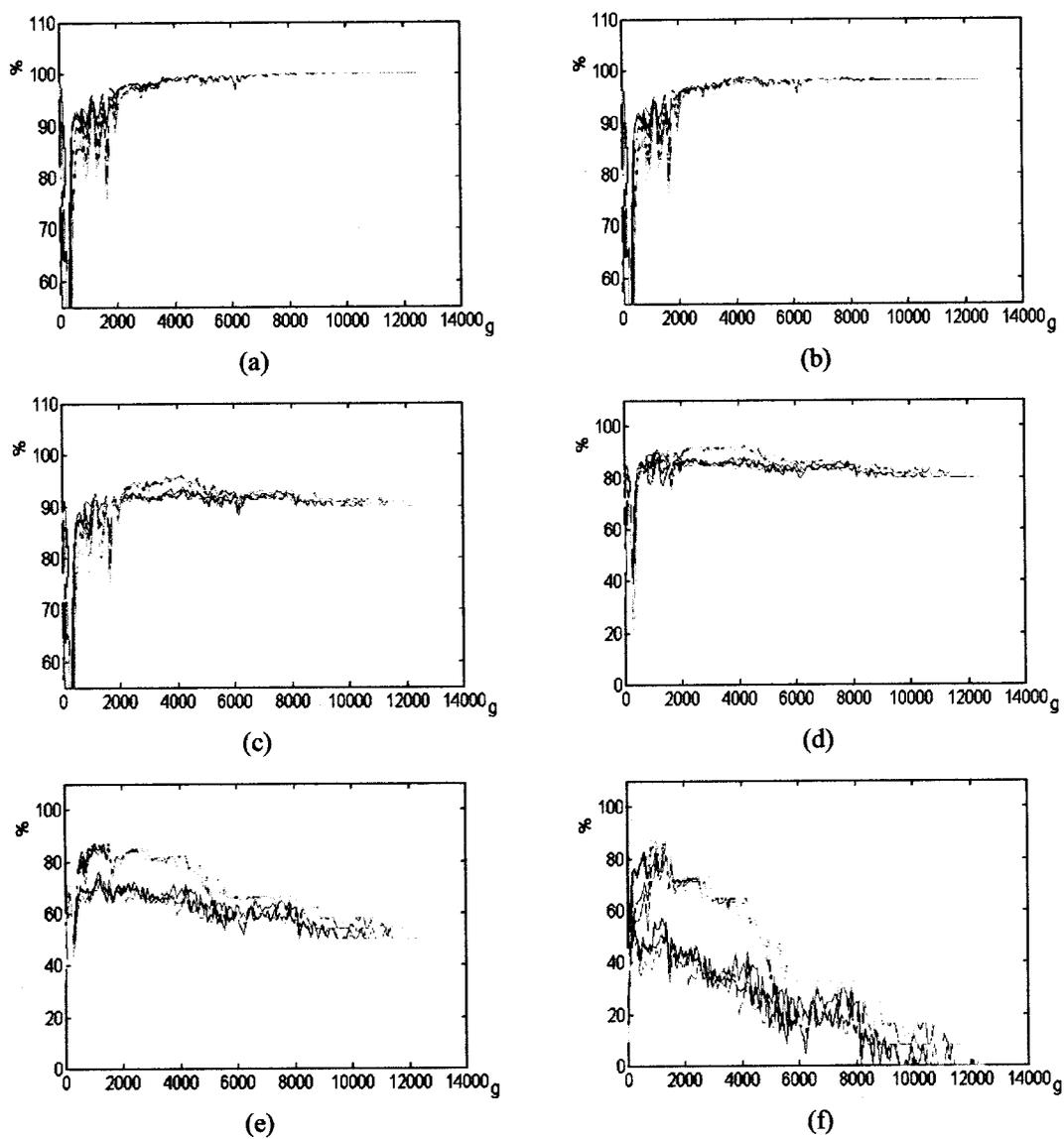


Figura F.9. Coeficiente proporcional do atributo  $g$  para a classe *externas*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

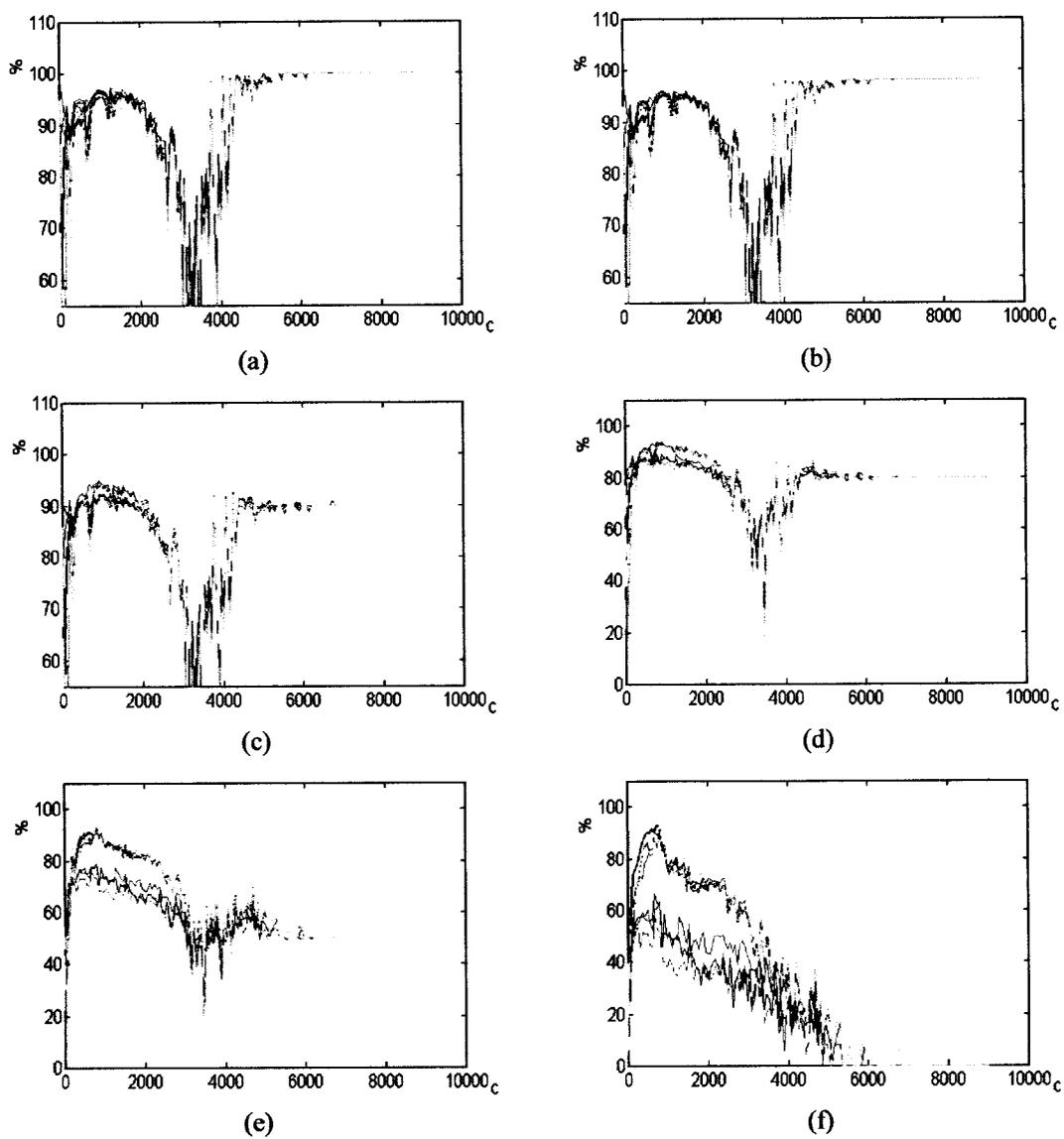


Figura F.10. Coeficiente proporcional do atributo  $c$  para a classe *externas*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

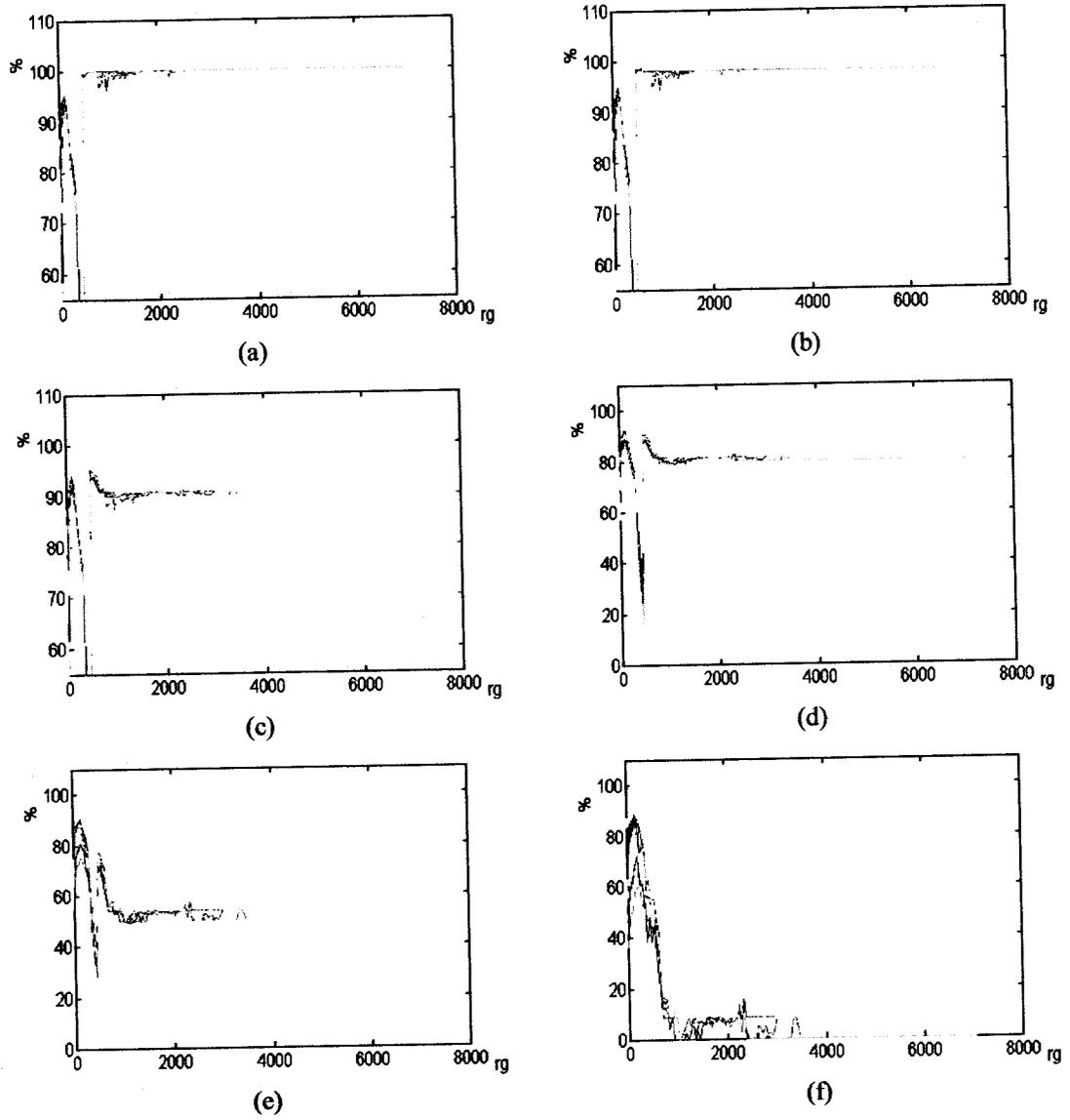


Figura F.11. Coeficiente proporcional do atributo *rg* para a classe *externas*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

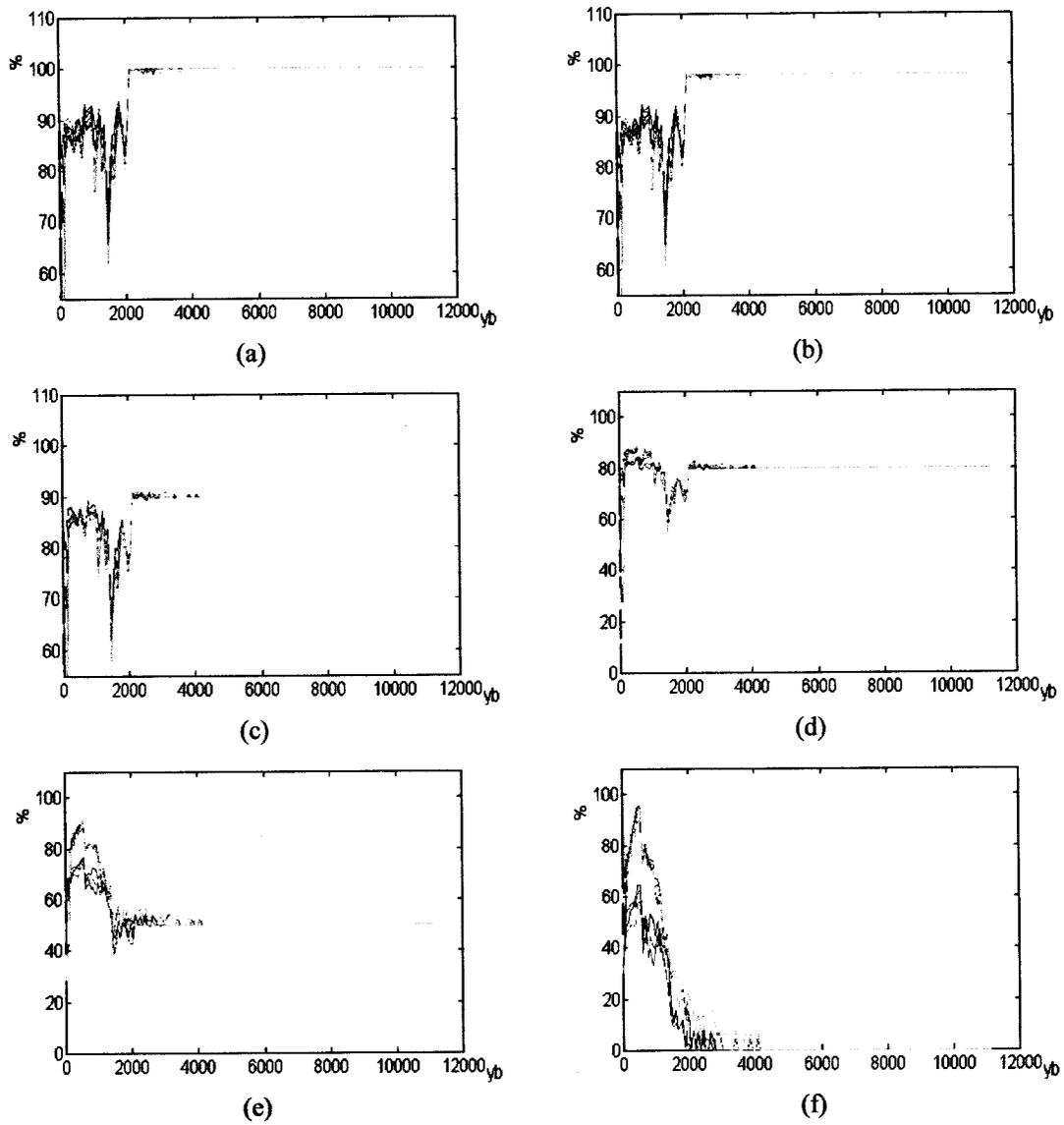


Figura F.12. Coeficiente proporcional do atributo *yb* para a classe *externas*. (a)-(f): *p* com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

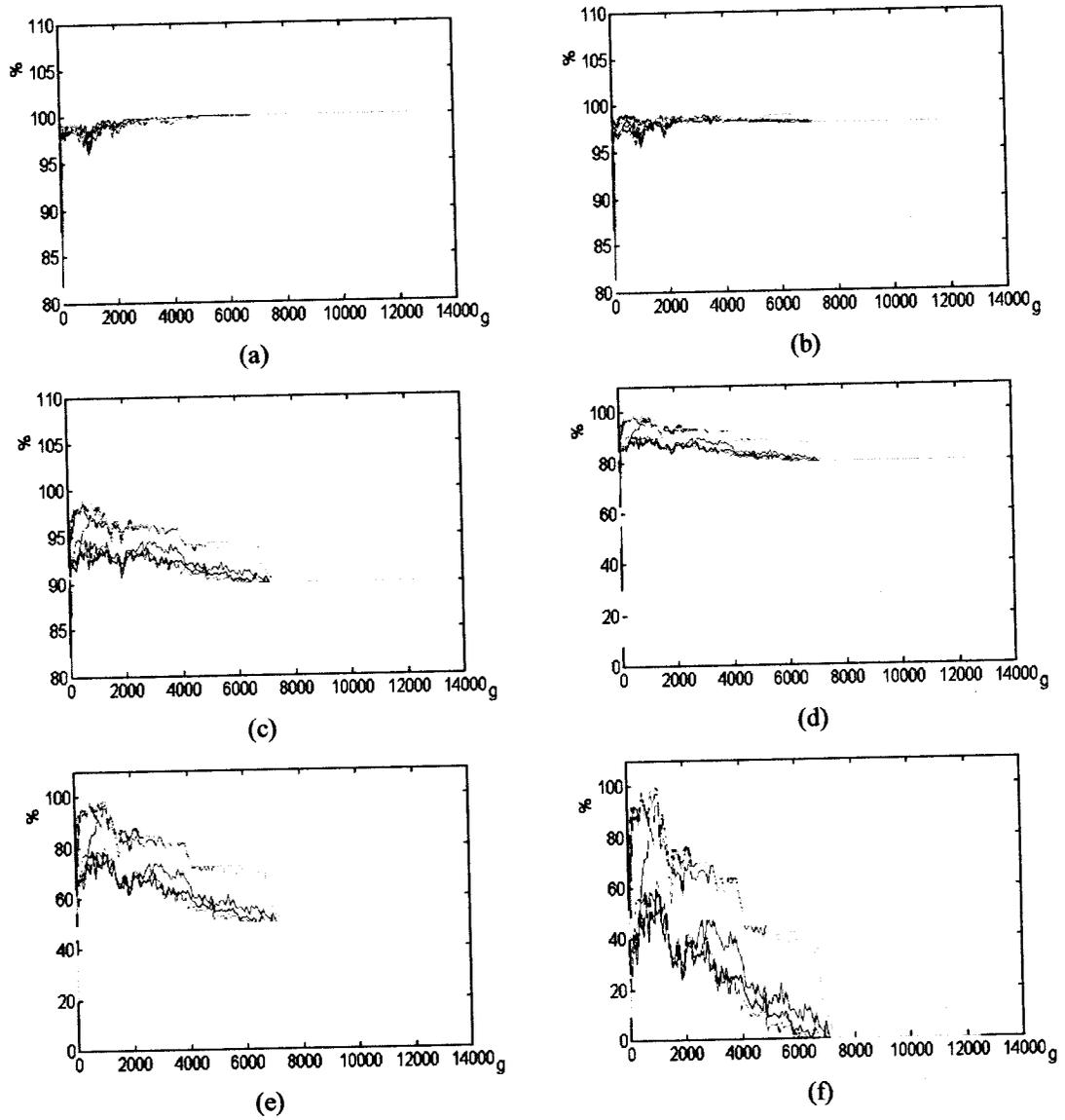
Classe: *laboratório*

Figura F.13. Coeficiente proporcional do atributo  $g$  para a classe *laboratório*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

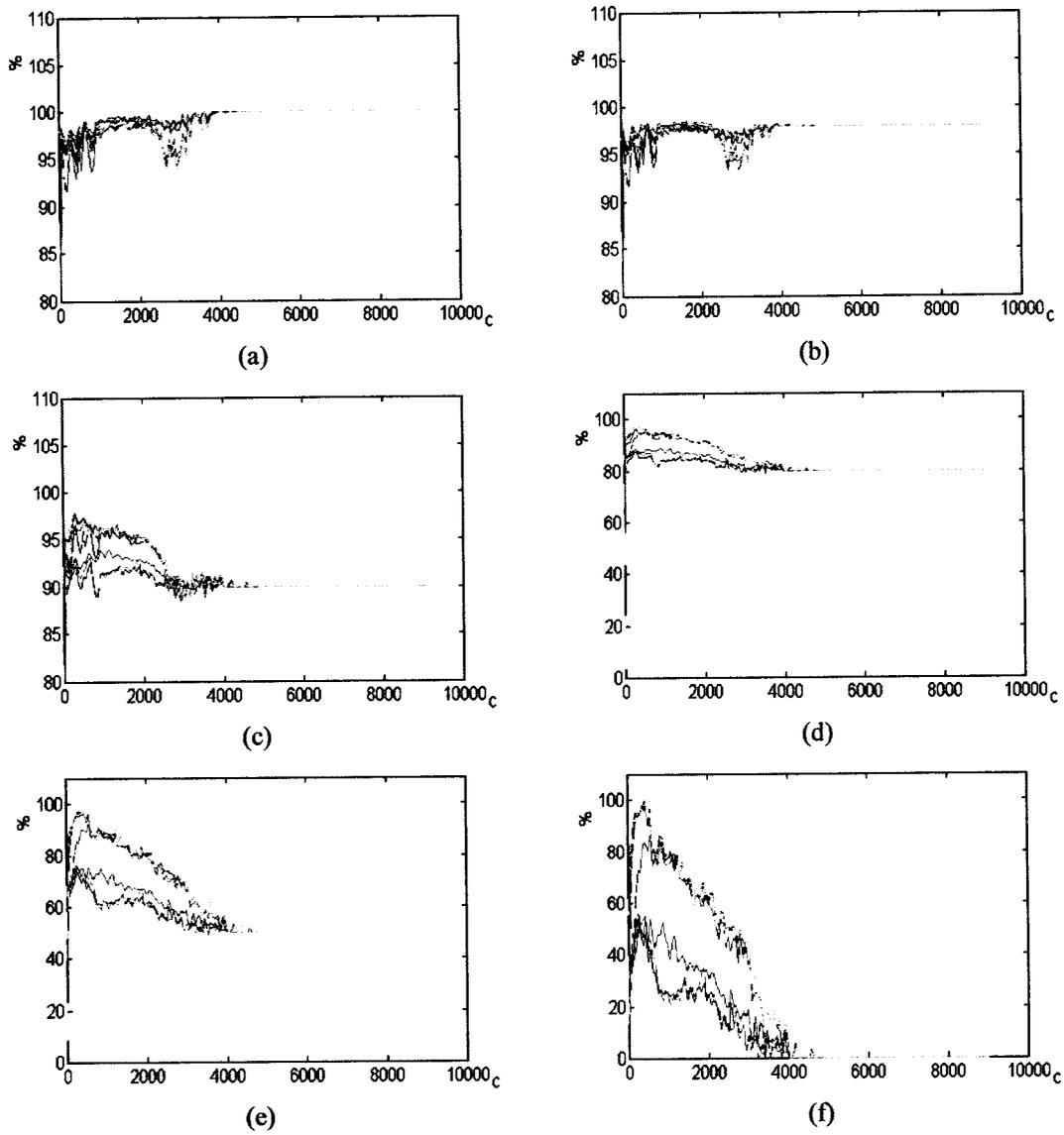


Figura F.14. Coeficiente proporcional do atributo  $c$  para a classe *laboratório*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

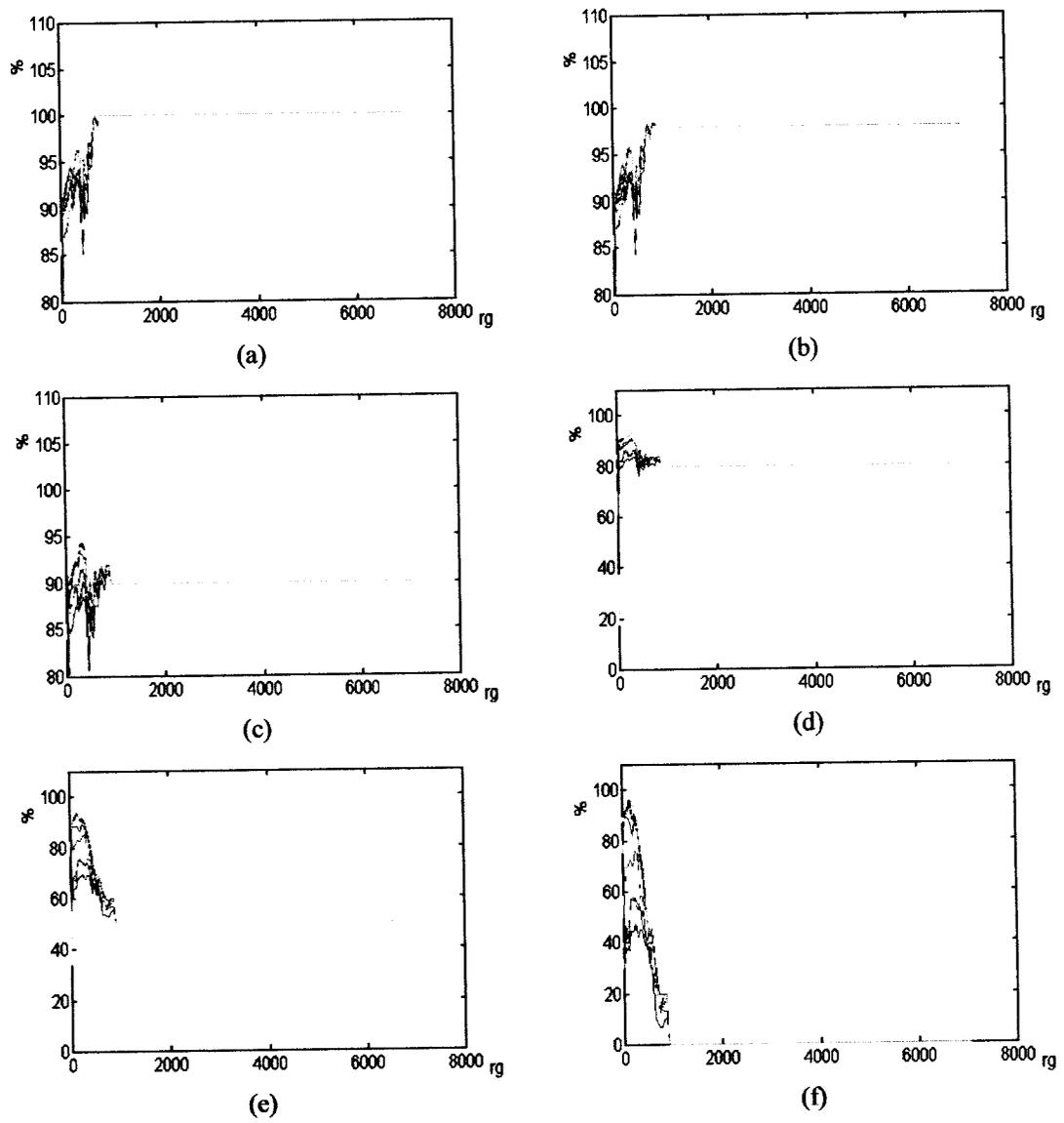


Figura F.15. Coeficiente proporcional do atributo *rg* para a classe *laboratório*. (a)-(f):  $p$  com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.

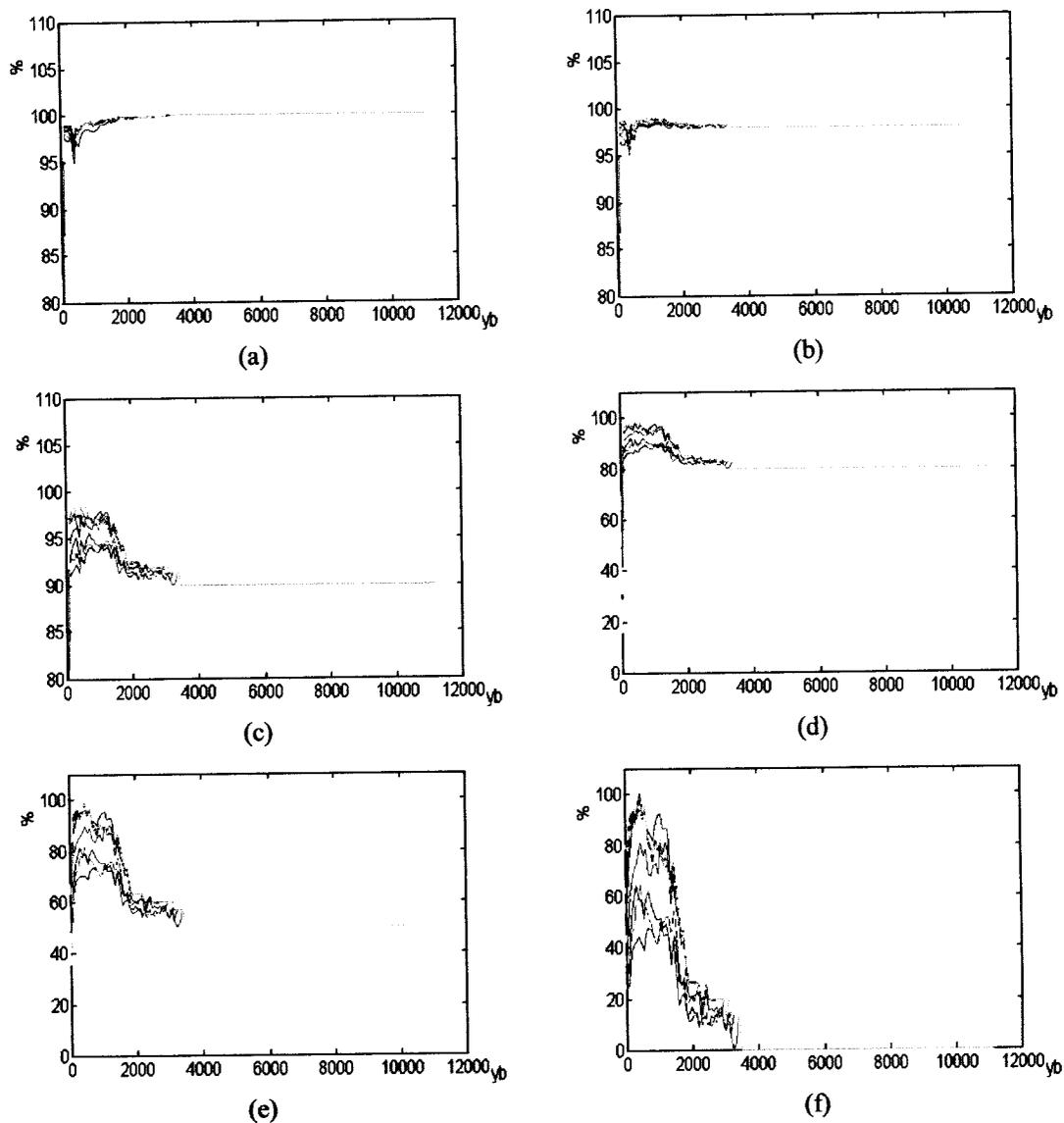


Figura F.16. Coeficiente proporcional do atributo *yb* para a classe *laboratório*. (a)-(f): *p* com valores de 0, 0,02, 0,1, 0,2, 0,5 e 1, respectivamente.