

**HORACIO ANDRÉS LEGAL AYALA**

**UMA ABORDAGEM DE SEGMENTAÇÃO DE  
IMAGENS POR APRENDIZAGEM**

Tese apresentada ao Programa de Pós-Graduação em  
Informática Aplicada da Pontifícia Universidade  
Católica do Paraná como requisito parcial para  
obtenção do título de Doutor em Informática  
Aplicada.

**CURITIBA**

**2004**

**HORACIO ANDRÉS LEGAL AYALA**

**UMA ABORDAGEM DE SEGMENTAÇÃO DE  
IMAGENS POR APRENDIZAGEM**

Tese apresentada ao Programa de Pós-Graduação em  
Informática Aplicada da Pontifícia Universidade  
Católica do Paraná como requisito parcial para  
obtenção do título de Doutor em Informática  
Aplicada.

Área de Concentração: *Metodologias e Técnicas de  
Computação*

Orientador: Prof. Dr. Jacques Facon

Co-orientador: Prof. Dr. Edson José Rodrigues Justino

**CURITIBA**

**2004**

Legal-Ayala, Horacio Andrés

Uma Abordagem de Segmentação de Imagens por Aprendizagem. Curitiba, 2004. 151p.

Tese – Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Informática Aplicada.

1. Segmentação 2. Aprendizagem Computacional 3. Binarização 4. kNN.  
I.Pontifícia Universidade Católica do Paraná. Centro de Ciências Exatas e de Tecnologia. Programa de Pós-Graduação em Informática Aplicada.

Esta página deve ser reservada à ata de defesa e termo de aprovação, que serão fornecidos pela secretaria após a defesa da dissertação e efetuadas as correções solicitadas.

À minha esposa Edith,  
pelo incentivo a este trabalho e principalmente,  
pela compreensão que teve nestes anos.

## Agradecimentos

Desejo agradecer em primeiro lugar a Deus, por acompanhar os meus passos durante todos estes anos, muitas vezes sem que eu perceba.

Aos meus pais, Andrés e Graciela (*in memoriam*), que me inculcaram desde pequeno que o único bem que perdura durante a vida, é o conhecimento adquirido.

Gostaria de agradecer de forma muito especial, ao Professor Orientador Jacques Facon, pela amizade, apóio e excelente orientação, que se remonta à época do mestrado, vários anos atrás.

Ao amigo e Professor Co-orientador Edson Justino, pela ajuda na correção do texto, assim como por contribuições ao trabalho e por disponibilizar a base de imagens de assinaturas.

Meu sincero agradecimento ao prof. Flávio Bortolozzi, pelo apóio financeiro concedido através da Pró-Reitoria de Pesquisa e Pós-Graduação (PRPPG), e pelos comentários dados na correção da tese.

Ao Dr. André F. Gomes, do DAPI – Diagnóstico Avançado por Imagem, por disponibilizar a base de imagens com tumor de ressonância magnética e pelas informações de grande utilidade, que foram fornecidas.

Aos meus ex-colegas e amigos, Eduardo Akira Yonekura, pela ajuda na etapa inicial do Visual C++, e David Menoti, pelos comentários dados na parte de algoritmos.

Aos demais Colegas de estudos, Professores, e Funcionários do PPGIA.

# Sumário

<b>Sumário</b>	viii
<b>Lista de Figuras</b>	xii
<b>Lista de Tabelas</b>	xvi
<b>Lista de Abreviaturas</b>	xix
<b>Resumo</b>	xii
<b>Abstract</b>	xiii
<b>Capítulo 1</b>	
<b>Introdução</b>	1
1.1. Desafios	1
1.2. Proposta	2
1.3. Contribuições	2
1.4. Organização	3
<b>Capítulo 2</b>	
<b>Segmentação</b>	5
2.1. Segmentação por Região	6
2.1.1. Binarização	6
2.1.1.1. Binarização Global	7
2.1.1.2. Binarização Local Adaptativa	7
2.1.2. Método de Divisão e Fusão	8
2.1.3. Multi-binarização	8
2.2. Segmentação por Contorno	8
2.3. Segmentação por Textura	9
2.4. Segmentação por Modelos de Contornos Ativos	10
2.5. Segmentação Morfológica	10
2.6. Classificação da Abordagem do Ponto de Vista de Segmentação	11

2.7. Comentários Finais	13
<b>Capítulo 3</b>	
<b>A Abordagem Proposta de Segmentação por Aprendizagem</b>	15
3.1. Aprendizagem Computacional	15
3.2. Abordagem Proposta	17
3.3. Estudo Preliminar de Abordagens Anteriores	18
3.4. Fundamentos da Abordagem Proposta	20
3.5. A Seleção das Características	21
3.6. A Etapa de Aprendizagem	25
3.7. A Escolha do Tamanho Ideal da Janela	28
3.8. A Construção da Matriz de Decisão	32
3.9. Segmentação de Novas Imagens por Aprendizagem	32
3.9.1. Classificador Supervisionado	33
3.9.2. A Regra do Vizinho Mais Próximo (1-NN)	34
3.9.3. Distância Euclidiana	34
3.9.4. Taxa de Erro da Classificação pela Regra do Vizinho Mais Próximo	35
3.9.5. A Regra dos $k$ Vizinhos Mais Próximos ( $k$ -NN)	35
3.9.6. O Valor de $k$ , o Número de Vizinhos Mais Próximos	36
3.9.7. Características da Regra dos $k$ -Vizinhos Mais Próximos	37
3.10. Comentários Finais	37
<b>Capítulo 4</b>	
<b>Considerações sobre as Bases de Imagens Utilizadas</b>	39
4.1. Imagens de Assinaturas	39
4.1.1. Introdução	39
4.1.2. Descrição da Base de Imagens de Assinaturas Utilizada	40
4.2. Imagens de Impressões Digitais	41
4.2.1. Introdução	41
4.2.2. Descrição da Base de Imagens de Impressões Digitais Utilizada	42
4.3. Imagens de Ressonância Magnética	42
4.3.1. Introdução	42
4.3.2. Descrição das Bases de Imagens de Ressonância Magnética (MRI) utilizadas	43

4.4. Imagens de Envelopes Postais	44
4.4.1. Introdução	44
4.4.1. Descrição da Base de Imagens de Envelopes Postais Utilizada	46
4.5. Comentários Finais	47
<b>Capítulo 5</b>	
<b>Experimentos de Aplicações</b>	49
5.1. Metodologia de Avaliação dos Resultados	49
5.1.1. Metodologia de Avaliação para Duas Classes	49
5.1.2. Metodologia de Avaliação para Quatro Classes	50
5.2. Metodologia de Avaliação do Desempenho de cada Característica	52
5.3. Testes de Segmentação em Imagens de Assinaturas	54
5.4. Testes de Segmentação em Imagens de Impressões Digitais	63
5.5. Testes de Segmentação por Multi-região de Imagens de Ressonância Magnética	70
5.5.1. Testes de Segmentação da Massa Branca de Imagens de Ressonância Magnética	71
5.5.2. Testes de Segmentação da Massa Cinza de Imagens de Ressonância Magnética	76
5.5.3. Testes de Segmentação do Líquor de Imagens de Ressonância Magnética	82
5.6. Testes de Segmentação em Imagens de Envelopes Postais	90
5.6.1. Calibração dos Histogramas das Imagens a serem Segmentadas	90
5.6.2. Testes de Segmentação na Base ECT-DB1	92
5.6.3. Testes de Segmentação na Base ECT-DB2	98
5.7. Discussão dos Resultados	105
5.8. Comentários Finais	108
<b>Capítulo 6</b>	
<b>Conclusões</b>	109
<b>Capítulo 7</b>	
<b>Trabalhos Futuros</b>	112
7.1. Pós-processamentos	112
7.2. Pesquisa de Características Adicionais	114

7.3. Implementação de Algoritmos Mais Eficientes de Busca dos $k$ Vizinhos Mais Próximos	114
7.4. Implementação de Outras Ferramentas de Similaridade	115
7.5. Implementação de Outras Métricas	115
7.6. Experimentos de Segmentação de Outras Bases de Imagens	115
7.6.1. Segmentação de Rostos	115
7.6.2. Segmentação de Caracteres Coloridos em Imagens da Bíblia de Gutemberg	115
7.6.3. Quantificação dos Resultados de Segmentação de MRI de Cérebro com tumor	118
<b>Referências Bibliográficas</b>	122
<b>Apêndice A</b>	
<b>Exemplo Prático de Resolução de Conflitos e Cálculo do Tamanho Adequado da Janela</b>	134
<b>Apêndice B</b>	
<b>Estudo e Implementação de Estruturas de Otimização do Processo de Aprendizagem</b>	140
<b>Apêndice C</b>	
<b>Publicações Decorrentes da Pesquisa</b>	150

## Lista de Figuras

Figura 2.1	Comparação de resultados de binarização em imagem de ressonância magnética	12
Figura 2.2	Comparação de resultados de binarização em imagem de impressões digitais	14
Figura 3.1	Esquema geral da abordagem proposta de Segmentação por Aprendizagem	18
Figura 3.2.	Exemplificação do coeficiente de assimetria	23
Figura 3.3.	Exemplificação do coeficiente de apontamento	24
Figura 3.4	Estrutura do vetor de características	26
Figura 3.5	A etapa de aprendizagem e definição automática do tamanho da janela	27
Figura 3.6	Exemplo de presença de inconsistência na imagem	29
Figura 3.7	Gráfico do total de erros ( $\sum \epsilon$ ) vs. iteração	31
Figura 3.8	Pseudo-código que ilustra o critério de parada do algoritmo de aprendizagem	31
Figura 3.9	Processo de Segmentação de Imagens	33
Figura 3.10	Regras do 1-NN e do 3-NN	36
Figura 4.1	Exemplos da base de imagens de assinaturas	40
Figura 4.2	Exemplos da base de imagens de impressões digitais	41
Figura 4.3	Exemplos da base de imagens de ressonância magnética	44
Figura 4.4	Exemplos da base de imagens de envelopes postais	48
Figura 5.1	Gráfico da Segmentação de imagens de Assinaturas e o correspondente Ruído para cada característica isolada e o número de vizinhos mais próximos	56
Figura 5.2	Modelo utilizado para a codificação das combinações de características	57
Figura 5.3	Gráfico da amostra de segmentação por binarização de Assinaturas para todas as combinações de características, k-NN =1	58
Figura 5.4	Gráfico da amostra de segmentação por binarização de Assinaturas para todas as combinações de características, k-NN =11	59

Figura 5.5	Gráfico da Segmentação de imagens de Assinaturas e o correspondente Ruído para as características Média-Skewness e o número de vizinhos mais próximos	60
Figura 5.6	Resultados comparativos de segmentação em imagem de assinatura	61-62
Figura 5.7	Gráfico da Segmentação de imagens de Impressões digitais e o correspondente Ruído para cada característica isolada e o número de vizinhos mais próximos	64-65
Figura 5.8	Gráfico da amostra de segmentação por binarização de Impressões digitais para todas as combinações de características, k-NN =1	66
Figura 5.9	Gráfico da amostra de segmentação por binarização de Impressões digitais para todas as combinações de características, k-NN =15	67
Figura 5.10	Gráfico da Segmentação de imagens de Impressões digitais e o correspondente Ruído para a característica Média e o número de vizinhos mais próximos	68
Figura 5.11	Resultados comparativos de segmentação em imagem de impressão digital	69-70
Figura 5.12	Gráfico da Segmentação de imagens de MRI-WM e o correspondente Ruído para cada característica isolada e o número de vizinhos mais próximos	72
Figura 5.13	Gráfico da amostra de segmentação por multi-região de MRI-WM para todas as combinações de características, k-NN =1	73
Figura 5.14	Gráfico da amostra de segmentação por multi-região de MRI-WM para todas as combinações de características, k-NN =3	74
Figura 5.15	Gráfico da Segmentação multi-região de MRI-WM e o correspondente Ruído para as características Média-Variância e o número de vizinhos mais próximos	75
Figura 5.16	Gráfico da Segmentação por multi-região de MRI-GM e o correspondente Ruído para cada característica isolada e o número de vizinhos mais próximos	77-78
Figura 5.17	Gráfico da amostra de segmentação por multi-região de MRI-GM para todas as combinações de características, k-NN =1	79
Figura 5.18	Gráfico da amostra de segmentação por multi-região de MRI-GM para todas as combinações de características, k-NN =11	80
Figura 5.19	Gráfico da Segmentação multi-região de MRI-GM e o correspondente Ruído para as características Média-Variância-Skewness e o número de vizinhos mais próximos	81
Figura 5.20	Gráfico da Segmentação por multi-região de MRI-CSF e o correspondente Ruído para cada característica isolada e o número de vizinhos mais próximos	83-84
Figura 5.21	Gráfico da amostra de segmentação por multi-região de MRI-CSF para todas as combinações de características, k-NN =1	85
Figura 5.22	Gráfico da amostra de segmentação por multi-região de MRI-CSF para todas as combinações de características, k-NN =5	86
Figura 5.23	Gráfico da Segmentação multi-região de MRI-CSF e o correspondente Ruído para as características Média-Variância-Skewness-Curtosis e o número de vizinhos mais próximos	87

Figura 5.24	Segmentação da massa cinza GM, massa branca WM e líquido CSF de imagens de MR por aprendizagem	88-89
Figura 5.25	Exemplo prático do processo de correção de histograma nas imagens a serem segmentadas	91
Figura 5.26	Gráfico da Segmentação de imagens de Envelopes e o correspondente Ruído para cada característica isolada e o número de vizinhos mais próximos	94
Figura 5.27	Gráfico da amostra de segmentação por binarização de Envelopes para todas as combinações de características, $k\text{-NN} = 1$	95
Figura 5.28	Gráfico comparativo dos resultados de segmentação por binarização de Envelopes, utilizando-se cada uma das imagens de aprendizagem por vez, para todas as combinações de características, $k\text{-NN} = 1$	96
Figura 5.29	Gráfico da Segmentação de imagens de Envelopes e o correspondente Ruído para a características Média-Variância e o número de vizinhos mais próximos	97
Figura 5.30	Resultados de segmentação por binarização em imagem de envelope	101-4
Figura 7.1	Exemplo de pós-processamento por reconstrução morfológica	113
Figura 7.2	Resultados preliminares de segmentação de rosto	116
Figura 7.3	Resultados preliminares da segmentação em uma página da Bíblia de Gutemberg	117
Figura 7.4.	Exemplos da base de imagens de ressonância magnética	119
Figura 7.5	Segmentação de tumores em imagens MR cerebrais sem realce por aprendizagem	120
Figura 7.6	Segmentação da massa cinza (GM) de imagens cerebrais não realçadas com tumor por aprendizagem	121
Figura A.1	Imagens amostra e ideal submetidas à rotina de aprendizagem	134
Figura A.2	Detalhes na imagem amostra e ideal submetidas à rotina de aprendizagem	135
Figura A.3	Detalhes ampliados das áreas de conflito na Imagem amostra	136
Figura A.4	Detalhes ampliados das áreas de conflito na Imagem ideal	137
Figura A.5	Regiões cinza $R_{c1}$ , $R_{c2}$ e correspondentes regiões binarizadas $R_{b1}$ , $R_{b2}$	138
Figura B.1	Árvore binária	141
Figura B.2	Estrutura dos dados da árvore binária	142
Figura B.3	Modelo de uma estrutura indexada	142
Figura B.4	Arquitetura do cluster utilizado	144
Figura B.5	Imagens usadas na aprendizagem	145

Figura B.6	Resultados obtidos no caso da segmentação de imagens de envelopes postais manuscritos	146-7
Figura B.7	Tamanho da imagem Versus Tempo de Segmentação no caso da segmentação de imagens de envelopes postais manuscritos	148
Figura B.8	Tamanho da imagem Versus Tempo de Segmentação (escala logarítmica) no caso da segmentação de imagens de envelopes postais manuscritos	149

## Lista de Tabelas

Tabela 5.1	Resultados da segmentação por binarização em assinaturas para cada uma das características isoladas em função ao valor k-NN	55
Tabela 5.2	Resultados da amostra de segmentação por binarização de assinaturas para todas as combinações de características e k-NN =1	57
Tabela 5.3	Identificação utilizada para as características	57
Tabela 5.4	Resultados da amostra de segmentação por binarização de assinaturas para todas as combinações de características e k-NN =11	58
Tabela 5.5	Resultados da segmentação por binarização de assinaturas, para as características Média- <i>Skewness</i> em função ao valor k-NN	59
Tabela 5.6	Resultados da segmentação por binarização em impressões digitais para cada uma das características isoladas em função ao valor k-NN	64
Tabela 5.7	Resultados da amostra de segmentação por binarização de impressões digitais para todas as combinações de características e k-NN =1	65
Tabela 5.8	Resultados da amostra de segmentação por binarização de impressões digitais para todas as combinações de características e k-NN =15	66
Tabela 5.9	Resultados da segmentação por binarização de impressões digitais, para a característica Média em função ao valor k-NN	67
Tabela 5.10	Resultados da segmentação por binarização por multi-região em MRI-WM, para cada uma das características isoladas em função ao valor k-NN	71-72
Tabela 5.11	Resultados da amostra de segmentação por multi-região de MRI-WM, para todas as combinações de características e k-NN =1	73
Tabela 5.12	Resultados da amostra de segmentação por multi-região de MRI-WM, para todas as combinações de características e k-NN =3	74
Tabela 5.13	Resultados da segmentação por multi-região de MRI-WM, para a	75

	características Média-Variância em função ao valor k-NN	
Tabela 5.14	Resultados da segmentação por binarização por multi-região em MRI-GM, para cada uma das características isoladas em função ao valor k-NN	77
Tabela 5.15	Resultados da amostra de segmentação por multi-região de MRI-GM, para todas as combinações de características e k-NN =1	78
Tabela 5.16	Resultados da amostra de segmentação por multi-região de MRI-GM, para todas as combinações de características e k-NN =11	79
Tabela 5.17	Resultados da segmentação por multi-região de MRI-GM, para as características Média-Variância-Skewness em função ao valor k-NN	80
Tabela 5.18	Resultados da segmentação por binarização por multi-região em MRI-CSF, para cada uma das características isoladas em função ao valor k-NN	83
Tabela 5.19	Resultados da amostra de segmentação por multi-região de MRI-CSF, para todas as combinações de características e k-NN =1	84
Tabela 5.20	Resultados da amostra de segmentação por multi-região de MRI-CSF, para todas as combinações de características e k-NN =5	85
Tabela 5.21	Resultados da segmentação por multi-região de MRI-CSF, para as características Média-Variância-Skewness-Curtosis em função ao valor k-NN	86
Tabela 5.22	Resultados da segmentação por binarização em Envelopes para cada uma das características isoladas em função ao valor k-NN	93
Tabela 5.23	Resultados da amostra de segmentação por binarização de Envelopes para todas as combinações de características e k-NN =1	95
Tabela 5.24	Resultados da segmentação por binarização de Envelopes, para as características Média-Variância em função ao valor k-NN	97
Tabela 5.25	Resultados da segmentação por binarização de Envelopes, para as características Média-Variância para a base completa ECT-DB1 (n=200)	98
Tabela 5.26	Resultados da segmentação por classes de Envelopes, para as características Média-Variância para a base completa ECT-DB1 (n=200)	98

Tabela 5.27	Resultados da segmentação por binarização de Envelopes, para as características Média-Variância para a base completa ECT-DB2 (n=200)	98
Tabela 5.28	Resultados comparativos da segmentação por classes de Envelopes, para a base completa ECT-DB2 (n=200) com a abordagem por <i>wavelet</i>	99
Tabela 5.29	Resultados comparativos da segmentação por classes de Envelopes, para a base completa ECT-DB2 (n=200) com a abordagem por dimensão fractal, para diferentes parâmetros	100
Tabela 5.30	Comparação dos resultados obtidos para as diversas bases de imagens	107
Tabela 5.31	Tempos de processamento típicos na aprendizagem por estrutura de árvore binária	108

## Lista de Abreviaturas

$A_x$	Imagem de entrada ou amostra
$A_y$	Imagem de saída ou ideal
$Q_x$	Nova imagem amostra
$Q_p$	Imagem processada
PAC	<i>Probably Approximately Correct</i> , ou Provavelmente Aproximadamente Correto
MRI	<i>Magnetic Resonance Image</i> , ou Imagem de Ressonância Magnética
kNN	<i>k-Nearest Neighbors</i> , ou k-Vizinhos Mais Próximos
ABL	<i>Address Block Location</i> , ou Bloco de Endereçamento do Destinatário
RM	Ressonância Magnética
GM	<i>Gray Matter</i> , ou Massa Cinza
WM	<i>White Matter</i> , ou Massa Branca
CSF	<i>Cerebrospinal Fluid</i> , ou Líquor.
$R_{c1}$	Região cinza 1
$R_{c2}$	Região cinza 2
$R_{b1}$	Região binarizada 1
$R_{b2}$	Região binarizada 2
W	Janela de aprendizagem
T	Threshold, ou valor de limiar
$N_c$	Nível de cinza do <i>pixel</i>

$\mu$	Média da região
$\sigma^2$	Variância da região
$\sigma$	Desvio padrão
<i>Ske</i>	<i>Skewness</i> da região
$\gamma$	Coefficiente de assimetria
<i>Cur</i>	Curtosis
<i>C</i>	Coefficiente de apontamento
<i>grad</i>	Gradiente morfológico da região
<i>dil</i>	Dilatação morfológica
<i>ero</i>	Erosão morfológica
$\varepsilon$	Erro
<i>k</i>	Número de vizinhos mais próximos
$H_n$	Conjunto de treinamento na forma de amostras identificadas
$D(\underline{x}, \underline{y})$	Distância euclidiana entre dois vetores
$e_{1NN}$	Taxa de erro obtida através de uma classificação por 1-NN
$e_{Bayes}$	Taxa de erro obtida através de uma abordagem Bayesiana
<i>m</i>	Tamanho da população
<i>C</i>	Número de classes em que se deseja segmentar
<i>I</i>	Imagem
$I_S$	Imagem segmentada
$I_{GT}$	Imagem ideal ( <i>ground truth</i> )
$I_{SO}$	Classe objeto da imagem segmentada
$I_{SF}$	Classe fundo da imagem segmentada
$I_{GO}$	Classe objeto da imagem ideal

$I_{GF}$	Classe fundo da imagem ideal
$TxSegmentação$	Taxa de segmentação
$TxRuído$	Taxa de ruído
$I_{GTO}$	Classe objeto da imagem ideal
$I_{GTF}$	Classe fundo da imagem ideal
$I_{GT1}$	Classe Bloco-Endereço da imagem ideal
$I_{GT2}$	Classe Selo da imagem ideal
$I_{GT3}$	Classe Carimbo da imagem ideal
$TSClasse 1$	Taxa de segmentação da classe 1
$TSClasse 2$	Taxa de segmentação da classe 2
$TSClasse 3$	Taxa de segmentação da classe 3

## Resumo

Este trabalho propõe uma nova abordagem de segmentação de imagens por aprendizagem. O processo de aprendizagem utiliza duas imagens, a imagem original e sua versão ideal segmentada. A partir destas imagens, é gerada uma matriz de decisão que contempla cada pixel e cada nível de cinza das mesmas. Para segmentar outra imagem similar, é reutilizado o conhecimento adquirido que é armazenado na matriz de decisão. A nova imagem é segmentada através de uma estratégia baseada nos vizinhos mais próximos. O processo busca, para cada pixel desta nova imagem, a melhor solução na matriz de decisão. Os resultados dos experimentos realizados em imagens de assinaturas (taxas de segmentação e ruído em  $\% \mu \pm \sigma$ , respectivamente:  $94.19 \pm 6.64$  e  $1.00 \pm 0.92$ ), de envelopes postais manuscritos ( $80.31 \pm 14.30$  e  $1.86 \pm 1.68$ ), impressões digitais ( $94.15 \pm 4.10$  e  $17.55 \pm 7.28$ ) e imagens não realçadas de ressonância magnética (WM:  $96.00 \pm 1.03$  e  $0.53 \pm 0.10$ , GM:  $97.33 \pm 1.49$  e  $0.67 \pm 0.09$ , CSF:  $91.85 \pm 2.74$  e  $0.42 \pm 0.24$ ) mostraram que a abordagem de segmentação por binarização e por multi-região é promissora. As principais vantagens desta abordagem são: a reprodução fiel dos objetivos do usuário; o uso de somente duas imagens para a aprendizagem; o desnecessário uso de parâmetros heurísticos e interação com um usuário experiente após o processo de aprendizagem.

**Palavras-Chave:** 1. Segmentação. 2. Aprendizagem Computacional. 3. Binarização. 4. kNN.

## Abstract

This thesis proposes a new image segmentation by learning approach. The learning process uses two images, the original one and its ideal segmented version. From both images, a decision matrix for each pixel and each gray level is generated. To segment another similar image, the knowledge acquired in the decision matrix is re-used. The new image is segmented by means of a strategy based on the nearest neighbors, that seeks, for each pixel of this new image, the best solution in the decision matrix. Results of performed tests on signature images (segmentation and noise rates in %  $\mu \pm \sigma$ , respectively:  $94.19 \pm 6.64$  and  $1.00 \pm 0.92$ ), handwritten postal envelopes ( $80.31 \pm 14.30$  and  $1.86 \pm 1.68$ ), fingerprint images ( $94.15 \pm 4.10$  and  $17.55 \pm 7.28$ ) and magnetic resonance non-enhancing images (WM:  $96.00 \pm 1.03$  and  $0.53 \pm 0.10$ , GM:  $97.33 \pm 1.49$  and  $0.67 \pm 0.09$ , CSF:  $91.85 \pm 2.74$  and  $0.42 \pm 0.24$ ) showed that this approach is promising performing image bi-level and multi-level segmentation. The main advantages of this method are: it faithfully reproduces the objectives of the user, it needs only two images, it does not require the use of heuristic parameters neither the interaction from an experienced user after the learning process.

**Keywords:** 1. Segmentation. 2. Computational Learning. 3. Thresholding. 4. kNN .

# Capítulo 1

## Introdução

O Processamento e a Análise de Imagens caracterizam-se, de uma maneira geral, pelas fases de aquisição, pré-processamento, segmentação, representação/descrição, e reconhecimento de imagens. A segmentação consiste em extrair de uma imagem “primitivas”, ou características que são informações semânticas da imagem.

A segmentação, etapa primeira da Análise de Imagens, é de fundamental importância para as etapas posteriores do Processamento Digital de Imagens. Esta etapa é alvo de muitos estudos e pesquisas para desenvolvimento de métodos mais abrangentes e computacionalmente mais eficientes.

### 1.1. Desafios

A segmentação de imagens, pela sua característica, é uma tarefa de processamento de imagens desafiadora. O ser humano é capaz de, visualmente, saber o que extrair da imagem, porém é frequentemente observada a incapacidade do computador em executar a mesma tarefa.

A grande dificuldade na segmentação reside no fato de não se conhecer, de antemão, o número e tipo de estruturas que se encontram numa imagem. Essas estruturas são identificadas a partir da geometria, forma, topologia, textura, cor ou brilho. Portanto, são escolhidas aquelas que possibilitam a melhor aplicação.

Geralmente a segmentação de imagens é uma área árdua, difícil e de grande importância; deste modo propenso a expansão de novas pesquisas e estudos. Em particular, a segmentação automática é uma das tarefas mais difíceis em processamento de imagens. Esta etapa determina o eventual sucesso ou fracasso da análise das imagens. De fato, a segmentação efetiva quase sempre garante sucesso no reconhecimento de padrões.

## 1.2. Proposta

O presente estudo se propõe a atingir os seguintes objetivos:

- Definir uma abordagem que minimize a intervenção do operador humano. O objetivo é selecionar, automaticamente, a técnica mais adequada para cada caso;
- Definir uma abordagem que atinja as expectativas do usuário (em termos de segmentação);
- Definir uma abordagem que minimize artefatos ou confusões na aprendizagem;
- Definir uma abordagem eficiente.

## 1.3. Contribuições

De uma forma geral, a contribuição desse trabalho está no desenvolvimento de uma abordagem genérica de aprendizagem que permite transformar matematicamente o conhecimento de um especialista e permite solucionar processamentos complexos de segmentação em diferentes tipos de imagens.

De forma particular, as principais contribuições desse trabalho são sumarizadas a seguir, sendo o detalhamento de cada uma delas feito ao longo do trabalho:

- Escolha das características (*features*) relevantes para a análise das imagens. As características inicialmente propostas são: média, variância, média dos gradientes, gradiente mínimo, gradiente máximo, *skewness* e *curtosis*;
- Definição da Matriz de Decisão;

- Definição de similaridade pelo conceito dos vizinhos mais próximos dos vetores de características dentro da matriz de decisão;
- Tratamento de confusões no processo de aprendizagem;
- Algoritmo de escolha automática do tamanho da janela de aprendizagem.

## 1.4. Organização

A tese é organizada da seguinte maneira:

O Capítulo 2 (**Segmentação**) apresenta um resumo das técnicas de segmentação de imagens. Inclui-se também uma breve comparação de resultados de binarização obtidos por algoritmos tradicionais locais e globais.

No Capítulo 3 (**A abordagem Proposta de Segmentação por Aprendizagem**) são sumarizadas as diversas abordagens da aprendizagem computacional na segmentação de imagens. Também, são mostrados os diferentes detalhes da técnica desenvolvida ao longo desta pesquisa.

Uma descrição sobre cada uma das bases utilizadas na experimentação é apresentada no Capítulo 4 (**Considerações sobre as Bases de Imagens Utilizadas**).

No Capítulo 5 (**Experimentos de Aplicações**) discutem-se testes de segmentação realizados em bases de imagens de assinaturas, envelopes postais, impressões digitais e imagens de ressonância magnética.

Nos Capítulos 6, 7, e 8 são apresentados, respectivamente as **Conclusões**, os **Trabalhos Futuros** e as **Referências Bibliográficas**.

Um **Exemplo Prático de Resolução de Conflitos e Cálculo do Tamanho Adequado da Janela**, que ilustra passo a passo o algoritmo proposto foi colocado como Apêndice A. O Apêndice B apresenta o **Estudo e Implementação de Estruturas de Otimização do Processo de Aprendizagem**. Finalmente, as **Publicações Decorrentes da Pesquisa** são listadas no Apêndice C.

A notação matemática utilizada em toda a tese segue, sempre que possível, o padrão do livro *Pattern Recognition*, de Robert Schalkoff [SCH92].

## Capítulo 2

### Segmentação

Neste Capítulo apresenta-se um resumo da classificação das técnicas de segmentação de imagens. Inclui-se também uma breve comparação de resultados de binarização, obtidos por algoritmos tradicionais, que utilizam abordagens locais e globais. Ao final, é apresentada uma classificação da abordagem proposta do ponto de vista de segmentação.

Alguns psicólogos alemães, no início do século passado (Köhler, Wertheimer, Koffka) [KOF35] [KOH29] [WER44], mostraram que o ser humano, ao nível da visão, efetua agrupamentos sobre o que ele percebe, baseados na proximidade, na similaridade, e na continuidade [ROB97]. O objetivo da segmentação é obter, a partir de uma imagem, um conjunto de características ou segmentos significativos que contenha a informação semântica relativa à imagem. A grande dificuldade da segmentação reside no fato de não se conhecer, a priori, o número e tipo de estruturas que se encontram na imagem. Essas estruturas são identificadas a partir da geometria, forma, topologia, textura, cor ou brilho. Nesse caso, são escolhidas aquelas que possibilitam a melhor aplicação.

A etapa de segmentação de imagens está presente em qualquer projeto de processamento de imagens, na qual é definida a existência, a localização e os tipos de estruturas procuradas. Pela grande variedade de “primitivas” ou “segmentos significativos”, que contêm as informações semânticas, e pelas inúmeras aplicações, a etapa de segmentação talvez seja um dos maiores desafios da área de processamento e análise de imagens.

As abordagens existentes na literatura para solucionar este problema são múltiplas e basicamente levam à noção de região, de contorno, de textura ou ainda a uma abordagem mista. Essa última é preferível, mas é mais complicada porque a passagem região-textura-contorno ou contorno-região-textura é mais fácil na teoria que na prática. Muitas vezes, uma abordagem unilateral é escolhida em razão dos tempos de execução mais curtos.

Nesse trabalho será adotada a seguinte classificação de tipos de segmentação de imagens:

- Segmentação por Região;
- Segmentação por Contorno;
- Segmentação por Textura;
- Segmentação por Modelos de Contornos Ativos;
- Segmentação Morfológica.

## **2.1 Segmentação por Região**

Uma região pode ser definida como sendo um conjunto de pontos que respeitam um mesmo predicado de homogeneidade. Assim, deve sempre existir pelo menos um caminho inteiramente contido nessa região, ligando dois pontos. Para uma introdução neste assunto, recomenda-se [HAR85].

As abordagens de segmentação por região podem ser divididas basicamente nos três seguintes grupos:

- Binarização (*thresholding*);
- Divisão e fusão (*split and merge*);
- Multi-binarização

### **2.1.1. Binarização**

Uma das técnicas de segmentação por região mais tradicionais e populares é a segmentação por binarização. O princípio da binarização consiste em separar as regiões de

uma imagem, quando esta apresenta duas classes (o fundo e o objeto). Existem duas abordagens clássicas de binarização, a binarização global e a binarização local adaptativa. Uma revisão das diversas abordagens de binarização pode ser vista em [SAN01], ou ainda em [FAC96].

#### 2.1.1.1. Binarização Global

Na binarização global busca-se um único valor de limiar para toda a imagem. Assim, seja  $T$  o valor de limiar, para cada pixel  $P(x,y)$ :

$$P(x,y) \begin{cases} \in \text{primeiro plano} & \text{se o valor de } P(x,y) > T \\ \in \text{fundo} & \text{caso contrário} \end{cases} \quad (2.1)$$

A técnica é de princípio simples, porém, apresenta a desvantagem de nem sempre as imagens conterem intensidades de primeiro e segundo planos bem diferenciados.

Um levantamento feito recentemente mostra que existem pelo menos 200 publicações de binarização global [FAC03]. Basicamente a grande maioria das técnicas busca particionar a imagem em duas classes, maximizando a variância inter-classes e minimizando a variância intra-classes. Isto se faz, empregando funções critério específicas para cada abordagem.

#### 2.1.1.2. Binarização Local Adaptativa

Devido à dificuldade em selecionar um limiar global, definir valores diferentes de limiar para regiões diferentes da imagem provou ser uma abordagem interessante. Este tipo de binarização é chamado de binarização adaptativa ou ainda local. O problema principal deste tipo de abordagem é a escolha do tamanho da janela para a definição do limiar local. A literatura fornece várias abordagens, sendo talvez as mais interessantes na atualidade, os métodos de Bernsen [BER86], Eikvil [ETM91] e Wellner [WEL93].

### 2.1.2. Método de Divisão e Fusão

Segundo [CMV86], o método de Divisão e Fusão permite agrupar *pixels* para criar regiões sobre as quais um predicado de homogeneidade  $H$  é verificado. A imagem será manipulada através de divisões e fusões, até se obter regiões que satisfaçam o critério  $H$  (por ex., mesmo nível de cinza, mesma média, etc) [WNS97].

Este método é muitas vezes associado a uma estrutura QUADTREE [CAH84] que permite, de uma forma elegante, decompor e agrupar partes de uma imagem.

### 2.1.3. Multi-binarização

Dependendo do conteúdo da imagem, o simples particionamento em somente duas classes, o primeiro plano e o fundo, acarreta uma grande perda de informação. Neste caso, o particionamento em várias classes, recupera informações valiosas e relevantes. O maior problema da multi-binarização reside em saber em quantas classes uma imagem pode ser particionada. De fato, a avaliação do número de classes é um problema específico da multi-binarização. Infelizmente a maioria das abordagens precisa prever, a priori, o número de classes [MAR00], embora existam algumas abordagens que permitam realmente uma segmentação automática, entre outras pode –se citar a de [YCC95].

## 2.2. Segmentação por Contorno

Um contorno ou borda é uma mudança brusca do nível de cinza entre duas regiões relativamente homogêneas. A segmentação por contorno consiste, portanto, em localizar pontos nessas mudanças bruscas de níveis de cinza. Há uma grande variedade de métodos disponíveis na literatura, para detectar contornos. As técnicas de detecção de bordas podem ser classificadas em duas categorias: a seqüencial e a paralela.

Na abordagem seqüencial, o objetivo de decidir se um *pixel* pertence à borda ou não, depende do resultado do detector nos *pixels* examinados anteriormente. No caso da abordagem paralela, a decisão é tomada com base no ponto em consideração e alguns de seus

pontos vizinhos. O desempenho do método de detecção de bordas seqüencial depende da escolha de um ponto inicial apropriado, pois os resultados nos pontos anteriores influenciam na seleção e no resultado do próximo ponto.

A segmentação por contorno pode ser realizada com base na:

- Detecção de pontos isolados muito significativos, basicamente cantos (ex.: operador de Moravec [MOR87]);
- Detecção de bordas (ex.: gradiente [AND00] e laplaciano [KKR99])
- Detecção de curvas fechadas como círculo e elipse (ex.: transformada de Hough [BAL81]).

### **2.3. Segmentação por Textura**

A textura é uma propriedade das superfícies que o sistema de percepção visual humano utiliza para segmentar e classificar objetos. Se for possível desenvolver algoritmos de processamento de imagens próprios para textura, então as propriedades da textura poderão fornecer valiosas informações para resolver as difíceis tarefas de segmentação e de classificação.

A segmentação por textura deve permitir a colocação em evidência dos motivos da imagem, considerando-se as propriedades de regularidade e de repetição de um padrão [TAM78]. A noção de textura supõe que:

- O padrão seja um agrupamento aleatório ou não aleatório de subconjuntos da imagem;
- As entidades detectadas sejam uniformes e tenham aproximadamente as mesmas dimensões, em qualquer lugar da imagem;
- O padrão local seja repetido numa região comparativamente grande, em relação ao tamanho desse padrão.

Os principais métodos de segmentação por textura são [FAC01]:

- Métodos estatísticos que permitem caracterizar uma textura de tipo suave, irregular, granulada, etc (por ex. a matriz de co-ocorrência [COH80], [YON02]);
- Métodos estruturais que tratam da disposição das primitivas numa imagem, por exemplo, textura baseada em linhas paralelas dispostas regularmente;
- Métodos espectrais onde se busca detectar periodicidades por identificação de picos no espectro de Fourier [MAR87], [HEW88];
- Métodos baseados em filtragem (por ex.: *wavelets* [MBF03]);
- Métodos baseados em modelos (por ex. fractais [MOS92] ou Markov [CFA91]);
- Métodos mistos.

## 2.4. Segmentação por Modelos de Contornos Ativos

Um contorno ativo é um modelo de uma curva deformável. O modelo de contorno ativo é uma técnica de ajuste de uma curva deformável, em relação às bordas de conjuntos de uma imagem baseada na minimização local de uma função energia. Por ser ativo, está sempre procurando minimizar a sua função energia. E, este modelo difere da maioria das técnicas de detecção de contornos. Esse método é, portanto, dinâmico. Durante o processo de minimização, o contorno ativo tem um comportamento parecido ao de uma cobra que rasteja, daí originou-se a denominação de “*snakes*”. Para uma síntese da evolução do modelo de contornos ativos, sugerem-se como referência as abordagens de Kass, Witkin e Terzopoulos [KWT88] e Cohen [COH91].

## 2.5. Segmentação Morfológica

O princípio básico da morfologia matemática [SER82], consiste em extrair uma informação relativa à geometria e à topologia de um conjunto desconhecido de uma imagem. Isto é realizado pela transformação a partir de um outro conjunto completamente definido chamado elemento estruturante. O tipo e a natureza da informação extraída depende, necessariamente do tipo do elemento estruturante e do tipo de imagem estudada. A eficiência

e a dificuldade da morfologia matemática reside na escolha da deformação certa, para transformar a expectativa intelectual em aplicação prática.

Em geral, as abordagens de segmentação morfológica podem ser agrupadas nas seguintes classes:

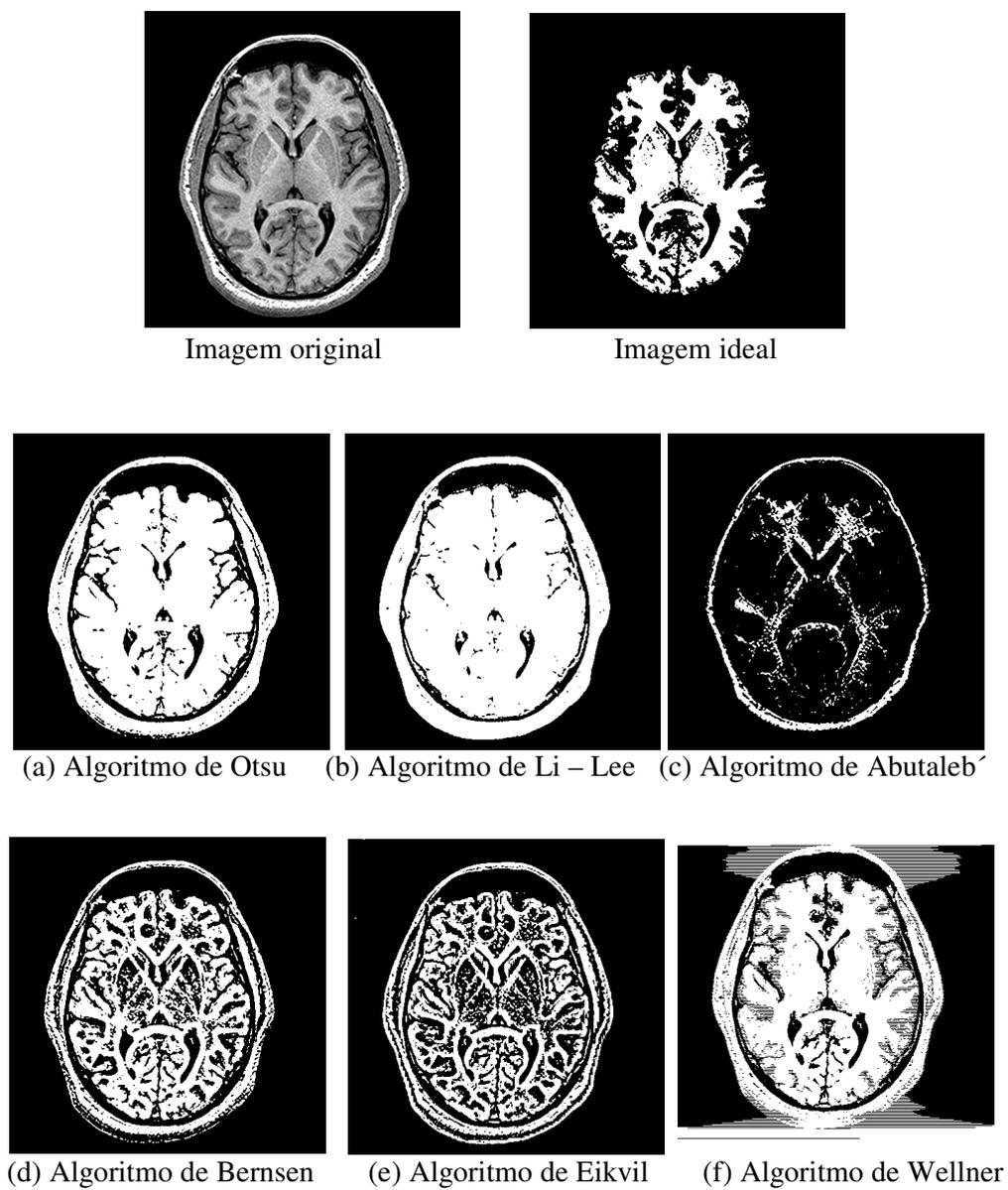
- Segmentação morfológica por contorno [GOW92];
- Segmentação morfológica por transformação top-hat [DOU92];
- Segmentação morfológica por abordagem *watershed* [VIS91][MEY91], *waterfall* [BEU90] e *dynamics* [GRI91] que são o estado da arte em segmentação morfológica [AND98].

## 2.6. Classificação da Abordagem Proposta do Ponto de Vista de Segmentação

Na área de segmentação de imagens, a abordagem proposta pode ser classificada como uma técnica de segmentação por região, especificamente como uma segmentação por binarização ou ainda, como uma segmentação multi-modal.

Deve ser observado que, em um contexto geral, não existe uma técnica de segmentação universal válida para qualquer tipo de imagens. A obtenção de imagens binarizadas, em uma mesma base de dados, nem sempre pode ser realizada com os mesmos algoritmos, nem com os mesmos parâmetros de um mesmo algoritmo. Por este motivo, é comum observar, na etapa de segmentação, o uso de heurística e a intervenção do operador humano com o objetivo de selecionar, manualmente, a técnica mais adequada para cada caso. As **Figuras 2.1** e **2.2** ilustram melhor o exposto anteriormente.

A **Figura 2.1** apresenta uma comparação de resultados obtidos da binarização de uma imagem de ressonância magnética, utilizando-se técnicas tradicionais globais (Otsu [OTS79], Le and Lee [LIL93], Abutaleb [ABU89]) e adaptativas (Bernsen [BER86], Eikvil [EIK91], Wellner [WEL93]). Pode-se observar; que além destas técnicas apresentarem resultados diferentes, não conseguiram produzir resultados desejáveis introduzindo distorções e artefatos.



**Figura 2.1.** Comparação de resultados de binarização em imagem de ressonância magnética

A **Figura 2.2** apresenta outra comparação ilustrativa, em uma imagem de impressão digital. O objetivo era binarizar as linhas da impressão, sem cortá-las, mesclá-las ou criando artefatos. Foram utilizadas as mesmas técnicas acima, e como podem ser observados, os resultados são muito dispares entre si.

## **2.7. Comentários Finais**

Neste Capítulo apresentou-se um resumo da classificação das técnicas de segmentação de imagens. Incluiu-se também uma breve comparação de resultados de binarização obtidos por algoritmos tradicionais locais e globais. Finalmente, apresentou-se uma classificação da abordagem proposta, sob o ponto de vista de segmentação. No Capítulo seguinte será feita uma breve revisão do conceito de Aprendizagem Computacional, assim como serão detalhados os fundamentos da abordagem proposta de segmentação por aprendizagem



Imagem original



Imagem ideal



(a) Algoritmo de Otsu



(b) Algoritmo de Li – Lee



(c) Algoritmo de Abutaleb'



(d) Algoritmo de Bernsen



(e) Algoritmo de Eikvil



(f) Algoritmo de Wellner

**Figura 2.2** Comparação de resultados de binarização em imagem de impressões digitais

## Capítulo 3

# A abordagem Proposta de Segmentação por Aprendizagem

Neste Capítulo será apresentada a abordagem proposta de Segmentação por Aprendizagem. Inicialmente, será feita uma breve revisão do conceito de Aprendizagem Computacional. Depois, um estudo das abordagens anteriores e a seguir, serão apresentados os fundamentos da abordagem proposta. Serão detalhadas as etapas de aprendizagem e segmentação, incluindo a seleção das características, as considerações sobre o tamanho da janela e sobre o modelo de aprendizagem baseada em instâncias, o  $k$ -NN.

### 3.1. Aprendizagem Computacional

Neste Capítulo será feita uma breve revisão do conceito de Aprendizagem Computacional. Será apresentado o agrupamento das abordagens de aprendizagem computacional: as baseadas em técnicas de inteligência artificial, as baseadas em modelagem de imagens e as baseadas em técnicas de indução a partir de exemplos. Será contextualizada a abordagem proposta, desde o ponto de vista de aprendizagem computacional, como sendo uma técnica baseada em indução.

A Teoria da Aprendizagem Computacional é uma área de pesquisa que estuda o desenvolvimento e a análise de algoritmos. Ela tem por objetivo, realizar previsões sobre o futuro baseadas nas experiências passadas.

Os estudos sobre a aprendizagem computacional não são novos. De fato, Valiant [VAL84] apresenta uma definição para aprendizagem computacional: “Diz-se que um programa de computador, para efetuar uma tarefa, foi adquirido por aprendizagem, se foi adquirido por qualquer meio exceto por programação explícita”. Este conceito é mais bem esclarecido pelo autor, pela exemplificação, através da seguinte comparação com o ser humano. “Dentre as habilidades humanas, algumas parecem possuir claramente um elemento pré-programado geneticamente, enquanto que algumas outras consistem de seqüências explícitas de instruções que foram memorizadas. Resta uma grande área de aquisição de habilidades onde não podem ser identificadas tais aquisições de programação explícita”. Valiant descreve esta última área como aprendizagem.

As diversas abordagens de aprendizagem computacional podem ser divididas em três grupos [HIR00]:

- As baseadas em técnicas de inteligência artificial;
- As baseadas em modelagem de imagens e;
- As baseadas em técnicas de indução a partir de exemplos.

A aplicação de técnicas de inteligência artificial inclui sistemas especialistas e prova automática de teoremas, com o objetivo de automatizar a geração de procedimentos eficientes para análise de imagens [SCH89], [VOG89], [JOO90], [JOO91].

Abordagens baseadas em modelagem de imagens estudam a caracterização formal das imagens e sua manipulação analítica. Essas abordagens são interessantes quando as imagens e os tipos de ruído ou de textura encontrados podem ser modelados e simulados [SIV97], [DOL92], [DOU99], [SCH94].

Nas abordagens baseadas em indução, uma transformação de imagens é expressa através de pares de imagens, onde cada par é formado por uma imagem anterior e outra posterior ao processamento desejado [BRE98], [BHA00]. A imagem a ser processada é denominada imagem amostra, enquanto a imagem desejada, como resultado do processamento é denominada imagem ideal. A aprendizagem é realizada a partir de exemplos

obtidos dos pares de imagens amostras-ideais. Existem várias técnicas de aprendizagem empregando o conceito de pares de imagens amostras-ideais:

- Técnicas no contexto de estimação estatística, consideram as imagens como realizações de processos aleatórios e portanto, a aprendizagem é o resultado de estimadores estatísticos [BAR97], [DOU97], [DOU98];
- Técnicas adaptativas tais como algoritmos genéticos e redes neurais. No caso das redes neurais, para o treinamento é necessário um grande conjunto de pares de imagens de amostra entrada/saída ideal [WIL93], [SAR98], [SAR99], [SOU00], [SCA00]. As técnicas que utilizam de algoritmos genéticos, tentam obter o melhor conjunto de características nestes treinamentos [HAR96], [OLI98], [YOD99]. Por exemplo, no trabalho desenvolvido por Bhanu, [BHA00] para garantir uma boa segmentação de imagens, o processo de avaliação e recombinação deve ser repetido para um grande número de gerações.

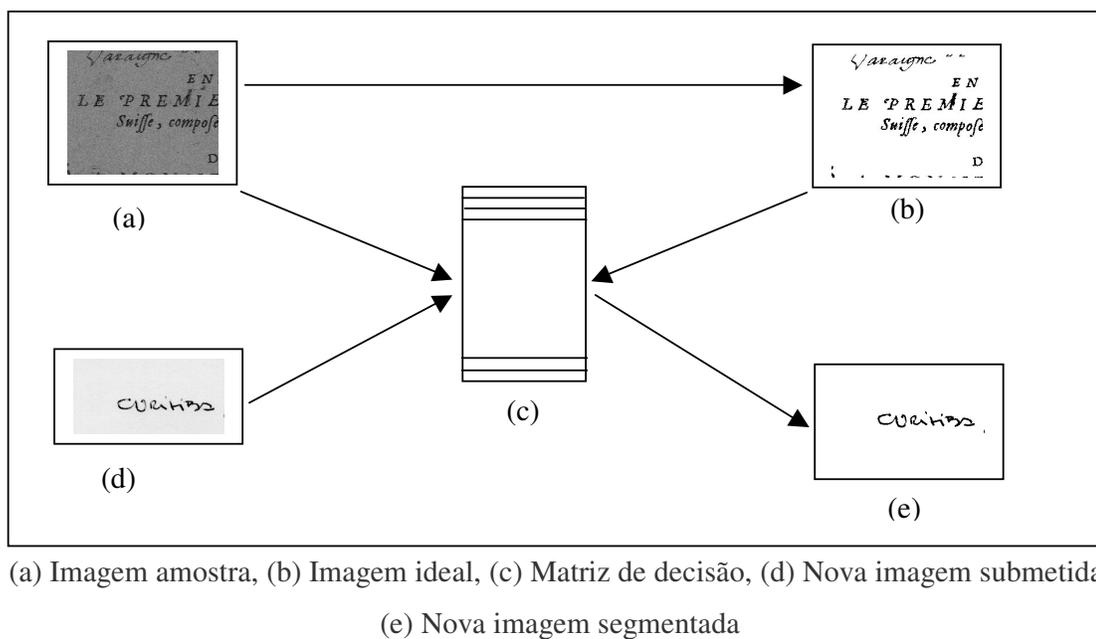
Do ponto de vista conceitual, a nova abordagem proposta neste trabalho é multidisciplinar. A mesma, mistura, de forma dinâmica, as áreas de segmentação por região, estimação estatística, e aprendizagem computacional. No campo de aprendizagem computacional, a abordagem proposta pertence ao grupo de técnicas baseadas em indução.

### **3.2. A Abordagem Proposta**

A idéia básica da abordagem proposta é utilizar o conhecimento e/ou a meta do usuário, em transformar uma imagem em níveis de cinza numa imagem binária. Se o usuário sabe o que extrair de uma imagem, porque não compartilhar este conhecimento com o computador?

A abordagem apresentada consiste em ensinar o computador a segmentar, empregando para isto imagens cujas soluções são conhecidas. Diferentemente das técnicas que utilizam redes neurais ou algoritmos genéticos na aprendizagem, a presente abordagem aprende a partir de um único par de imagens amostra/ideal.

O processo consiste em duas etapas (**Figura 3.1**), a aprendizagem e a segmentação. Na primeira etapa, submete-se uma imagem para a extração das características relevantes, na vizinhança de cada *pixel*, junto com a sua versão segmentada ideal. A segunda etapa, consiste em segmentar novas imagens, empregando uma denominada matriz de decisão construída (ou conjunto de treinamento).



**Figura 3.1.** Esquema geral da abordagem proposta de Segmentação por Aprendizagem

### 3.3. Estudo preliminar de Abordagens Anteriores

Do estudo bibliográfico realizado, percebeu-se que existem poucas abordagens disponíveis na literatura, associadas à segmentação de imagens em níveis de cinza, que empregam o conceito de aprendizagem através de um único par de imagens amostras-ideais. Os trabalhos de Barrera *et al* [BAT95], [BDT97] e Hirata [HIR00] utilizaram o modelo de estimação estatística para formalizar a aprendizagem de operadores morfológicos binários estatisticamente ótimos. Estas técnicas baseiam-se no modelo de aprendizado PAC (do inglês, “*Probably Approximately Correct*”), trabalhando com imagens e operadores binários.

Por outro lado, Kim [KIM97], [KIM99] trabalhando já com imagens em níveis de cinza, projetou filtros digitais automáticos utilizando imagens amostras-ideais.

A técnica abordada por Kim consiste basicamente no seguinte: Para cada ponto da imagem desconhecida ( $Q_x$ ) analisa-se uma área de vizinhança (por exemplo, uma janela 3 x 3 ou 5 x 5); procura-se na imagem de entrada ( $A_x$ ), a janela “mais parecida” com a janela em estudo. Obtém-se essa janela “mais parecida”, calculando-se a soma dos módulos das diferenças dos pixels entre as duas janelas em análise. A imagem ( $A_x$ ) é varrida pela janela, com o correspondente cálculo do módulo das diferenças para cada caso. Finalizado este processo, seleciona-se a janela que tiver menor valor na soma dos módulos. Assim, a janela em  $A_x$  selecionada é a janela “mais parecida” a  $Q_x$ . Para a obtenção da imagem  $Q_p$ , procura-se então o valor do pixel em  $A_y$ , correspondente ao pixel em  $A_x$  com a janela “mais parecida”, e atribui-se a  $Q_p$ . Repetindo este processo, para toda a imagem em análise, obtém-se a nova imagem  $Q_p$ .

Na abordagem de Kim, o tamanho da janela deve ser escolhido manualmente. O resultado da aplicação de segmentação de textura, apresentada para imagens de isotermas em mapas [KIY97], não mostra uma comparação estatística, sendo a avaliação realizada de forma visual.

Brun e colaboradores [BRU01], também utilizaram a técnica de aprendizagem a partir de pares de imagens amostras / ideais, para o projeto de filtros não lineares aplicados em imagens em níveis de cinza, não tratando o caso de segmentação.

Hirata *et al* [HID00], [HDB00] e Flores *et al* [FHB0], desenvolveram pesquisas sobre a utilização de operadores morfológicos em níveis de cinza projetados por aprendizagem computacional. Nestes casos, a segmentação realizada, no contexto da morfologia matemática, depende da obtenção dos marcadores para os objetos especificados.

Recentemente, Leandro, Soares e César-Jr [LSC03] utilizaram classificadores supervisionados para segmentar, automaticamente, vasos sanguíneos de imagens de retina.

Nesta abordagem, cada *pixel* é representado por um vetor de características. Este vetor é composto por informação da cor e medições em diferentes escalas, tomadas a partir da transformada *wavelet* contínua (de Morlet), assim como de filtros da média e de ordem aplicados ao canal verde. A imagem resultante é ainda pós-processada utilizando ferramentas morfológicas. Nesse artigo não são apresentadas avaliações estatísticas dos resultados obtidos com imagens *ground-thruth*.

Do estudo bibliográfico das pesquisas existentes, em segmentação por aprendizagem, observa-se que a maioria dos trabalhos envolve operadores binários ou contemplam imagens em níveis de cinza para filtragem. Também, não prevêem mecanismos para solucionar um dos maiores desafios da aprendizagem que é de evitar ambigüidades, ou exigem a intervenção manual para a identificação de um marcador. Esta pesquisa insere-se no contexto das novas abordagens de segmentação por aprendizagem.

### **3.4. Fundamentos da Abordagem Proposta**

Esta abordagem propõe o aproveitamento do conhecimento e a experiência do usuário, para segmentar uma imagem em níveis de cinza, em uma ou várias classes. Desta maneira, o usuário que conhece exatamente o que deseja extrair da imagem, interage com o computador através da transmissão deste conhecimento.

O método consta de duas etapas, a aprendizagem e a segmentação. O processo de aprendizagem utiliza somente duas imagens, a original e a ideal. A partir destas duas imagens é montada uma matriz de decisão para cada um dos *pixels*, do par de imagens analisado. No processo de segmentação de outras imagens é utilizado o conhecimento adquirido e armazenado na matriz de decisão. Cada uma das novas imagens é segmentada através da estratégia dos vizinhos mais próximos, que procura, para cada *pixel* a ser processado, a melhor solução na matriz de decisão.

No processo da aprendizagem do par de imagens original/ideal, são implementados mecanismos para solucionar os problemas das confusões existentes.

Nos itens a seguir, serão explicados, de forma detalhada, todos os processos e parâmetros utilizados nesta abordagem. Entre eles, a seleção e estrutura do vetor de características, a escolha automática do tamanho apropriado da janela de processamento e a construção da matriz de decisão.

### 3.5. A Seleção das Características

Nessa etapa, o objetivo é selecionar aquelas características que melhor definam os relacionamentos entre o *pixel* amostra e seu correspondente *pixel* ideal. A disposição dos *pixels* vizinhos, em relação ao ponto em estudo, é de fundamental importância, principalmente considerando-se que:

- a) Em geral, o valor limiar que separa um objeto do fundo é específico, ao redor de uma vizinhança, e varia significativamente dependendo da região em análise. A informação do nível de cinza não é suficiente para uma efetiva segmentação, pois este mesmo nível pode fazer parte de objeto, em uma região e do fundo em outra;
- b) A noção de fundo e objeto é dependente, em vários casos, da subjetividade do ser humano que segmenta visualmente a imagem. Deve ser considerado, além dos níveis de cinza, a proximidade dos pontos, continuidade de áreas, etc. Esta noção subjetiva deve ser levada em consideração no momento da aprendizagem.

Pode-se obter na literatura as mais variadas características para extrair informação de uma região, porém para este trabalho, procurou-se selecionar aqueles que cumprem os seguintes critérios:

- 1) Escolha de características independentes (ou não correlacionadas) de menor custo computacional, devido ao processamento que é realizado em toda a imagem para detecção da janela mais parecida;

- 2) A extração das características deve ser independente dos dados obtidos em cada imagem. Isto é, a informação armazenada na matriz de decisão tem que estar limitada àquelas que podem ser extraídas também da imagem desconhecida.

Tendo em vista o precedente, são propostas as seguintes características:

- **Nível de cinza do pixel:** por ser uma característica natural do *pixel*, este é o ponto inicial a partir do qual será analisada a informação;

$$N_c = x_i \quad \text{sendo } x_i \in [0 \dots 255] \text{ no domínio discreto}$$

- **Média da região ( $\mu$ ):** permite a caracterização do comportamento local, assim como a redução da influência do ruído. Nos casos de homogeneidade absoluta da vizinhança, a média é igual ao valor do pixel;

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i$$

sendo  $N$  o número de *pixels* da região

- **Variância da região ( $\sigma^2$ ):** fornece a noção da distribuição dos níveis de cinza ao redor do *pixel*. Permite a caracterização do aspecto homogêneo, junto com a detecção da influência do ruído;

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \quad (3.1)$$

- **Skewness da região (*Ske*):** é definido como o quociente do momento de terceira ordem e a terceira potência do desvio padrão ( $\sigma$ )

$$Ske = \frac{1}{\sigma^3} \sum_{i=1}^N (x_i - \mu)^3 \quad (3.2)$$

Com o objetivo de diminuir a quantidade dos valores calculados, é utilizado neste trabalho o *Skewness*, dividido pelo total de *pixels* da região. Esta expressão é também conhecida por coeficiente de assimetria ( $\gamma$ )

$$\gamma = \frac{Ske}{N} \quad (3.3)$$

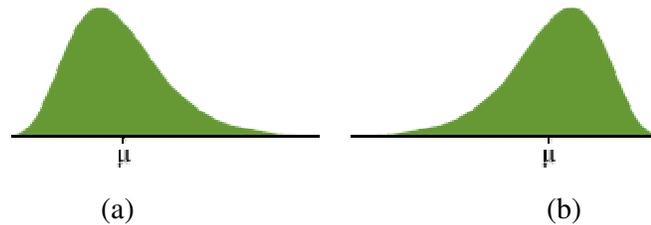
$$\gamma = \frac{1}{N\sigma^3} \sum_{i=1}^N (x_i - \mu)^3 \quad (3.4)$$

O coeficiente de assimetria fornece as seguintes informações (**Figura 3.2**):

$\gamma = 0$ ; a distribuição é simétrica ao redor da média.

$\gamma > 0$ ; indicam distribuições com maior sesgo à direita.

$\gamma < 0$ ; indicam distribuições com maior sesgo à esquerda.



**Figura 3.2.** Exemplificação do coeficiente de assimetria.

Ambas as funções de densidade de probabilidade acima possuem mesma média e variância, porém em (a) o sesgo é à direita,  $\gamma > 0$  e em (b) o sesgo é à esquerda,  $\gamma < 0$

- **Curtosis da região (*Cur*):** é definido como o quociente do momento de quarta ordem em torno da média  $\mu$  e a quarta potência do desvio padrão ( $\sigma$ ).

$$Cur = \frac{1}{\sigma^4} \sum_{i=1}^N (x_i - \mu)^4 \quad (3.5)$$

De forma análoga ao caso anterior, com o objetivo de diminuir o tamanho dos valores calculados, foi utilizado neste trabalho o Curtosis, dividido pelo total de *pixels* da região. Esta expressão é também conhecida por coeficiente de apontamento (c)

$$c = \frac{Cur}{N} \quad (3.6)$$

$$c = \frac{1}{N\sigma^4} \sum_{i=1}^N (x_i - \mu)^4 \quad (3.7)$$

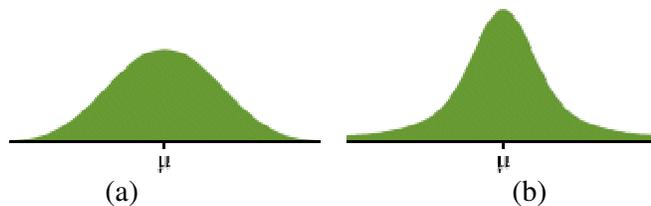
O coeficiente de apontamento fornece as seguintes informações (**Figura 3.3**):

$c = 0$ ; a distribuição é normal (mesocúrtica) comparada com o modelo de distribuição normal.

$c > 0$ ; a distribuição tem maior pico (leptocúrtica) que a distribuição do modelo normal.

$c < 0$ , a distribuição é mais plana (platicúrtica) que a distribuição do modelo normal.

São utilizados nesta tese os conceitos genéricos de *skewness* e curtosis, respectivamente, para as duas características definidas anteriormente.



**Figura 3.3.** Exemplificação do coeficiente de apontamento.

A função de densidade de probabilidade em (a) é mais plana,  $c < 0$  do que em (b); onde tem um pico maior,  $c > 0$

- **Média dos gradientes da região:** fornece uma noção adicional da dispersão dos níveis de cinza ao redor do *pixel* em análise. Para cada *pixel* na vizinhança, é calculado o gradiente morfológico, armazenando-se em cada caso, a média destes valores. Do ponto de vista digital, o gradiente morfológico pode ser calculado por erosão, por dilatação ou por erosão-dilatação, sendo este último utilizado na presente abordagem. O gradiente morfológico  $\text{grad}_{\text{ero-dil}}^g$  de uma imagem  $f(x)$  por um elemento estruturante  $g$  pode ser implementado de forma prática da seguinte maneira:

$$\mathbf{grad}_{ero-dil}^g(f) = \mathbf{dil}^g(f) - \mathbf{ero}^g(f) \quad (3.8)$$

sendo:  $\mathbf{dil}^g(f)$  a dilatação de uma imagem  $f(x)$  por um elemento estruturante  $g$ , e  $\mathbf{ero}^g(f)$  a erosão de uma imagem  $f(x)$  por um elemento estruturante  $g$ .

A média dos gradientes da vizinhança é então definida:

$$E(\mathbf{grad}_{ero-dil}^g) = \frac{1}{N} \sum_{i=1}^N (\mathbf{grad}_{ero-dil}^g)_i \quad (3.9)$$

- **Gradiente mínimo da região ( $grad_{MIN}$ ):** é o valor mínimo dos gradientes morfológicos da região.

Seja o gradiente morfológico por erosão-dilatação definido na Eq 3.8 ( $\mathbf{grad}_{ero-dil}^g$ ), simplificado pela notação  $\mathbf{grad}$ . Para um conjunto de  $n$  gradientes calculados em uma região:

$$\{\mathbf{grad}_1, \dots, \mathbf{grad}_i, \dots, \mathbf{grad}_n\}$$

Define-se gradiente mínimo como:

$$\mathbf{grad}_{MIN} = \min \{\mathbf{grad}_i\} \quad (3.10)$$

- **Gradiente máximo da região ( $grad_{MAX}$ ):** é o valor máximo dos gradientes morfológicos da região.

De forma análoga à definição anterior, para um conjunto de  $n$  gradientes calculados em uma região:

$$\{\mathbf{grad}_1, \dots, \mathbf{grad}_i, \dots, \mathbf{grad}_n\}$$

Define-se gradiente máximo como:

$$\mathbf{grad}_{MAX} = \max \{\mathbf{grad}_i\} \quad (3.11)$$

### 3.6. A etapa de Aprendizagem

A imagem digital é definida, no contexto desta tese, como sendo o conjunto discreto de números correspondentes a cada *pixel* que conforma esta imagem. A imagem em níveis de cinza é composta por um conjunto discreto de números no intervalo de 0 até 255.

Seja uma imagem digital, e  $x_i$  o valor correspondente a cada *pixel*, no domínio discreto, definem-se:

Imagem em níveis de cinza:  $x_i \in [0, \dots, 255]$ , podendo assumir 256 valores.

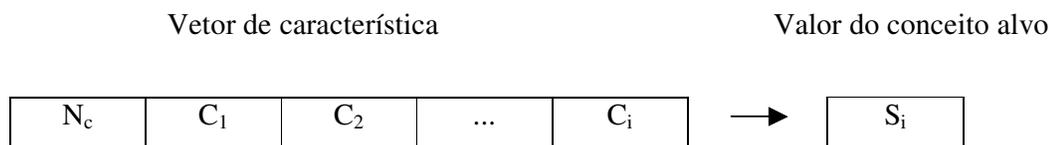
Imagem binarizada:  $x_i \in [0, 255]$ , podendo assumir 2 valores

Em ambos os casos:

$x_i=0$ , o nível preto e  $x_i=255$ , o nível branco.

Na etapa de aprendizagem são processadas uma imagem em níveis de cinza, chamada amostra (ou imagem observada), e sua correspondente solução ideal segmentada. O termo “solução ideal” ou ainda, imagem ideal, neste contexto, significa o resultado esperado pelo usuário ou especialista.

São analisadas as vizinhanças, de cada par de *pixels* da imagem amostra e a correspondente imagem ideal, extraindo-se um conjunto de características. Nesta etapa é gerado um vetor de características (**Figura 3.4**), para cada pixel e seu nível de cinza associado, junto a seu valor do conceito alvo, o qual será reutilizado no momento da segmentação das novas imagens. O conjunto de características é normalizado no intervalo  $[0, \dots, 255]$ .



$N_c$ : Nível de cinza do *pixel*

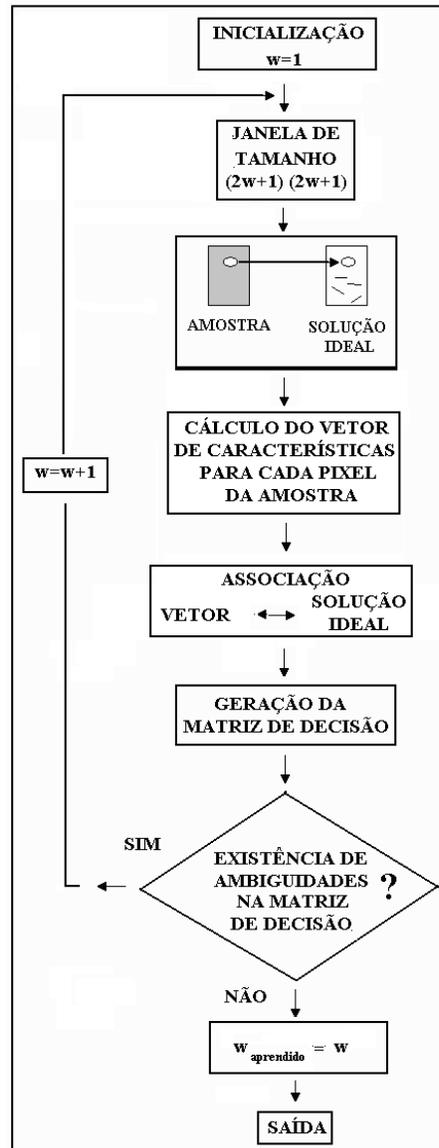
$C_i$ :  $i$ -ésima característica

$S_i$ : saída ideal do  $i$ -ésimo vetor de característica

**Figura 3.4.** Estrutura do vetor de características

Observa-se também, na estrutura da **Figura 3.4**, que cada vetor possui um respectivo valor de saída ideal. O conjunto obtido de todos os vetores de características, ordenados de acordo com um critério (por exemplo, em ordem crescente dos níveis de cinza dos *pixels*), foi

chamado de “matriz de decisão”. Esta matriz abrange todos os níveis de cinza da imagem amostra.



**Figura 3.5.** A etapa de aprendizagem e definição automática do tamanho da janela.

Os passos da etapa de aprendizagem são, portanto:

- O cálculo do vetor de características, normalizado, para cada *pixel*, e sua correspondente vizinhança (de janela centralizada), na imagem amostra;
- A definição automática do tamanho de janela, a partir da análise de cada *pixel* na imagem amostra e sua saída ideal. No caso de existirem dois *pixels* de entrada apresentando o mesmo nível de cinza, com vetores de características iguais, porém com saídas ideais diferentes, o tamanho da janela é aumentado até que estas ambigüidades sejam eliminadas e solucionadas.

Um esquema mais detalhado dos passos da etapa de aprendizagem é apresentado na **Figura 3.5**.

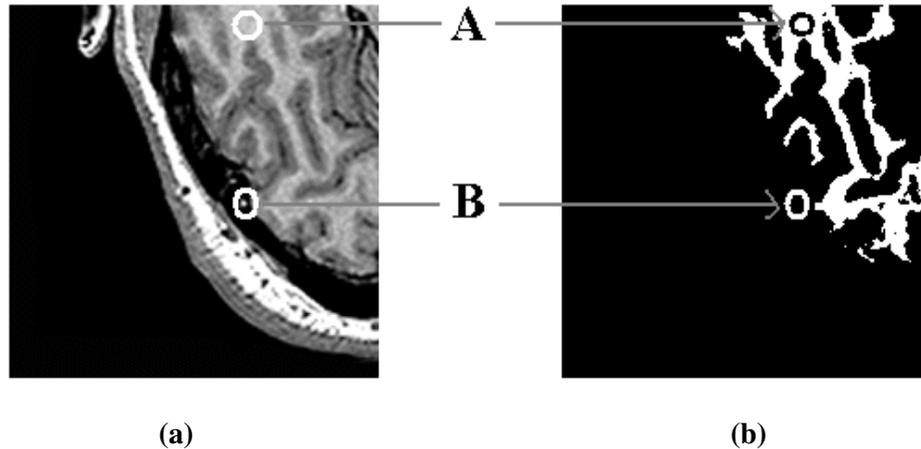
### 3.7. A Escolha do Tamanho Ideal da Janela

Essa abordagem baseia-se na extração e no armazenamento das características de cada *pixel*, considerando sua vizinhança. Portanto, a escolha do tamanho de janela apropriado é fundamental para o desempenho geral do processo. Por outro lado, o procedimento aqui apresentado faz parte da estratégia para solucionar as confusões na aprendizagem.

A janela é definida, nessa abordagem, como o conjunto de *pixels* vizinhos centralizados no *pixel* em análise. A janela é de forma quadrada e os lados possuem sempre um número ímpar de *pixels*. Assim, tem-se a janela 3x3, 5x5, 7x7, etc.

Para evitar inconsistências na decisão, deve-se garantir que não existam *pixels* com saídas ideais diferentes das que tenham o mesmo nível de cinza e vetores de características iguais.

A **Figura 3.6** ilustra um exemplo da presença de conflitos, no caso de uma imagem de ressonância magnética (MRI): A área cinza **A**, em círculo deve ser segmentada como branco. Porém, a outra área **B** em círculo, segmentada como preto. As duas áreas **A** e **B**, na imagem (a), possuem o mesmo nível de cinza no *pixel* central.



**Figura 3.6.** Exemplo de presença de inconsistência na imagem.

O cálculo dos vetores de características depende diretamente do tamanho da janela. Portanto, é necessário descobrir o tamanho de janela adequado, que soluciona essas inconsistências. Porém, deve exigir o menor tempo computacional possível. Ou seja:

- Quanto maior a janela, maior o tempo de processamento. Isso ocorre pois, para o cálculo das características, utiliza-se um maior número de pontos ou *pixels*. Portanto, é desejável utilizar sempre a menor janela possível;
- Janelas muito pequenas, em alguns casos, podem gerar vetores de características similares, porém com saídas ideais diferentes, criando ambigüidades;
- Como o tamanho da janela apropriada varia com o tipo de imagem utilizado, é recomendável que este tamanho seja ajustado, sem que haja intervenção manual do operador.

Devido às considerações anteriores, foi projetado um algoritmo de seleção automática de janela, no qual é processada a imagem inteira, iniciando-se num tamanho 3x3. No final do processamento da imagem, verifica-se a existência de inconsistências na Matriz de Decisão. Caso exista alguma, reinicia-se o processamento da imagem, com o tamanho de janela ímpar superior (no caso, 5x5; logo 7x7; etc). Este algoritmo encontra-se esquematizado na **Figura**

**3.5.** Por motivos práticos, o algoritmo de ajuste de janela foi limitado ao tamanho máximo de 15x15.

Cada inconsistência achada é interpretada como a presença de um erro ( $\epsilon$ ) na aprendizagem. A inconsistência, ou ambigüidade, foi definida como a presença de dois vetores de característica idênticos, porém com saídas diferentes. O algoritmo de aprendizagem e ajuste automático do tamanho de janela foi projetado para admitir três condições de parada: No final do processo de aprendizagem, com um dado tamanho de janela, é computado o total de inconsistências achadas, ou seja o total de erros. Este total de erros  $(\sum\epsilon)_k$ , é comparado com o valor total da iteração anterior  $(\sum\epsilon)_{k-1}$ . Assim, foram estabelecidas as três seguintes condições de parada:

- a) Total de erros na aprendizagem da iteração atual igual a zero;

$$(\sum\epsilon)_k = 0 \quad (3.12)$$

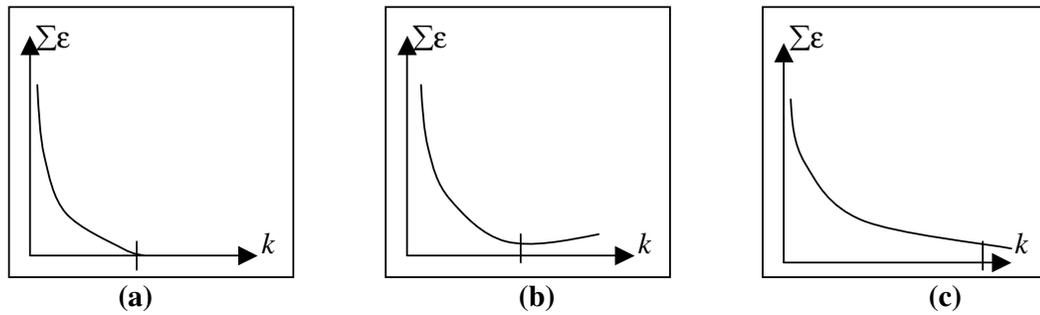
- b) Total de erros na aprendizagem da iteração atual maior ao total calculado na iteração anterior;

$$(\sum\epsilon)_k > (\sum\epsilon)_{k-1} \quad (3.13)$$

- c) Número de 15 iterações na aprendizagem;

$$(\sum\epsilon)_{15} \quad (3.14)$$

Estes três critérios podem ser melhor compreendidos de forma gráfica na **Figura 3.7** a seguir:



**Figura 3.7.** Gráfico de total de erros ( $\Sigma\epsilon$ ) vs. iteração

(a)  $(\Sigma\epsilon)_k = 0$ ; (b)  $(\Sigma\epsilon)_k > (\Sigma\epsilon)_{k-1}$  (c)  $(\Sigma\epsilon)_{15}$

A **Figura 3.8** mostra um pseudo-código que ilustra o critério de parada do algoritmo de aprendizagem.

```

Inicializa
  Tamanho Janela = 3x3
  Erro_Anterior = NÚMERO GRANDE
  Executar = VERDADEIRO
  Iteração = 1;
enquanto Executar = VERDADEIRO
  aprendizagem;
  calcular Erro_Atual;
  se
    Erro_Atual = 0 ou
    Erro_Atual > Erro_Anterior ou
    Iteração = 15
  então Executar = FALSO
  fim-se
fim-enquanto
  
```

**Figura 3.8** Pseudo-código que ilustra o critério de parada do algoritmo de aprendizagem

No **Apêndice A** é apresentado um exemplo prático detalhado do cálculo do tamanho de janela ideal para solucionar o problema de inconsistências na aprendizagem.

### 3.8. A construção da matriz de decisão

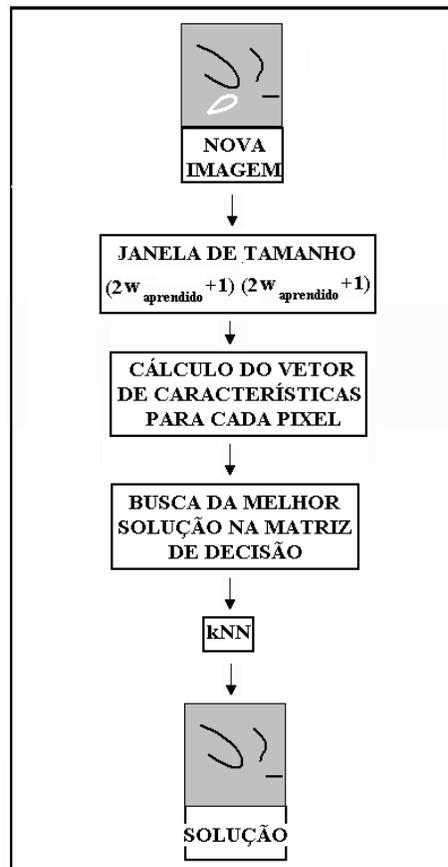
A matriz de decisão armazena, de acordo com os níveis de cinza da imagem amostra e com uma ordenação particular, o conjunto de vetores de características obtidos para todos os pares de *pixels* do par de imagens amostra/ideal. O número de vetores de características na matriz de decisão é igual ao número de pares de *pixels* aprendido da imagem submetida da amostra e sua correspondente imagem esperada ideal.

A matriz de decisão é construída com todos os vetores de características ordenados seqüencialmente, pelo respectivo nível de cinza dos pixels. Esta estrutura permite, a identificação das faixas de níveis de cinza não contemplados e, portanto não aprendidos. Na etapa de segmentação das novas imagens, a estratégia dos vizinhos mais próximos (*k*-NN) seleciona o vetor de características mais próximo ao nível de cinza em análise. Nos casos de níveis de cinza não contemplados, são incluídos os vetores de características, com níveis de cinza imediatamente anterior e posterior, como sendo pertencentes à faixa válida.

### **3.9. Segmentação de Novas Imagens por Aprendizagem**

Depois de concluída a etapa de aprendizagem, novas imagens do mesmo tipo podem ser segmentadas. Os passos da etapa de segmentação por aprendizagem são (**Figura 3.9**):

- a) Cálculo de vetor de características, normalizado, para cada *pixel* na nova imagem. Para isto, utiliza-se a informação do tamanho da janela, obtida durante o processo de aprendizagem;
- b) Escolha do vetor de características mais similar na matriz de decisão;



**Figura 3.9.** Processo de Segmentação de Imagens.

- c) Atribuição do valor de saída do *pixel* armazenado desse vetor na matriz de decisão. A estratégia do vizinho mais próximo é utilizada para seleccionar o vetor de características “mais similar” dentro da matriz de decisão, atribuindo-se assim o valor do *pixel* de saída armazenado.

### 3.9.1 Classificador Supervisionado

O armazenamento da solução ideal, dentro da Matriz de Decisão, junto com os vetores de características, permite a utilização de um classificador supervisionado.

Os classificadores supervisionados são métodos nos quais o usuário interage com o algoritmo, fornecendo informações como o número de classes e amostras de áreas conhecidas da imagem, para cada classe.

Um classificador supervisionado muito utilizado é o k-NN. Trata-se de um algoritmo baseado em instâncias para aproximar funções alvo, de valor real ou de valor discreto, assumindo que as instâncias correspondem a pontos no espaço  $d$ -dimensional [MIT97].

### 3.9.2. A Regra do Vizinho Mais Próximo (1-NN)

A regra do vizinho mais próximo é o modelo de aprendizagem baseada em instâncias utilizada nesta abordagem. O modelo NN (do inglês *nearest neighbor*), tem sido amplamente estudado e aplicado em diversas áreas. A formalização desse modelo e os resultados teóricos de convergência estatística, podem ser encontrados em [COH67]. Outros artigos teóricos que tratam do modelo NN são [SHF81], [FUF85] e [LOM87].

A regra do vizinho mais próximo pode ser definida da seguinte maneira. Dado um novo vetor para classificação, denotado por  $\underline{x}$ , e um conjunto de treinamento na forma de amostras identificadas:

$$H_n = \{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\}, \quad (3.15)$$

sendo cada vetor  $\underline{x}_i$  identificado como pertencendo a uma classe. Suponha que  $\underline{x}'$  é a amostra identificada em  $H_n$ , mais próxima a  $\underline{x}$ . A regra do 1-NN consiste em: classificar  $\underline{x}$ , atribuindo-o a mesma identificação que a classe  $\underline{x}'$ .

### 3.9.3. Distância Euclidiana

A regra do vizinho mais próximo implica na definição de uma distância entre os elementos, sendo utilizada nesta abordagem a distância euclidiana. O valor desta distância entre dois vetores  $\underline{x}$  e  $\underline{y}$  está dado por:

$$D(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\| = \sqrt{(\underline{x} - \underline{y})^T (\underline{x} - \underline{y})} \quad (3.16)$$

$$= + \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3.17)$$

Sendo  $n$ , a quantidade de atributos usados para descrever as instâncias.

### 3.9.4. Taxa de Erro da Classificação pela Regra do Vizinho Mais Próximo

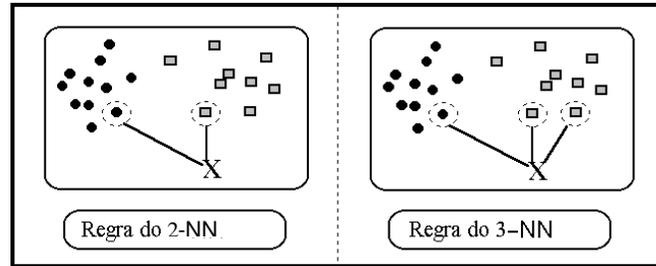
A relação entre a taxa de erro obtida através de uma classificação, por 1-NN ( $e_{1NN}$ ) e aquela obtida através de uma abordagem Bayesiana ( $e_{Bayes}$ ) está definida pela inequação a seguir [SCH92]:

$$e_{Bayes} < e_{1NN} < 2 * e_{Bayes} \quad (3.18)$$

Sabe-se que a teoria de decisão de Bayes fornece os melhores resultados de classificação, se for conhecida a informação *à priori* sobre os padrões. Porém, em situações práticas, é difícil obter informação completa *à priori*, devido a que se possui somente um número limitado de amostras. Neste caso, de acordo com [KIH95], o método de decisão de Bayes não apresenta os melhores resultados, sendo portanto preferível, o método dos vizinhos mais próximos.

### 3.9.5. A Regra dos $k$ Vizinhos Mais Próximos ( $k$ -NN)

A abordagem 1-NN pode ser estendida para a regra dos  $k$ -vizinhos mais próximos, ou  $k$ -NN. Neste caso, examinam-se as identificações nas  $k$ -amostras mais próximas em  $H$  e classifica-se seguindo um esquema de votos. A regra dos  $k$ -NN pode ser formulada: “O elemento  $x$  pertence à classe  $\underline{x}'$  se  $\underline{x}'$  é representada majoritariamente, entre os  $k$  vizinhos mais próximos de  $x$ ”. A **Figura 3.10**. ilustra um exemplo de aplicação das regras do 2-NN e 3-NN.



**Figura 3.10:** Regras do 2-NN e do 3-NN

### 3.9.6. O Valor de $k$ , o Número de Vizinhos Mais Próximos

Verificou-se que até o momento, não existe um consenso na literatura sobre o valor de  $k$  que minimize o erro da classificação, pela regra dos vizinhos mais próximos. Cohen afirma no artigo [COH67], que a estratégia 1-NN é estritamente melhor que qualquer estratégia  $k$ -NN,  $k \neq 1$ .

Por outro lado, diversos autores, como por exemplo, Cornuéjols e Miclet [COM02], sugerem a utilização da seguinte fórmula heurística para a seleção do valor de  $k$ :

$$k \approx \sqrt{\frac{m}{C}} \quad (3.19)$$

onde:  $k$  é o número de vizinhos,

$m$  é o tamanho da população, e

$C$  é o número de classes em que se deseja segmentar.

Todas as imagens objeto de pesquisa nesta abordagem são cinza com 256 níveis. A aplicação desta fórmula heurística para os dados  $m = 256$  (a população, contendo os níveis de cinza possíveis) e  $C = 2$  (as classes a segmentar, objeto e fundo), dá como resultado o valor:  $k \approx 11$

Kim [KIY97] afirma, no entanto, que na prática, o valor de  $k$  que minimiza o erro depende da aplicação.

No **Capítulo 4** será feito um estudo comparativo da influência do valor de  $k$ , nos resultados de segmentação de imagens, no intervalo ímpar  $[1, \dots, 15]$ . O objetivo é de verificar a influência da escolha do parâmetro  $k$ , no desempenho do algoritmo, aplicado aos tipos de imagens testados nesta abordagem.

### **3.9.7. Características da Regra dos $k$ -Vizinhos Mais Próximos**

A abordagem dos  $k$ -vizinhos mais próximos apresenta vantagens e desvantagens, das quais são citadas as mais relevantes [JAI00], [KIM96]:

Vantagens:

- Simplicidade de formulação e aplicação;
- É intuitiva;
- Não requer do conhecimento da distribuição das classes;
- Desempenho robusto.

Desvantagens:

- Tempo de processamento elevado para dimensões de espaços muito grandes.
- Desempenho dependente da escala (ou métrica) utilizada.

### **3.10. Comentários Finais**

Neste Capítulo foi apresentada a abordagem proposta de Segmentação por Aprendizagem. Inicialmente foi feita uma breve revisão do conceito de Aprendizagem Computacional, depois, um estudo das abordagens anteriores, e a seguir, foram apresentados os fundamentos da abordagem proposta. Foram detalhadas as etapas de aprendizagem e segmentação, incluindo a seleção das características, as considerações sobre o tamanho da

janela e sobre a medida de similaridade utilizada, o  $k$ -NN. No próximo Capítulo, será apresentada uma descrição de cada uma das bases de imagens utilizadas nos experimentos.

## Capítulo 4

# Considerações sobre as Bases de Imagens Utilizadas

Neste Capítulo será apresentada uma descrição das diversas bases utilizadas: assinaturas, impressões digitais, imagens de ressonância magnética e envelopes. É descrita também, de forma sucinta, a utilidade da segmentação destas imagens no contexto prático.

### 4.1. Imagens de Assinaturas

#### 4.1.1. Introdução

O problema em verificação automática de assinaturas é, em teoria, um processo de reconhecimento de padrões usado para discriminar as classes de assinaturas originais e as falsas. Segundo Justino [JUS01], apesar dos muitos esforços para desenvolver novas técnicas de verificação de assinaturas estáticas, a influência dos vários tipos de falsificações ainda não foi extensamente estudada, sob a luz do grafismo.

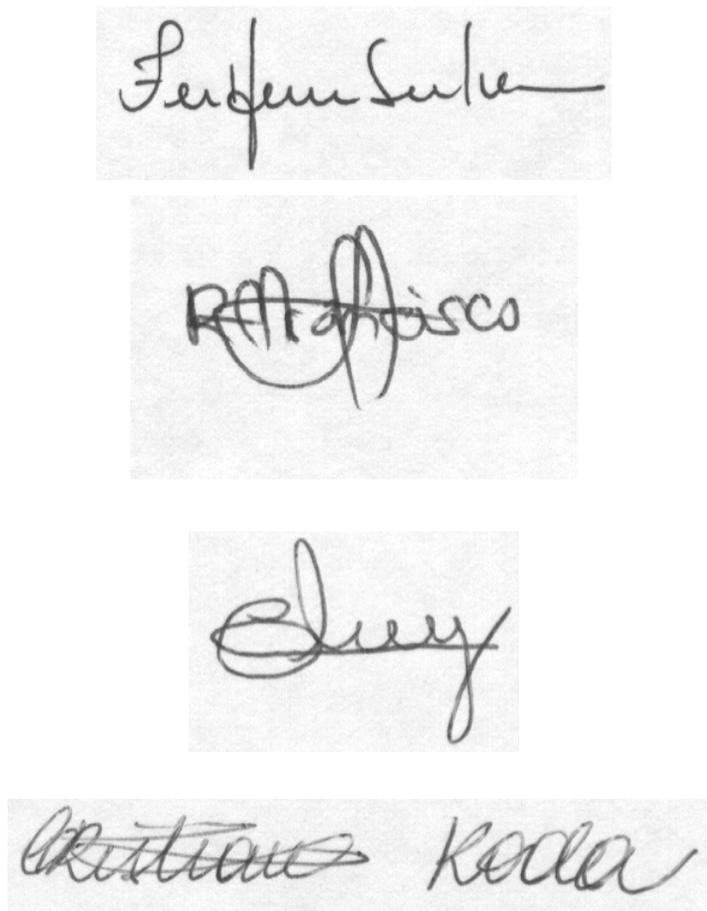
Na composição da base criada por [JUS01], foram coletadas 5.200 imagens de assinaturas. Desse montante, existem 4.000 assinaturas genuínas, oriundas de 100 autores diferentes (40 amostras de assinaturas para cada autor) e 1.200 falsificações. As assinaturas foram digitalizadas numa resolução de 300dpi, e em 256 níveis de cinza. Em seguida, as imagens foram recortadas e gravadas em um arquivo no formato BMP, num retângulo de tamanho de 3x10 cm ou (400x1.000 *pixels*), correspondendo à área destinada à assinatura no cheque bancário. Para maiores detalhes sobre a montagem desta base de assinaturas, sugere-se a referência original.

Foi realizada também, a partir da referida base, um processo de binarização das imagens em níveis de cinza, usando o método de Abutaleb [ABU89]. De acordo com Justino,

a escolha desse método teve como base os resultados satisfatórios obtidos com a base de laboratório.

#### 4.1.2. Descrição da Base de Imagens de Assinaturas Utilizada

A base de imagens de assinaturas, utilizada nos experimentos, é composta de uma amostra de cada autor num total de 100 imagens de assinaturas em níveis de cinza retiradas da base [JUS01] disponível no Laboratório de Análise e Reconhecimento de Documentos (LARD) PPGIA/PUCPR. Para cada uma destas imagens, foi gerada de forma manual, a correspondente imagem binarizada ideal (*ground-truth*). As imagens ideais geradas, não são coincidentes com aquelas binarizadas por [JUS01]. A **Figura 4.1** apresenta alguns exemplos das imagens desta base.



**Figura 4.1.** Exemplos da base de imagens de assinaturas.

## 4.2. Imagens de Impressões Digitais

### 4.2.1. Introdução

Uma etapa muito importante, para o reconhecimento e classificação de imagens de impressões digitais, é a segmentação por binarização destas impressões, ou seja, a separação das linhas dos sulcos característicos. Isto implica na consideração de um fator que influencia na qualidade da imagem digitalizada: o *threshold* ou limiar de sensibilidade do software de binarização. Este limiar de sensibilidade ótimo varia de impressão para impressão. Caso este limiar não seja posicionado adequadamente, pode-se ter o desaparecimento de trechos de linhas, o afinamento das mesmas, ou no caso contrário, a conexão entre linhas muito próximas ou até a formação de manchas. Outro problema típico, na aquisição de impressões digitais latentes, é a digitalização de áreas com manchas devido ao espalhamento de tinta na hora da coleta da impressão.



**Figura 4.2.** Exemplos da base de imagens de impressões digitais.

Assim, um processo de binarização ideal de impressões digitais deve, imperativamente [LEA91]:

- Manter a continuidade de uma linha;
- Restituir a continuidade no caso desta apresentar variações no nível de cinza;
- Restituir a espessura constante das linhas;
- Preservar as minúcias, evitando, por exemplo, a separação de duas linhas conectadas;
- Impedir a geração de minúcias falsas, impedindo, por exemplo, a reunião de duas linhas originalmente separadas;
- Conservar a simetria.

#### **4.2.2. Descrição da Base de Imagens de Impressões Digitais Utilizada**

Na montagem da base de teste de imagens de impressões digitais, foram extraídas 80 imagens em níveis de cinza, de uma base de impressões digitais do *Finger-print Verification Competition – FVC200*, disponível no *site* da BioLab, Universidade de Bologna [FVC02] e geradas as correspondentes imagens ideais (*ground-truth*). Esta base foi escolhida por ser de acesso público, portanto, de fácil verificação. Na **Figura 4.2**, apresentam-se exemplos da base de imagens de impressões digitais.

### **4.3. Imagens de Ressonância Magnética**

#### **4.3.1. Introdução**

O estudo de ressonância magnética é uma técnica avançada de avaliação médica por imagens, que fornece uma grande quantidade de informação sobre a anatomia do tecido humano mole.

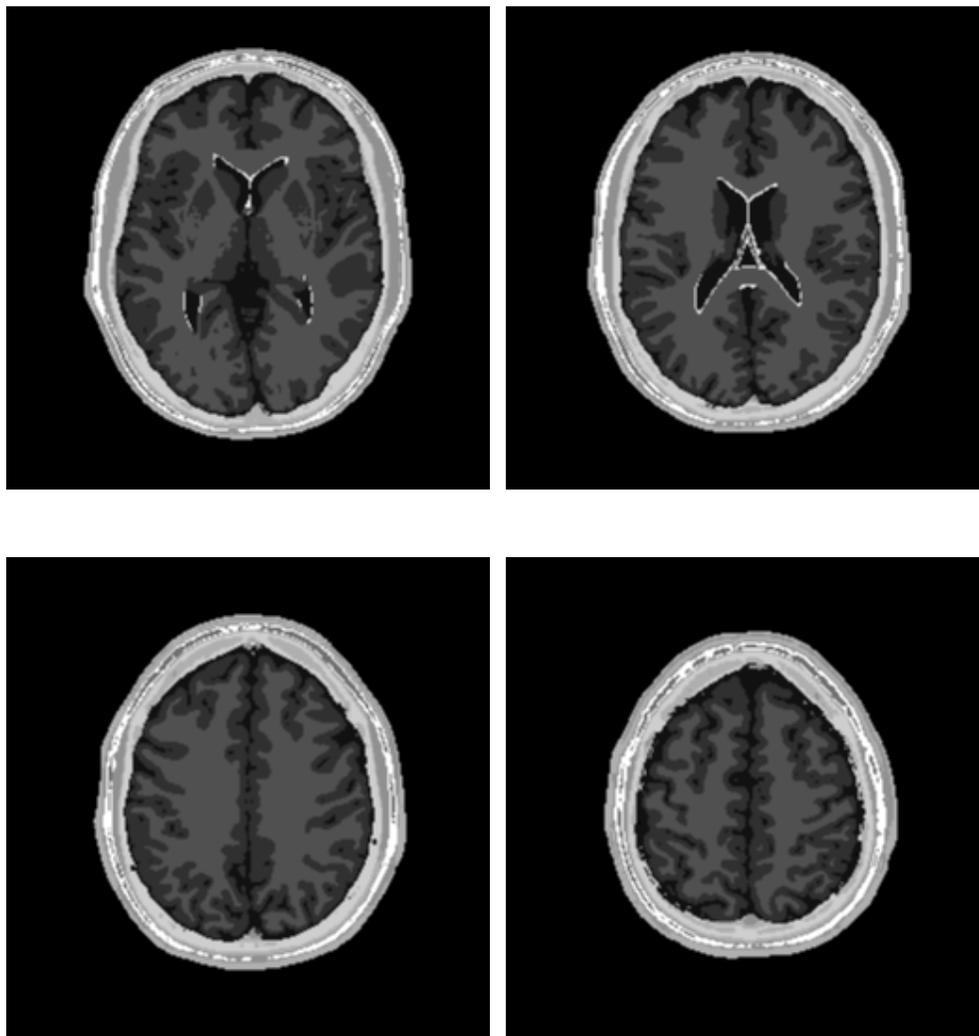
A segmentação de imagens de ressonância magnética (MRI) do cérebro, em diferentes classes de tecidos, especialmente a massa cinza, a massa branca e o líquido é um auxílio importante na avaliação de várias desordens neurológicas, incluindo a esclerose múltipla [GGU00] ou a doença de Alzheimer [FFR96]. A técnica de segmentação mais amplamente

utilizada, em imagem médica, é a segmentação manual [SVV00], porém ele requer uma experiência sobre o tamanho, localização e forma das várias estruturas cerebrais. Trata-se de um procedimento que consome uma grande quantidade de tempo e é de difícil reprodução. Por este motivo, são necessárias técnicas de análise de imagens assistidas por computador.

A maioria das técnicas de segmentação de MRI, pode ser dividida em *clustering* e métodos baseados em região e borda. Outra classificação, divide estas técnicas em automáticas e semi-automáticas. As abordagens mais utilizadas incluem, *fuzzy c-means*, *clustering*, *growing-based*, morfologia matemática, histograma bi-dimensional e redes neurais. A maioria destas técnicas requer médicos experientes para controlar ativamente o processo de segmentação e corrigir interativamente os resultados. Assim, a exatidão depende da precisão e repetibilidade do operador, necessários para a intervenção. Um método amplamente utilizado de segmentação de MRI, o *fuzzy c-means clustering*, é de acordo com Fletcher-Heath *et al* [FHG01], sensível à inicialização. Um estudo aprofundado de todos os métodos anteriormente citados está fora do escopo deste trabalho, portanto, para uma revisão sobre segmentação de MRI, sugere-se, por ex., [CVC95], [BHC93] e [NIE97].

#### 4.3.2. Descrição da Base de Imagens de Ressonância Magnética (MRI) Utilizada

A base de imagens de ressonância magnética (MRI) utilizada foi obtida do *Serial Sections Brain MRI database*, disponível no *site* do *Montréal Neurological Institute* [UNW03]. Esta base pública está disponível para a comunidade de neuro-imagem para avaliar o desempenho das abordagens de análise de imagens. Contém um conjunto de volumes de dados de MRI, produzidos por um Simulador de MRI, desenvolvido no *McConnell Brain Imaging Centre*. A grande vantagem da utilização desta base é a disponibilidade das classes *ground truth* massa cinza ou *gray matter* (GM), massa branca ou *white matter* (WM) e líquido ou *cerebrospinal fluid* (CSF) para cada imagem. Com o auxílio deste simulador, foi gerada uma base de 44 imagens de ressonância magnética (MRI), composta por 11 imagens MRI T1 de cérebro saudáveis, e as correspondentes imagens *groundtruth* GM, WM e CSF. Os parâmetros utilizados para a geração das imagens no simulador foram: formato *grayscale*, separação entre os cortes 5 mm e espessura de cada corte 1 mm. A **Figura 4.3.** apresenta exemplos da base de imagens de ressonância magnética.



**Figura 4.3.** Exemplos da base de imagens de ressonância magnética.

## **4.4. Imagens de Envelopes Postais**

### **4.4.1. Introdução**

A Automação Postal é uma área de grande pesquisa nos últimos tempos. Isto é devido, ao barateamento e à facilidade de aquisição das imagens de envelopes postais, assim como à evolução das unidades de processamento dos computadores, que viabilizaram a utilização de algoritmos complexos. A localização do Bloco Endereço em envelopes postais ainda é o

ponto crucial no desenvolvimento de um sistema de Análise de Imagens para a Automação Postal.

A Empresa de Correios e Telégrafos (ECT) recomenda certas regras para o preenchimento correto dos envelopes postais brasileiros:

- Os selos e os carimbos devem ocupar a parte superior à direita;
- A parte manuscrita e/ou pré-impressa deve ser preenchida de tal forma que, não se sobreponha aos selos e carimbos;

Na maior parte dos casos, as mesmas não se verificam, inviabilizando, portanto, uma técnica de segmentação baseada na posição das classes que compõem o envelope postal brasileiro.

O processo de automação postal consiste em um sistema de Visão, que seja capaz de segmentar uma imagem de envelope postal, extraíndo as informações do destinatário, e interpretá-las, possibilitando, dessa forma, o envio direto da correspondência.

Além do descumprimento das regras de preenchimento da ECT, outros problemas dificultam a implementação de um sistema eficiente de segmentação de imagens de envelopes postais, tais como:

- Imperfeições dos próprios envelopes devidas em geral, ao manuseio, como amassados e marcas.
- Presença de desenhos feitos à mão, que torna o fundo mais complexo;
- Problema de aquisição, gerando bordas adicionais que não fazem parte dos envelopes postais.

Os experimentos a serem apresentados visam separar nos envelopes o fundo, o selo, o carimbo e o bloco endereço. Sendo estes três últimos, considerados objetos de segmentação. Segundo Menoti [MEN03], apesar de haver vários trabalhos que tratam da segmentação de envelopes postais, tal tarefa ainda é um problema desafiador, considerando a grande variedade

de selos, fundos, e tipos e tamanhos de manuscritos/pré-impressos, que compõem o bloco endereço.

Dos trabalhos sobre segmentação de imagens de envelopes postais realizados no PPGIA/PUCPR, sobre a base descrita na seção seguinte, citamos: Yonekura [YON02] por abordagem morfológica, utilizando matriz de co-ocorrência e histograma 2D; Menoti [MEN03] utilizando uma abordagem baseada em seleção de características no espaço *wavelet*; e ainda Eiterer [EIT04] através de uma abordagem baseada na dimensão fractal.

#### **4.4.2. Descrição da Base de Imagens de Envelopes Postais Utilizada**

A base de envelopes postais é composta de 200 imagens reais, em 256 níveis de cinza e digitalizadas a 200 *dpi*. Os mesmos foram retirados de uma base de 55.000 envelopes adquiridos da ECT - Empresa Brasileira de Correios e Telégrafos.

Em princípio, pode parecer muito pequena, um conjunto de 200 imagens, de uma base de 55.000 existentes. Os critérios para a seleção das 200 imagens foram as características, tais como, envelope grande, amassado, fundo complexo e o texto manuscrito. Cabe salientar que os envelopes de tamanho grande foram utilizados, por conterem o fundo mais complexo. Uma inspeção visual da base completa da ECT apresenta os seguintes números:

- Envelopes pequenos de fundo simples (branco): 45.732 (83,1%)
- Envelopes grandes (inclui revistas e objetos): 9.268 (16,9%)

Do total de 55.000 envelopes, 2.204 (4%) são manuscritos e destes, 375 (0,7%) de tamanho grande. Finalmente, 204 (0,4%) envelopes cumpriram os critérios de seleção expostos anteriormente.

Assim, a sub-base de 200 imagens utilizada nos experimentos, não representa uma amostra estatística da base completa adquirida da ECT, e sim o grupo de imagens considerado mais complexo.

As dimensões das imagens selecionadas são ( $\mu \pm \sigma$ ): largura  $1857 \pm 20,5$  *pixels* e altura  $1340 \pm 41,9$  *pixels*. O tamanho médio das imagens é de 2,49 MB.

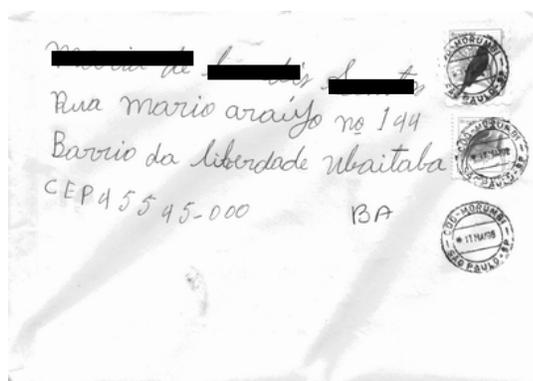
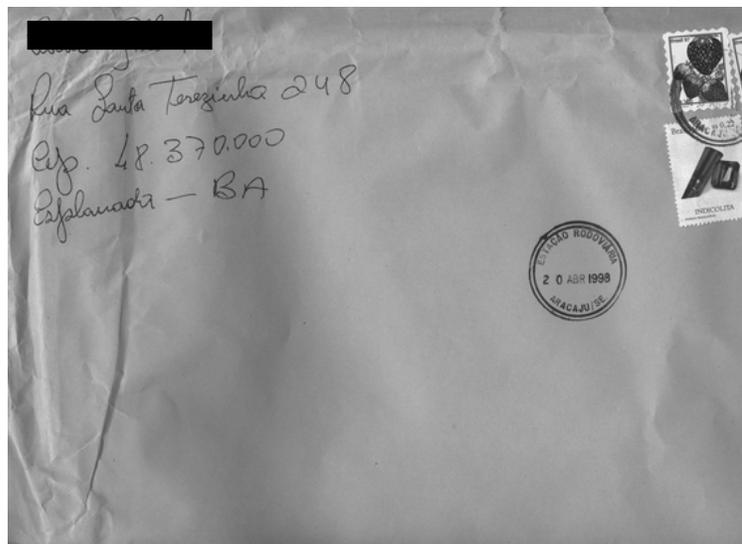
Devido ao tamanho das imagens a serem processadas, o esforço computacional requerido para todos os experimentos previstos e às limitações devidas ao compartilhamento do equipamento servidor, com outros processos concorrentes dos diversos projetos do PPGIA-PUCPR, optou-se por reduzir em 25% o tamanho das 200 imagens de envelopes. Esta base, chamada de ECT-DB1 possui as seguintes dimensões: largura  $465 \pm 5,1$  *pixels* e altura  $335 \pm 10,5$  *pixels*. O tamanho médio das imagens é de 155,79 KB. A base de 200 imagens com a dimensão original foi chamada de ECT-DB2. Considerações adicionais a respeito da otimização dos algoritmos de aprendizagem podem ser vistos no **APÊNDICE B**.

Para cada uma dessas duzentas imagens, das bases ECT-DB1 e ECT-DB2, foram segmentados manualmente o bloco endereço, os selos e os carimbos, gerando assim as imagens ideais ou soluções teóricas esperadas da segmentação (imagens *ground truth*). A segmentação manual da classe selo, foi desenhada como uma área inteiramente preta.

A **Figura 4.4**. apresenta exemplos da base de imagens de envelopes postais utilizada.

## 4.5. Comentários Finais

Neste Capítulo foi apresentada uma descrição das bases de imagens utilizadas nos testes. No próximo Capítulo serão apresentados experimentos com a aplicação da abordagem proposta, em diferentes bases de dados. Também, serão avaliadas as características propostas.



**Figura 4.4.** Exemplos da base de imagens de envelopes postais.

# Capítulo 5

## Experimentos de Aplicações

Neste capítulo serão avaliadas as características que foram propostas na abordagem de segmentação por aprendizagem. Também, serão apresentados de forma objetiva, os resultados de segmentação, através de uma comparação *pixel a pixel* entre a imagem obtida e o resultado ideal da mesma.

A abordagem de segmentação por aprendizagem, apresentada até o momento, foi desenvolvida utilizando-se características previamente selecionadas: nível de cinza, média, variância, *skewness*, curtosis, média dos gradientes e mínimo e máximo dos gradientes. É preciso verificar se estas características são relevantes para uma correta segmentação por binarização das diferentes bases de imagens e qual das combinações produziu o melhor resultado.

### 5.1. Metodologia de Avaliação dos Resultados

#### 5.1.1. Metodologia de Avaliação para Duas Classes

A necessidade de contar com uma quantificação objetiva dos resultados obtidos nas imagens pela abordagem proposta, levou à adoção de uma metodologia de avaliação dos resultados. Esta metodologia será utilizada para as bases de assinaturas, impressões digitais, imagens de ressonância magnética e envelopes.

Seja uma imagem  $I$ , que se deseja segmentar em duas classes, objeto e fundo. A metodologia consiste na comparação *pixel a pixel* da imagem segmentada  $I_S$ , com a correspondente imagem ideal (*ground truth*)  $I_G$ . As classes objeto e fundo, da imagem

segmentada, são  $I_{SO}$  e  $I_{SF}$ . As correspondentes classes na imagem ideal são  $I_{GO}$  e  $I_{GF}$ , respectivamente. Têm-se as seguintes relações:

$$I_S = I_{SO} \cap I_{SF} \quad (5.1)$$

$$I_G = I_{GO} \cap I_{GF} \quad (5.2)$$

No processo de avaliação, caso o *pixel* de  $I_S$  seja encontrado em  $I_G$ , o *pixel* é classificado como sendo pertencente ao objeto de segmentação. Caso contrário, o *pixel* é classificado como ruído e pertencente ao fundo da imagem. A taxa de segmentação é a razão entre o número de *pixels* existente na imagem segmentada  $I_{SO}$  e na imagem ideal  $I_{GO}$ . O ruído do processo é medido como, a razão entre o número de *pixels* que foram segmentados em  $I_{SO}$  e não pertencem a  $I_{GO}$ , ou seja, é o número de *pixels* que pertencem ao fundo do envelope  $I_{GF}$ . Essas taxas de segmentação e de ruído podem ser formalmente definidas como:

$$TxSegmentação = \frac{\text{n}^\circ \text{ pixel } (I_{SO} \cap I_{GO})}{\text{n}^\circ \text{ pixel } (I_{GO})} \quad (5.3)$$

$$TxRuido = \frac{\text{n}^\circ \text{ pixel } (I_{SO} \cap I_{GF})}{\text{n}^\circ \text{ pixel } (I_{GF})} \quad (5.4)$$

### 5.1.2. Metodologia de Avaliação para Quatro Classes

Para o caso de envelopes postais, será também apresentada uma avaliação comparativa da abordagem proposta, com duas outras. A primeira refere-se à segmentação do bloco endereço baseada na seleção de características no espaço *wavelet*, e a segunda abordagem é baseada em abordagem por fractais. As técnicas, objeto de comparação, encontram-se detalhadas em [MEN03] e [EIT04], e utilizam como uma das bases de experimentos a mesma base de 200 imagens ECT-DB2. A abordagem de [MEN03], segmenta as imagens em quatro classes, chamadas de bloco endereço, carimbo, selo e fundo; e propõe uma avaliação particular dos resultados. Esta avaliação, também foi adotada por [EIT04]. Visando a objetividade na comparação a ser apresentada, implementou-se a mesma metodologia, a qual é resumida a seguir:

Todas as imagens que são submetidas ao algoritmo de segmentação possuem sua respectiva imagem esperada, isto é, o resultado da segmentação desejado  $I_{GT}$  e está composta pelos objetos  $I_{GTO}$  e pelo fundo  $I_{GTF}$ . Esta imagem esperada ou desejada é a imagem de referência (*ground truth*), previamente separada em três classes  $I_{GT1}$ ,  $I_{GT2}$  e  $I_{GT3}$ . A imagem resultante  $I_S$  (imagem segmentada), do algoritmo de segmentação, é comparada com as imagens de referência  $I_{GT1}$ ,  $I_{GT2}$  e  $I_{GT3}$  (imagens esperadas). Cada *pixel* de  $I_S$ , que foi segmentado (*pixel* preto), é comparado com o respectivo *pixel* em  $I_{GT1}$ ,  $I_{GT2}$  e  $I_{GT3}$ . Caso esse *pixel* de  $I_S$  seja encontrado, em uma dessas três imagens (e só em uma delas), o *pixel* é classificado como sendo pertencente ao objeto de segmentação da respectiva classe. Caso contrário, o *pixel* é classificado como ruído e pertencente ao fundo do envelope. Têm-se as seguintes relações entre as imagens:

$$I_{GT} = I_{GTO} \cup I_{GTF} \quad (5.5)$$

$$I_{GTO} = I_{GT1} \cup I_{GT2} \cup I_{GT3} \quad (5.6)$$

$$I_{GTF} = \overline{I_{GT1} \cup I_{GT2} \cup I_{GT3}} \quad (5.7)$$

Sendo  $I_{GTF}$  o complemento de  $I_{GTO}$

As taxas de segmentação e de ruído podem ser formalmente definidas como:

$$TsClasse\ 1 = \frac{\text{n}^\circ \text{ pixel}(I_S \cap I_{GT1})}{\text{n}^\circ \text{ pixel}(I_{GT1})} \quad (5.8)$$

$$TsClasse\ 2 = \frac{\text{n}^\circ \text{ pixel}(I_S \cap I_{GT2})}{\text{n}^\circ \text{ pixel}(I_{GT2})} \quad (5.9)$$

$$TsClasse\ 3 = \frac{\text{n}^\circ \text{ pixel}(I_S \cap I_{GT3})}{\text{n}^\circ \text{ pixel}(I_{GT3})} \quad (5.10)$$

$$Ruído = \frac{\text{n}^\circ \text{ pixel}(I_S - (I_{GT1} \cup I_{GT2} \cup I_{GT3}))}{\text{n}^\circ \text{ pixel}(\overline{(I_{GT1} \cup I_{GT2} \cup I_{GT3})})} = \frac{I_S - I_{GTO}}{I_{GTF}} \quad (5.11)$$

onde  $TsClasse\ 1$ ,  $TsClasse\ 2$  e  $TsClasse\ 3$  são as taxas de segmentação das classes 1 (bloco endereço), 2 (selo) e 3 (carimbo), respectivamente;  $I_S \cap I_{GT1}$ ,  $I_S \cap I_{GT2}$  e  $I_S \cap I_{GT3}$  são os conjuntos de intersecção entre a imagem segmentada  $I_S$  e as imagens *ground truth* (de

referência)  $I_{GT1}$ ,  $I_{GT2}$  e  $I_{GT3}$ , respectivamente, e  $(I_{GT1} \cup I_{GT2} \cup I_{GT3})$  é o conjunto de *pixels* que representam todos os objetos de segmentação;  $\overline{(I_{GT1} \cup I_{GT2} \cup I_{GT3})}$  é o complemento do conjunto  $(I_{GT1} \cup I_{GT2} \cup I_{GT3})$ , ou seja, o fundo da imagem *ground truth*  $I_{GTF}$ . Para uma revisão mais detalhada desta metodologia de avaliação, sugere-se a referência original [MEN03].

## 5.2. Metodologia de Avaliação do Desempenho de cada Característica

Nesta Seção, é apresentada a metodologia de avaliação quantitativa do desempenho de cada característica de forma isolada e combinada. Objetiva-se, através desta metodologia, verificar qual o melhor conjunto ou conjuntos de características, que segmenta as bases de imagens de assinaturas, impressões digitais, imagens de ressonância magnética e envelopes, descritas no **Capítulo 4**.

Em experimentos preliminares do autor, publicados em [LAa03], [LAb03] e [LAc03], foram utilizadas as características, Média, Variância e Média dos gradientes, nas diversas aplicações das bases de dados. No entanto, resta a pergunta: É esta a melhor combinação de características para todas as bases de imagens? Com o objetivo de responder a esta pergunta e de validar os resultados da abordagem proposta, com todas as características descritas na **Seção 4.4**, foi implementada a metodologia descrita a seguir, para cada uma das bases de imagens.

Inicialmente, foram selecionadas 10 imagens junto com suas respectivas versões ideais binarizadas, de cada uma das 4 bases. Cada par (imagem amostra, imagem *ground-truth*), foi submetida ao processo de aprendizagem da segmentação, gerando desta maneira, a correspondente matriz de decisão. Foi utilizada, separadamente, cada uma das características. Ou seja, a aprendizagem foi realizada com o nível de cinza da imagem e cada uma das seguintes características: Média, Variância, Média dos gradientes, Mínimo dos gradientes, Máximo dos gradientes, *Skewness* e *Curtosis*

A seguir, foram segmentadas as outras 9 imagens (método *Ten-Fold Cross Validation*). Este processo foi repetido com as subseqüentes imagens realizando-se assim um total de 90 testes por característica. Para comparar os resultados, foi calculado o número médio de pixels iguais, entre as imagens segmentadas e as correspondentes *ground-truth*.

Com o objetivo de avaliar a influência do parâmetro  $k$ , na técnica dos vizinhos mais próximos, assim como para verificar a validade da fórmula empírica da **seção 3.9.6**, os testes anteriormente citados, foram realizados para  $k = \{1, 3, 5, 7, 9, 11, 13, 15\}$ . O teste até o parâmetro  $k=15$ , foi realizado para permitir a inclusão de pelo menos 2  $k$ 's a mais ( $k=13$ ,  $k=15$ ) do valor de  $k = 11$ , obtido pela aplicação da fórmula empírica (3.19) da **Seção 3.9.6**.

Verificou-se que em todos os casos, o percentual de *pixels* iguais, entre as imagens segmentadas e as correspondentes *ground-truth*, obtidas com a utilização das características isoladas como Média dos gradientes, Gradiente mínimo e Gradiente máximo foi inferior a 60%. Por considerar estes resultados insatisfatórios, decidiu-se descartar as três características nos demais testes combinatórios e análises.

Assim, as características objeto de análise estatístico foram:

- Média,
- Variância,
- *Skewness*
- Curtosis

Nos cálculos de avaliação quantitativa, apresentados nas seguintes seções, além do cálculo da taxa de segmentação de *pixels* iguais entre as imagens segmentadas e as correspondentes *ground-truth*, foi também computado, para cada caso, a taxa de ruído conforme descrito na **Seção 5.1.1**.

Após a apresentação dos resultados de segmentação, para cada base de imagens calculadas, com cada uma das características isoladas, buscou-se qual a combinação de características de melhor desempenho. A estratégia geral é descrita a seguir.

Dadas as 4 características (Média, Variância, *Skewness* e *Curtosis*), o total de combinações possíveis entre elas é de 15. Assim, foram calculadas para cada uma das 10 imagens todas as matrizes de decisão com as 15 combinações de características, obtendo-se assim, 150 matrizes de decisão. Logo, escolhendo-se uma das dez imagens, aplicou-se a segmentação, com o algoritmo 1-NN, utilizando-se estas 150 matrizes de decisão, previamente geradas.

Assim, validou-se cada combinação de características realizando-se 10 segmentações de imagens diferentes, num total de 1.500 testes. Desta maneira, foi feita uma amostragem que relacionou as 10 imagens com as 15 combinações de características e as 150 matrizes de decisão. Para cada caso, foram computados o percentual de *pixels* coincidentes (taxa de segmentação) e o percentual de ruído. Foram calculadas também, para cada combinação de características, as médias destes dois valores anteriores e apresentados os resultados em um gráfico.

Assim, pelo método anterior, foi identificada a combinação de características que melhor segmenta imagens. A seguir, foram re-feitos os testes de aprendizagem, com cada uma das 10 imagens, e de segmentação, com as restantes 9, perfazendo um total de 90 testes. O processo foi repetido para os  $k$  vizinhos mais próximos  $\{1, 3, 5, 7, 9, 11, 13, 15\}$ , de maneira a verificar a validade da fórmula empírica (4.19) do parâmetro  $k$ . Todos estes resultados foram incluídos em gráficos comparativos.

### 5.3. Testes de Segmentação em Imagens de Assinaturas

Foram realizados os testes sobre base de assinaturas, detalhada previamente, sendo selecionado o subconjunto de teste de 10 imagens. A **Tabela 5.1** e a **Figura 5.1** mostram que, para o conjunto das imagens de assinaturas testado, a melhor característica isolada é a Média (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $91.62 \pm 6.66$  e  $0.99 \pm 0.93$ ) para o parâmetro  $k = 11$ .

Tendo-se achado o melhor resultado para uma característica isolada e parâmetro  $k$ , foi realizada a seguir a segmentação de imagens, com todas as combinações de características, porém com  $k=1$ . Esses resultados são apresentados na **Tabela 5.2** e de forma gráfica na

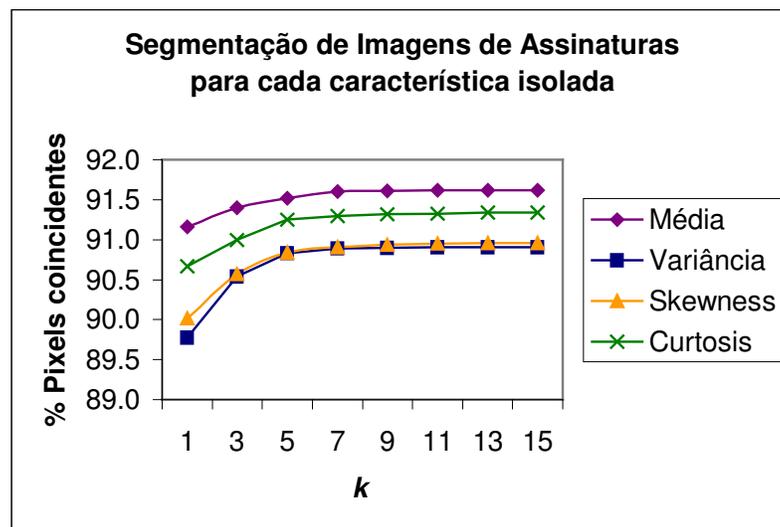
**Figura 5.3.** De forma comparativa, também foi realizada a segmentação de imagens, com todas as combinações de características, porém com  $k=11$ . Esses resultados são apresentados na **Tabela 5.4** e de forma gráfica na **Figura 5.4**. Isso foi feito, devido a que a melhor segmentação isolada foi obtida com  $k=11$ , para a característica Média. Observa-se que a combinação de características, que obteve os melhores resultados foi Média e Skewness, identificada como 1010, seja para  $k=1$  (taxa de segmentação e ruído em %  $\mu\pm\sigma$ , respectivamente:  $89.26\pm 5.71$  e  $0.12\pm 0.11$ ), como para  $k=11$  (%  $\mu\pm\sigma$ , respectivamente:  $91.23\pm 4.95$  e  $0.15\pm 0.13$ ). A identificação utilizada, para cada combinação de características, e um modelo desta identificação, estão na **Tabela 5.3** e **Figura 5.2**. Finalmente, tendo-se identificada a melhor combinação de características, vários experimentos foram feitos variando-se os valores do parâmetro  $k$  num intervalo ímpar [1, 15]. Verificou-se que, os melhores resultados, foram obtidos com o valor  $k=11$  (taxa de segmentação e ruído em %  $\mu\pm\sigma$ , respectivamente:  $94.19\pm 6.64$  e  $1.00\pm 0.92$ ). Apesar da taxa de segmentação, com  $k=9$ , ser igual à taxa obtida com  $k=11$ , este último parâmetro possibilitou um erro percentual 0,01 menor, considerando-se portanto, o melhor resultado. Os resultados da segmentação de imagens para as características Média e Skewness e o parâmetro  $k$ , no intervalo ímpar [1, 15], são apresentados na **Tabela 5.5** e de forma gráfica na **Figura 5.5**.

**Tabela 5.1** - Resultados da segmentação por binarização em assinaturas para cada uma das características isoladas em função ao valor k-NN

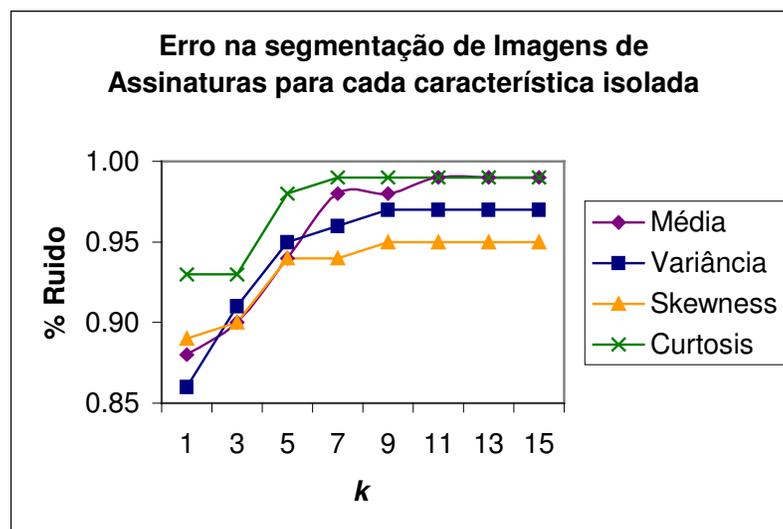
Característica: Média			Característica: Variância		
k-NN	Segmentação (% $\mu\pm\sigma$ )	Ruído (% $\mu\pm\sigma$ )	k-NN	Segmentação (% $\mu\pm\sigma$ )	Ruído (% $\mu\pm\sigma$ )
1	91.16±6.86	0.88±0.83	1	89.78±9.20	0.86±0.82
3	91.40±6.71	0.90±0.86	3	90.54±7.82	0.91±0.88
5	91.52±6.68	0.94±0.88	5	90.83±7.58	0.95±0.92
7	91.60±6.66	0.98±0.92	7	90.89±7.52	0.96±0.92
9	91.61±6.66	0.98±0.92	9	90.90±7.51	0.97±0.92
11	91.62±6.66	0.99±0.93	11	90.91±7.50	0.97±0.92
13	91.62±6.66	0.99±0.93	13	90.91±7.50	0.97±0.92
15	91.62±6.66	0.99±0.93	15	90.91±7.50	0.97±0.92

Característica: Skewness			Característica: Curtosis		
k-NN	Segmentação (% $\mu\pm\sigma$ )	Ruído (% $\mu\pm\sigma$ )	k-NN	Segmentação (% $\mu\pm\sigma$ )	Ruído (% $\mu\pm\sigma$ )
1	90.02±8.41	0.89±0.87	1	90.67±8.07	0.93±0.90
3	90.57±7.80	0.90±0.88	3	91.00±7.81	0.93±0.92
5	90.84±7.52	0.94±0.92	5	91.25±7.53	0.98±0.96
7	90.91±7.34	0.94±0.92	7	91.30±7.44	0.99±0.96
9	90.94±7.27	0.95±0.92	9	91.32±7.35	0.99±0.96
11	90.95±7.22	0.95±0.92	11	91.33±7.30	0.99±0.96
13	90.96±7.20	0.95±0.92	13	91.34±7.27	0.99±0.96
15	90.96±7.18	0.95±0.92	15	91.34±7.27	0.99±0.96



(a)



(b)

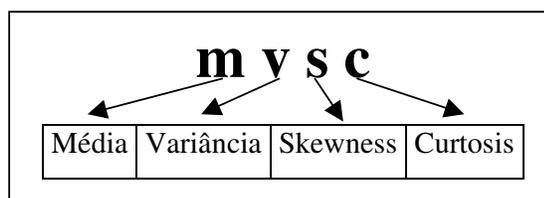
**Figura 5.1** – Gráfico da Segmentação de imagens de Assinaturas (a) e o correspondente Ruído (b), para cada característica isolada e o número de vizinhos mais próximos. Observa-se em ambos os gráficos que as taxas de segmentação e o ruído aumentam com os primeiros valores de  $k$ , e estabilizando-se com valores superiores de  $k$ .

**Tabela 5.2** - Resultados da amostra de segmentação por binarização de assinaturas para todas as combinações de características e  $k$ -NN = 1.

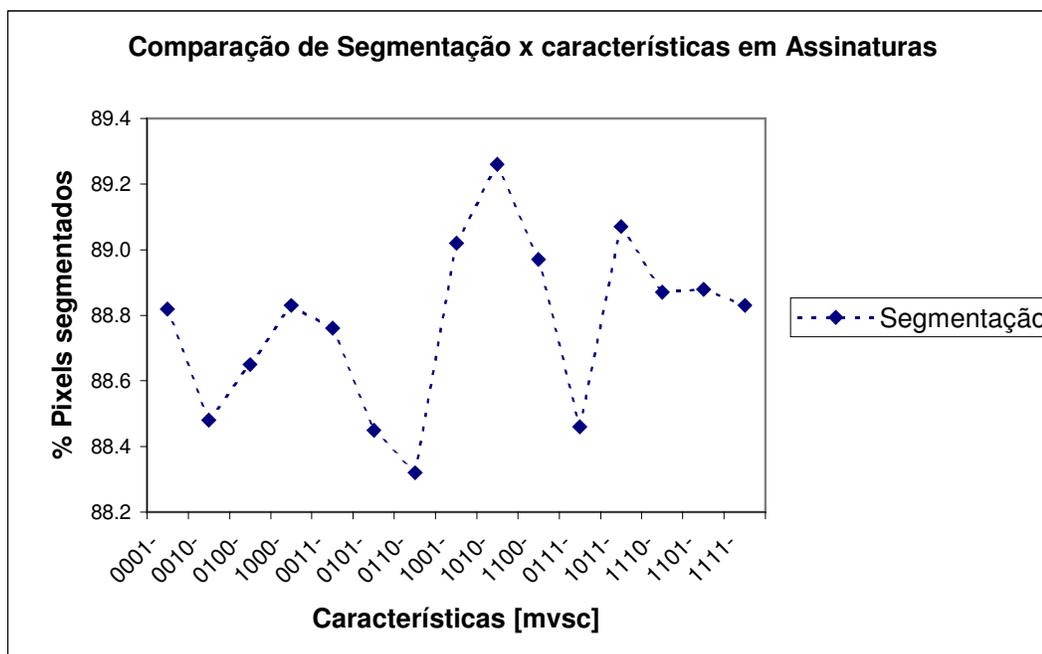
Caract	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
0001-	88.82 $\pm$ 5.71	0.12 $\pm$ 0.12
0010-	88.48 $\pm$ 5.71	0.15 $\pm$ 0.11
0100-	88.65 $\pm$ 5.71	0.13 $\pm$ 0.13
1000-	88.83 $\pm$ 5.71	0.12 $\pm$ 0.11
0011-	88.76 $\pm$ 5.71	0.12 $\pm$ 0.11
0101-	88.45 $\pm$ 5.71	0.11 $\pm$ 0.11
0110-	88.32 $\pm$ 5.71	0.13 $\pm$ 0.12
1001-	89.02 $\pm$ 5.71	0.13 $\pm$ 0.12
1010-	89.26 $\pm$ 5.71	0.12 $\pm$ 0.11
1100-	88.97 $\pm$ 5.71	0.13 $\pm$ 0.13
0111-	88.46 $\pm$ 5.71	0.12 $\pm$ 0.11
1011-	89.07 $\pm$ 5.71	0.12 $\pm$ 0.11
1110-	88.87 $\pm$ 5.71	0.13 $\pm$ 0.12
1101-	88.88 $\pm$ 5.71	0.12 $\pm$ 0.11
1111-	88.83 $\pm$ 5.71	0.12 $\pm$ 0.11

**Tabela 5.3** – Identificação utilizada para as características.

0001 – Curtosis	1010 – Média, Skewness
0010 – Skewness	1100 – Média, Variância
0100 – Variância	0111 – Variância, Skewness, Curtosis
1000 – Média	1011 – Média, Skewness, Curtosis
0011 – Skewness, Curtosis	1110 – Média, Variância, Skewness
0101 – Variância, Curtosis	1101 – Média, Variância, Curtosis
0110 – Variância, Skewness	1111 – Média, Variância, Skewness, Curtosis
1001 – Média, Curtosis	



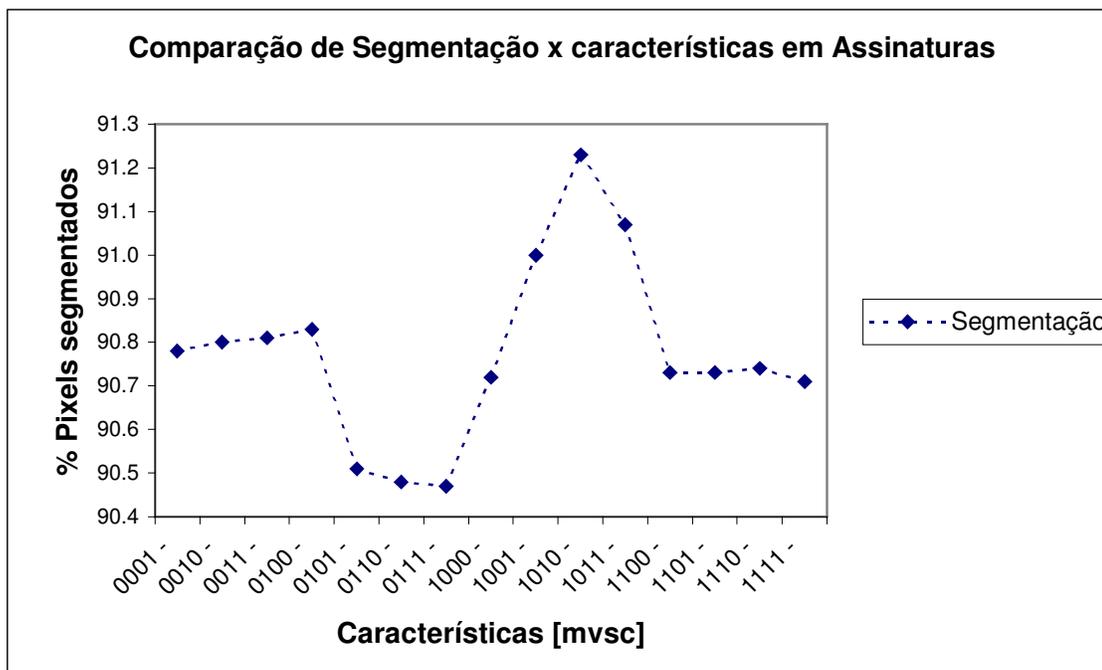
**Figura 5.2** – Modelo utilizado para a codificação das combinações de características.



**Figura 5.3** – Gráfico da amostra de segmentação por binarização de Assinaturas, para todas as combinações de características,  $k\text{-NN} = 1$ . Observa-se o maior pico para as características Média-Skewness e um segundo pico para Média-Skewness-Curtosis.

**Tabela 5.4** - Resultados da amostra de segmentação por binarização de assinaturas para todas as combinações de características e  $k\text{-NN} = 11$ .

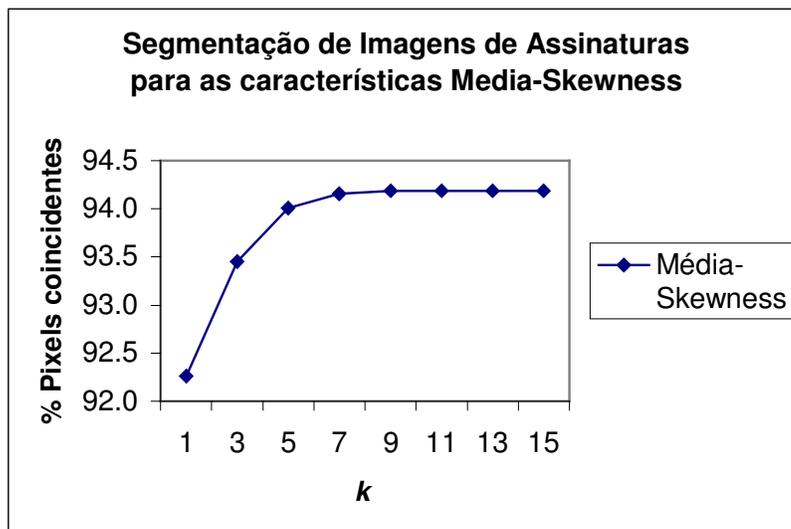
Caract	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
0001 -	90.78 $\pm$ 5.22	0.15 $\pm$ 0.11
0010 -	90.80 $\pm$ 5.16	0.16 $\pm$ 0.12
0011 -	90.81 $\pm$ 4.98	0.15 $\pm$ 0.11
0100 -	90.83 $\pm$ 4.83	0.14 $\pm$ 0.12
0101 -	90.51 $\pm$ 5.18	0.14 $\pm$ 0.12
0110 -	90.48 $\pm$ 5.10	0.15 $\pm$ 0.12
0111 -	90.47 $\pm$ 5.05	0.15 $\pm$ 0.12
1000 -	90.72 $\pm$ 5.54	0.14 $\pm$ 0.12
1001 -	91.00 $\pm$ 5.15	0.15 $\pm$ 0.12
1010 -	91.23 $\pm$ 4.95	0.15 $\pm$ 0.13
1011 -	91.07 $\pm$ 4.97	0.14 $\pm$ 0.12
1100 -	90.73 $\pm$ 5.08	0.15 $\pm$ 0.12
1101 -	90.73 $\pm$ 5.02	0.15 $\pm$ 0.12
1110 -	90.74 $\pm$ 4.97	0.15 $\pm$ 0.12
1111 -	90.71 $\pm$ 4.91	0.15 $\pm$ 0.12



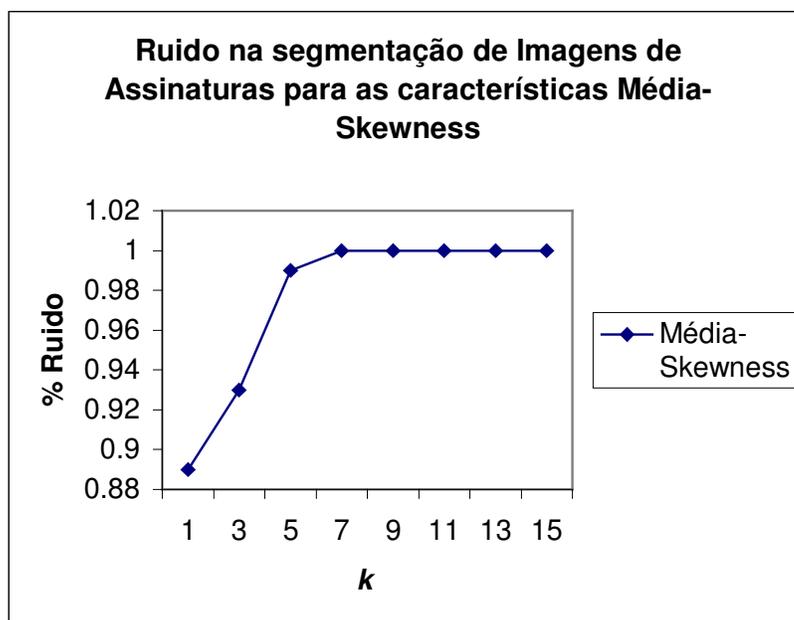
**Figura 5.4** – Gráfico da amostra de segmentação por binarização de Assinaturas, para todas as combinações de características,  $k\text{-NN} = 11$ . Observa-se um pico principal para as características Média-Skewness.

**Tabela 5.5** - Resultados da segmentação por binarização de Assinaturas, para as características Média – *Skewness* em função ao valor  $k\text{-NN}$ .

<b>Características: Média - Skewness</b>		
<b>k-NN</b>	<b>Segmentação (% <math>\mu \pm \sigma</math>)</b>	<b>Ruído (% <math>\mu \pm \sigma</math>)</b>
1	92.26 $\pm$ 7.49	0.89 $\pm$ 0.84
3	93.45 $\pm$ 6.93	0.93 $\pm$ 0.87
5	94.01 $\pm$ 6.75	0.99 $\pm$ 0.91
7	94.16 $\pm$ 6.66	1.00 $\pm$ 0.92
9	94.19 $\pm$ 6.65	1.00 $\pm$ 0.92
11	94.19 $\pm$ 6.64	1.00 $\pm$ 0.92
13	94.19 $\pm$ 6.64	1.00 $\pm$ 0.92
15	94.19 $\pm$ 6.64	1.00 $\pm$ 0.92

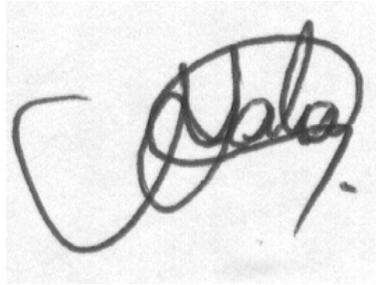


(a)



(b)

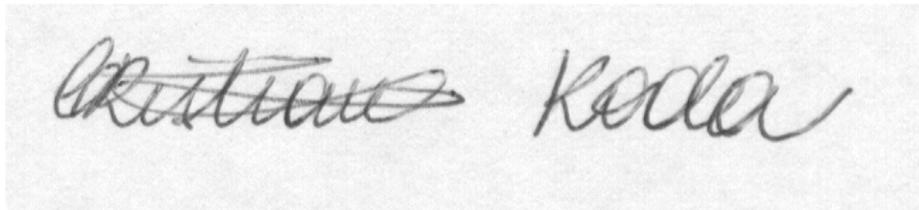
**Figura 5.5** – Gráfico da Segmentação de imagens de Assinaturas (a) e o correspondente Ruído (b), para as características Média-Skewness e o número de vizinhos mais próximos. Observa-se em ambos os gráficos que a taxa de segmentação e o ruído aumentam com os primeiros valores de  $k$ , e estabilizando-se com valores superiores de  $k$ .



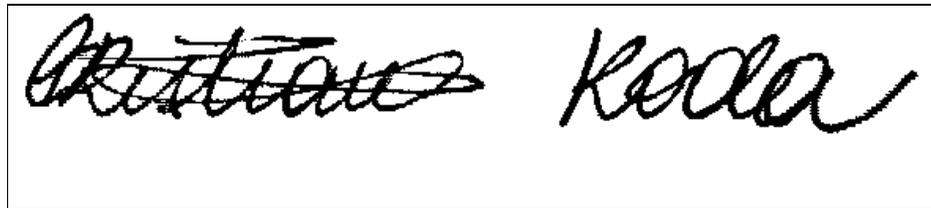
(a) Imagem amostra em níveis de cinza



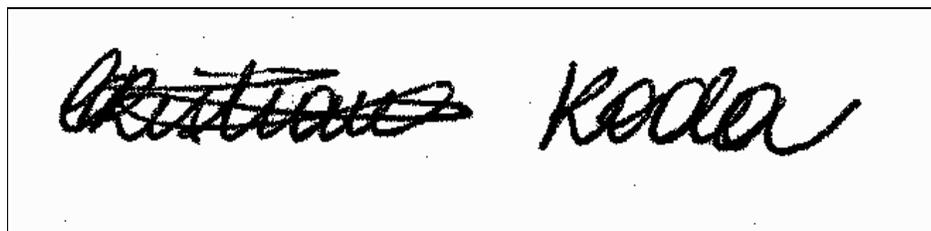
(b) Imagem ideal binarizada



(c) Nova imagem amostra em níveis de cinza

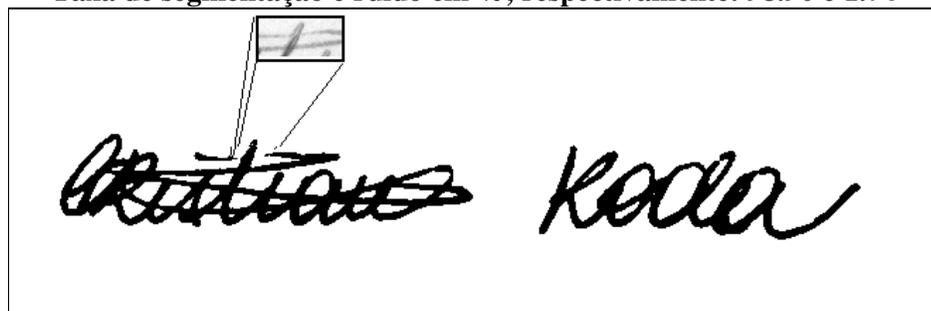


(d) Imagem ideal



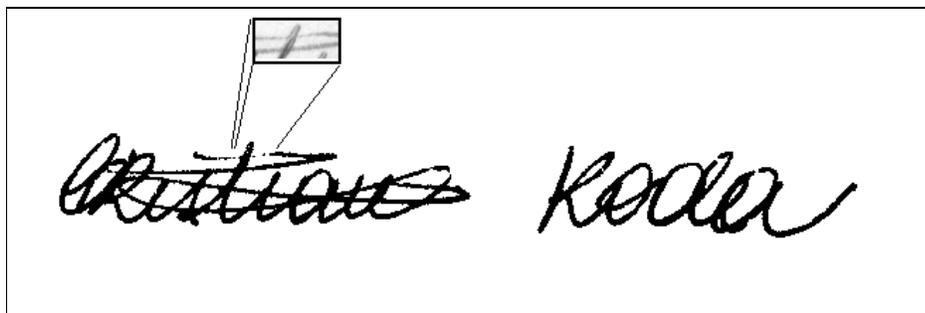
(e) Binarização obtida por aprendizagem

Taxa de segmentação e ruído em %, respectivamente: 98.90 e 1.74

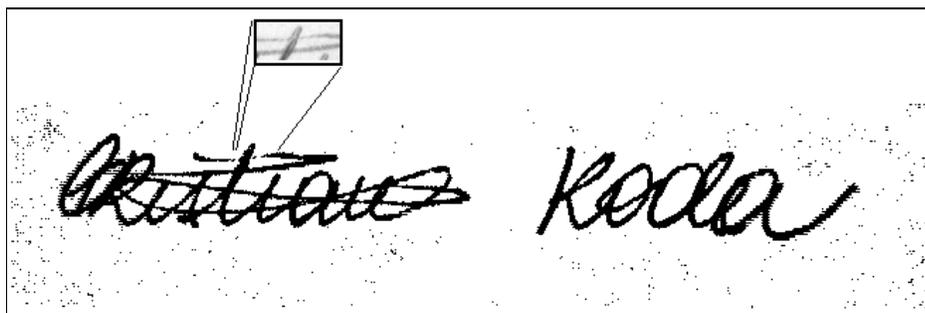


(f) Binarização obtida pelo método de Abutaleb

Taxa de segmentação e ruído em %, respectivamente: 91.62 e 1.34



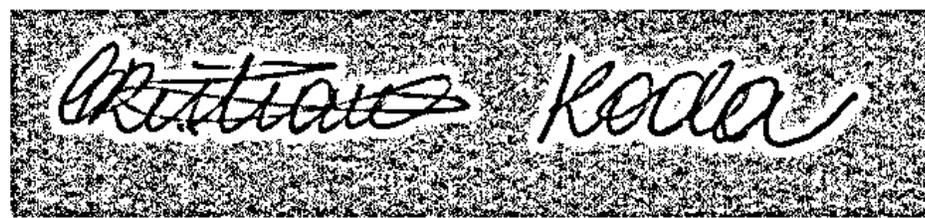
(g) Binarização obtida pelo método de Otsu  
Taxa de segmentação e ruído em %, respectivamente: 92.03 e 0.75



(h) Binarização obtida pelo método de Wellner  
Taxa de segmentação e ruído em %, respectivamente: 96.23 e 1.83



(i) Binarização obtida pelo método de Bernsen  
Taxa de segmentação e ruído em %, respectivamente: 72.95 e 13.55



(j) Binarização obtida pelo método de Niblack  
Taxa de segmentação e ruído em %, respectivamente: 87.68 e 33.64

**Figura 5.6.** Resultados comparativos de segmentação em imagem de assinatura.

A **Figura 5.6** apresenta um exemplo de resultado de aplicação da abordagem por aprendizagem e uma comparação de resultados obtidos por métodos tradicionais de

binarização globais e locais. Em (a) foi submetida a imagem amostra e em (b) a correspondente imagem ideal para aprendizagem. Em (c) está a nova imagem em níveis de cinza e em (d) a correspondente imagem ideal. Em (e) apresenta-se o resultado da segmentação por binarização da abordagem proposta. Em (f) e (g), os algoritmos de Abutaleb e Otsu apresentam falhas na segmentação da parte superior da letra “t”. Em (h), (i) e (j), os algoritmos de Wellner, Bernsen e Niblack, respectivamente, introduzem artefatos no fundo da imagem segmentada.

#### 5.4. Testes de Segmentação em Imagens de Impressões Digitais

Foram realizados os testes sobre base de impressões digitais detalhada previamente, sendo selecionado o subconjunto de teste de 10 imagens. A **Tabela 5.6** e a **Figura 5.7** mostram que, para o conjunto testado de imagens de impressões digitais, a melhor característica isolada é a Média (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $94.15 \pm 4.10$  e  $17.55 \pm 07.28$ ) para o parâmetro  $k = 15$ .

Tendo-se achado o melhor resultado para uma característica isolada e parâmetro  $k$ , foi realizada a seguir a segmentação de imagens, com todas as combinações de características, porém com  $k=1$ . Esses resultados são apresentados na **Tabela 5.7** e de forma gráfica, na **Figura 5.8**. De forma comparativa, também foi realizada a segmentação de imagens, com todas as combinações de características, porém com  $k=15$ . Esses resultados são apresentados na **Tabela 5.8** e de forma gráfica na **Figura 5.9**. Isso foi feito, devido a que a melhor segmentação isolada foi obtida com  $k=15$  para a característica Média. Observa-se que a combinação de características, que obteve os melhores resultados foi a a Média, identificada como 1010, seja para  $k=1$  (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $91.60 \pm 04.83$  e  $03.66 \pm 02.32$ ), como para  $k=15$  (%  $\mu \pm \sigma$ , respectivamente:  $93.61 \pm 4.40$  e  $9.46 \pm 4.26$ ). Finalmente, tendo-se identificado a melhor combinação de características, como sendo a característica isolada Média, o melhor resultado médio obtido foi taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $94.15 \pm 4.10$  e  $17.55 \pm 07.28$  para o parâmetro  $k = 15$ . Os resultados da segmentação de imagens para a característica Média e o parâmetro  $k$  no intervalo ímpar  $[1, 15]$  são apresentados na **Tabela 5.9** e de forma gráfica na **Figura 5.10**.

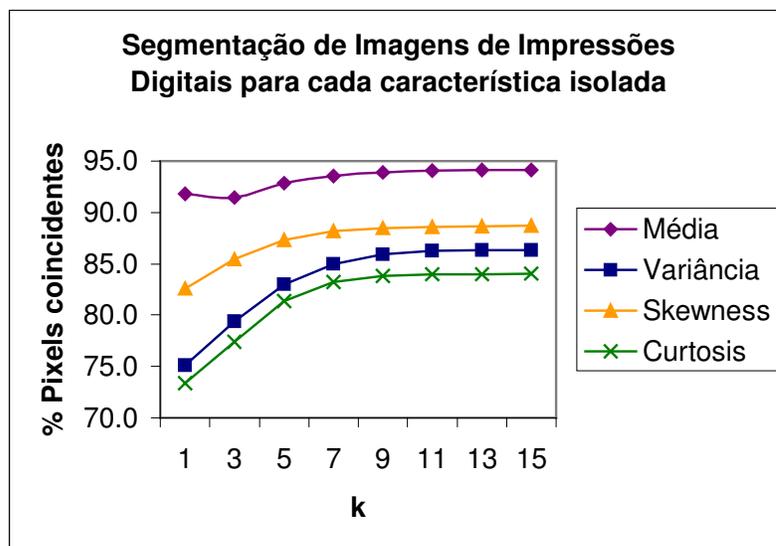
**Tabela 5.6** - Resultados da segmentação por binarização em impressões digitais para cada uma das características isoladas em função ao valor k-NN

Característica: Média		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	91.79 $\pm$ 5.75	20.31 $\pm$ 10.97
3	91.46 $\pm$ 4.27	10.16 $\pm$ 05.24
5	92.84 $\pm$ 4.09	12.08 $\pm$ 05.80
7	93.55 $\pm$ 4.02	14.02 $\pm$ 06.26
9	93.89 $\pm$ 4.05	15.57 $\pm$ 06.62
11	94.05 $\pm$ 4.06	16.56 $\pm$ 06.87
13	94.12 $\pm$ 4.08	17.21 $\pm$ 07.10
15	94.15 $\pm$ 4.10	17.55 $\pm$ 07.28

Característica: Variância		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	75.13 $\pm$ 18.97	20.01 $\pm$ 12.17
3	79.38 $\pm$ 17.38	21.15 $\pm$ 12.93
5	83.02 $\pm$ 15.69	22.88 $\pm$ 13.55
7	85.00 $\pm$ 14.80	24.01 $\pm$ 13.94
9	85.92 $\pm$ 14.46	24.56 $\pm$ 14.17
11	86.23 $\pm$ 14.43	24.75 $\pm$ 14.32
13	86.32 $\pm$ 14.45	24.87 $\pm$ 14.38
15	86.32 $\pm$ 14.46	24.89 $\pm$ 14.40

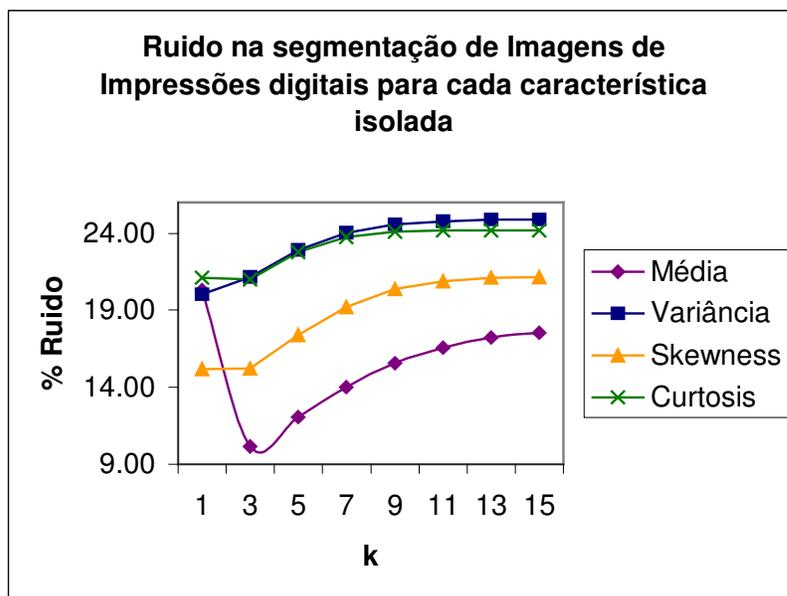
Característica: Skewness		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	82.58 $\pm$ 11.77	15.18 $\pm$ 09.82
3	85.47 $\pm$ 11.73	15.21 $\pm$ 11.03
5	87.33 $\pm$ 11.82	17.39 $\pm$ 12.28
7	88.16 $\pm$ 12.03	19.22 $\pm$ 13.23
9	88.48 $\pm$ 12.16	20.37 $\pm$ 13.82
11	88.60 $\pm$ 12.20	20.89 $\pm$ 14.06
13	88.64 $\pm$ 12.21	21.10 $\pm$ 14.14
15	88.70 $\pm$ 12.35	21.13 $\pm$ 14.18

Característica: Curtosis		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	73.37 $\pm$ 19.02	21.10 $\pm$ 12.55
3	77.42 $\pm$ 10.05	20.98 $\pm$ 13.89
5	81.36 $\pm$ 17.84	22.76 $\pm$ 14.83
7	83.20 $\pm$ 17.21	23.73 $\pm$ 15.33
9	83.81 $\pm$ 17.01	24.08 $\pm$ 15.51
11	83.97 $\pm$ 16.96	24.17 $\pm$ 15.56
13	84.01 $\pm$ 16.95	24.18 $\pm$ 15.57
15	84.05 $\pm$ 16.96	24.19 $\pm$ 15.59



(a)

**Figura 5.7** – Gráfico da Segmentação de imagens de Impressões digitais (a) e o correspondente Ruído (b) para cada característica isolada e o número de vizinhos mais próximos

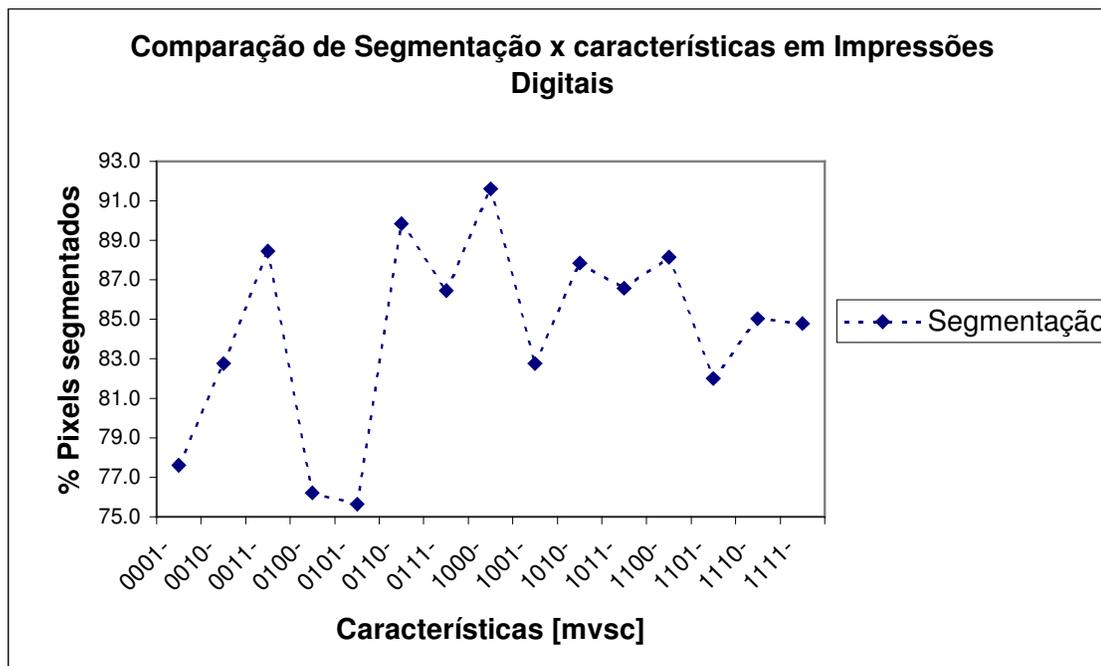


(b)

**Figura 5.7 (cont.)** – Gráfico da Segmentação de imagens de Impressões digitais (a) e o correspondente Ruído (b) para cada característica isolada e o número de vizinhos mais próximos. Observa-se em (a) que as taxas de segmentação aumentam com os primeiros valores de  $k$ , estabilizando-se com valores superiores de  $k$ . Em (b) também existe uma tendência similar em relação ao ruído, com exceção da característica Média, que apresenta uma grande diminuição no ruído para  $k=3$ .

**Tabela 5.7** - Resultados da amostra de segmentação por binarização de impressões digitais para todas as combinações de características e  $k$ -NN = 1

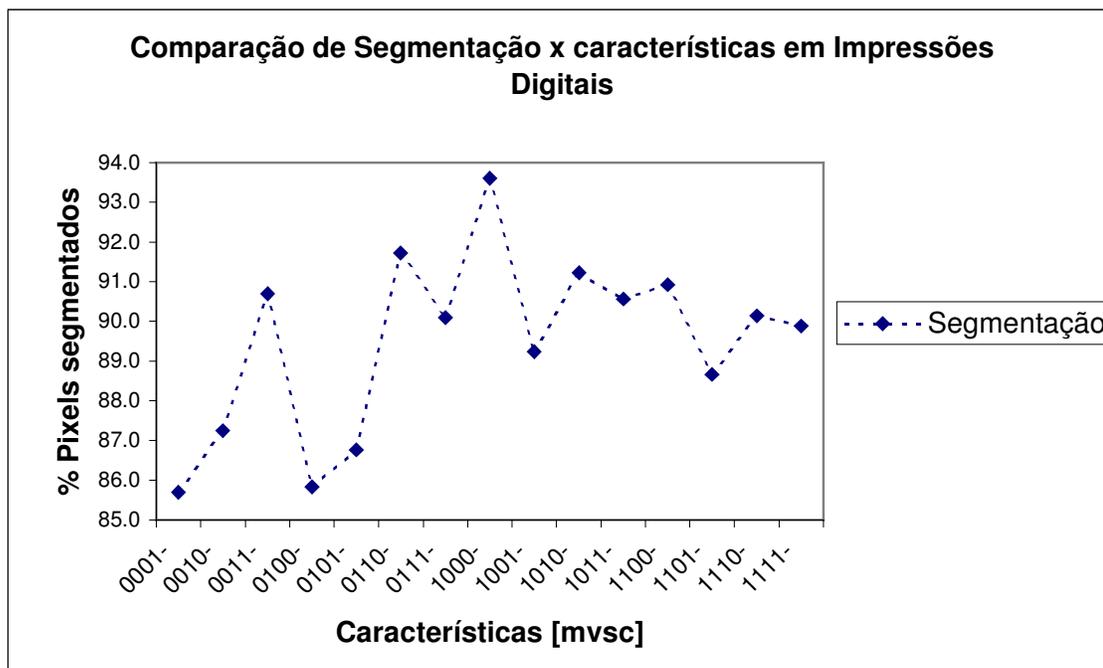
Caract	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
0001-	77.61 $\pm$ 11.49	17.43 $\pm$ 10.44
0010-	82.77 $\pm$ 11.62	22.01 $\pm$ 08.65
0011-	88.46 $\pm$ 05.34	19.67 $\pm$ 11.05
0100-	76.21 $\pm$ 10.16	14.16 $\pm$ 07.29
0101-	75.63 $\pm$ 11.58	16.90 $\pm$ 07.71
0110-	89.85 $\pm$ 04.87	20.39 $\pm$ 11.75
0111-	86.44 $\pm$ 07.21	16.23 $\pm$ 09.26
1000-	91.60 $\pm$ 04.83	03.66 $\pm$ 02.32
1001-	82.77 $\pm$ 06.62	11.15 $\pm$ 05.77
1010-	87.85 $\pm$ 06.08	11.56 $\pm$ 05.05
1011-	86.59 $\pm$ 07.09	13.49 $\pm$ 07.52
1100-	88.15 $\pm$ 05.47	06.47 $\pm$ 02.88
1101-	82.01 $\pm$ 07.32	12.25 $\pm$ 06.39
1110-	85.04 $\pm$ 07.37	12.00 $\pm$ 06.24
1111-	84.80 $\pm$ 07.59	12.82 $\pm$ 07.34



**Figura 5.8** – Gráfico da amostra de segmentação por binarização de Impressões digitais para todas as combinações de características,  $k$ -NN = 1. Observa-se o maior pico para a característica Média e um segundo pico para as características Variância-Skewness.

**Tabela 5.8** - Resultados da amostra de segmentação por binarização de impressões digitais para todas as combinações de características e  $k$ -NN = 15

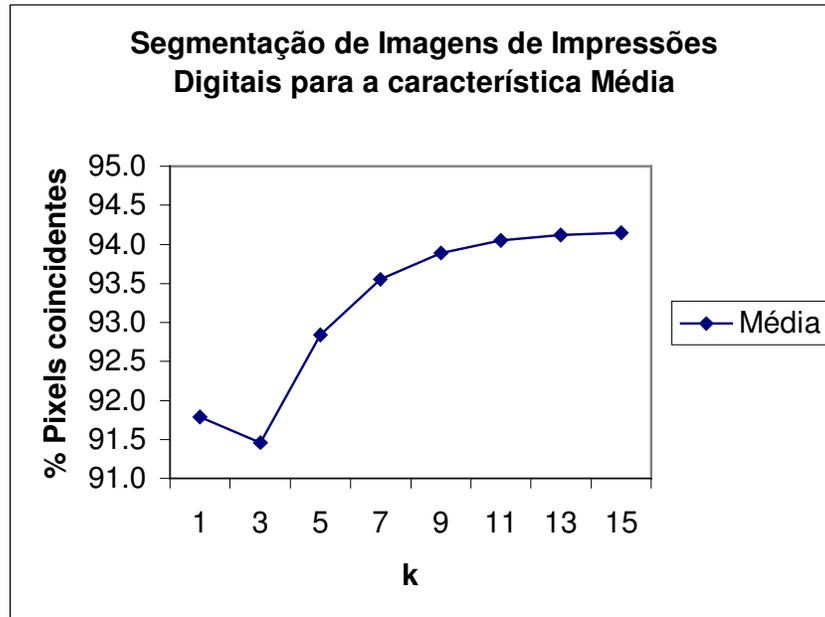
Caract	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
0001-	85.69 $\pm$ 6.66	24.38 $\pm$ 8.73
0010-	87.25 $\pm$ 6.01	23.17 $\pm$ 5.93
0011-	90.70 $\pm$ 4.10	23.66 $\pm$ 7.53
0100-	85.83 $\pm$ 6.08	22.58 $\pm$ 7.34
0101-	86.77 $\pm$ 4.14	24.81 $\pm$ 5.93
0110-	91.72 $\pm$ 3.63	25.40 $\pm$ 7.08
0111-	90.10 $\pm$ 3.64	22.11 $\pm$ 5.63
1000-	93.61 $\pm$ 4.40	09.46 $\pm$ 4.26
1001-	89.23 $\pm$ 3.95	20.64 $\pm$ 6.90
1010-	91.22 $\pm$ 4.24	19.77 $\pm$ 5.72
1011-	90.56 $\pm$ 4.02	22.11 $\pm$ 6.41
1100-	90.92 $\pm$ 3.87	17.03 $\pm$ 4.60
1101-	88.66 $\pm$ 4.04	21.87 $\pm$ 6.98
1110-	90.14 $\pm$ 3.79	20.81 $\pm$ 6.17
1111-	89.88 $\pm$ 3.61	21.64 $\pm$ 5.85



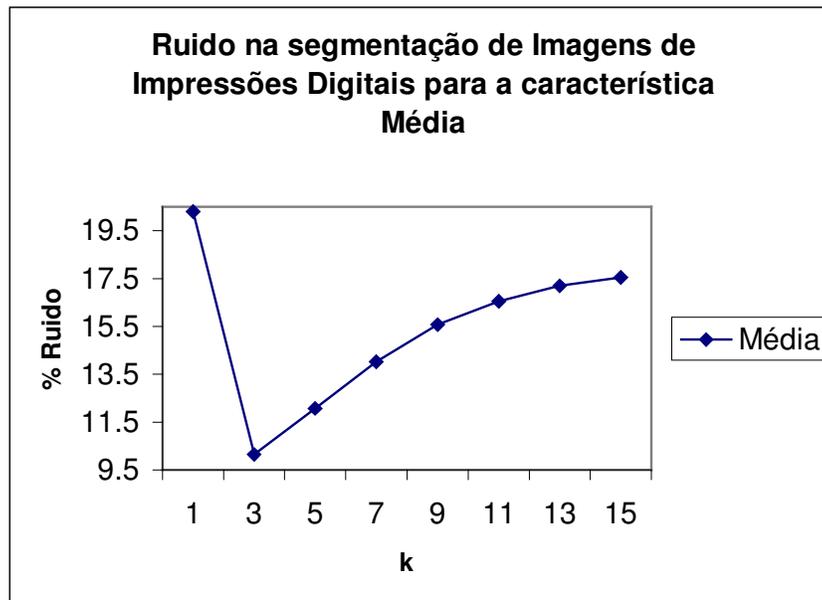
**Figura 5.9** – Gráfico da amostra de segmentação por binarização de Impressões digitais para todas as combinações de características, k-NN = 15. Observa-se o maior pico para a característica Média e um segundo pico para as características Variância-Skewness.

**Tabela 5.9** - Resultados da segmentação por binarização de Impressões digitais para a característica Média em função ao valor k-NN.

Característica: Média		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	91.79 $\pm$ 5.75	20.31 $\pm$ 10.97
3	91.46 $\pm$ 4.27	10.16 $\pm$ 05.24
5	92.84 $\pm$ 4.09	12.08 $\pm$ 05.80
7	93.55 $\pm$ 4.02	14.02 $\pm$ 06.26
9	93.89 $\pm$ 4.05	15.57 $\pm$ 06.62
11	94.05 $\pm$ 4.06	16.56 $\pm$ 06.87
13	94.12 $\pm$ 4.08	17.21 $\pm$ 07.10
15	94.15 $\pm$ 4.10	17.55 $\pm$ 07.28



(a)



(b)

**Figura 5.10** – Gráfico da Segmentação de imagens de Impressões digitais (a) e o correspondente Ruído (b), para a característica Média e o número de vizinhos mais próximos. Observa-se em ambos os gráficos que a taxa de segmentação e o ruído diminuem até o valor de  $k=3$  e logo aumentam com o valor de  $k$ .



(a) Amostra em nível de cinza



(b) Imagem ideal binarizada



(c) Nova imagem amostra



(d) Binarização obtida pelo algoritmo de aprendizagem

Taxa de segmentação e ruído em %,  
respectivamente: 99.74 e 31.33



(e) Binarização obtida pelo algoritmo de Otsu

Taxa de segmentação e ruído em %,  
respectivamente: 93.11 e 18.29



(f) Binarização obtida pelo algoritmo de Li and Lee

Taxa de segmentação e ruído em %,  
respectivamente: 91.03 e 15.44

**Figura 5.11.** Resultados comparativos de segmentação em imagem de impressão digital.



**Figura 5.11 (cont.).** Resultados comparativos de segmentação em imagem de impressão digital.

A **Figura 5.11** apresenta um exemplo de resultado de aplicação da abordagem por aprendizagem e uma comparação de resultados obtidos, por métodos tradicionais de binarização, globais e locais. Em (a), é apresentada a imagem amostra submetida e em (b), a correspondente imagem ideal para aprendizagem. Em (c), está a nova imagem em níveis de cinza. Em (d) apresenta-se o resultado da segmentação por binarização da abordagem proposta. Em (e) e (f), os algoritmos de Abutaleb e Otsu apresentam manchas na segmentação da parte central superior, unindo vários sulcos. Em (g), e (h), os algoritmos de Bernsen e Eikvil, introduzem artefatos principalmente no fundo da imagem segmentada.

## 5.5. Testes de Segmentação por Multi-região em Imagens de Ressonância Magnética

No caso das imagens ressonância magnética de cérebros, o objetivo foi testar a abordagem proposta, para segmentar as imagens cerebrais em três classes: massa cinza (GM), massa branca (WM) e líquido (CSF).

### 5.5.1. Testes de Segmentação da Massa Branca de Imagens de Ressonância Magnética

Foram realizados os testes de segmentação da massa branca (WM), sobre a base de MRI detalhada no **Capítulo 4**, sendo selecionado o subconjunto de teste de 10 imagens. A **Tabela 5.10** e a **Figura 5.12** mostram que, para o conjunto das imagens de ressonância magnética testado, a melhor característica isolada é a Variância (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $95.61 \pm 1.21$  e  $0.55 \pm 0.12$ ) para o parâmetro  $k=1$ .

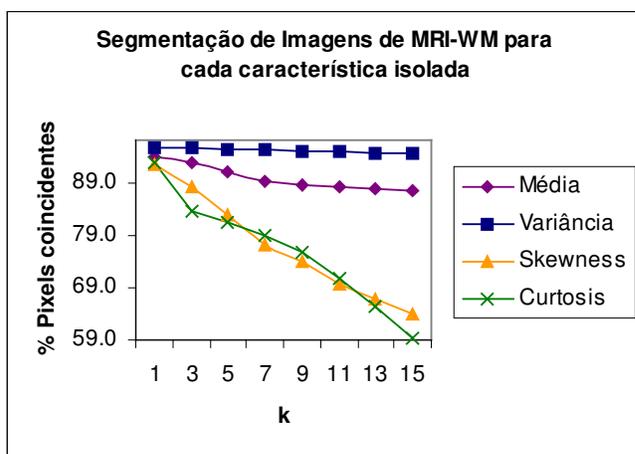
Tendo-se achado o melhor resultado para uma característica isolada e parâmetro  $k$ , foi realizada a seguir a segmentação de imagens, com todas as combinações de características, porém com  $k=1$ . Esses resultados são apresentados na **Tabela 5.11** e de forma gráfica, na **Figura 5.13**. De forma comparativa, também foi realizada a segmentação de imagens, com todas as combinações de características, porém com  $k=3$ . Esses resultados são apresentados na **Tabela 5.12** e de forma gráfica na **Figura 5.14**. Isso foi feito, devido a que a melhor segmentação isolada foi obtida com  $k=3$ , para a característica Variância. Observa-se que a combinação de características, que obteve os melhores resultados foi Média-Variância, identificada como 1100, seja para  $k=1$  (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $97.25 \pm 1.04$  e  $0.46 \pm 0.16$ ), como para  $k=3$  (%  $\mu \pm \sigma$ , respectivamente:  $96.95 \pm 1.03$  e  $0.46 \pm 0.09$ ). Finalmente, tendo-se identificada a melhor combinação de características, foram feitos experimentos variando-se os valores do parâmetro  $k$ , no intervalo ímpar [1, 15]. Verificou-se que os melhores resultados foram obtidos com o valor  $k=1$  (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $96.00 \pm 1.03$  e  $0.53 \pm 0.10$ ). Os resultados da segmentação de imagens, para as características Média-Variância, e o parâmetro  $k$ , no intervalo ímpar [1, 15] são apresentados na **Tabela 5.13** e de forma gráfica na **Figura 5.15**.

**Tabela 5.10** - Resultados da segmentação por multi-região em MRI-WM, para cada uma das características isoladas, em função ao valor k-NN.

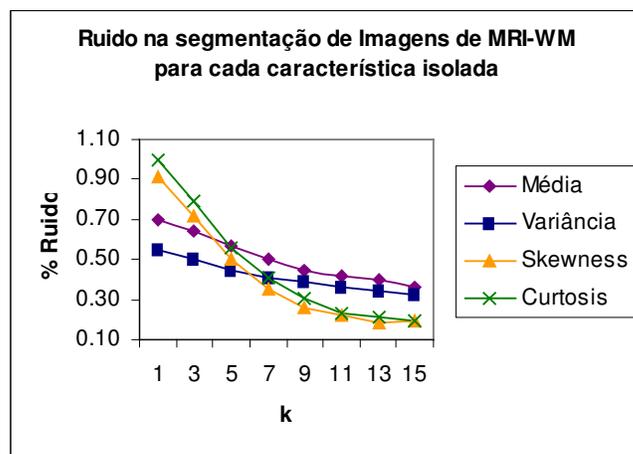
Característica: Média			Característica: Variância		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )	k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	$93.75 \pm 2.59$	$0.70 \pm 0.18$	1	$95.61 \pm 1.21$	$0.55 \pm 0.12$
3	$92.59 \pm 3.87$	$0.64 \pm 0.16$	3	$95.61 \pm 1.12$	$0.50 \pm 0.12$
5	$90.91 \pm 5.13$	$0.57 \pm 0.16$	5	$95.36 \pm 1.25$	$0.45 \pm 0.13$
7	$89.22 \pm 6.10$	$0.50 \pm 0.18$	7	$95.11 \pm 1.47$	$0.41 \pm 0.16$
9	$88.58 \pm 6.35$	$0.45 \pm 0.21$	9	$94.93 \pm 1.68$	$0.39 \pm 0.19$
11	$88.09 \pm 6.52$	$0.42 \pm 0.25$	11	$94.73 \pm 1.92$	$0.36 \pm 0.22$
13	$87.59 \pm 6.80$	$0.40 \pm 0.28$	13	$94.59 \pm 2.18$	$0.34 \pm 0.25$
15	$87.30 \pm 6.67$	$0.36 \pm 0.30$	15	$94.46 \pm 2.21$	$0.32 \pm 0.26$

**Tabela 5.10 (cont.)**- Resultados da segmentação por multi-região em MRI-WM, para cada uma das características isoladas, em função ao valor k-NN.

Característica: <i>Skewness</i>			Característica: <i>Curtosis</i>		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )	k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	92.51 $\pm$ 2.12	0.91 $\pm$ 0.18	1	92.85 $\pm$ 1.55	1.00 $\pm$ 0.17
3	88.14 $\pm$ 6.02	0.72 $\pm$ 0.15	3	83.66 $\pm$ 7.59	0.79 $\pm$ 0.16
5	82.76 $\pm$ 7.30	0.50 $\pm$ 0.13	5	81.42 $\pm$ 7.66	0.56 $\pm$ 0.18
7	76.97 $\pm$ 6.48	0.35 $\pm$ 0.12	7	78.81 $\pm$ 7.57	0.41 $\pm$ 0.18
9	73.89 $\pm$ 6.25	0.26 $\pm$ 0.12	9	75.56 $\pm$ 7.27	0.31 $\pm$ 0.18
11	69.60 $\pm$ 5.45	0.22 $\pm$ 0.12	11	70.59 $\pm$ 6.69	0.23 $\pm$ 0.15
13	66.76 $\pm$ 4.93	0.18 $\pm$ 0.11	13	65.29 $\pm$ 7.11	0.21 $\pm$ 0.16
15	63.95 $\pm$ 5.19	0.19 $\pm$ 0.11	15	59.25 $\pm$ 7.92	0.19 $\pm$ 0.15



(a)

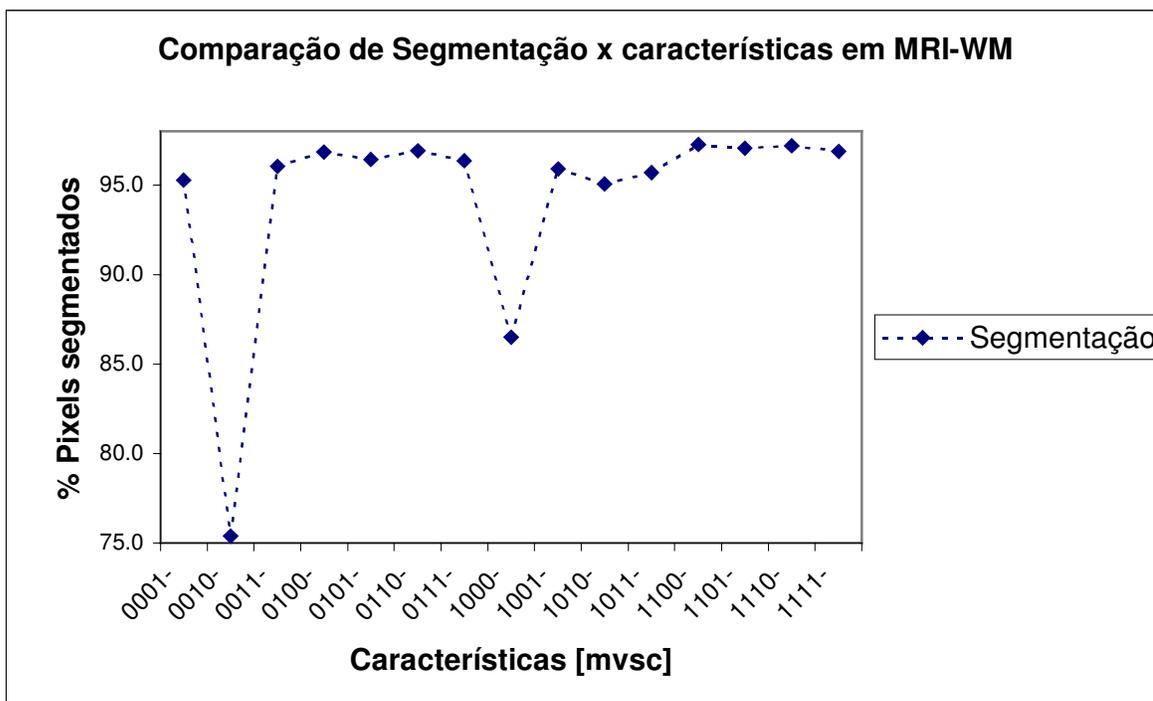


(b)

**Figura 5.12** – Gráfico da Segmentação de imagens de MRI-WM (a) e o correspondente Ruído (b), para cada característica isolada e o número de vizinhos mais próximos. Em (a) observa-se que as taxas de segmentação para as características Média e Variância, diminuem muito pouco com o aumento de k, enquanto que as taxas de segmentação diminuem drasticamente com o valor de k. Em (b) observa-se que as taxas de ruído, diminuem com o valor de k.

**Tabela 5.11** - Resultados da amostra de segmentação por multi-reigão de MRI-WM, para todas as combinações de características e  $k\text{-NN} = 1$ .

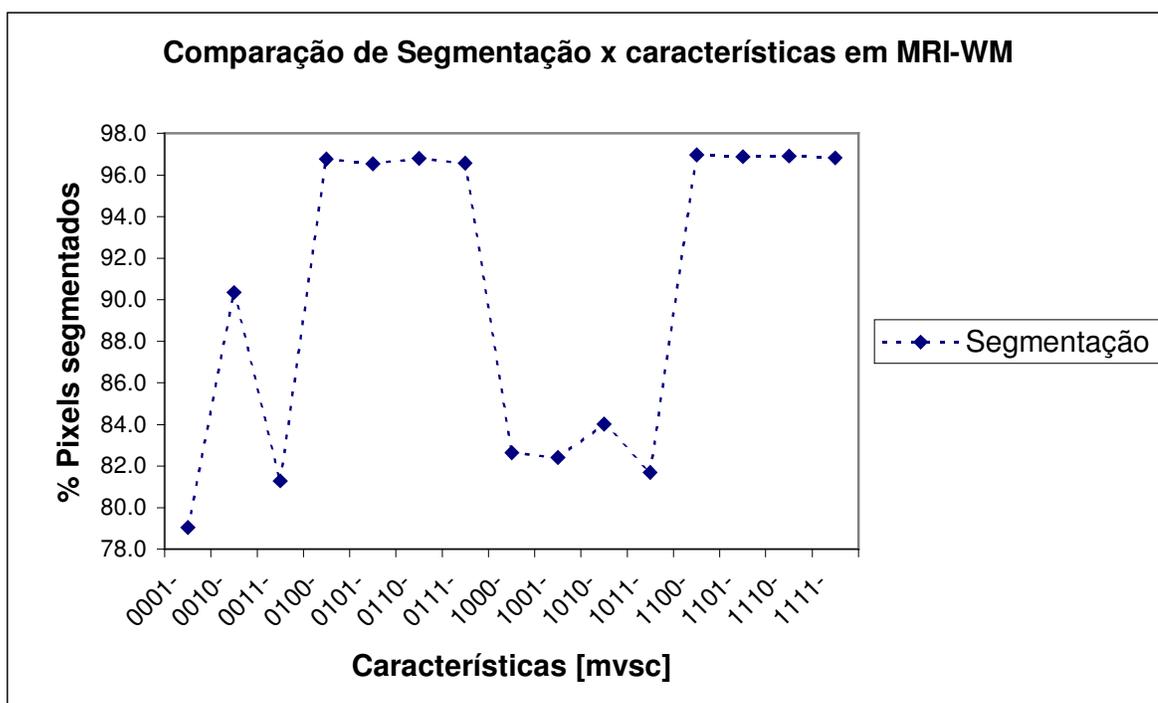
Caract	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
0001	95.29±1.65	0.90±0.32
0010	75.39±3.41	0.42±0.08
0011	96.06±0.80	0.69±0.15
0100	96.86±1.23	0.49±0.17
0101	96.41±1.83	0.46±0.16
0110	96.93±1.17	0.47±0.11
0111	96.37±1.57	0.46±0.11
1000	86.50±7.21	0.63±0.24
1001	95.90±2.49	0.46±0.16
1010	95.06±3.10	0.49±0.13
1011	95.70±2.53	0.45±0.11
1100	97.25±1.04	0.46±0.16
1101	97.07±1.24	0.45±0.15
1110	97.20±0.91	0.46±0.13
1111	96.87±1.17	0.45±0.11



**Figura 5.13** – Gráfico da amostra de segmentação por multi-região de MRI-WM, para todas as combinações de características,  $k\text{-NN} = 1$ . Observam-se vários picos próximos, sendo o maior para as características Média-Variância.

**Tabela 5.12** - Resultados da amostra de segmentação por multi-reigão de MRI-WM, para todas as combinações de características e  $k\text{-NN} = 3$ .

Caract	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
0001	79.03 $\pm$ 7.33	0.75 $\pm$ 0.15
0010	90.34 $\pm$ 7.67	0.70 $\pm$ 0.15
0011	81.27 $\pm$ 8.21	0.53 $\pm$ 0.11
0100	96.77 $\pm$ 0.77	0.46 $\pm$ 0.09
0101	96.54 $\pm$ 0.93	0.43 $\pm$ 0.07
0110	96.80 $\pm$ 1.11	0.47 $\pm$ 0.09
0111	96.56 $\pm$ 1.18	0.45 $\pm$ 0.08
1000	82.64 $\pm$ 5.91	0.61 $\pm$ 0.15
1001	82.39 $\pm$ 8.18	0.47 $\pm$ 0.09
1010	84.01 $\pm$ 6.44	0.47 $\pm$ 0.10
1011	81.68 $\pm$ 7.70	0.45 $\pm$ 0.09
1100	96.95 $\pm$ 1.03	0.46 $\pm$ 0.09
1101	96.89 $\pm$ 0.94	0.45 $\pm$ 0.08
1110	96.90 $\pm$ 1.20	0.46 $\pm$ 0.10
1111	96.81 $\pm$ 1.10	0.44 $\pm$ 0.09

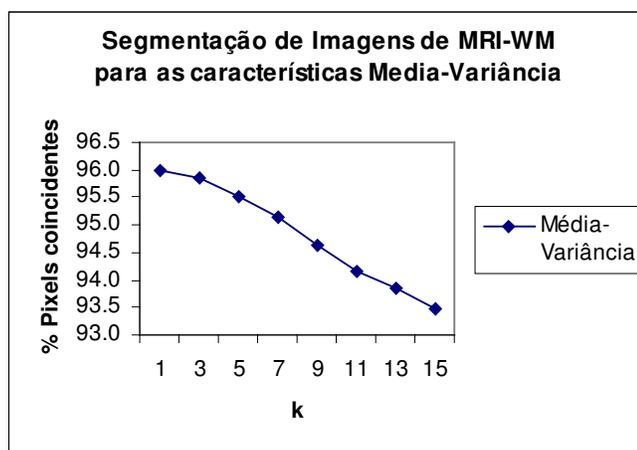


**Figura 5.14** – Gráfico da amostra de segmentação por multi-região de MRI-WM, para todas as combinações de características,  $k\text{-NN} = 3$ . Observam-se vários picos próximos, sendo o maior para as características Média-Variância.

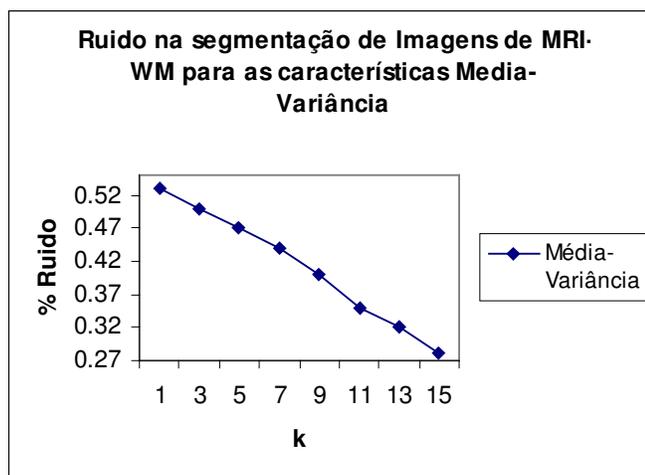
**Tabela 5.13** - Resultados da segmentação por multi-região de MRI-WM, para as características Média-Variância em função ao valor k-NN.

Características: Média-Variância		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	96.00 $\pm$ 1.03	0.53 $\pm$ 0.10
3	95.84 $\pm$ 1.26	0.50 $\pm$ 0.11
5	95.53 $\pm$ 1.58	0.47 $\pm$ 0.13
7	95.13 $\pm$ 1.78	0.44 $\pm$ 0.16
9	94.63 $\pm$ 2.02	0.40 $\pm$ 0.18
11	94.17 $\pm$ 2.34	0.35 $\pm$ 0.20
13	93.84 $\pm$ 2.61	0.32 $\pm$ 0.23
15	93.46 $\pm$ 2.88	0.28 $\pm$ 0.25

(a)



(b)



**Figura 5.15** – Gráfico da Segmentação por multi-região de MRI-WM (a) e o correspondente Ruído (b), para as características Média-Variância e o número de vizinhos mais próximos. Observam-se em ambos os gráficos, que a taxa de segmentação e o ruído diminuem com o valor do k.

### 5.5.2. Testes de Segmentação da Massa Cinza de Imagens de Ressonância Magnética

Foram realizados os testes de segmentação da massa cinza (GM), sobre a base de MRI detalhada no **Capítulo 4**, sendo selecionado o subconjunto de teste de 10 imagens. A **Tabela 5.14** e a **Figura 5.16** mostram que, para o conjunto das imagens de ressonância magnética testado, a melhor característica isolada é a Variância (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $97.00 \pm 2.07$  e  $1.74 \pm 1.45$ ) para o parâmetro  $k = 11$ .

Tendo-se achado o melhor resultado para uma característica isolada e parâmetro  $k$ , foi realizada a seguir a segmentação de imagens, com todas as combinações de características, porém com  $k=1$ . Esses resultados são apresentados na **Tabela 5.15** e de forma gráfica, na **Figura 5.17**.

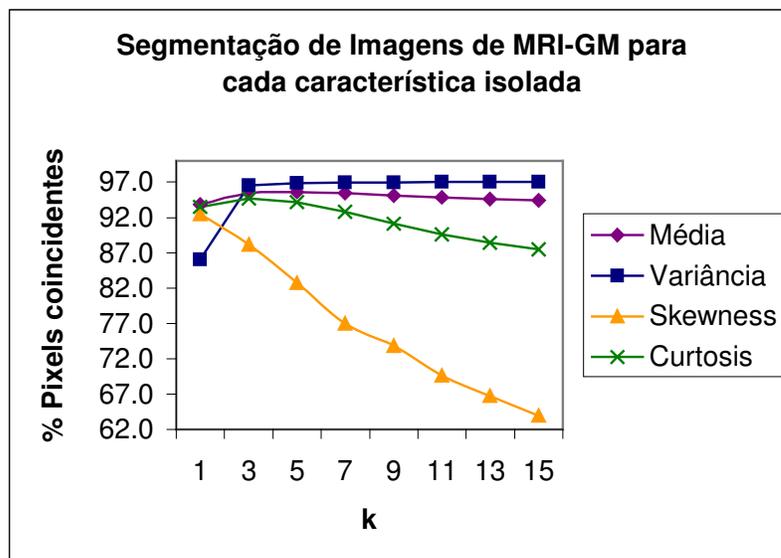
De forma comparativa, também foi realizada a segmentação de imagens, com todas as combinações de características, porém com  $k=11$ . Esses resultados são apresentados na **Tabela 5.16** e de forma gráfica na **Figura 5.18**. Isso foi feito, devido a que a melhor segmentação isolada foi obtida com  $k=11$ , para a característica Variância. Observa-se que a combinação de características, que obteve os melhores resultados foi Média-Variância-Curtosis, identificada como 1101, para  $k=1$  (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $95.83 \pm 1.44$  e  $0.62 \pm 0.24$ ), porém, para  $k=11$  foi Média-Variância-Skewness, identificada como 1110 (%  $\mu \pm \sigma$ , respectivamente:  $97.70 \pm 1.31$  e  $0.57 \pm 0.02$ ). Finalmente, tendo-se identificado a melhor combinação de características, foram feitos experimentos variando-se os valores do parâmetro  $k$ , no intervalo ímpar [1, 15]. Verificou-se que os melhores resultados foram obtidos com o valor  $k=11$  (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $97.33 \pm 1.49$  e  $0.67 \pm 0.09$ ). Os resultados da segmentação de imagens para as características Média-Variância-Skewness e o parâmetro  $k$ , no intervalo ímpar [1, 15] são apresentados na **Tabela 5.17** e de forma gráfica, na **Figura 5.19**.

**Tabela 5.14** - Resultados da segmentação por multi-região de MRI-GM, para cada uma das características isoladas, em função ao valor k-NN.

Característica: Média			Característica: Variância		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )	k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	93.80 $\pm$ 1.12	0.83 $\pm$ 0.13	1	86.04 $\pm$ 27.21	1.13 $\pm$ 1.31
3	95.40 $\pm$ 1.31	0.74 $\pm$ 0.10	3	96.50 $\pm$ 01.36	1.78 $\pm$ 1.44
5	95.64 $\pm$ 1.86	0.71 $\pm$ 0.12	5	96.87 $\pm$ 01.46	1.76 $\pm$ 1.44
7	95.45 $\pm$ 2.24	0.68 $\pm$ 0.12	7	96.95 $\pm$ 01.84	1.75 $\pm$ 1.45
9	95.12 $\pm$ 2.47	0.64 $\pm$ 0.11	9	96.98 $\pm$ 02.02	1.74 $\pm$ 1.45
11	94.82 $\pm$ 2.53	0.62 $\pm$ 0.11	11	97.00 $\pm$ 02.07	1.74 $\pm$ 1.45
13	94.55 $\pm$ 2.57	0.60 $\pm$ 0.11	13	97.00 $\pm$ 02.07	1.74 $\pm$ 1.45
15	94.39 $\pm$ 2.60	0.60 $\pm$ 0.11	15	97.00 $\pm$ 02.07	1.74 $\pm$ 1.45

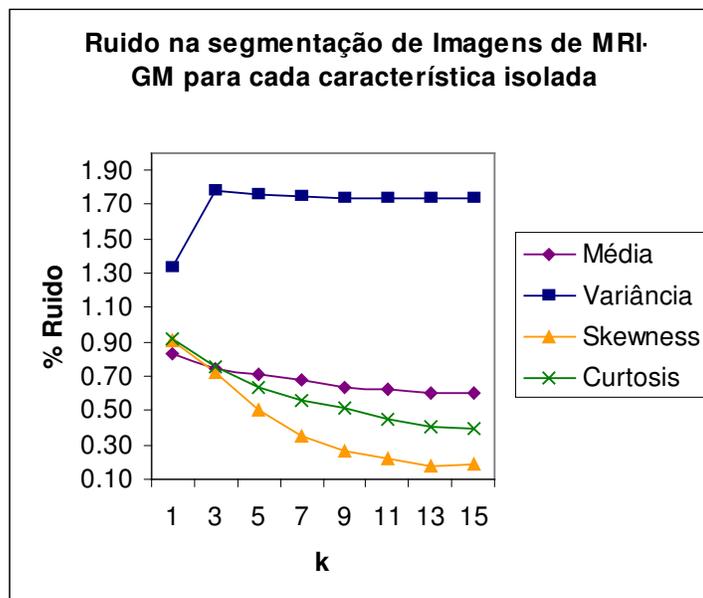
  

Característica: Skewness			Característica: Curtosis		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )	k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	92.98 $\pm$ 0.72	1.02 $\pm$ 0.15	1	93.47 $\pm$ 0.80	0.92 $\pm$ 0.14
3	93.35 $\pm$ 0.93	0.81 $\pm$ 0.10	3	94.65 $\pm$ 1.16	0.75 $\pm$ 0.10
5	89.55 $\pm$ 2.28	0.52 $\pm$ 0.08	5	94.17 $\pm$ 1.77	0.64 $\pm$ 0.09
7	87.95 $\pm$ 2.89	0.45 $\pm$ 0.08	7	92.81 $\pm$ 2.85	0.56 $\pm$ 0.08
9	91.62 $\pm$ 1.48	0.63 $\pm$ 0.08	9	91.14 $\pm$ 4.12	0.51 $\pm$ 0.08
11	87.04 $\pm$ 2.91	0.41 $\pm$ 0.08	11	89.56 $\pm$ 4.98	0.45 $\pm$ 0.10
13	86.51 $\pm$ 2.83	0.39 $\pm$ 0.08	13	88.40 $\pm$ 5.28	0.41 $\pm$ 0.11
15	86.23 $\pm$ 2.78	0.38 $\pm$ 0.08	15	87.50 $\pm$ 5.48	0.39 $\pm$ 0.11



(a)

**Figura 5.16** – Gráfico da Segmentação por multi-região de MRI-GM (a) e o correspondente Ruído (b), para cada característica isolada e o número de vizinhos mais próximos.

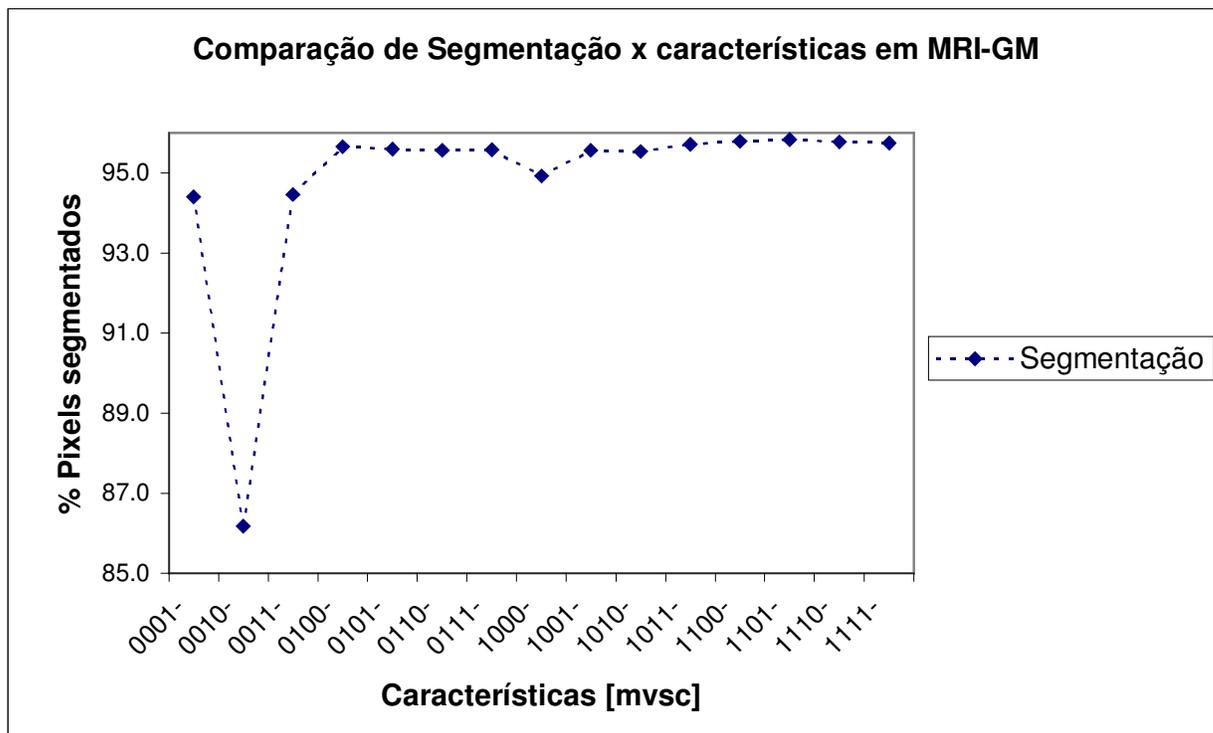


(b)

**Figura 5.16 (cont.)**– Gráfico da Segmentação por multi-região de MRI-GM (a) e o correspondente Ruído (b), para cada característica isolada e o número de vizinhos mais próximos. Observa-se em (a) que as taxas de segmentação com as características Média e Variância permanecem aproximadamente estáveis com o aumento de k. No entanto, as taxas para Skewness e Curtosis diminuem com o valor do k. Em (b) com exceção da característica Variância, em todos os outros casos, o ruído diminui com o valor de k.

**Tabela 5.15** - Resultados da amostra de segmentação por multi-região de MRI-GM, para todas as combinações de características e  $k$ -NN = 1.

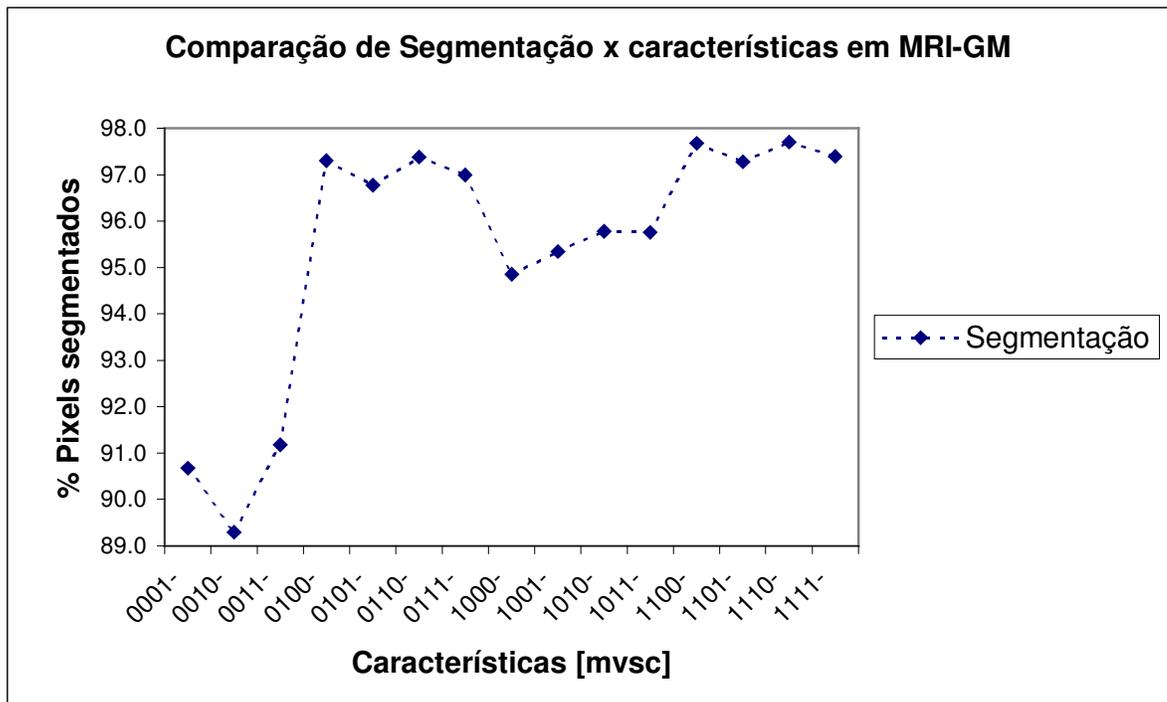
Caract	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
0001	94.40 $\pm$ 1.99	0.72 $\pm$ 0.27
0010	86.18 $\pm$ 4.94	0.50 $\pm$ 0.10
0011	94.47 $\pm$ 1.07	0.75 $\pm$ 0.15
0100	95.66 $\pm$ 1.48	0.62 $\pm$ 0.23
0101	95.59 $\pm$ 1.52	0.80 $\pm$ 0.62
0110	95.56 $\pm$ 0.91	0.64 $\pm$ 0.15
0111	95.58 $\pm$ 0.99	0.64 $\pm$ 0.16
1000	94.92 $\pm$ 1.82	0.66 $\pm$ 0.23
1001	95.56 $\pm$ 1.55	0.60 $\pm$ 0.23
1010	95.54 $\pm$ 0.99	0.64 $\pm$ 0.19
1011	95.71 $\pm$ 1.08	0.63 $\pm$ 0.20
1100	95.79 $\pm$ 1.44	0.61 $\pm$ 0.23
1101	95.83 $\pm$ 1.44	0.62 $\pm$ 0.24
1110	95.77 $\pm$ 1.13	0.62 $\pm$ 0.19
1111	95.74 $\pm$ 1.15	0.63 $\pm$ 0.20



**Figura 5.17** – Gráfico da amostra de segmentação por multi-região de MRI-GM, para todas as combinações de características,  $k$ -NN = 1. Observam-se vários picos muito próximos, sendo o maior para as características Média-Variância-Curtosis.

**Tabela 5.16** - Resultados da amostra de segmentação por multi-região de MRI-GM, para todas as combinações de características e  $k$ -NN = 11.

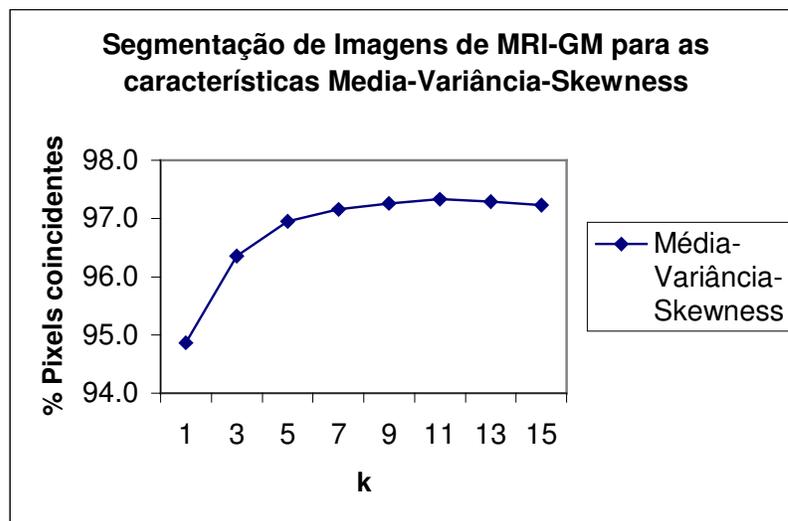
Caract	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
0001	90.67 $\pm$ 4.66	0.38 $\pm$ 0.10
0010	89.30 $\pm$ 1.83	0.33 $\pm$ 0.09
0011	91.18 $\pm$ 2.40	0.38 $\pm$ 0.08
0100	97.30 $\pm$ 1.47	0.56 $\pm$ 0.03
0101	96.78 $\pm$ 2.22	0.55 $\pm$ 0.03
0110	97.38 $\pm$ 1.52	0.57 $\pm$ 0.03
0111	96.99 $\pm$ 2.07	0.56 $\pm$ 0.03
1000	94.86 $\pm$ 2.92	0.51 $\pm$ 0.08
1001	95.34 $\pm$ 2.20	0.50 $\pm$ 0.07
1010	95.78 $\pm$ 1.42	0.51 $\pm$ 0.05
1011	95.76 $\pm$ 1.30	0.51 $\pm$ 0.05
1100	97.68 $\pm$ 1.46	0.58 $\pm$ 0.03
1101	97.28 $\pm$ 2.14	0.57 $\pm$ 0.03
1110	97.70 $\pm$ 1.31	0.57 $\pm$ 0.02
1111	97.39 $\pm$ 1.93	0.57 $\pm$ 0.02



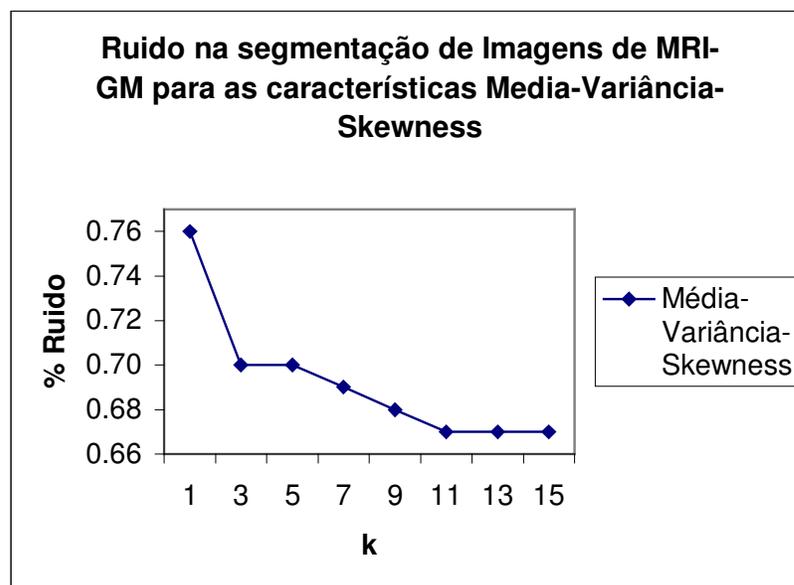
**Figura 5.18** – Gráfico da amostra de segmentação por multi-região de MRI-GM, para todas as combinações de características,  $k\text{-NN} = 11$ . Observam-se vários picos próximos, sendo o maior para as características Média-Variância-Skewness.

**Tabela 5.17** - Resultados da segmentação por multi-região de MRI-GM, para as características Média-Variância-Skewness em função ao valor  $k\text{-NN}$ .

<b>Características: Média-Variância-Skewness</b>		
<b>k-NN</b>	<b>Segmentação (% <math>\mu \pm \sigma</math>)</b>	<b>Ruído (% <math>\mu \pm \sigma</math>)</b>
1	94.86 $\pm$ 0.78	0.76 $\pm$ 0.13
3	96.35 $\pm$ 0.77	0.70 $\pm$ 0.11
5	96.95 $\pm$ 0.93	0.70 $\pm$ 0.10
7	97.16 $\pm$ 1.22	0.69 $\pm$ 0.09
9	97.26 $\pm$ 1.37	0.68 $\pm$ 0.09
11	97.33 $\pm$ 1.49	0.67 $\pm$ 0.09
13	97.29 $\pm$ 1.66	0.67 $\pm$ 0.09
15	97.23 $\pm$ 1.84	0.67 $\pm$ 0.09



(a)



(b)

**Figura 5.19** – Gráfico da Segmentação por multi-região de MRI-GM (a) e o correspondente Ruído (b), para as características Média-Variância-Skewness, e o número de vizinhos mais próximos. Observa-se que enquanto a taxa de segmentação em (a) aumenta praticamente com os valores de  $k$ , o inverso acontece em (b) com o ruído.

### 5.5.3. Testes de segmentação do líquor de imagens de ressonância magnética

Foram realizados os testes de segmentação líquor (CSF), sobre a base de MRI detalhada no **Capítulo 4**, sendo selecionado o subconjunto de teste de 10 imagens. A **Tabela 5.18** e a **Figura 5.20** mostram que, para o conjunto das imagens de ressonância magnética testado, a melhor característica isolada é a Variância (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $90.99 \pm 3.23$  e  $0.43 \pm 0.20$ ) para o parâmetro  $k = 5$ .

Tendo-se achado o melhor resultado para uma característica isolada e parâmetro  $k$ , foi realizada a seguir a segmentação de imagens, com todas as combinações de características, porém com  $k=1$ . Esses resultados são apresentados na **Tabela 5.19** e de forma gráfica, na **Figura 5.21**. De forma comparativa, também foi realizada a segmentação de imagens, com todas as combinações de características, porém com  $k=5$ . Esses resultados são apresentados na **Tabela 5.20** e de forma gráfica na **Figura 5.24**. Isso foi feito, devido a que a melhor segmentação isolada foi obtida com  $k=5$ , para a característica Variância. Observa-se que a combinação de características, que obteve os melhores resultados foi Média-Variância-Skewness-Curtosis, identificada como 1111, seja para  $k=1$  (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $88.48 \pm 4.31$  e  $0.32 \pm 0.18$ ), como para  $k=5$  (%  $\mu \pm \sigma$ , respectivamente:  $90.30 \pm 2.593$  e  $0.29 \pm 0.15$ ). Finalmente, tendo-se identificado a melhor combinação de características, foram feitos experimentos variando-se os valores do parâmetro  $k$  no intervalo ímpar [1,15]. Verificou-se que os melhores resultados foram obtidos com o parâmetro  $k=7$  (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $91.85 \pm 2.74$  e  $0.42 \pm 0.24$ ). Os resultados da segmentação de imagens para as características Média-Variância-Skewness-Curtosis e o parâmetro  $k$  no intervalo ímpar [1, 15] são apresentados na **Tabela 5.21** e de forma gráfica na **Figura 5.23**. Alguns exemplos de resultados de segmentação obtidos para as três classes são apresentados na **Figura 5.24**.

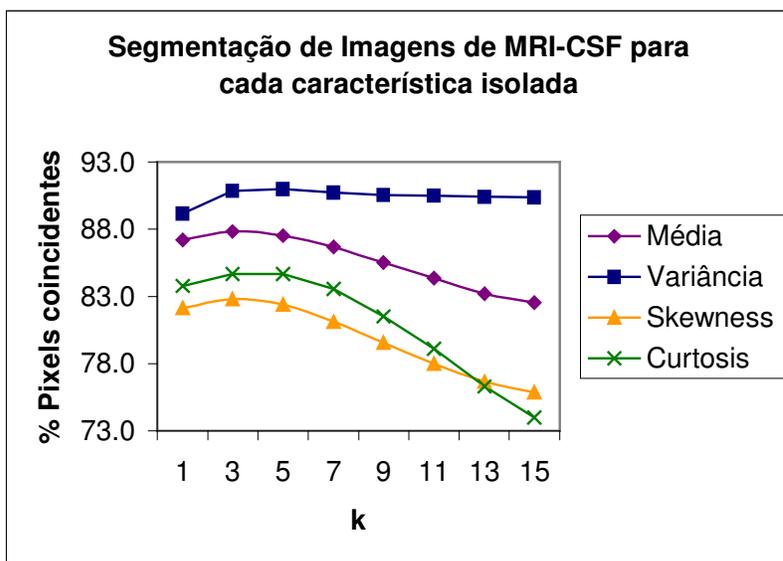
**Tabela 5.18** - Resultados da segmentação por multi-região de MRI-CSF para cada uma das características isoladas em função ao valor k-NN.

Característica: Média		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	87.22 $\pm$ 2.46	0.61 $\pm$ 0.19
3	87.84 $\pm$ 2.80	0.43 $\pm$ 0.20
5	87.50 $\pm$ 3.23	0.42 $\pm$ 0.25
7	86.68 $\pm$ 3.59	0.41 $\pm$ 0.28
9	85.50 $\pm$ 3.84	0.40 $\pm$ 0.30
11	84.35 $\pm$ 4.12	0.39 $\pm$ 0.30
13	83.20 $\pm$ 4.67	0.37 $\pm$ 0.30
15	82.52 $\pm$ 4.97	0.36 $\pm$ 0.30

Característica: Variância		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	89.17 $\pm$ 2.32	0.55 $\pm$ 0.26
3	90.84 $\pm$ 2.51	0.45 $\pm$ 0.20
5	90.99 $\pm$ 3.23	0.43 $\pm$ 0.20
7	90.72 $\pm$ 3.83	0.43 $\pm$ 0.21
9	90.54 $\pm$ 3.99	0.44 $\pm$ 0.21
11	90.51 $\pm$ 4.07	0.44 $\pm$ 0.21
13	90.41 $\pm$ 4.21	0.44 $\pm$ 0.21
15	90.39 $\pm$ 4.30	0.44 $\pm$ 0.21

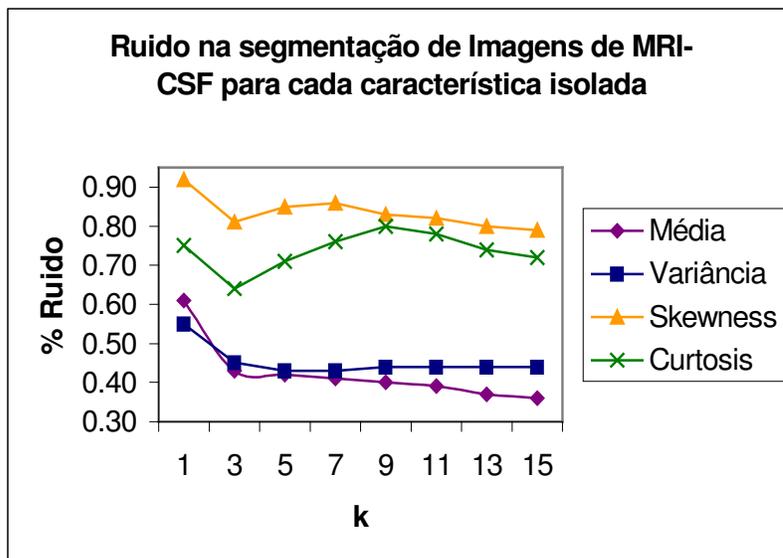
Característica: Skewness		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	82.11 $\pm$ 2.86	0.92 $\pm$ 0.31
3	82.81 $\pm$ 2.85	0.81 $\pm$ 0.31
5	82.40 $\pm$ 3.13	0.85 $\pm$ 0.35
7	81.11 $\pm$ 3.65	0.86 $\pm$ 0.34
9	79.56 $\pm$ 4.23	0.83 $\pm$ 0.27
11	77.98 $\pm$ 4.69	0.82 $\pm$ 0.28
13	76.65 $\pm$ 4.85	0.80 $\pm$ 0.28
15	75.87 $\pm$ 4.85	0.79 $\pm$ 0.28

Característica: Curtosis		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	83.76 $\pm$ 2.10	0.75 $\pm$ 0.16
3	84.69 $\pm$ 2.32	0.64 $\pm$ 0.17
5	84.69 $\pm$ 2.60	0.71 $\pm$ 0.26
7	83.56 $\pm$ 2.98	0.76 $\pm$ 0.32
9	81.49 $\pm$ 3.35	0.80 $\pm$ 0.38
11	79.10 $\pm$ 3.66	0.78 $\pm$ 0.38
13	76.31 $\pm$ 3.83	0.74 $\pm$ 0.38
15	73.96 $\pm$ 3.80	0.72 $\pm$ 0.38



(a)

**Figura 5.20** – Gráfico da Segmentação por multi-região de MRI-CSF (a) e o correspondente Ruído (b), para cada característica isolada e o número de vizinhos mais próximos.

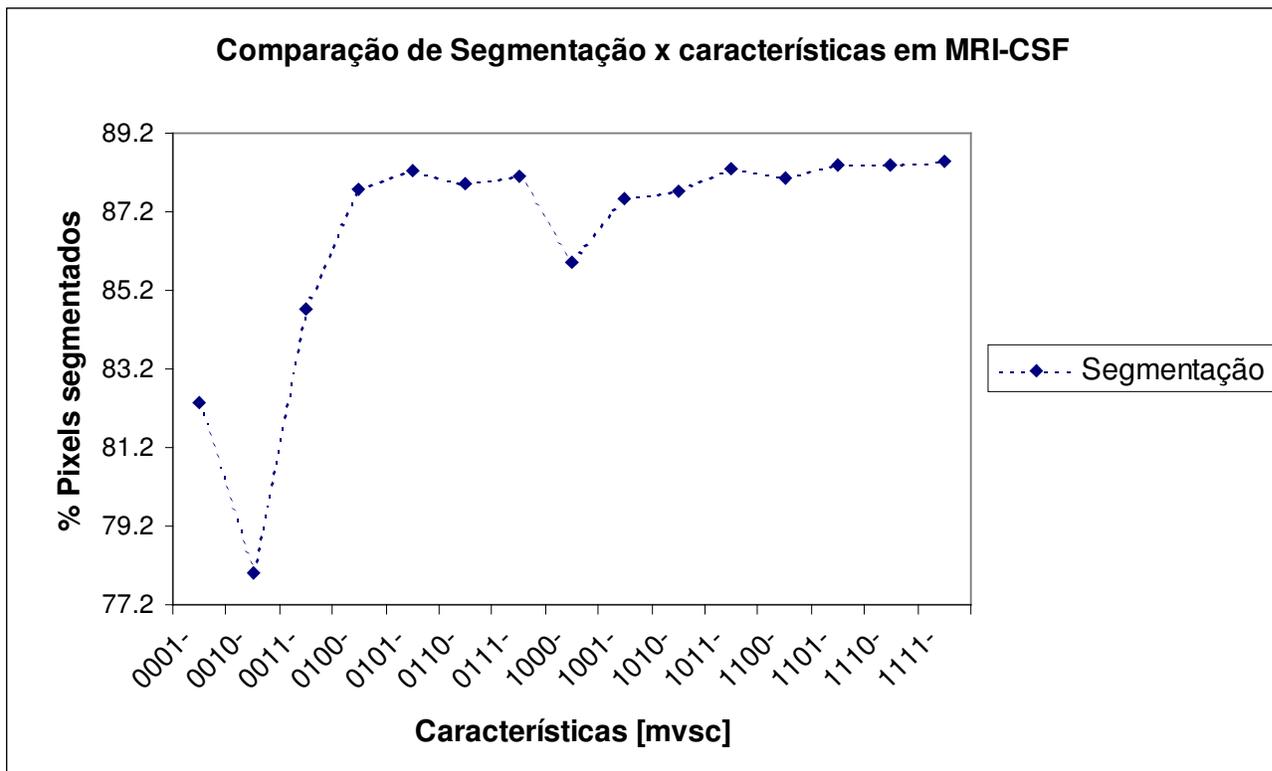


(b)

**Figura 5.20 (cont.)**– Gráfico da Segmentação por multi-região de MRI-CSF (a) e o correspondente Ruído (b), para cada característica isolada e o número de vizinhos mais próximos. Observa-se em (a) que com exceção da característica Variância, as taxas de segmentação diminuem com o valor de k. Em (b), não existe um padrão definido do ruído com o valor de k.

**Tabela 5.19** - Resultados da amostra de segmentação por multi-região de MRI-CSF para todas as combinações de características e  $k\text{-NN} = 1$ .

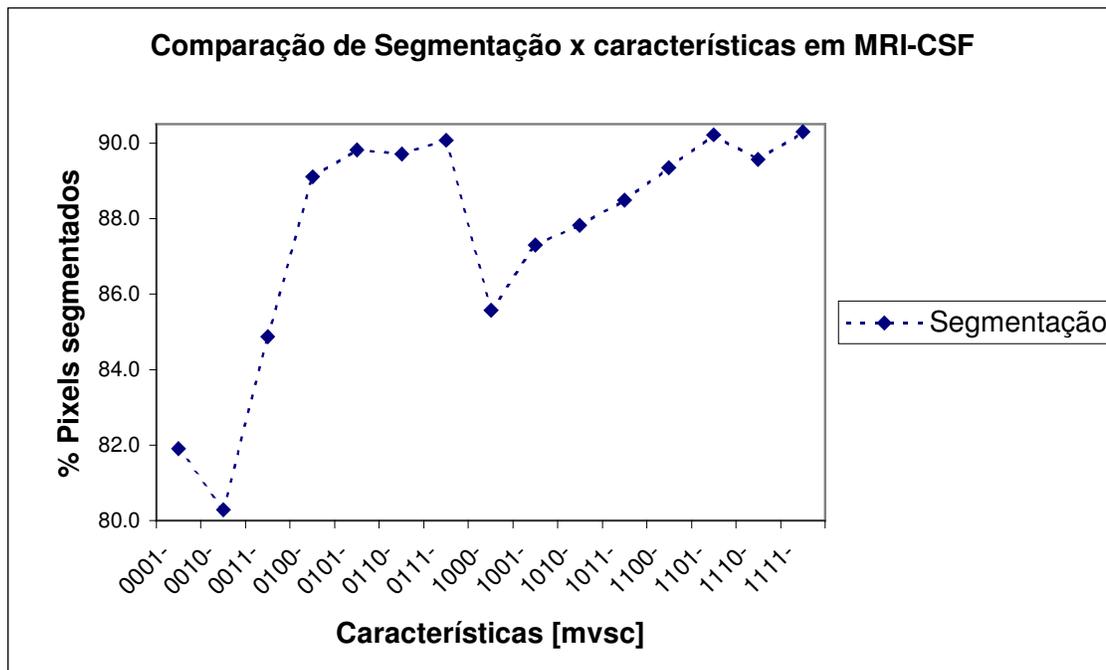
Caract	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
0001	82.36 $\pm$ 6.06	0.55 $\pm$ 0.19
0010	78.00 $\pm$ 2.97	1.36 $\pm$ 0.78
0011	84.71 $\pm$ 3.78	0.55 $\pm$ 0.19
0100	87.75 $\pm$ 4.47	0.41 $\pm$ 0.34
0101	88.24 $\pm$ 4.28	0.46 $\pm$ 0.35
0110	87.93 $\pm$ 3.49	0.40 $\pm$ 0.29
0111	88.11 $\pm$ 3.91	0.45 $\pm$ 0.31
1000	85.92 $\pm$ 4.87	0.42 $\pm$ 0.16
1001	87.55 $\pm$ 4.61	0.35 $\pm$ 0.19
1010	87.73 $\pm$ 3.78	0.44 $\pm$ 0.31
1011	88.30 $\pm$ 3.98	0.38 $\pm$ 0.25
1100	88.08 $\pm$ 4.46	0.33 $\pm$ 0.21
1101	88.40 $\pm$ 4.39	0.33 $\pm$ 0.21
1110	88.37 $\pm$ 4.22	0.30 $\pm$ 0.18
1111	88.48 $\pm$ 4.31	0.32 $\pm$ 0.18



**Figura 5.21** – Gráfico da amostra de segmentação por multi-região de MRI-CSF para todas as combinações de características,  $k$ -NN = 1. Observam-se vários picos muito próximos, sendo o maior para as características Média-Variância-Skewness-Curtosis.

**Tabela 5.20** - Resultados da amostra de segmentação por multi-região de MRI-CSF para todas as combinações de características e  $k$ -NN = 5.

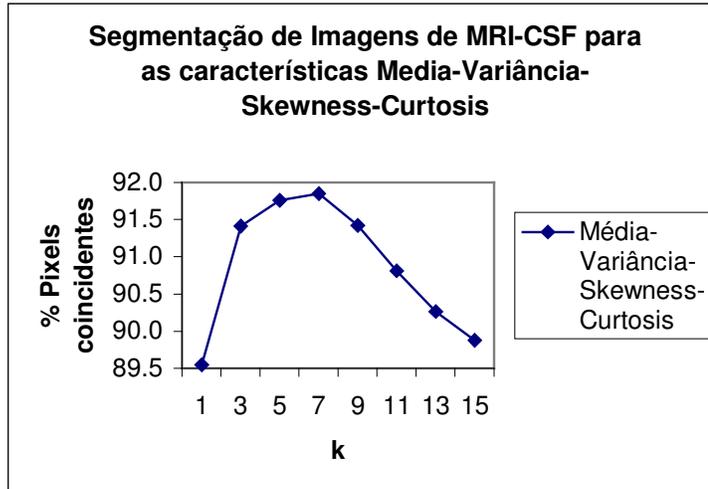
Caract	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
0001	81.90 $\pm$ 3.65	0.57 $\pm$ 0.20
0010	80.29 $\pm$ 3.16	0.71 $\pm$ 0.42
0011	84.87 $\pm$ 2.61	0.52 $\pm$ 0.20
0100	89.10 $\pm$ 3.36	0.32 $\pm$ 0.22
0101	89.82 $\pm$ 2.49	0.35 $\pm$ 0.20
0110	89.70 $\pm$ 3.30	0.30 $\pm$ 0.19
0111	90.07 $\pm$ 2.62	0.34 $\pm$ 0.20
1000	85.57 $\pm$ 2.80	0.30 $\pm$ 0.23
1001	87.30 $\pm$ 2.91	0.26 $\pm$ 0.12
1010	87.82 $\pm$ 2.81	0.31 $\pm$ 0.22
1011	88.48 $\pm$ 2.72	0.30 $\pm$ 0.20
1100	89.34 $\pm$ 3.11	0.28 $\pm$ 0.14
1101	90.22 $\pm$ 2.64	0.28 $\pm$ 0.14
1110	89.57 $\pm$ 2.91	0.29 $\pm$ 0.15
1111	90.30 $\pm$ 2.59	0.29 $\pm$ 0.15



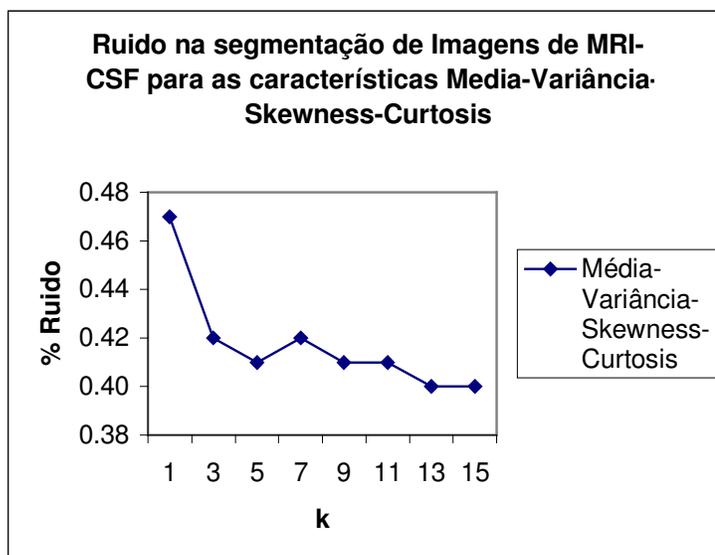
**Figura 5.22** – Gráfico da amostra de segmentação por multi-região de MRI-CSF para todas as combinações de características,  $k\text{-NN} = 5$ . Observam-se vários picos próximos, sendo o maior para as características Média-Variância-Skewness-Curtosis.

**Tabela 5.21** - Resultados da segmentação por multi-região de MRI-CSF para as características Média-Variância-Skewness-Curtosis em função ao valor  $k\text{-NN}$ .

<b>Características: Média-Variância-Skewness-Curtosis</b>		
<b>k-NN</b>	<b>Segmentação (% <math>\mu \pm \sigma</math>)</b>	<b>Ruído (% <math>\mu \pm \sigma</math>)</b>
1	89.55 $\pm$ 2.83	0.47 $\pm$ 0.25
3	91.41 $\pm$ 2.43	0.42 $\pm$ 0.26
5	91.76 $\pm$ 2.65	0.41 $\pm$ 0.24
7	91.85 $\pm$ 2.74	0.42 $\pm$ 0.24
9	91.42 $\pm$ 3.27	0.41 $\pm$ 0.22
11	90.81 $\pm$ 3.98	0.41 $\pm$ 0.21
13	90.26 $\pm$ 4.55	0.40 $\pm$ 0.22
15	89.88 $\pm$ 4.97	0.40 $\pm$ 0.22

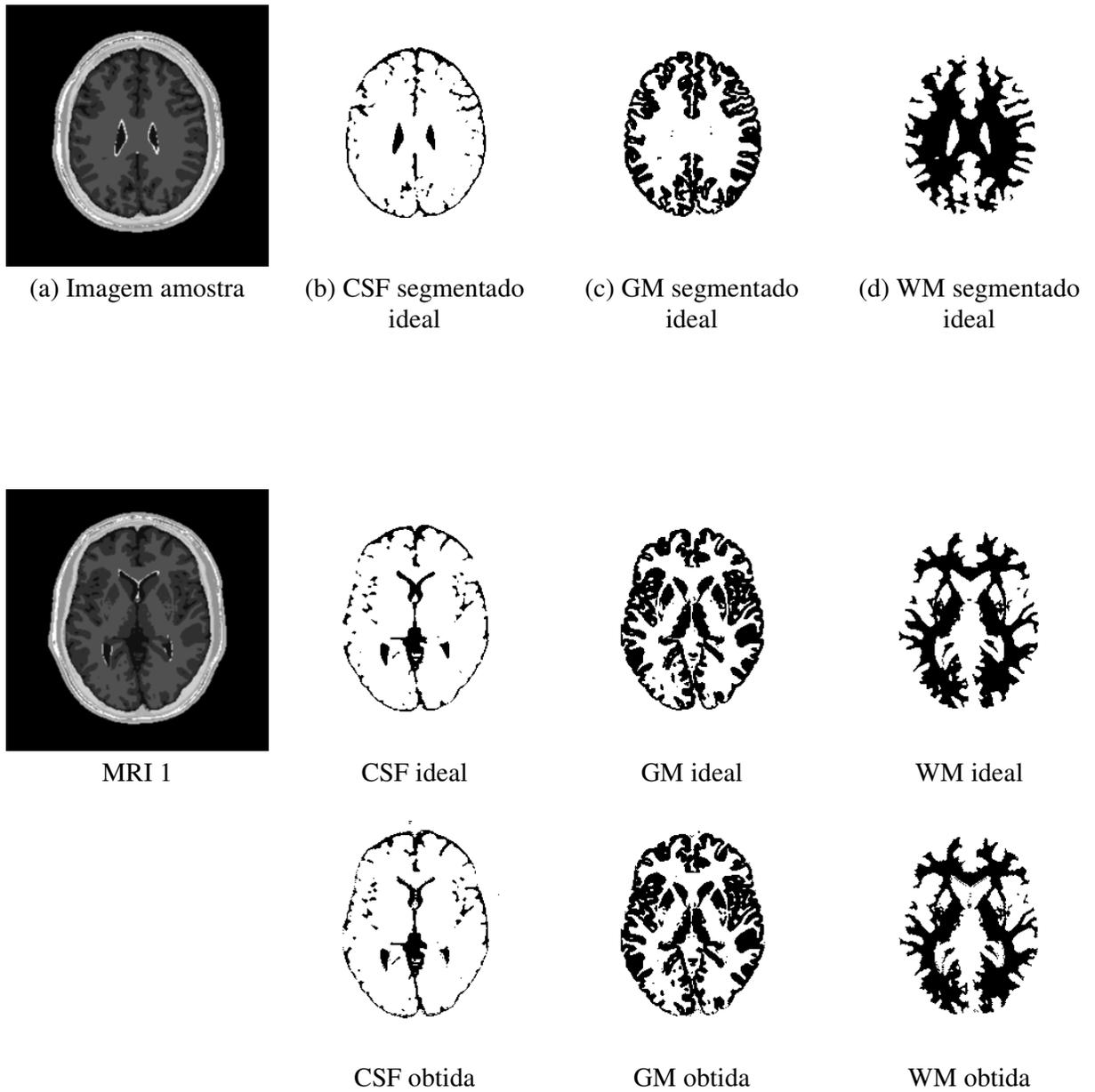


(a)

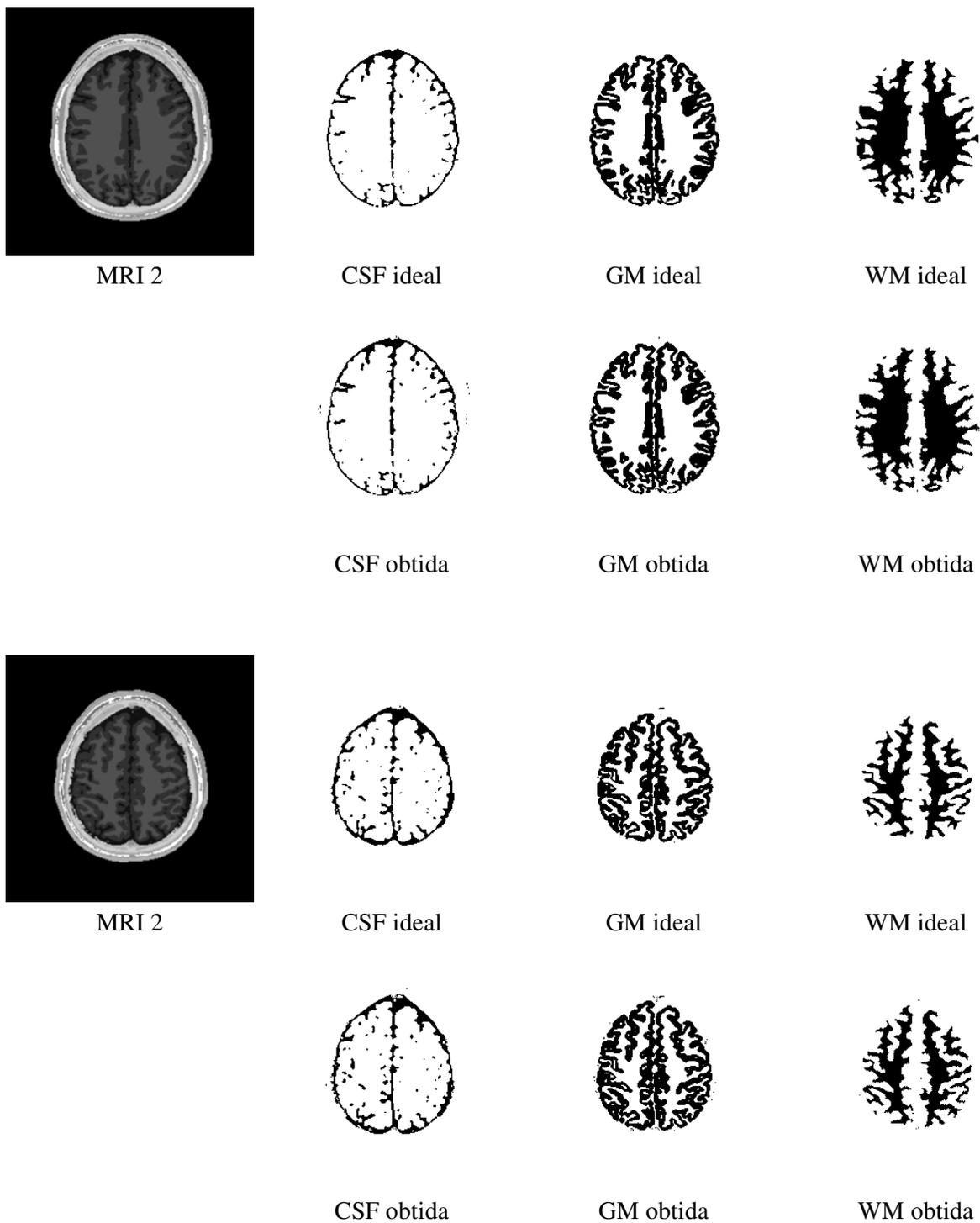


(b)

**Figura 5.23** – Gráfico da Segmentação por multi-região de MRI-CSF (a) e o correspondente Ruído (b), para a característica Média-Variância-Skewness-Curtosis e o número de vizinhos mais próximos. Em (a) observa-se que a taxa de segmentação aumenta até o valor de  $k=7$ ; logo diminuindo com os valores maiores de  $k$ . Em (b) o ruído diminui à medida que o  $k$  aumenta, com exceção de  $k=7$ .



**Figura 5.24.** Segmentação da massa cinza GM, massa branca WM e líquido CSF de imagens de MR por aprendizagem.



**Figura 5.24 (cont.).** Segmentação da massa cinza GM, massa branca WM e líquido CSF de imagens de MR por aprendizagem.

## 5.6. Testes de segmentação em imagens de envelopes postais

### 5.6.1. Calibração dos histogramas das imagens a serem segmentadas

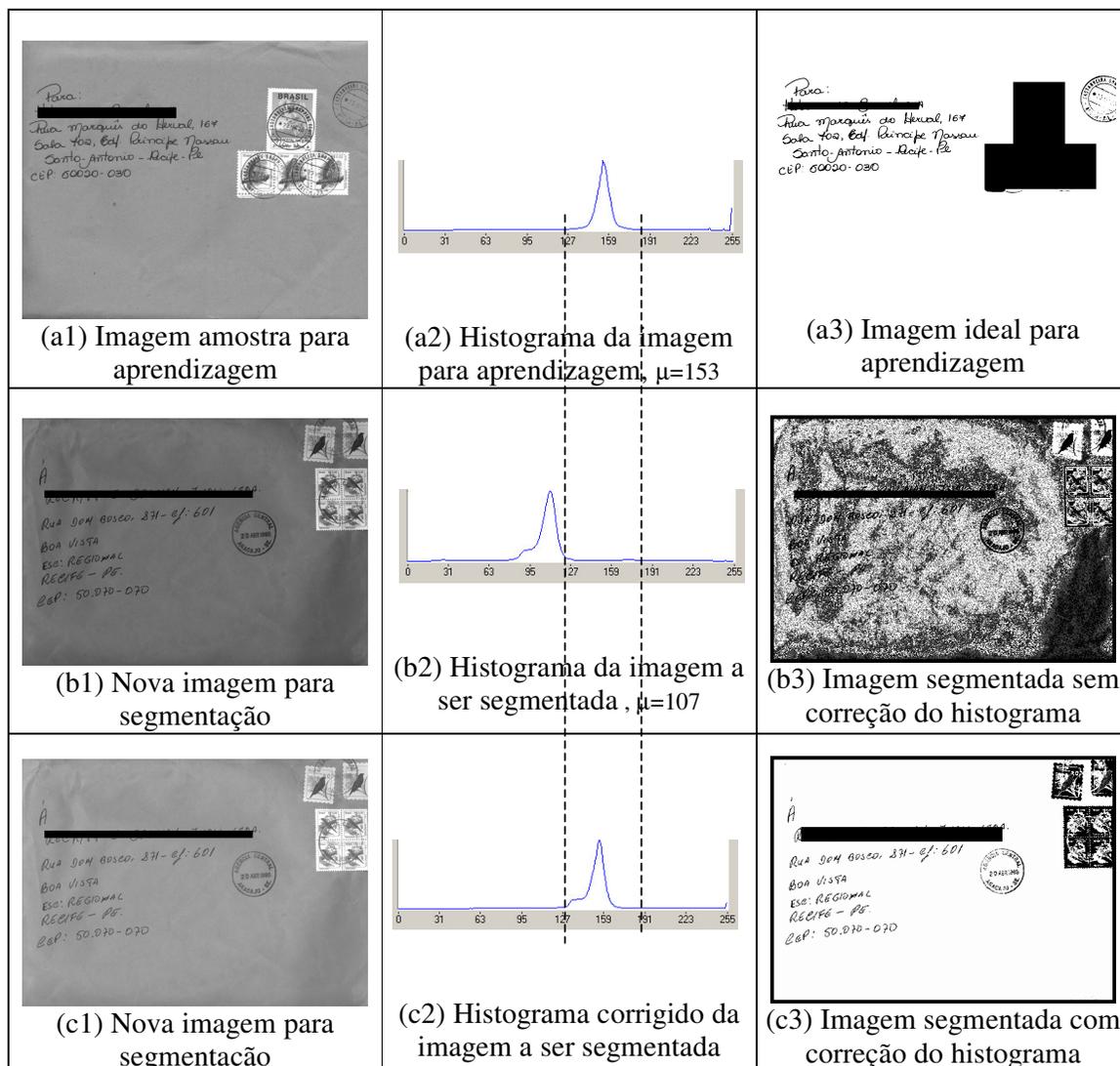
A abordagem de segmentação proposta baseia-se no nível de cinza do *pixel* objeto de análise, do qual são extraídas as características. Portanto, se a base a ser segmentada apresenta uma grande variação na iluminação ou na aquisição, pode acontecer que o histograma da imagem aprendida encontre-se em uma faixa diferente e não coincidente com o histograma objeto de segmentação.

A situação anteriormente exposta pode ser melhor explicada com a **Figura 5.25**. A média do histograma (a2) da imagem para aprendizagem é  $\mu=153$ , enquanto que a média do histograma (b2), da nova imagem para segmentação, é  $\mu=107$ . Observe que os *pixels* mais escuros da imagem (a1), têm um nível de cinza maior que os *pixels* mais claros de (b1). O resultado da aplicação direta do algoritmo de segmentação apresentado em (b3) portanto, não é satisfatório. Os *pixels* da imagem em (c3) foram corrigidos previamente adicionando-se o valor de 46 a cada um. Apresenta-se assim um novo histograma (c2). Logo após a segmentação pelo algoritmo, obtém-se a imagem resultante em (c3). A comparação visual da imagem segmentada em (c3) mostra que a mesma apresenta uma qualidade superior à obtida em (b3).

O processo de correção de histograma da imagem a ser segmentada pode ser expresso de maneira formal:

Sejam  $\mu_A$  e  $\mu_S$  as médias dos histogramas das imagens aprendidas e a serem segmentadas, respectivamente;  $P(x,y)_S$  e  $P(x,y)_C$  o nível de cinza do *pixel* da imagem a ser segmentada e da imagem corrigida respectivamente, e  $\mu_A \neq \mu_S$ .

$$P(x,y)_C = \begin{cases} 0 & \text{se } [P(x,y)_S + (\mu_A - \mu_S)] < 0 \\ [P(x,y)_S + (\mu_A - \mu_S)] & \text{se } 0 < [P(x,y)_S + (\mu_A - \mu_S)] < 255 \\ 255 & \text{se } [P(x,y)_S + (\mu_A - \mu_S)] > 255 \end{cases} \quad (5.12)$$



**Figura 5.25.** Exemplo prático do processo de correção de histograma nas imagens a ser segmentadas. Os nomes dos destinatários foram escondidos propositalmente.

A calibração do histograma da imagem a ser segmentada é um processo automático e não requer a intervenção manual do operador. Armazena-se o valor aprendido da média do histograma da imagem no cabeçalho da Matriz de Decisão, juntamente com o tamanho da janela. Por conseguinte, no momento da segmentação de novas imagens, é realizada primeiramente uma verificação das médias dos histogramas. Caso estas médias sejam diferentes, é feita a conversão em tempo de processamento, sem modificar a imagem submetida.

### 5.6.2. Testes de segmentação na base ECT-DB1

Foram realizados os testes de segmentação sobre a base de envelopes postais de tamanho reduzido ECT-DB1 detalhada no **Capítulo 4**, sendo selecionado o subconjunto de teste de 10 imagens. A **Tabela 5.22** e a **Figura 5.26** mostram que, para o conjunto das imagens de envelopes testado, a melhor característica isolada é a Média (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $79.16 \pm 13.86$  e  $1.88 \pm 1.83$ ), para o parâmetro  $k = 1$ .

Tendo-se achado o melhor resultado para uma característica isolada e parâmetro  $k$ , foi realizada a seguir a segmentação de imagens, com todas as combinações de características, e  $k=1$ . Esses resultados são apresentados na **Tabela 5.23** e de forma gráfica, na **Figura 5.27**. Observa-se que a combinação de características que obteve melhores resultados foi Média-Variância, identificada como 1100, (taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente  $86.50 \pm 12.23$  e  $1.74 \pm 1.15$ ).

Foi feito um experimento adicional de verificação do melhor conjunto de características. À diferença dos casos anteriores, onde era escolhida uma amostra de imagem com todas as combinações de características e de Matrizes de Decisão, nesta ocasião foram utilizadas cada uma das imagens de aprendizagem por vez, para todas as combinações de características,  $k\text{-NN} = 1$ . Na **Figura 5.28** observa-se que, em todos os casos, a melhor combinação de características é Média-Variância, com exceção da aprendizagem com a imagem c003, onde houve um empate nos resultados de segmentação entre dois conjuntos de características a) Média-Variância e b) Média-Variância-Curtosis. Porém, neste caso de empate, os valores de desvio-padrão e ruído são diferentes. Os valores obtidos, para o conjunto a), são (taxa de segmentação e ruído em %  $\mu \pm \sigma$ ), respectivamente  $82.68 \pm 13.4$  e  $4.61 \pm 2.65$ ) e b) são (taxa de segmentação e ruído em %  $\mu \pm \sigma$ ), respectivamente  $82.68 \pm 12.6$  e  $5.04 \pm 2.81$ ).

Tendo-se identificado a melhor combinação de características, como sendo o conjunto Média-Variância, o melhor resultado médio obtido foi taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente:  $80.31 \pm 14.3$  e  $1.86 \pm 1.68$  para o parâmetro  $k = 1$ . Os resultados da

segmentação de imagens para a característica Média-Variância e o parâmetro  $k$  no intervalo ímpar [1, 15] são apresentados na **Tabela 5.24** e de forma gráfica na **Figura 5.29**.

Finalmente, foi realizado o teste de segmentação da base completa de imagens ECT-DB1 ( $n=200$ ). Para a aprendizagem selecionou-se uma das 10 imagens utilizadas na **Seção 5.6.1**. Após a aprendizagem, foi realizada a segmentação nas 199 imagens restantes da base ECT-DB1, utilizando-se as características Média-Variância, e o parâmetro  $k=1$ , que apresentaram melhor desempenho. As taxas de segmentação e ruído obtidos foram, em  $\% \mu \pm \sigma$ , respectivamente:  $87.88 \pm 11.76$  e  $03.70 \pm 02.74$ . Observa-se que estes resultados, apresentados na **Tabela 5.25**, são melhores que os obtidos com a amostra de 10 imagens de envelopes. A **Tabela 5.26** apresenta os resultados (em  $\% \mu \pm \sigma$ ) computados pela metodologia adotada por [MEN03] nas classes Bloco-Endereço:  $96.74 \pm 09.47$ , Selo:  $85.30 \pm 05.75$ , Carimbo:  $93.54 \pm 14.94$  e Ruído:  $03.70 \pm 02.74$ .

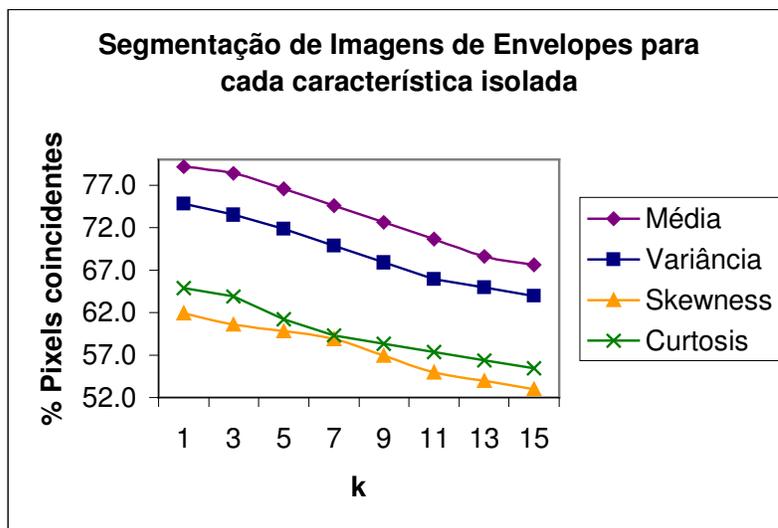
**Tabela 5.22** - Resultados da segmentação por binarização para cada uma das características isoladas em função ao valor k-NN

Característica: Média		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	79.16 $\pm$ 13.86	1.88 $\pm$ 1.83
3	78.40 $\pm$ 12.71	1.87 $\pm$ 1.86
5	76.52 $\pm$ 12.68	1.85 $\pm$ 1.88
7	74.60 $\pm$ 11.66	1.86 $\pm$ 1.92
9	72.61 $\pm$ 12.63	1.86 $\pm$ 1.92
11	70.62 $\pm$ 12.65	1.87 $\pm$ 1.93
13	68.62 $\pm$ 12.46	1.87 $\pm$ 1.93
15	67.62 $\pm$ 12.06	1.88 $\pm$ 1.93

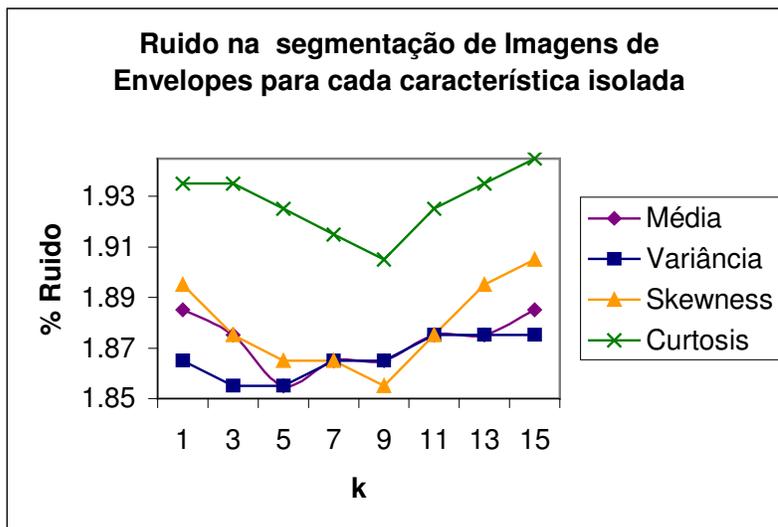
Característica: Variância		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	74.78 $\pm$ 14.20	1.86 $\pm$ 1.82
3	73.54 $\pm$ 13.82	1.85 $\pm$ 1.88
5	71.83 $\pm$ 12.58	1.85 $\pm$ 1.92
7	69.89 $\pm$ 12.52	1.86 $\pm$ 1.89
9	67.90 $\pm$ 13.51	1.86 $\pm$ 1.83
11	65.91 $\pm$ 13.50	1.87 $\pm$ 1.92
13	64.91 $\pm$ 13.55	1.87 $\pm$ 1.72
15	63.94 $\pm$ 13.50	1.87 $\pm$ 1.82

Característica: Skewness		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	61.93 $\pm$ 13.41	1.89 $\pm$ 1.87
3	60.57 $\pm$ 12.80	1.87 $\pm$ 1.88
5	59.84 $\pm$ 12.52	1.86 $\pm$ 1.92
7	58.91 $\pm$ 12.34	1.86 $\pm$ 1.93
9	56.94 $\pm$ 12.27	1.85 $\pm$ 1.90
11	54.95 $\pm$ 12.22	1.87 $\pm$ 1.89
13	53.96 $\pm$ 12.20	1.89 $\pm$ 1.88
15	52.95 $\pm$ 12.18	1.90 $\pm$ 1.75

Característica: Curtosis		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	64.87 $\pm$ 13.07	1.93 $\pm$ 1.90
3	63.91 $\pm$ 12.81	1.93 $\pm$ 1.92
5	61.25 $\pm$ 12.53	1.92 $\pm$ 1.96
7	59.30 $\pm$ 12.44	1.91 $\pm$ 1.96
9	58.32 $\pm$ 12.35	1.90 $\pm$ 1.96
11	57.33 $\pm$ 12.30	1.92 $\pm$ 1.96
13	56.34 $\pm$ 12.27	1.93 $\pm$ 1.96
15	55.46 $\pm$ 12.25	1.94 $\pm$ 1.96



(a)

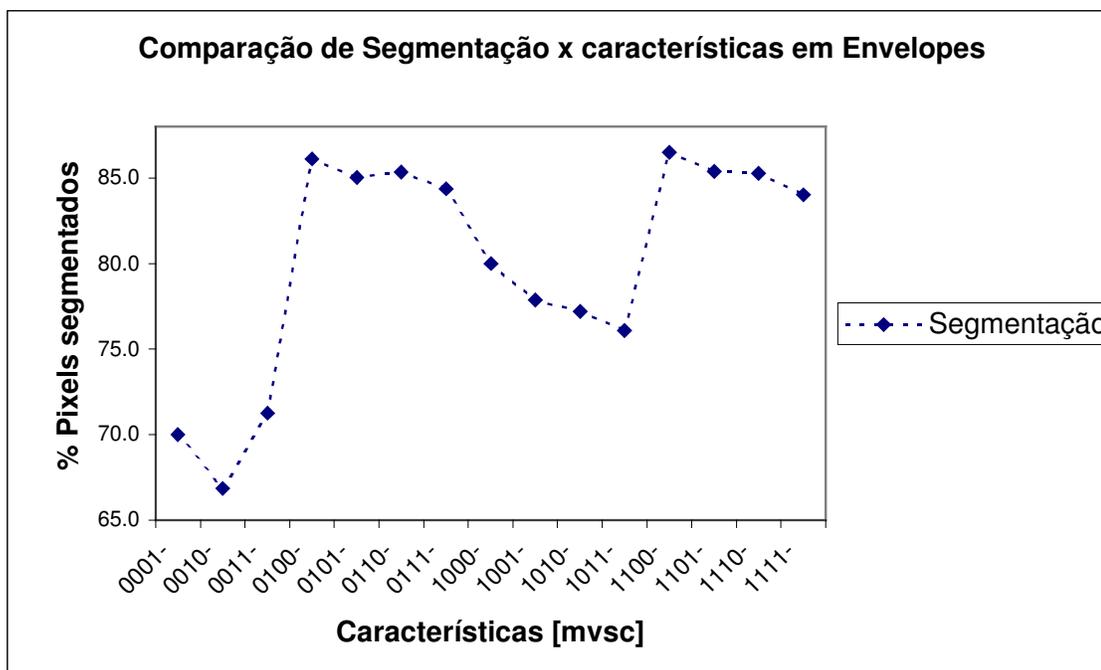


(b)

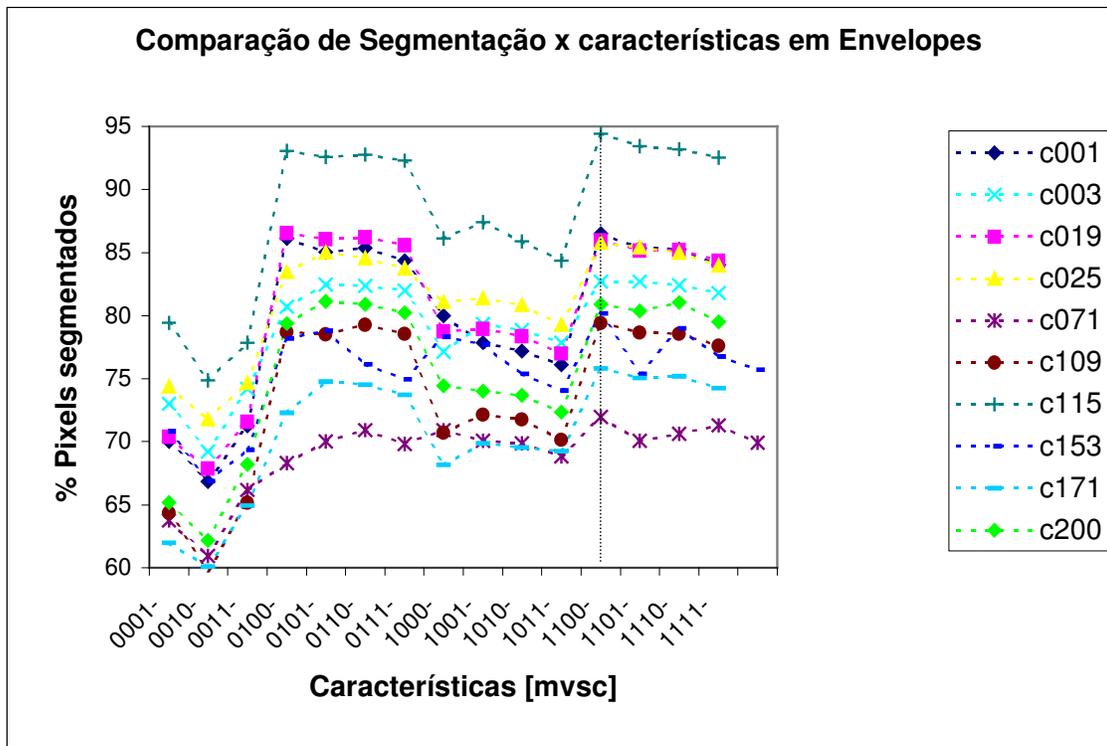
**Figura 5.26** – Gráfico da Segmentação de imagens de Envelopes (a) e o correspondente Ruído (b), para cada característica isolada e o número de vizinhos mais próximos. Em (a) observa-se que as taxas de segmentação diminuem com o aumento do k, para todas as características. Em (b) não existe um padrão definido do ruído com o valor de k.

**Tabela 5.23** - Resultados da amostra de segmentação por binarização de Envelopes para todas as combinações de características k-NN = 1.

Caract	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
0001	70.01 $\pm$ 15.08	3.52 $\pm$ 2.61
0010	66.84 $\pm$ 12.44	3.11 $\pm$ 2.28
0011	71.25 $\pm$ 14.66	3.09 $\pm$ 1.67
0100	86.10 $\pm$ 12.00	2.23 $\pm$ 1.24
0101	85.04 $\pm$ 12.23	1.94 $\pm$ 1.08
0110	85.35 $\pm$ 12.29	1.74 $\pm$ 1.01
0111	84.37 $\pm$ 12.55	1.76 $\pm$ 1.08
1000	79.99 $\pm$ 11.75	3.91 $\pm$ 2.86
1001	77.86 $\pm$ 12.59	2.17 $\pm$ 1.41
1010	77.19 $\pm$ 13.51	2.09 $\pm$ 1.30
1011	76.09 $\pm$ 14.09	2.03 $\pm$ 1.35
1100	86.50 $\pm$ 12.23	1.74 $\pm$ 1.15
1101	85.37 $\pm$ 12.28	1.47 $\pm$ 0.86
1110	85.28 $\pm$ 12.14	1.48 $\pm$ 0.89
1111	84.03 $\pm$ 12.39	1.55 $\pm$ 1.04



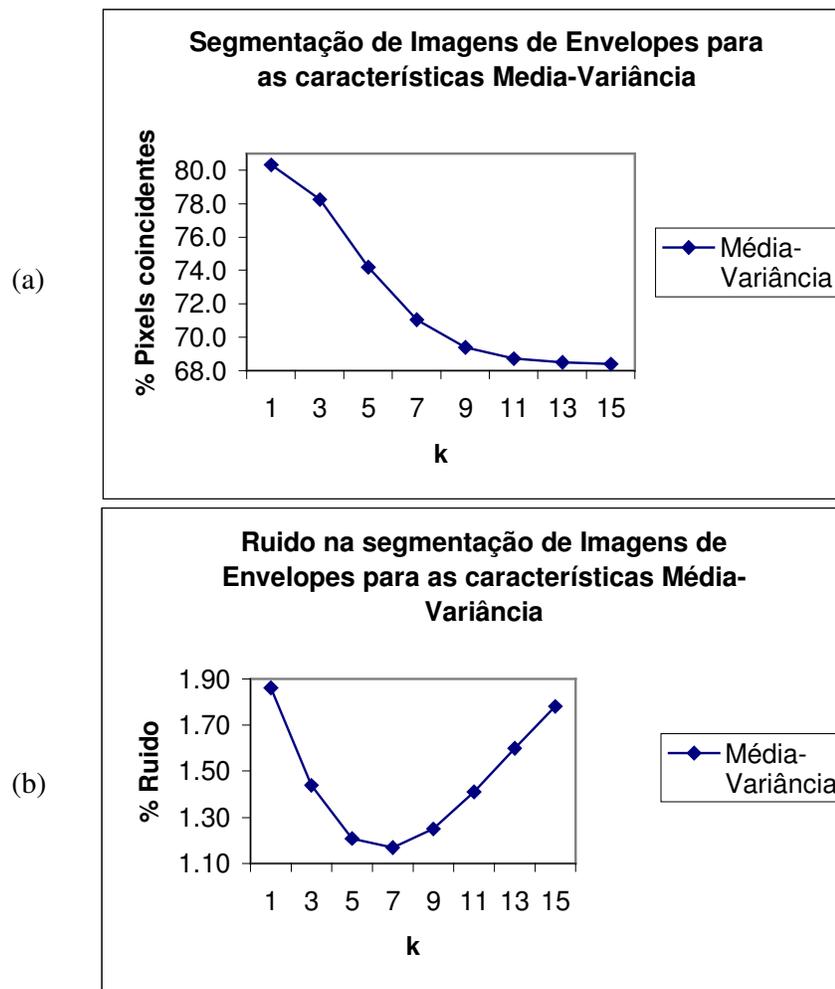
**Figura 5.27** – Gráfico da amostra de segmentação por binarização de Envelopes para todas as combinações de características, k-NN = 1. Observam-se dois picos significativos, sendo o maior, para as características Média-Variância.



**Figura 5.28** – Gráfico comparativo dos resultados da segmentação por binarização de Envelopes. Para tanto, utilizando-se cada uma das imagens de aprendizagem por vez, para todas as combinações de características,  $k\text{-NN} = 1$ . Observa-se que, em todos os casos, a melhor combinação de características é Média-Variância, com exceção da aprendizagem com a imagem c003, onde houve um empate nos resultados de segmentação, entre dois conjuntos de características, a) Média-Variância e b) Média-Variância-Curtosis. Porém, neste caso de empate, os valores de desvio-padrão e ruído são diferentes. Os valores obtidos para o conjunto a) são (taxa de segmentação e ruído em %  $\mu \pm \sigma$ ), respectivamente,  $82.68 \pm 13.4$  e  $4.61 \pm 2.65$ ) e b) são (taxa de segmentação e ruído em %  $\mu \pm \sigma$ ), respectivamente,  $82.68 \pm 12.6$  e  $5.04 \pm 2.81$ ).

**Tabela 5.24** - Resultados da segmentação por binarização de Envelopes, para as características Média-Variância, em função ao valor k-NN.

Características: Média - Variância		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	80.31 $\pm$ 14.3	1.86 $\pm$ 1.68
3	78.26 $\pm$ 11.6	1.44 $\pm$ 1.61
5	74.18 $\pm$ 11.9	1.21 $\pm$ 1.58
7	71.04 $\pm$ 12.2	1.17 $\pm$ 1.73
9	69.39 $\pm$ 12.2	1.25 $\pm$ 1.95
11	68.70 $\pm$ 12.1	1.41 $\pm$ 2.33
13	68.49 $\pm$ 12.1	1.60 $\pm$ 2.86
15	68.40 $\pm$ 12.0	1.78 $\pm$ 3.46



**Figura 5.29** – Gráfico da Segmentação de imagens de Envelopes (a) e o correspondente Ruído (b), para as características Média-Variância e o número de vizinhos mais próximos. Observa-se em (a) que a taxa de segmentação diminui com o valor de k. Em (b) o ruído tem um mínimo para k = 7.

**Tabela 5.25** - Resultados da segmentação por binarização de Envelopes, para as características Média – Variância, para a base completa ECT-DB1 (n= 200).

Características: Média - Variância		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	87.88 $\pm$ 11.76	03.70 $\pm$ 02.74

**Tabela 5.26** - Resultados da segmentação por classes de Envelopes, para as características Média – Variância para a base completa ECT-DB1 (n= 200).

Características: Média – Variância (% $\mu \pm \sigma$ )	
Bloco-Endereço	96.74 $\pm$ 09.47
Selo	85.30 $\pm$ 05.75
Carimbo	93.54 $\pm$ 14.94
Ruído	03.70 $\pm$ 02.74

### 5.6.3. Testes de segmentação na base ECT-DB2

Foram realizados dois experimentos de segmentação, sobre a base de envelopes postais de tamanho real ECT-DB2 detalhada no **Capítulo 4**. Neste caso, sobre o conjunto de 200 imagens. No primeiro experimento, foi realizada a aprendizagem, selecionando-se uma das 10 imagens utilizadas na **Seção 5.6.1**, porém com o tamanho real. Após a aprendizagem, foi realizada a segmentação nas 199 imagens restantes da base ECT-DB2, utilizando-se as características, Média-Variância e o parâmetro  $k=1$ . Essas, apresentaram melhor desempenho nos testes anteriores com envelopes reduzidos. A **Tabela 5.27** mostra os resultados obtidos, de taxa de segmentação e ruído em %  $\mu \pm \sigma$ , respectivamente.

**Tabela 5.27** - Resultados da segmentação por binarização de Envelopes, para as características Média – Variância para a base completa ECT-DB2 (n= 200).

Características: Média - Variância		
k-NN	Segmentação (% $\mu \pm \sigma$ )	Ruído (% $\mu \pm \sigma$ )
1	90.59 $\pm$ 15.05	02.06 $\pm$ 01.32

No segundo experimento, foi realizado o teste de segmentação, nas quatro classes (Bloco Endereço, carimbo, selo e fundo), sobre a base de envelopes postais de tamanho real ECT-DB2, porém seguindo-se a metodologia de avaliação de resultados por classe descrita na **Seção 5.2.2**. Esta metodologia apresentada em [MEN03] e logo utilizada em [EIT04], foi implementada neste teste de maneira a obter um resultado comparativo da abordagem por aprendizagem proposta com outras abordagens aplicadas sobre uma mesma base de imagens. Neste teste foi realizada a aprendizagem selecionando-se uma das 10 imagens utilizadas na **Seção 5.6.1**, porém com o tamanho real. Após a aprendizagem, foi realizada a segmentação nas 199 imagens restantes da base ECT-DB2, utilizando-se as características, Média-Variância e o parâmetro  $k=1$ . Essas apresentaram melhor desempenho nos testes anteriores com envelopes reduzidos. A **Tabela 5.28** mostra um comparativo dos resultados obtidos da taxa de segmentação das três classes e ruído em  $\% \mu \pm \sigma$ , pela presente abordagem proposta por aprendizagem e pela abordagem por *wavelet* apresentada em [MEN03]. Observa-se que a abordagem por aprendizagem apresentou resultados melhores nas três classes de objetos (Bloco-endereço 1.49%, Selo 63.05% e Carimbo 9.35%), sendo que o ruído foi maior em 1.43%.

A **Tabela 5.29** mostra um comparativo dos resultados obtidos com a abordagem baseada na dimensão fractal, apresentada em [EIT04]. Neste caso, o melhor resultado obtido pela abordagem por fractal utiliza os parâmetros [3,13]. Observa-se, no entanto, que o ruído é 5.79% superior ao obtido por aprendizagem.

**Tabela 5.28** - Resultados comparativos da segmentação por classes de Envelopes, para a base completa ECT-DB2 (n= 200) com a abordagem por *wavelet*.

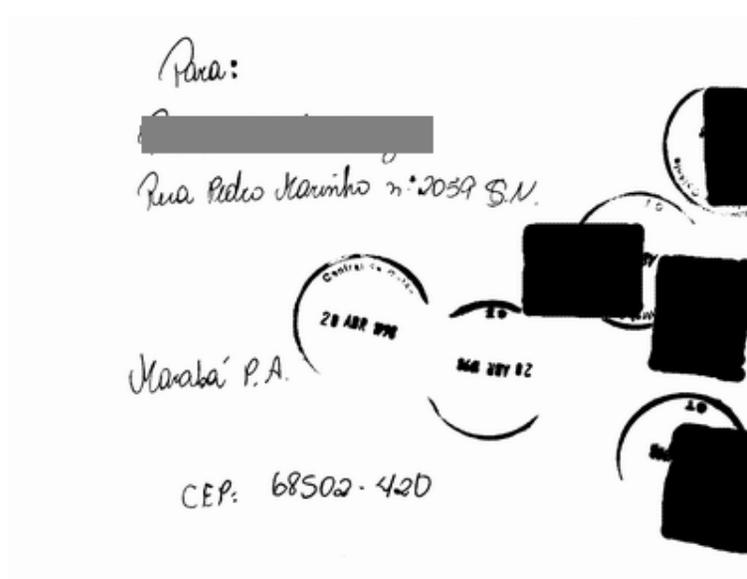
(% $\mu \pm \sigma$ )	Abordagem <i>wavelet</i>	Abordagem por aprendizagem
<b>Bloco-Endereço</b>	97.13 $\pm$ 05.86	98.62 $\pm$ 06.91
<b>Selo</b>	23.99 $\pm$ 18.65	87.04 $\pm$ 19.67
<b>Carimbo</b>	86.01 $\pm$ 23.31	95.36 $\pm$ 13.64
<b>Ruído</b>	00.63 $\pm$ 00.93	02.06 $\pm$ 01.32

**Tabela 5.29** - Resultados comparativos da segmentação por classes de Envelopes, para a base completa ECT-DB2 (n= 200) com a abordagem por dimensão fractal, para diferentes parâmetros.

<b>Abordagem por dimensão fractal</b>	<b>Bloco-Endereço (%)</b>	<b>Selo (%)</b>	<b>Carimbo (%)</b>	<b>Ruído (%)</b>
<b>[3,5]</b>	92.61	48.86	82.02	3.97
<b>[3,7]</b>	96.05	59.76	88.97	6.20
<b>[3,9]</b>	97.24	66.34	91.89	7.37
<b>[3,11]</b>	99.02	73.39	95.64	7.22
<b>[3,13]</b>	99.25	77.19	96.84	7.85
<b>Abordagem por aprendizagem</b>	98.62	87.04	95.36	2.06

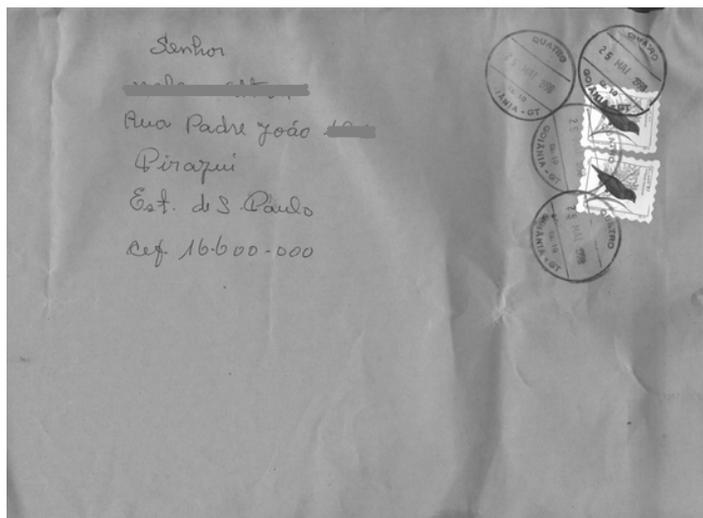


(a) Imagem amostra em níveis de cinza

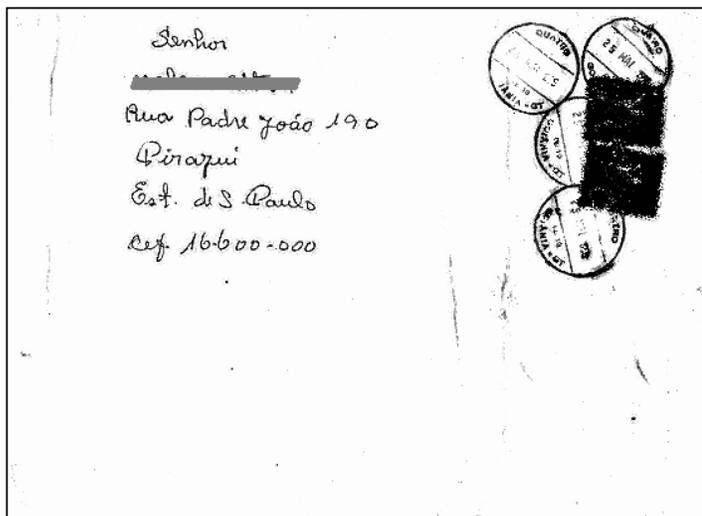


(b) Imagem segmentada ideal

**Figura 5.30.** Resultados de segmentação por binarização em imagem de envelope.



(c) Nova imagem amostra em níveis de cinza

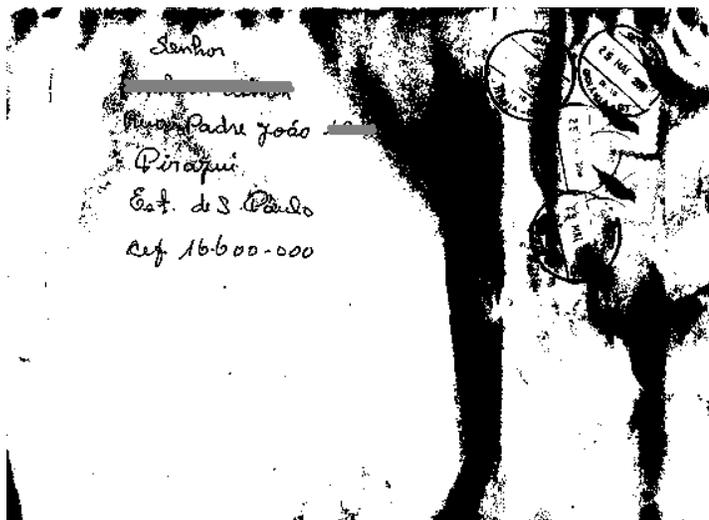


(d) Segmentação por binarização obtida por aprendizagem

Taxa de segmentação: 92,64% - Taxa de ruído 1,08%

Segmentação por classes - Bloco-Endereço: 95,77%, Carimbo: 96,67%, Selo: 84,01% -  
Ruído: 1,08%

**Figura 5.30 (cont.).** Resultados de segmentação por binarização em imagem de envelope.

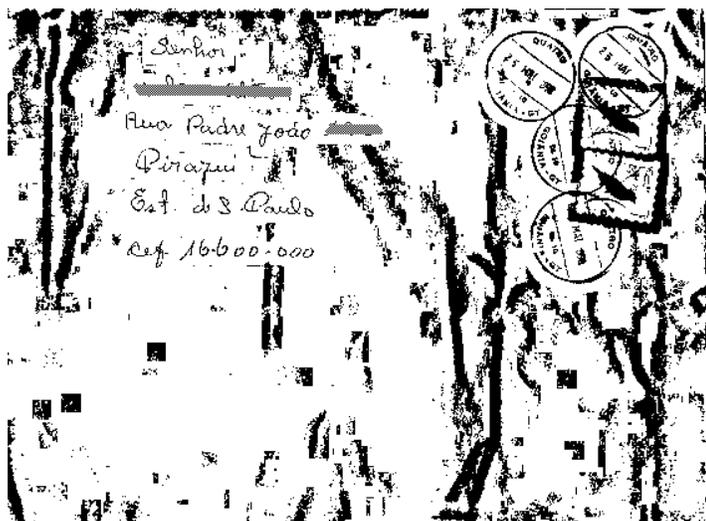


(e) Binarização obtida pelo método de Otsu

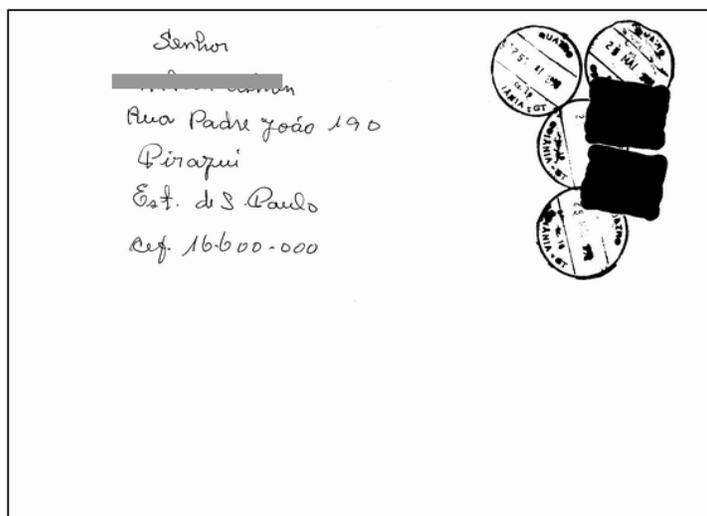


(f) Binarização obtida pelo método de Eikvil

**Figura 5.30 (cont.).** Resultados de segmentação por binarização em imagem de envelope.



(g) Binarização obtida pelo método de Bernsen



(b) Imagem segmentada ideal

**Figura 5.30 (cont.).** Resultados de segmentação por binarização em imagem de envelope. Os nomes dos destinatários foram escondidos propositalmente.

## 5.7. Discussão dos resultados

Os experimentos realizados nas seções anteriores mostraram que não existe um conjunto de características único que apresenta o melhor resultado para todas as bases de imagens testadas. Da mesma forma, o parâmetro  $k$  relativo ao número de vizinhos mais próximos que resultou em melhor desempenho foi diferente para cada caso. Portanto, a validade da fórmula empírica do parâmetro  $k$  não foi comprovada, com base nos testes realizados, como sendo um critério geral.

Nos testes com a base de imagens de assinaturas, o conjunto de características que apresentou o melhor resultado foi o conjunto Média-Variância, junto com o parâmetro  $k=11$ . Na **Tabela 5.5** observa-se que as taxas de segmentação e ruído foram em %  $\mu \pm \sigma$ , respectivamente:  $94.19 \pm 6.64$  e  $1.00 \pm 0.92$ . Este resultado foi superior ao obtido em experimentos anteriores sobre a mesma base, publicados em [LAa03], onde foram utilizadas sem uma análise de desempenho, as características Média-Variância-Média dos Gradientes e tinha se obtido uma taxa de segmentação de 90.3%.

Nos testes com a base de imagens de impressões digitais, a característica que apresentou o melhor resultado foi Média, junto com o parâmetro  $k=15$ . Na **Tabela 5.9** observa-se que as taxas de segmentação e ruído foram em %  $\mu \pm \sigma$ , respectivamente:  $94.15 \pm 4.10$  e  $17.55 \pm 0.28$ .

Nos testes com a base de imagens de imagens de ressonância magnética, os conjuntos de características que apresentaram melhor segmentação por multi-região para cada classe (WM, GM, CSF) foram diferentes. De fato, para a segmentação da massa branca (WM), o conjunto de características que apresentou o melhor resultado foi o conjunto Média-Variância, junto com o parâmetro  $k=1$ . Na **Tabela 5.13** observa-se que as taxas de segmentação e ruído foram em %  $\mu \pm \sigma$ , respectivamente:  $96.00 \pm 1.03$  e  $0.53 \pm 0.10$ . Para a segmentação da massa cinza (GM), o melhor conjunto foi o conjunto Média-Variância-Skewness, junto com o parâmetro  $k=11$  ( $97.33 \pm 1.49$  e  $0.67 \pm 0.09$ , **Tabela 5.18**) e para a segmentação do líquido (CSF), o melhor desempenho foi obtido com o conjunto Média-Variância-Skewness-Curtosis, junto com o parâmetro  $k=7$  ( $91.85 \pm 2.74$  e  $0.42 \pm 0.24$ , **Tabela 5.21**).

Nos testes com a base de imagens de envelopes de tamanho reduzido (ECT-DB1), o conjunto de características que apresentou o melhor resultado foi o conjunto Média-Variância, junto com o parâmetro  $k=1$ . Na **Tabela 5.24** observa-se que as taxas de segmentação e ruído foram em %  $\mu \pm \sigma$ , respectivamente:  $80.31 \pm 12.43$  e  $1.86 \pm 1.68$ . Os resultados das taxas de segmentação e ruído sobre a base completa (**Tabela 5.25**,  $n=200$ ) foram em %  $\mu \pm \sigma$ , respectivamente:  $87.88 \pm 11.76$  e  $3.70 \pm 2.74$ . Os resultados obtidos da segmentação por classe sobre a base ECT-DB1 completa (**Tabela 5.26**) foram, em %  $\mu \pm \sigma$ , respectivamente, Bloco-endereço:  $96.74 \pm 9.47$ , Selo:  $85.30 \pm 5.75$ , Carimbo:  $93.54 \pm 14.94$  e Ruído:  $3.70 \pm 2.74$ .

Os resultados das taxas de segmentação e de ruído sobre a base completa ECT-DB2 (**Tabela 5.27**) foram em %  $\mu \pm \sigma$ , respectivamente:  $90.59 \pm 15.05$  e  $2.06 \pm 1.32$ . Os resultados obtidos de segmentação por classe foram (**Tabela 5.28**), em %  $\mu \pm \sigma$ , respectivamente, Bloco-endereço:  $98,62 \pm 06,9$ ; Selo:  $87,04 \pm 19,67$ ; Carimbo:  $95,36 \pm 13,64$  e Ruído:  $02,06 \pm 01,32$ . A abordagem por aprendizagem apresentou melhores resultados que a abordagem por *wavelet* nas três classes de objetos (Bloco-endereço 1.49%, Selo 63.05% e Carimbo 9.35%), sendo que o ruído foi maior em 1.43%. Na comparação dos resultados com a abordagem por dimensão fractal (**Tabela 5.29**), observa-se que a mesma apresentou resultados com ruído 5.79% superior ao obtido por aprendizagem.

Um resumo comparativo dos resultados obtidos para as diversas bases de imagens é apresentado na **Tabela 5.30**. Observa-se que as características Média e Variância encontram-se presentes em quase todas as bases, o que sugere que estas características são as mais significativas.

A diferença entre as taxas de segmentação obtidas pela metodologia objeto-ruído e pela metodologia classes-ruído está na composição das classes na base de imagens ECT-DB2. De acordo com [MEN03], a grande maioria dos *pixels* dos envelopes desta base pertence ao fundo (aproximadamente 93,50% dos *pixels*). Os objetos perfazem um total de 6,50% dos *pixels*, sendo que: o bloco endereço fica com 1,50%; o selo com 4,00%; e o carimbo com 1,00% dos *pixels*. Portanto, como a classe selo é a majoritária entre os objetos e esta classe obteve a menor taxa de segmentação, faz com que a média de segmentação pela metodologia objeto-ruído caia ao valor de 90.59%.

Na **Figura 5.25** apresenta-se um exemplo de resultado de segmentação obtida por aprendizagem (d). Na mesma figura são apresentados, resultados comparativos obtidos por outros métodos de binarização conhecidos, globais e locais. Observa-se que estes métodos de binarização, Otsu (e), Eikvil (f) e Bernsen (g), não conseguem segmentar apropriadamente o fundo e introduzem artefatos na imagem segmentada. Observa-se ainda que, a abordagem de segmentação proposta, consegue segmentar a região correspondente ao selo como área preta, de acordo à aprendizagem do par de imagens amostra (a) / ideal (b).

Os resultados obtidos nos experimentos são promissores, considerando-se que não foi aplicado pós-processamento às imagens resultantes, tais como filtragens ou operações morfológicas. Os resultados mostram que a abordagem proposta é robusta para diversos tipos de imagens.

As imagens segmentadas foram avaliadas segundo uma metodologia objetiva (*pixel a pixel*) de forma quantitativa, sem a necessidade de avaliações visuais, que são subjetivas.

**Tabela 5.30.** Comparação dos resultados obtidos para as diversas bases de imagens.

<b>Melhores conjuntos de características, parâmetros <math>k</math> obtidos e resultados dos testes para as diversas bases de imagens (% m <math>\pm</math> s)</b>				
<b>Base</b>	<b>Características</b>	<b><math>k</math></b>	<b>Taxa de Segmentação</b>	<b>Taxa de Ruído</b>
Assinaturas	Med-Var	11	94.19 $\pm$ 6.64	1.00 $\pm$ 0.92
Impressões Digitais	Média	15	94.15 $\pm$ 4.10	17.55 $\pm$ 7.28
MRI – WM	Med-Var	1	96.00 $\pm$ 1.03	0.53 $\pm$ 0.10
MRI – GM	Med-Var-Ske	11	97.33 $\pm$ 1.49	0.67 $\pm$ 0.09
MRI - CSF	Med-Var-Ske-Cur	7	91.85 $\pm$ 2.74	0.42 $\pm$ 0.24
Envelopes Postais	Med-Var	1	80.31 $\pm$ 14.30	1.86 $\pm$ 1.68

O tempo de processamento da aprendizagem e segmentação foi variável nas diferentes bases. Este tempo não é dependente só do tamanho da imagem, mas também do número de

iterações até a solução dos conflitos, ou o cumprimento dos critérios de parada e do tamanho das janelas. Tempos de processamento típicos na aprendizagem por estrutura de árvore binária, na execução remota em um servidor Intel Pentium III Dual 1.2. GHz, 3 GB de memória e com Sistema Operacional Linux RedHat são apresentados na **Tabela 5.31** e de forma gráfica na **Figura B.7** do **Apêndice B**. Os tempos de processamento típicos na segmentação são aproximadamente 10% da aprendizagem.

**Tabela 5.31.** Tempos de processamento típicos na aprendizagem por estrutura de árvore binária

Tamanho da imagem (em pixels)	188 x 188	376 x 376	1500 x 1500	3000 x 3000
Tempo de processamento (em segundos)	14	36	474	1921

## 5.8. Comentários Finais

Neste capítulo foi avaliada, cada uma das características que foram propostas na abordagem de segmentação por aprendizagem. Das sete características inicialmente propostas, as que apresentaram melhor desempenho foram, Média, Variância, *Skewness* e Curtosis. Essas foram utilizadas junto com o Nível de cinza do *pixel*. No entanto, o conjunto de características de melhor resultado variou para cada base. Também, foram apresentados de forma objetiva, os resultados de segmentação, através de uma comparação *pixel a pixel* entre a imagem obtida e o resultado ideal da mesma (*ground truth*). No próximo Capítulo serão discutidas as conclusões do trabalho, verificando se a abordagem proposta e os resultados obtidos encontram-se dentro do escopo da proposta e do desafio inicial.

## Capítulo 6

### Conclusões

Este trabalho apresenta uma abordagem de segmentação baseada em aprendizagem computacional. A aprendizagem que constitui a base da Proposta entre o par de imagens amostra e ideal, faz-se calculando um vetor de características para cada *pixel* em uma janela de tamanho variável adaptada automaticamente para eliminar possíveis redundâncias e conflitos. O “sucesso” do processo de aprendizagem é medido pelo algoritmo de seleção automática do tamanho da janela, através de três critérios: erro zero, erro mínimo ou erro constante.

O principal desafio, foi desenvolver uma abordagem genérica de segmentação de imagens. Os resultados dos experimentos realizados em imagens de assinaturas (taxas de segmentação e ruído em  $\% \mu \pm \sigma$ , respectivamente:  $94.19 \pm 6.64$  e  $1.00 \pm 0.92$ ), de envelopes postais manuscritos ( $80.31 \pm 14.30$  e  $1.86 \pm 1.68$ ), impressões digitais ( $94.15 \pm 4.10$  e  $17.55 \pm 7.28$ ) e imagens não realçadas de ressonância magnética (WM:  $96.00 \pm 1.03$  e  $0.53 \pm 0.10$ , GM:  $97.33 \pm 1.49$  e  $0.67 \pm 0.09$ , CSF:  $91.85 \pm 2.74$  e  $0.42 \pm 0.24$ ) mostraram que a abordagem proposta é eficiente e promissora na segmentação por binarização para diversos tipos de imagens.

Foi feito também um estudo do desempenho das diversas características que foram inicialmente propostas. Os experimentos demonstraram que não existe um único conjunto de características eficiente para todas as bases de dados testadas. De fato, na **Tabela 5.30** pode-se observar que os melhores conjuntos de características foram, para cada base, respectivamente:

- Assinaturas, Média-Variância;
- Impressões Digitais, Média;
- MRI-WM, Média-Variância;
- MRI-GM, Média-Variância-Skewness;
- MRI-CSF, Média-Variância-Skewness-Curtosis e

- Envelopes Postais, Média-Variância.

Os resultados obtidos com os diversos valores testados do parâmetro  $k$  sugerem o mesmo comportamento. Assim, na **Tabela 5.30** também se observa que os melhores conjuntos de características foram, para cada base, respectivamente: Assinaturas, Média-Variância; Impressões Digitais, Média; MRI-WM, Média-Variância; MRI-GM, Média-Variância-Skewness; MRI-CSF, Média-Variância-Skewness-Curtosis e Envelopes Postais, Média-Variância pode-se observar que os melhores valores de  $k$  foram, para cada base, respectivamente:

- Assinaturas,  $k = 11$ ;
- Impressões Digitais,  $k = 15$ ;
- MRI-WM,  $k = 1$ ;
- MRI-GM,  $k = 11$ ;
- MRI-CSF,  $k = 7$  e
- Envelopes Postais,  $k = 1$ .

Essas evidências levam à conclusão, que no projeto de um sistema especialista de segmentação de imagens, deverá ser feita uma análise prévia das características relevantes e o parâmetro  $k$  para aquela base específica. No entanto, a abordagem proposta, permite que essas análises prévias, na etapa de aprendizagem, possam ser automatizadas.

O desempenho da abordagem de segmentação por aprendizagem foi medido de forma objetiva através de comparações quantitativas entre os resultados das imagens obtidas e as desejadas. Foram também feitas, comparações de desempenho com outras abordagens (*wavelet* e dimensão fractal) para uma mesma base de imagens de envelopes postais. Em comparação com a abordagem *wavelet*, a abordagem por aprendizagem mostrou melhor desempenho na segmentação das três classes de objetos, e principalmente, na classe selo. Na comparação com a abordagem por dimensão fractal, a abordagem por aprendizagem resultou em uma taxa média de ruído inferior.

De maneira geral, as principais vantagens do método proposto são:

- A aprendizagem pode ser realizada a partir de um único par de imagens amostra/ideal, o que confere simplicidade à abordagem.
- As imagens amostra/ideais para aprendizagem podem ser parciais.
- O método é invariante à translação das imagens. Ele não armazena os padrões dos pixels na janela nem a posição relativa da janela dentro da imagem e sim as características que relacionam os pixels entre si.
- A escolha do tamanho da janela é completamente automática e não requer o uso de parâmetros heurísticos.
- O processo de segmentação é automático e não requer a interação de um usuário experiente, após o estágio de aprendizagem.
- Essa abordagem também permite a segmentação em várias classes.
- A abordagem é genérica e permite a aplicação em vários tipos de imagens.
- É possível transformar o conhecimento subjetivo do operador humano num processo automático e reproduzi-lo.

As desvantagens do método proposto são:

- O elevado custo computacional requerido pela abordagem apresentada, principalmente para a fase de aprendizagem.
- O par de imagens amostra / ideal para escolhido para a aprendizagem, deve ser do mesmo tipo das imagens de segmentação.
- A qualidade dos resultados é dependente das características que foram escolhidas.
- É ainda preciso de um usuário experiente para a definição da imagem ideal na fase de aprendizagem. Esta desvantagem é discutível, considerando-se que o objetivo da abordagem é a reprodução pelo sistema do que o especialista gostaria de obter.

# Capítulo 7

## Trabalhos futuros

Apresentam-se a seguir, propostas de trabalhos futuros (alguns já em andamento) objetivando o aproveitamento da pesquisa até agora realizada.

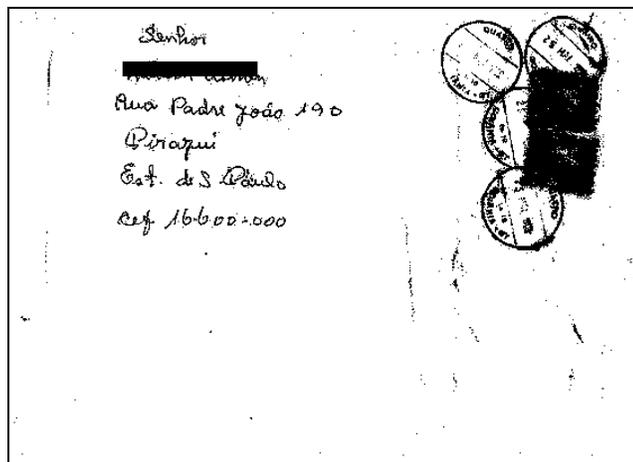
### 7.1. Pós-processamentos

Os resultados apresentados ao longo deste trabalho não têm nenhum pós-processamento para realçar as imagens. Portanto, é válida a hipótese de que filtragens ou realces nas imagens resultantes, poderiam melhorar a qualidade destas.

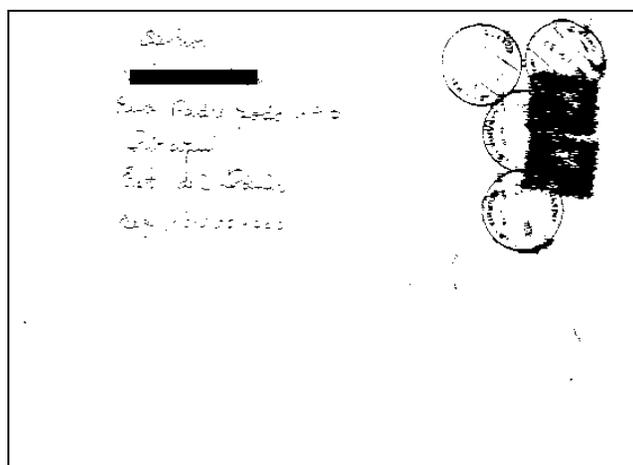
A título de exemplo do anteriormente exposto, na **Figura 7.1** foi realizado um simples pós-processamento da imagem resultante, utilizando-se duas ferramentas de morfologia matemática: a erosão binária e a reconstrução morfológica binária. O processo realizado foi o seguinte:

- 1) Erosão da imagem em (a), utilizando-se o elemento estruturante vertical com 3x1 *pixels* e 1 iteração.
- 2) Reconstrução morfológica da imagem (b) erodida, utilizando-se como máscara a imagem (a).

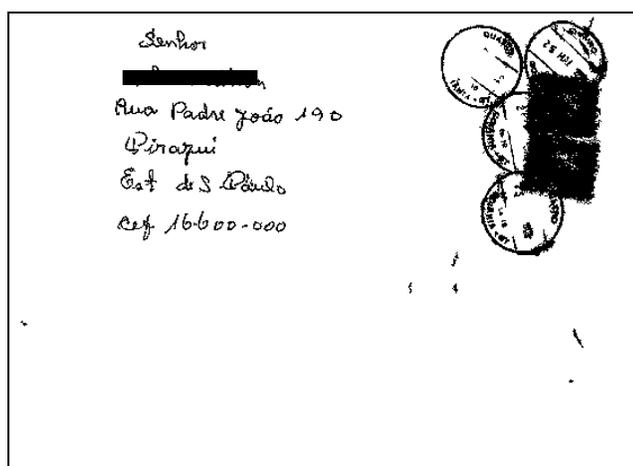
Com o exemplo anterior, pretende-se justificar a necessidade de incluir uma etapa de pós-processamento da imagem segmentada. Portanto, são necessárias pesquisas que identifiquem quais os melhores conjuntos de filtragem ou realce para cada uma das bases de imagens utilizadas.



(a) Imagem segmentada por aprendizagem



(a) Erosão da imagem segmentada com EE 3x1, 1 iteração

(b) Reconstrução da imagem erodida com a imagem segmentada como máscara  
**Figura 7.1.** Exemplo de pós-processamento por reconstrução morfológica.

## 7.2. Pesquisa de Características Adicionais

No presente trabalho, foram inicialmente propostas 8 características para a segmentação de imagens e logo comprovadas que 3 delas não apresentaram resultados eficientes para as bases utilizadas. Seguindo o mesmo raciocínio, as pesquisas devem ser continuadas na busca e validação de novas e melhores características para a aprendizagem.

Algumas das características propostas, para avaliação futura são: o laplaciano, os coeficientes de Fourier, os momentos de Zernike e a dimensão fractal.

## 7.3. Implementação de Algoritmos Mais Eficientes de Busca dos $k$ Vizinhos Mais Próximos

O elevado custo computacional requerido pela abordagem apresentada, principalmente para a fase de aprendizagem, exige a continuidade de pesquisas que objetivem a diminuição dos tempos de processamento.

No decorrer da pesquisa, o autor deparou-se em certos momentos, com a necessidade de decidir pela simplificação de algoritmos em detrimento da qualidade dos resultados, a fim de diminuir os tempos de processamento envolvidos. Porém decidiu-se primar, em todas as situações, pela qualidade dos resultados, levando-se em conta as seguintes considerações:

- O permanente avanço na tecnologia da micro-eletrônica, fazendo com que de acordo com a lei de Moore, a capacidade de processamento dos computadores dobre a cada 18 meses.
- Existem na literatura, diversos algoritmos já desenvolvidos para melhorar os tempos de busca em diferentes estruturas de dados. No **APÊNDICE B**, apresenta-se um estudo comparativo do tempo de processamento do algoritmo de busca da abordagem original, com outras técnicas e que foram realizadas como parte do presente projeto.

É recomendável, continuar as pesquisas que objetivem a diminuição dos tempos de processamento da abordagem apresentada.

#### **7.4. Implementação de Outras Ferramentas de Similaridade**

Propõe-se a pesquisa e avaliação de outras ferramentas de classificação, como por exemplo, o coeficiente de correlação entre as características.

#### **7.5. Implementação de Outras Métricas**

A presente abordagem utiliza como métrica a distância euclidiana. Porém, faz-se necessária, a avaliação dos resultados incorporando novas métricas, por exemplo, a distância de Mahalanobis.

#### **7.6 Experimentos de Segmentação de Outras Bases de Imagens**

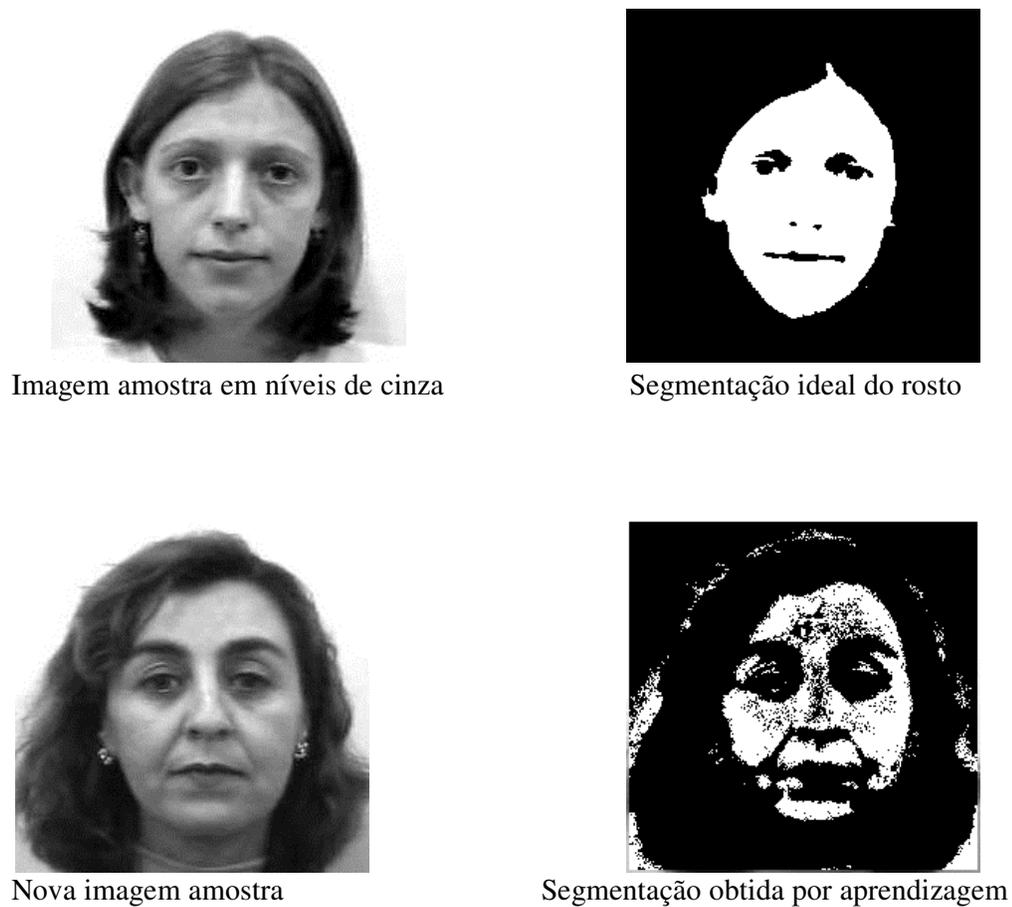
##### **7.6.1. Segmentação de Rostos**

Foram iniciados testes de aplicação do algoritmo de segmentação por aprendizagem, na exclusão dos olhos, nariz, boca e cabelos, do restante do rosto. Os resultados preliminares sobre imagens obtidas no site web do FBI [FBI03], sugerem que a abordagem é promissora também para esta aplicação. Espera-se que as melhoras nos processos sejam decorrentes da utilização de novas características a serem estudadas. Um exemplo do estágio atual dos testes é mostrado na **Figura 7.2**.

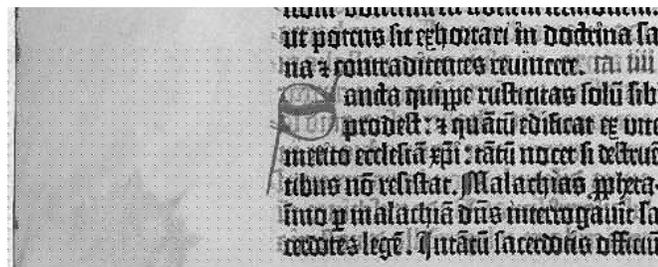
##### **7.6.2. Segmentação de Caracteres Coloridos em Imagens da Bíblia de Gutemberg**

Foram iniciadas pesquisas, para a segmentação de caracteres coloridos em imagens da Bíblia de Gutemberg. O objetivo é a separação destes caracteres que foram coloridos manualmente após a impressão das folhas, dos caracteres originais. O trabalho visa validar o desempenho da abordagem em imagens de documentos antigos.

Um exemplo dos resultados preliminares obtidos neste trabalho está na **Figura 7.3** e foi publicado em [LAa04].



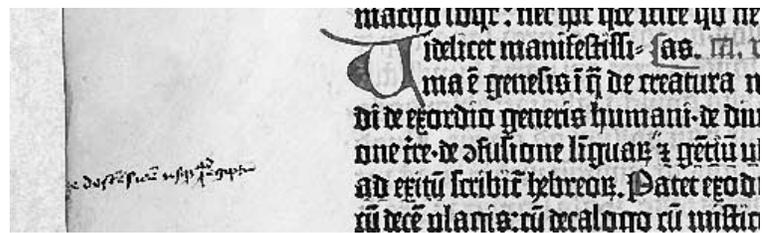
**Figura 7.2** Resultados preliminares de segmentação de rosto.



(a) Imagem amostra em níveis de cinza



(b) Imagem ideal segmentada



(c) Nova imagem amostra em níveis de cinza



(d) Segmentação da imagem obtida por aprendizagem

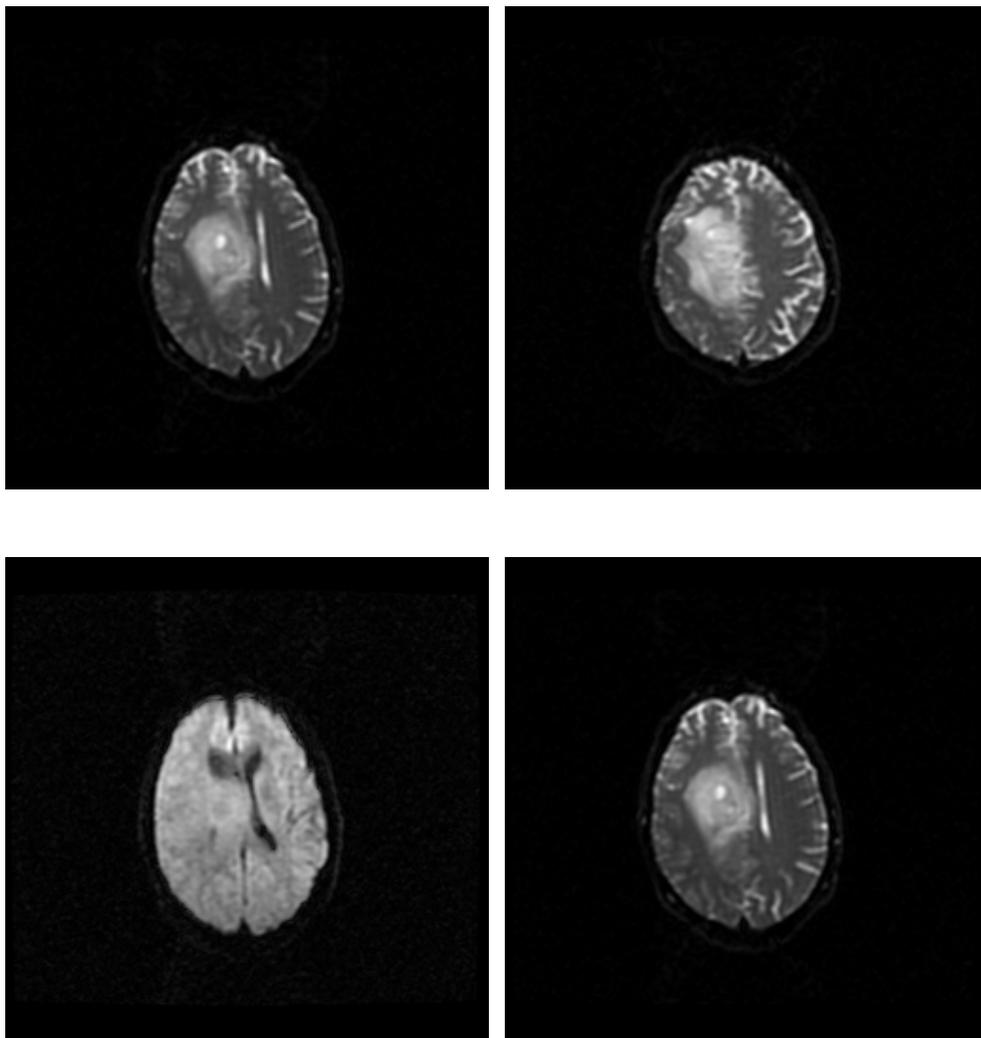
**Figura 7.3.** Resultados preliminares da segmentação em uma página da Bíblia de Gutemberg

### 7.6.3. Quantificação de Resultados de Segmentação de MRI de Cérebro com Tumor

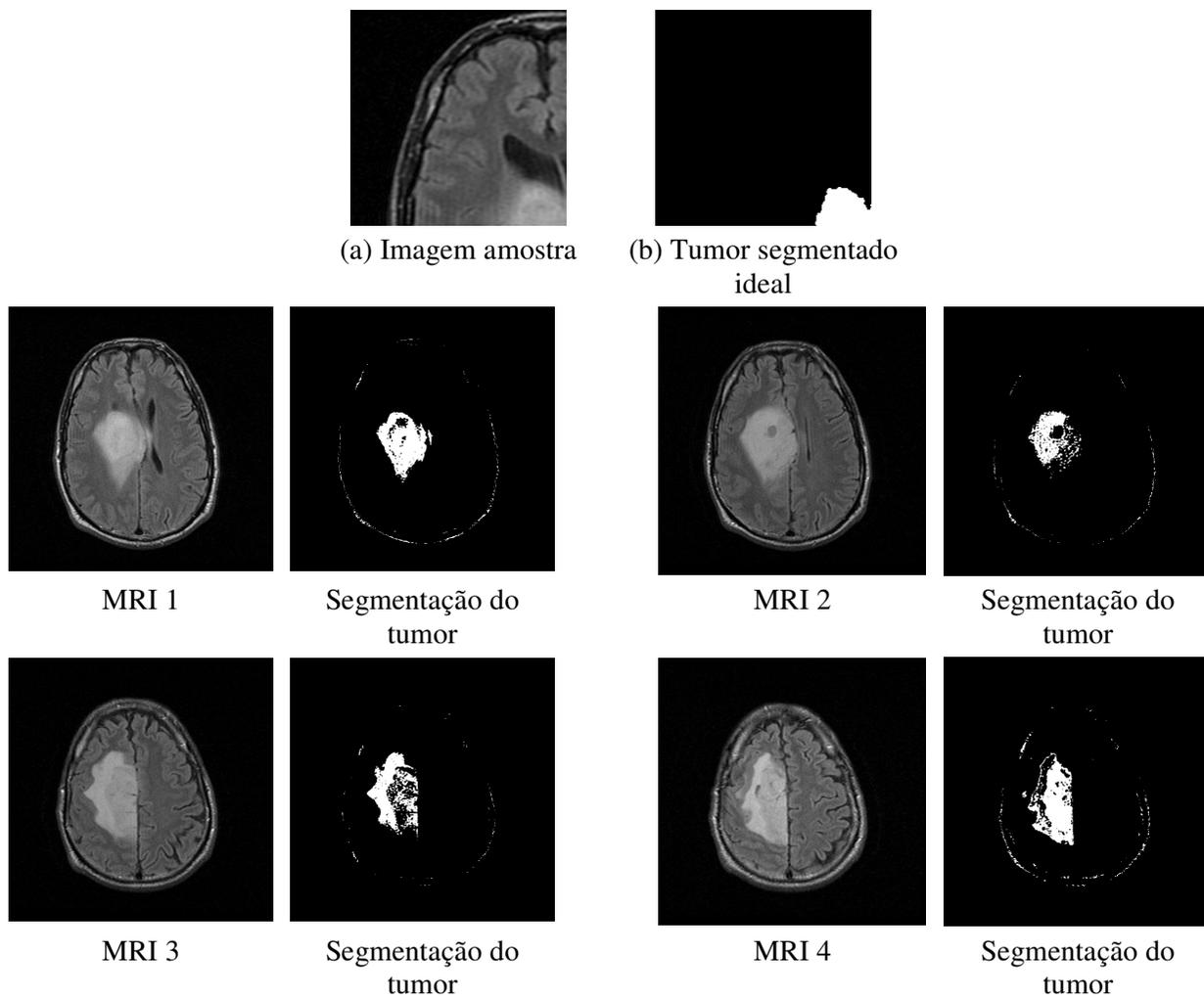
Foram realizados testes preliminares de segmentação de MRI de cérebro com tumor. A base utilizada é composta de 500 imagens de RM disponíveis no Laboratório Unificado de Ciência da Imagem do PPGIA/PUCPR e que foram fornecidas pelo Dr. André F Gomes, do Centro de Diagnóstico DAPI em Curitiba, PR. A base contém imagens reais de MRI com tumor das quais não se têm as classes *ground truth* correspondente. A **Figura 7.4.** apresenta exemplos da base de imagens de ressonância magnética de cérebro com tumor.

A avaliação feita nos testes preliminares foi visual, por não se contar com as imagens ideais *ground truth*. O objetivo foi extrair as classes correspondentes à anomalia e à massa cinza, as quais representam os tecidos mais doentes. A etapa de aprendizagem foi realizada, utilizando-se uma imagem de uma amostra parcial e sua correspondente imagem ideal segmentada, para cada classe (**Figura 7.5** (a) à (b) e **Figura 7.6** (a) à (b)). Nos testes preliminares, o conjunto de características utilizado para a aprendizagem foi o conjunto Média-Variância-Média dos Gradientes. Os resultados da segmentação das classes GM e tumor são apresentados nas **Figuras 7.5 e 7.6**, para quatro imagens de MR completas do cérebro. É possível observar visualmente que o método proposto recuperou as duas classes como era esperado.

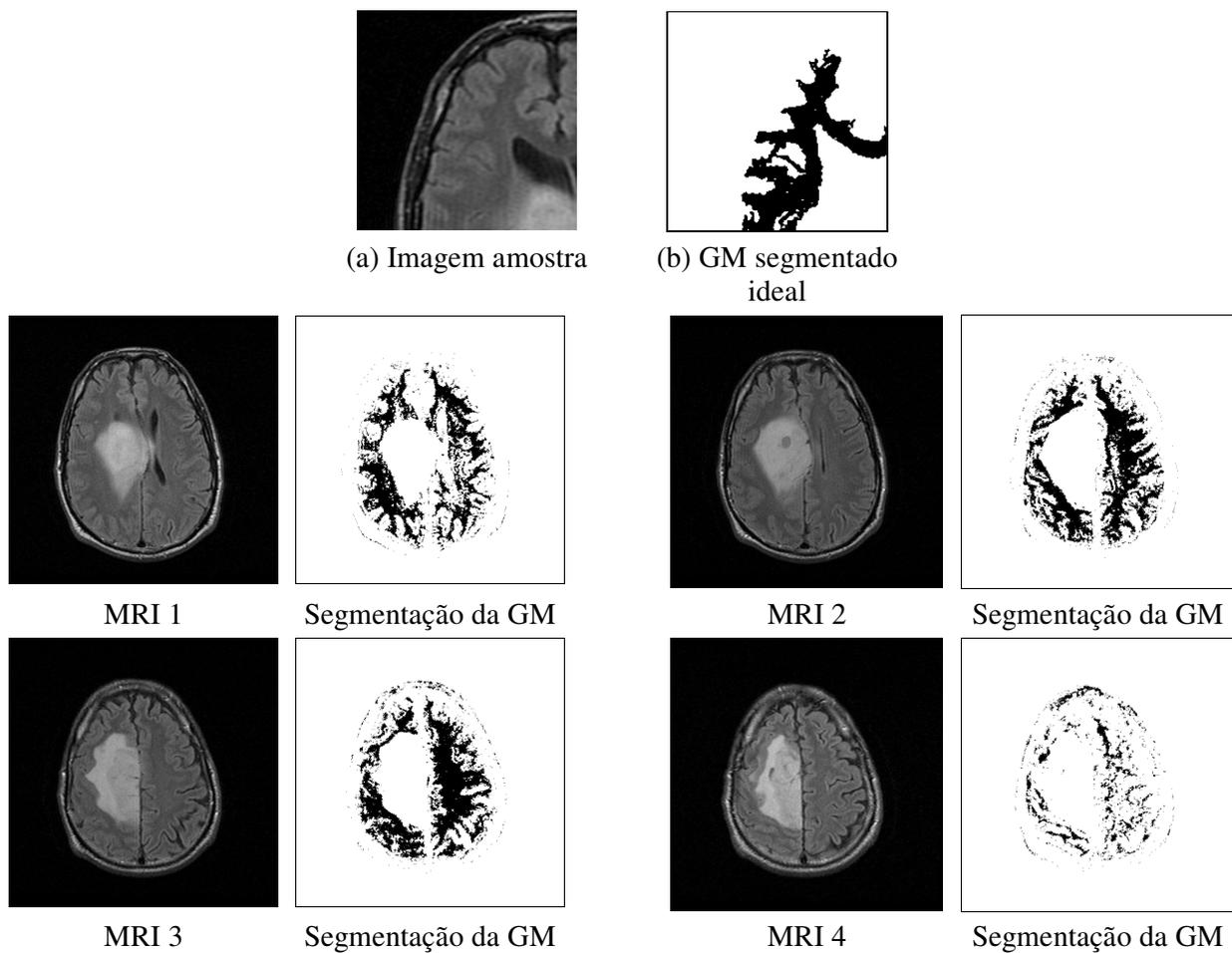
A quantificação dos resultados da segmentação de MRI de cérebro com tumor é proposta como trabalho futuro. É necessária a participação de um médico radiologista para segmentar manualmente as classes tumor e massa cinza. Essas imagens segmentadas manualmente pelo especialista serão consideradas *ground truth*. A medição *pixel a pixel* entre os resultados obtidos pela abordagem proposta e as imagens ideais, poderá indicar qual a melhor combinação de características e parâmetro  $k$ .



**Figura 7.4.** Exemplos da base de imagens de ressonância magnética de cérebro com tumor.



**Figura 7.5.** Segmentação de tumores em imagens MR cerebrais, sem realce por aprendizagem.



**Figura 7.6.** Segmentação da massa cinza GM de imagens cerebrais não realçadas com tumor por aprendizagem.

## Referências Bibliográficas

- [ABU89] ABUTALEB, A. S. *Automatic Thresholding of Gray-Level Pictures Using Two Dimensional Entropy*. Computer Graphics & Image Processing, 47(1) – pp 22-32, July 1989.
- [AND98] ANDRADE, M. C. *Um método Topológico de Segmentação de imagens por Atributos*. Ph. D. Thesis, Universidade Federal de Minas Gerais, DCC-UFMG, 1998.
- [AND00] ANDO, S. *Consistent gradient operators*. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-22(3):252-264, March, 2000.
- [BAL81] BALLARD, D. H. *Generalizing the Hough Transform to detect arbitrary shapes*. Pattern Recognition, 13(2): 111-122, 1981.
- [BAR97] BARRERA, J.; DOUGHERTY, E. R.; TOMITA, N. S. *Automatic Programming of Binary Morphological Machines by Design of Statistically Optimal Operators in the Context of Computational Learning Theory*. Electronic Imaging, 6(1):54-67, January 1997.
- [BAT95] BARRERA J.; TOMITA N. S.; SILVA F. S. C.; TERADA R. *Automatic Programming of Binary Morphological Machines by PAC Learning*. SPIE, Proceedings Neural, Morphological and Stochastic Methods in Image and Signal Processing, (2568), pp 233-244, 1995.
- [BDT97] BARRERA J.; DOUGHERTY, E. R.; TOMITA N. S. *Automatic Programming of Binary Morphological Machines by Design of Statistically Optimal Operators in the Context of Computational Learning Theory*. Electronic Imaging, 6(1): 54-67, January 1997.

- [BER86] BERNSEN, J. *Dynamic thresholding of gray-level images*. Proc. Eight Int'l Conf. On Pattern Recognition, Paris, France, pp. 1251-1256, 1986.
- [BER98] BERCHTOLD, S.; KEIM, D. A. *High Dimensional Index Structure*, Tutorial, p56, 1998.
- [BEU90] BEUCHER, S. *Segmentation d'image et Morphologie Mathématique*. Centre de Morphologie Mathématique, École Nationale Supérieure de Mines de Paris - ENSMP, Ph. D. Thesis, 1990.
- [BHA00] BHANU, B.; FONDER, S. *Learning Based Interactive Image: Segmentation*. International Conference on Pattern Recognition, Barcelona, Spain, 1, pp. 299-302, 2000
- [BHC93] BEZDEK, J.; HALL, C. L. O.; CLARKE, L. P. *Review of MR image segmentation techniques using pattern recognition*. Med Phys 20, 1033-1048, 1993.
- [BRE98] BREJL, M.; SONKA, M. *Edge-based Image Segmentation: Machine Learning from Examples*. Proceedings - IEEE World Congress on Computational Intelligence, pp. 814-819, 1998.
- [BRU01] BRUN, M.; HIRATA JR, R.; BARRERA, J.; DOUGHERTY, E. R. *Hybrid human-machine Non-linear Filter Design Using Envelopes*. Proceedings of the SIBGRAPI 2001, V1 pp. 106-111, 2001.
- [CFA91] COHEN, F. S.; FAN, Z.; ATTALI, S. *Automated inspection of textile fabrics using textural models*. PAMI V13, PP. 192-202, 1991.
- [CHA84] CHIEN, C. H.; AGGARWAL, J. K. *A normalized quadree representation*. Computer Vision, Graphics and Image Processing, 26:331-346, 1984.

- [CMV86] CHEEVASUVIT, F.; MAITRE, H., & VIDAL-MADJAR, D. *A robust method for picture segmentation based on a split and merge procedure*. Computer Vision, Graphics and Image Processing, 34: 268-281, 1986.
- [COH67] COVER, T. M. & HART P. E. *Nearest Neighbor Pattern Classification*. IEEE T. Information Theory, vol. IT-13, no. 1, pp. 21-27, 1967.
- [COM02] CORNUÉJOLS, A. & MICLET, L., *Apprentissage Artificiel: Méthodes et Algorithmes*. Eyrolles, 2002.
- [COH91] COHEN, L. D. *On active contour models and balloons*. CGVIP: Image Understanding, 53(2): 211-218, March, 1991.
- [CVC95] CLARKE L. P.; VELTHUIZEN R. P.; CAMACHO M. A.; HEINE J. J.; VALDYANATHAN M.; HALL L. O.; THATCHER R. W. & SILBINGER M. L. *MRI Segmentation: Methods and Applications*. Magnetic Resonance Imaging 13, 343-368, 1995.
- [DOL92] DOUGHERTY, E. R. & LOCE, R. P. *Optimal Binary Differencing Filters: Design, Logic Complexity, Precision Analysis, and Application to Digital Document Processing*. Electronic Imaging, 5(1):66-86, January 1996.
- [DOU92] DOUGHERTY, E. R. *An Introduction to Morphological Image Processing*. SPIE, Vol. TT09, 1992.
- [DOU97] DOUGHERTY, E. R. & BARRERA, J. *Bayesian Design of Optimal Morphological Operators Based on Prior Distributions for Conditional Probabilities*. Acta Stereologica, Bellingham, 16(3):167-174,1997.
- [DOU98] DOUGHERTY, E. R. *Random Processes for Image and Signal Processing*. SPIE and IEEE Presses, Bellingham, 1998.

- [DOU99] DOUGHERTY, E. R. *The Granulometric Size Density in Filter Optimization*. J. Stolfi and C.L. Tozzi, editors, Proc. Of Sibgrapi'99, pp. 257-264, Campinas, SP, Brazil, 1999.
- [EIK91] EIKVIL L., TAXT T. & MOEN K.. *A Fast Adaptive Method for Binarization of Document Images*. Proc First Int'l Conf Document Analysis and Recognition, Saint Malo, France, pp. 435-443, 1991.
- [EIT04] EITERER, L. F., FACON, J. & MENOTI, D. *Segmentation of Envelope Address Blocks through Fractal-Based Approach*. 5<sup>th</sup> Iberoamerican Meeting on Optics (RIO) and 8<sup>th</sup> Latin American Meeting on Optics, Lasers and Their Applications (OPTILAS), Proceedings SPIE Vol. 5622, pp. 1265-1270, 2004.
- [ETM91] EIKVIL, L, TAXT. T. & MOEN, K. *A fast adaptive method for binarization of document images*. Proc First Int'l Conf Document Analysis and Recognition, Saint Malo, France, pp. 435-443, 1991.
- [FAC01] FACON, J. *Metodologia de Avaliação de Abordagens de Segmentação de Imagens*. Technical Report, Curitiba: PUCPR, 138p, 2001.
- [FAC03] FACON, J. *Binarização de imagens*. Technical Report, PUCPR, <http://www.ppgia.pucpr.br/~facon/IndexPrincipalBrBinarizacao.htm>, 2003.
- [FBI03] FEDERAL BUREAU OF INVESTIGATION. <http://www.fbi.gov/mostwant.htm>, 2003, acessado em 17 de Maio de 2003
- [FFR96] FOX N. C., FREEBOROUGH P. A. & ROSSOR M. N. *Visualization and quantification of rates of atrophy in Alzheimer's*. Lancet 348, 94-7, 1996
- [FHB00] FLORES, F. C.; HIRATA JR., R.; BARRERA, J.; LOTUFO, R. A. & MEYER, F. *Morphological Operators for Segmentation of Color Sequences*. In: SIBGRAPI'2000, 2000, Gramado. SIBGRAPI'2000. IEEE, v.1. p.300 – 307, 2000.

- [FHG01] FLETCHER-HEATH L. M.; HALL L. O.; GOLDFOG D. B. & MURTAGH F. R. *Automatic segmentation of non-enhancing brain tumors in magnetic resonance images*. Artificial Intelligence in Medicine 21, 43-63, 2001.
- [FUF85] FUKUNAGA K. & FLICK T. E. *The 2-NN Rule for More Accurate NN Risk Estimation*. IEEE T. Pattern Anal. Machine Intell. Machine Intell., vol. PAMI-7, no. 1, pp. 107-112, 1985.
- [FVC02] <http://bias.csr.unibo.it/fvc2002/databases.asp>. *Finger-print Verification Competition – FVC2002* – BioLab – University of Bologna, acessado em 17 de Maio de 2003
- [GGU00] GE Y.; GROSSMAN R. I.; UDUPA J. K.; BABB J. S.; KOLSON D. L. & MCGOWAN J. C. *Magnetization transfer ratio histogram analysis of gray matter in relapsing-remitting multiple sclerosis*. AJNR Am J Neuroradiol 22, 470 –5, 2000.
- [GOW92] GONZALEZ, R. C. & WOODS, R. E. *Digital Image Processing*. Addison Wesley, 1992
- [GRI91] GRIMAUD, M. *La géodésie numérique en morphologie mathématique. Application a la détection automatique de micro calcifications en mammographie numérique*. Centre de Morphologie Mathématique, ENSMP, Ph. D. Thesis, 1991.
- [HAR85] HARALICK, R.M. & SHAPIRO, L.G. *Image Segmentation Techniques*. CVGIP, Vol 29, No. 1, 100-132, 1985.
- [HAR96] HARVEY, N. R. & MARSHAL, S. *The Use of Genetic Algorithms in Morphological Filter Design*. Signal Processing: Image Communication, 8(1):55-71, January 1996.
- [HDB00] HIRATA JR., R.; DOUGHERTY, E. R. & BARRERA, J. *Some applications of aperture filters*. In: ISMM'2000, 2000, San Jose, Mathematical Morphology and its Applications to Image Processing. Dordrecht: Kluwer Academic, 2000.

- [HEW88] HE, D-C.; WANG, L. & GUIBERT, J. *Texture discrimination based on an optimal utilization of texture features*. Pattern Recognition V21, No. 2, pp. 141-146, 1988.
- [HID00] HIRATA JR., R.; DOUGHERTY, E. R. & BARRERA, J. *Design of gray-scale nonlinear filters via multiresolution aperture*. In: EUSIPCO 2000, 2000, Tampere. Proceedings of EUSIPCO 2000, v.IV. p.1905 – 1908, 2000.
- [HIR00] HIRATA, N. S. T. *Projeto Automático de Operadores: Explorando Conhecimentos a Priori*, Ph. D. Thesis, IME/USP, São Paulo, 2000.
- [JAI00] JAIN, A. K. *et al. Statistical Pattern Recognition: A Review*. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 22, No. 1, 2000.
- [JOO90] JOO, H.; HARALICK, R. M. & SHAPIRO, L. G. *Toward the Automatic Generation of Mathematical Morphology Procedures Using Predicate logic*. Proc. Of the Third International Conference on Computer Vision, pp. 156-165, Osaka, Japan, 1990.
- [JOO91] JOO, H. *Automatic Morphology*. Ph. D. Thesis, University of Washington, 1991.
- [JUS01] JUSTINO, E. J. R. *O grafismo e os modelos escondidos de Markov na verificação automática de Assinaturas*. Ph. D. Thesis, PUCPR, Curitiba, 2001.
- [KIH97] KIM, Y. K. & HAN, J. H. *Fuzzy K-NN Algorithm using Modified K-Selection*. Proceedings of FUZZ-IEEE/IFES'95, pp.1673-1680, 1995.
- [KIM97] KIM, H. Y. *Quick Construction of Efficient Morphological Operators by Computational Learning*. Electronics Letters, 33(4), pp. 286-287, 1997.
- [KIY97] KIM, H. Y. *Construção Automática de Operadores Morfológicos por Aprendizagem Computacional*. Ph. D. Thesis, EP/USP, São Paulo, 1997.

- [KIM99] KIM, H. Y. *Segmentation-Free Printed Character Recognition by Relaxed Nearest Neighbor Learning of Windowed Operator*. Proceedings of XII Brazilian Symposium on Computer Graphics and Image Processing, SIBGRAPI'99, Campinas, Brazil, pp. 198-204, 1999.
- [KKR99] KANGAR-PARSI, B. & ROSENFELD A. *Optimally isotropic laplacian operator*, IEEE Transactions on Image Processing, 10(8): 1467-1472, October, 1999.
- [KOF35] KOFFKA, K. *Principles of Gestalt Psychology*. Publ. Lund Humphries, London. Chapter 1, 1935.
- [KOH29] KÖHLER, W. *Gestalt Psychology*. Liversight, New York, 1929.
- [KWT88] KASS, M.; WITKIN, A. & TERZOPOULOS, D. *Snakes: Active contour models*. International Journal of Computer Vision, 1(4): 321-331, 1988.
- [LAa03] LEGAL-AYALA, H. A. & FACON, J. *Image Segmentation by Learning Approach*. Proceedings of ICDAR 2003 - International Conference on Document Analysis and Recognition, Edinburgh, Scotland, v2 pp. 819-823, 2003.
- [LAb03] LEGAL-AYALA, H. A. & FACON, J. *Comparative Results of Image Segmentation By Learning in Different Types of Images*. Proceedings of ICISP'2003 - IAPR International Conference on Image and Signal Processing, Agadir, Morocco, v2 pp 381-388, 2003.
- [LAc03] LEGAL-AYALA, H. A. & FACON, J. *Segmentation Approach by Learning: Different Image Applications*. Proceedings of ICIAP 2003 - IAPR 12th International Conference on IMAGE ANALYSIS AND PROCESSING, Mantova, Italy, pp. 600-604, 2003.
- [LAa04] LEGAL-AYALA, H. A. & FACON, J. *Unified Approach for Document Segmentation*. 5<sup>th</sup> Iberoamerican Meeting on Optics (RIO) and 8<sup>th</sup> Latin American

- Meeting on Optics, Lasers and Their Applications (OPTILAS), Proceedings SPIE Vol. 5622, pp. 1277-1282, 2004.
- [LAb04] LEGAL-AYALA, H. A. & FACON, J. *Automatic Segmentation of Brain MRI Through Learning by Example*. In: Proceedings of 11<sup>th</sup> International Conference on Image Processing (ICIP), IEEE Signal Processing Society, Singapore, 24-27 October, 2004.
- [LEA91] LEGAL-AYALA, H. A. *Análise e Classificação Automática de Impressões Digitais*. Master Thesis, CPGEII, CEFET-PR, 1991.
- [LIL93] LI C. H. & LEE C.K.. *Minimum Cross Entropy Thresholding*. Pattern Recognition, v.26, no 4, pp 616 -626, 1993.
- [LOM87] LOIZOU G. & MAYBANK S. J. *The Nearest Neighbor and the Bayes Error Rates*. IEEE T. Pattern Anal. Machine Intell., vol. PAMI-9, no. 2, pp. 254-262, 1987.
- [LSC03] LEANDRO, J., J., G.; SOARES, J., V., B. & CESAR Jr., R., M. *Blood Vessels Segmentation in Non-Mydriatic Images Using Wavelets and Statistical Classifiers*. Proc. Brazilian Conference on Computer Graphics, Image Processing - Sibgrapi03, IEEE Computer Society Press, p.262 – 269, 2003.
- [MAC04] MACHADO L., B. *Estudo e Implementação de Estruturas de Otimização do Processo De Aprendizagem*. Relatório Projeto PIBIC/CNPq, PUCP-PR, 2004.
- [MAR87] MARPLE S. L. *Digital Spectral Analysis with Applications*. Prentice Hall, 1987.
- [MAR00] MARTINS, P. R. *Segmentação de Histogramas Multimodais: Simulações, Implementações e Aplicação em Cheques Bancários Brasileiros*. Master Thesis, PUCPR, Curitiba, Brazil, November, 2000.

- [MBF03] MENOTI, D.; BORGES, D. L.; FACON, J. & BRITTO JR, A. S. *Segmentation of Postal Envelopes for Address Block Location: an approach based on features selection in wavelet space*. Proceedings of the 7<sup>th</sup> International Conference on Document Analysis and Recognition (ICDAR) 2003, Edimburgo, IEEE, 2003.
- [MEN03] MENOTI, D. *Segmentação de Envelopes Postais para Localização do Bloco Endereço: uma abordagem baseada em seleção de características no espaço Wavelet*. Master's Dissertation, PUCPR, Curitiba, 2003.
- [MIT97] MITCHEL, T., M. *Machine Learning*. Mc-Graw-Hill Science/Engineering/Math, 1997.
- [MOS92] MOSQUERA, A.; CABELLO, D.; CARREIRA, M. J. & PENEDO, M. G. *A fractal-based approach to segmentation*. Proceedings of IEE International Conference on Image Processing and its Applications, Maastrich, Netherlands, pp. 450-453, 1992.
- [NIE97] NIESEN, W. J. *Multiscale Medical Image Analysis*. Phd Thesis, Utrecht University, The Netherlands, 1997.
- [MEY91] MEYER, F. *Un algorithme optimal de ligne de partage des eaux*. Congress AFCET, RFIA, Lyon, 1991.
- [OLI98] OLIVEIRA, J. R. F. *O uso de algoritmos genéticos na decomposição morfológica de operadores invariantes em translação aplicados a imagens digitais*. Ph. D. Thesis, Instituto Nacional de Pesquisas Espaciais – INPE, São José dos Campos, Brasil, December 1998.
- [OTS79] OTSU N. *A Threshold Selection Method from Gray-level Histograms*. IEEE Transactions on Systems, Man and Cybernetics, v. SMC 9, no 1, pp.62-66, 1979.
- [ROB97] ROBERT, A. *From Contour Completion to Image Schemas: A Modern Perspective on Gestalt Psychology*. Technical Report CogSci.UCSD-97.02, February 1997.

- [SAN01] SANKUR, B. & SEZGIN, M. *Image Thresholding Techniques: A Survey over Categories*. Pattern Recognition, 2001. (under review).
- [SAR98] SARCA, O. V.; DOUGHERTY, E. R. & ASTOLA, J. *Optimal Binary Filters with Linearly Separable Preprocessing*. Nonlinear Image Processing, Proc. Of SPIE, January 1998.
- [SAR99] SARCA, O. V.; DOUGHERTY, E. R. & ASTOLA, J. *Two-stage Binary Filters*. Electronic Imaging, 8(3):219-232, July 1999.
- [SCA00] SOUSA, R. P.; CARVALHO, J. M.; ASSIS, F. M. & PESSOA, L. F. C. *Designing translation invariant operations via neural network training*. ICIP2000, IEEE Signal Processing Society, 2000.
- [SCH89] SCHMITT, M. *Mathematical morphology and artificial intelligence: an automatic programming system*. Signal Processing, 16(4): 389-401, April, 1989.
- [SCH92] SCHALKOFF, R. J. *Pattern Recognition: Statistical, Structural and Neural Approaches*. John Wiley & Sons, Inc, 1992.
- [SCH94] SCHONFELD, D. *Optimal Structuring Elements for the Morphological Pattern Restoration of Binary Images*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(6):589-601, June 1994.
- [SER82] SERRA, J. *Image Analysis and Mathematical Morphology*. Academic Press, London, 1982.
- [SHF81] SHORT R. D. and FUKUNAGA K., *The Optimal Distance Measure for Nearest Neighbor Classification*, IEEE T. Information Theory, vol. IT-27, no. 5, pp. 622-627, 1981.

- [SIV97] SIVAKUMAR, K. *Morphological Analysis of Random Fields: Theory and Applications*. Ph. D. Thesis, Dept. of Electrical Engineering, The Johns Hopkins University, Baltimore, MD, 1997.
- [SOU00] SOUZA, R. P. *Projetos de Operadores Invariantes a Translação via treinamento de redes neurais*. Ph. D. Thesis, Universidade Federal da Paraíba, Campina Grande, Paraíba, Brasil, fevereiro 2000.
- [SVV00] STOKKING R.; VINCKEN K L. & VIERGEVER, M. A. *Automatic Morphology-Based Brain Segmentation (MBRASE) from MRI-T1 Data*. *NeuroImage*, 12, 726-738, 2000.
- [TAM78] TAMURA H.; MORI S. & YAMAWAKI T. *Textural features corresponding to visual perception*. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-8(6):460-472, 1978.
- [UNW03] <http://www9.biostr.washington.edu/da.html>. MRI Serial Sections. Digital Anatomist Project, Dept. Biological Structure, University of Washington, Seattle, acessado em 17 de Maio de 2003
- [VAL84]. VALIANT, L. G. *A Theory of Learnable*. *Comm. ACM*, (27):11, pp. 1134-1142, 1984.
- [VIS91] VINCENT, L. & SOILLE, P. *Watershed in digital spaces: An efficient algorithm based on immersion simulations*. *Transactions on Pattern Analysis and Machine Intelligence*, vol 13, no 6, pp. 583-598, 1991.
- [VOG89] VOGT, R. *Automatic Generation of Morphological Set Recognition Algorithms*. Springer-Verlag, 1989.

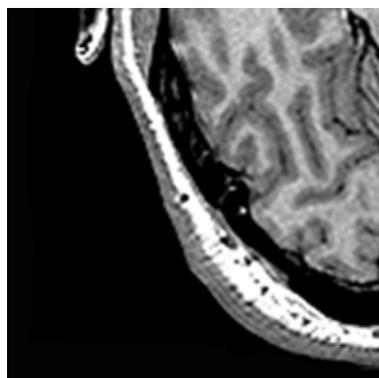
- [WEL93] WELLNER, P. D. *Adaptative thresholding for the digitaldesk*. Technical Report EPS-1993-110, Rank Xerox Ltd, July 1993.
- [WER44] WERTHEIMER, M. *Gestalt theory*. *Social Research*, J.J., 78-99, 1944
- [WIL93] WILSON, S. S. *Training Structuring Elements in Morphology Networks*. *Mathematical Morphology in Image Processing*, chapter 1, pp. 1-41, Marcel Dekker, 1993.
- [WNS97] WOLF, M.; NIEMAN, H. & SCHMIDT, W. *Fast Address Block Location on Handwritten and Machine Printed Mail-Piece Images*. ICDAR97 (Poste), 1997.
- [YCC95] YEN, J. C.; CHANG, F. J. & CHANG, S. *A new criterion for automatic multilevel thesholding*. *IEEE Transactions on Image Processing*, 4(3):370-378, March, 1995.
- [YOD99] YODA, I.; YAMAMOTO, K. & YAMADA, H. *Automatic Acquisition of Hierarchical Mathematical Morphology Procedures by Genetic Algorithms*. *Image and Vision Computing*, 17(10):749-760, August 1999.
- [YON02] YONEKURA, E. A. *Segmentação de Envelopes Postais Brasileiros por Abordagem Morfológica Utilizando Matriz de Co-Ocorrência e Histograma 2D*. Master's Dissertation, PUCPR, Curitiba, 2002.

## Apêndice A

### Exemplo Prático de Resolução de Conflitos e Cálculo Automático do Tamanho Adequado da Janela

Para uma melhor compreensão do algoritmo de resolução de conflitos na aprendizagem e do ajuste automático do tamanho da janela, apresenta-se a seguir o processo passo a passo.

Sejam as imagens amostra (a) e ideal (b) da **Figura A.1**, que são submetidas à rotina de aprendizagem, através da construção da matriz de decisão. Tendo sido calculados para todos os pares de pontos, os correspondentes vetores de característica, segue-se para a análise de conflitos. Da lista de pontos conflitantes, foi extraída como exemplo Ex.1. Pode-se observar que, para dois pontos com mesmo nível de cinza, médias e variâncias, existem saídas ideais diferentes. As posições nas imagens dos pontos em conflito encontram-se destacadas através dos quadrados em (c) e (d).



(a) Imagem amostra



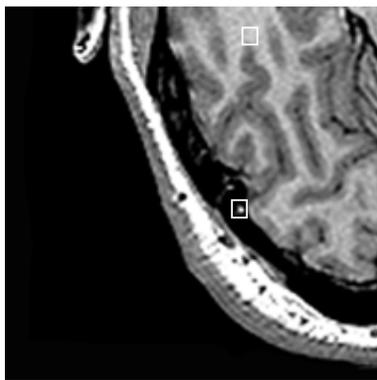
(b) Imagem ideal

**Figura A.1.** Imagens amostra e ideal submetidas à rotina de aprendizagem.

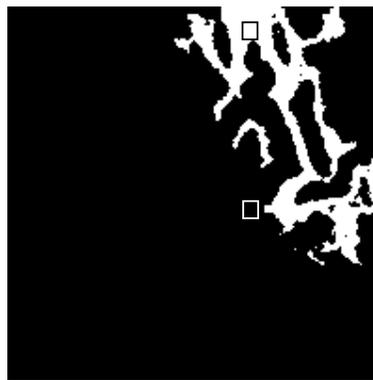
Ex.1- Exemplo de pontos conflitantes para uma janela de tamanho 3 x 3:

```
Linha_MatDecisao = 20881 [ 191 190.111111 0.098765 255 ]  
Linha_MatDecisao = 20864 [ 191 190.111111 0.098765 0 ]
```

Estes pontos correspondem ao centro das áreas salientadas na imagem abaixo (**Figura A.2**)

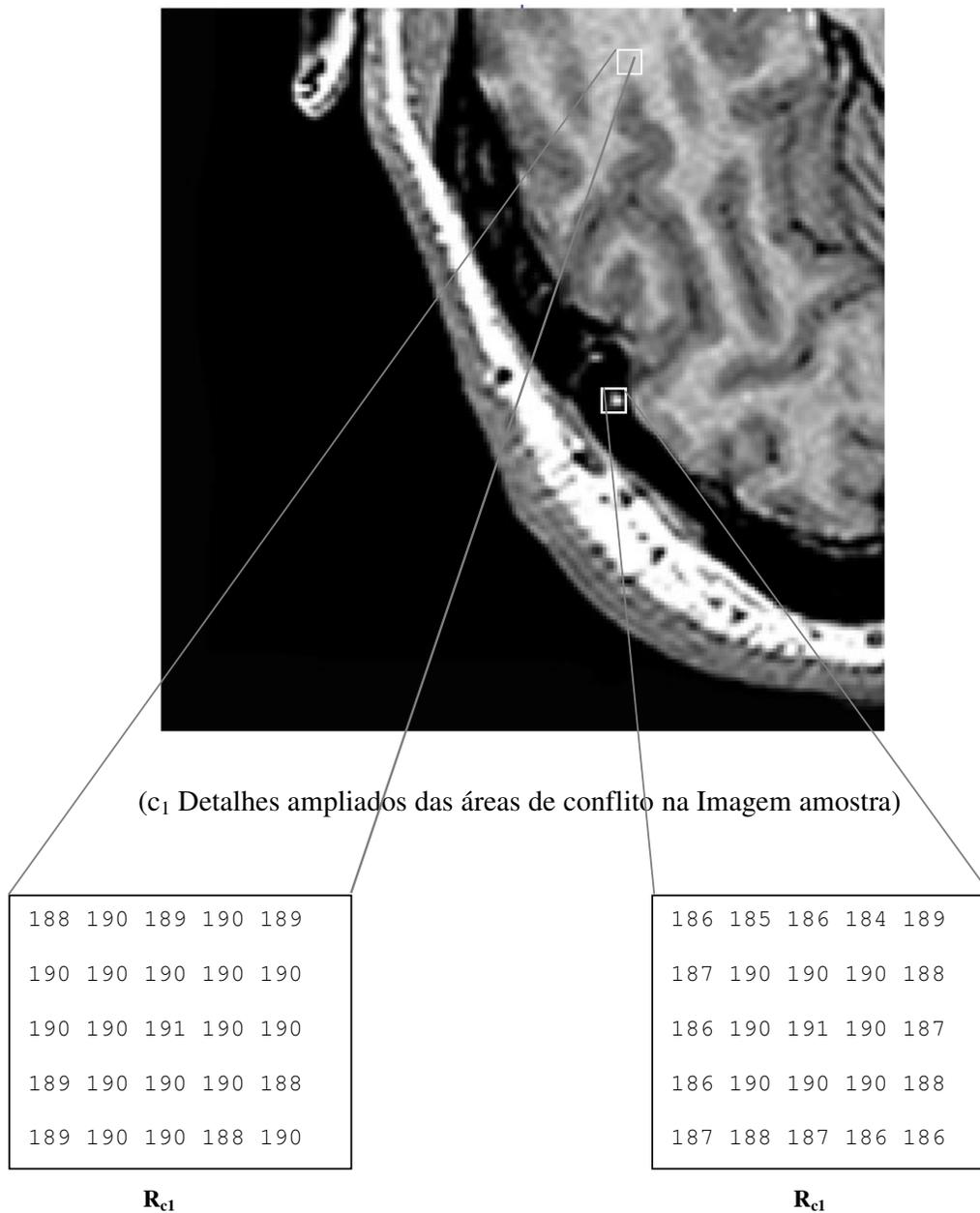


(c) Detalhes na Imagem amostra



(d) Detalhes na Imagem ideal

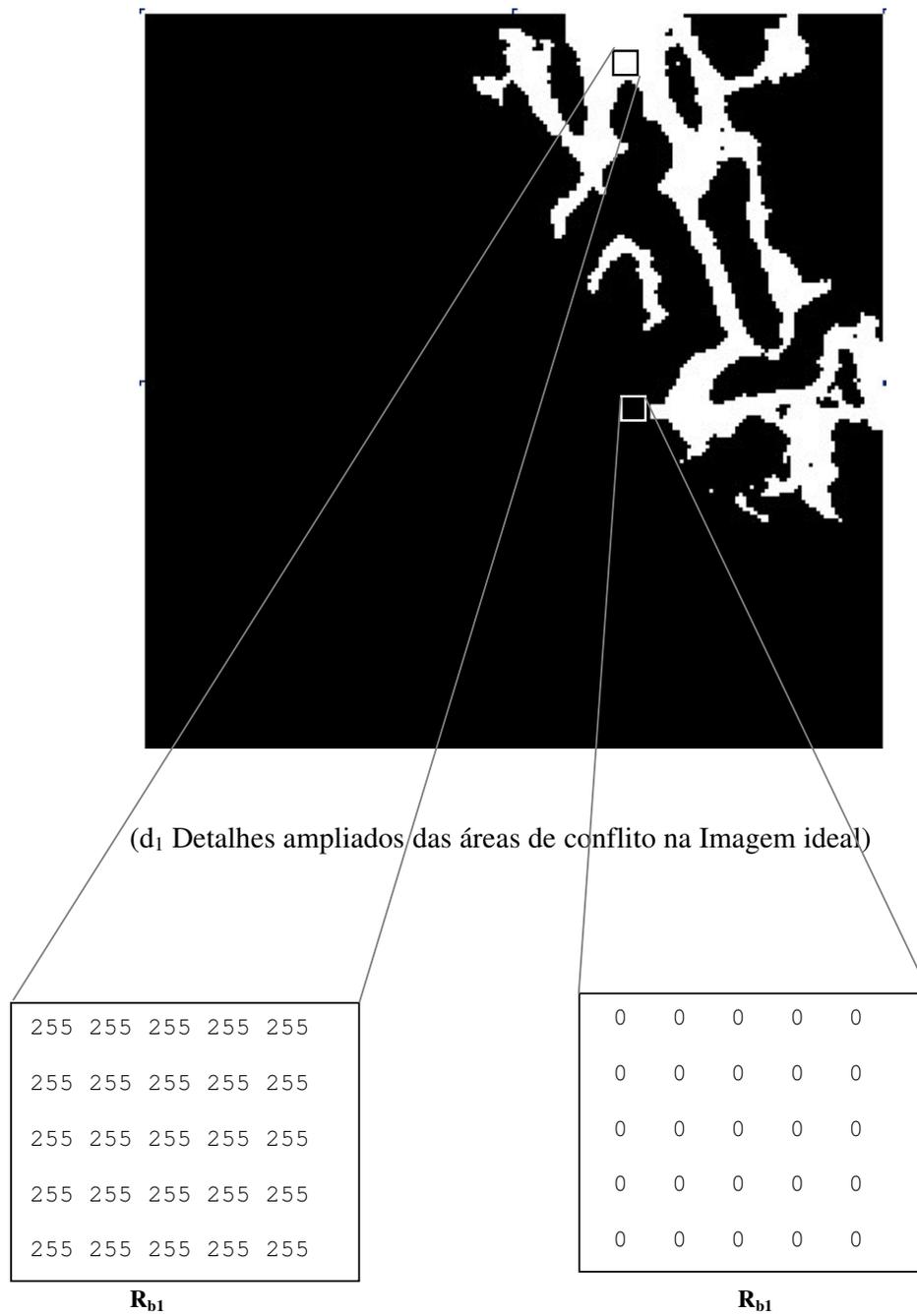
**Figura A.2.** Detalhes na imagem amostra e ideal submetidas à rotina de aprendizagem.



$R_{c1}$ : Região cinza 1,  $R_{c2}$ : Região cinza 2

**Figura A.3.** Detalhes ampliados das áreas de conflito na Imagem amostra

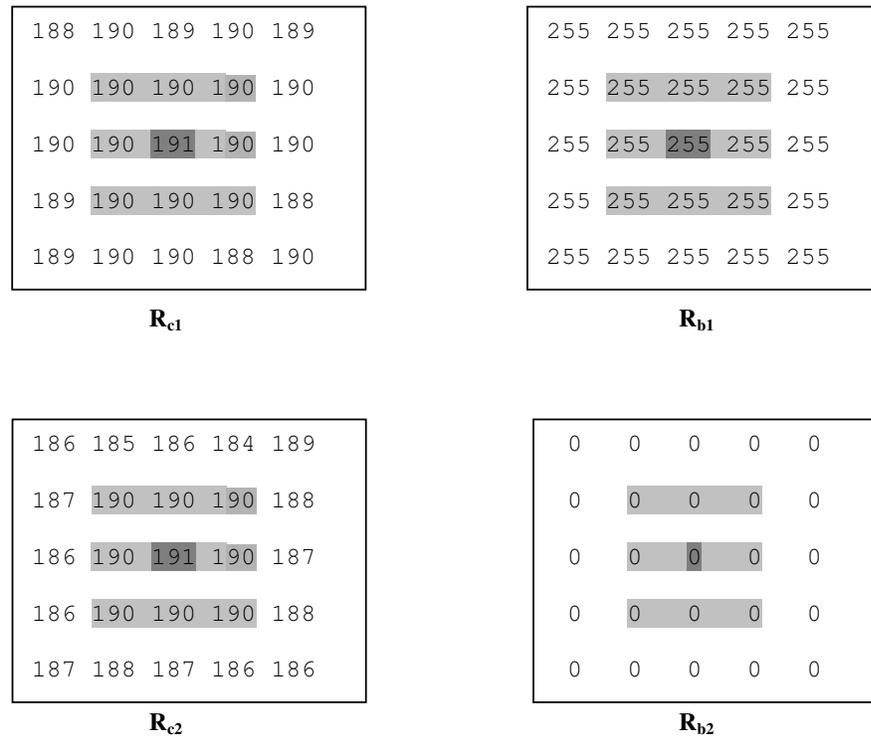
A **Figura A.3** apresenta os detalhes ampliados das áreas de conflito na Imagem amostra  $c_1$ , junto com os correspondentes valores dos níveis de cinza de cada ponto da região.



R<sub>b1</sub>: Região binarizada 1, R<sub>b2</sub>: Região binarizada 2

**Figura A.4.** Detalhes ampliados das áreas de conflito na Imagem ideal

A **Figura A.4** apresenta os detalhes ampliados das correspondentes áreas de conflito na Imagem ideal  $d_1$ , junto com os valores dos níveis binarizados de cada ponto da região.



**Figura A.5.** Regiões cinza  $R_{c1}$ ,  $R_{c2}$  e correspondentes regiões binarizadas  $R_{b1}$ ,  $R_{b2}$

### Janela de tamanho 3x3 – Vetores de característica

$R_{c1}$ : [ 191    190.111111    0.098765    255 ]  
 $R_{c2}$ : [ 191    190.111111    0.098765    0 ]

### Janela de tamanho 5x5 – Vetores de característica

$R_{c1}$ : [ 191    189.640000    0.550400    255 ]  
 $R_{c2}$ : [ 191    187.880000    3.785600    0 ]

Da análise das regiões acima, e dos vetores de características correspondentes, observa-se que para o tamanho da janela  $3 \times 3$ , como todos os pontos são iguais, estes geram uma mesma média e mesma variância. Aumentando-se o tamanho da janela para o número ímpar seguinte ( $5 \times 5$ ) o recálculo dos valores dos vetores de características para estes mesmos pontos fornece resultados diferentes.

No caso em que para toda a matriz de decisão não se encontrem outras inconsistências, utiliza-se este valor ( $5 \times 5$ ) como tamanho de aprendizagem e segmentação de novas imagens.

## Apêndice B

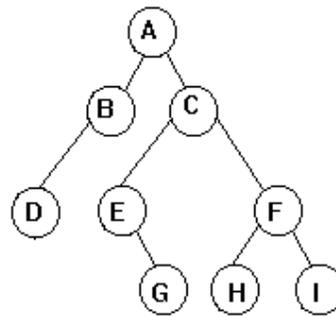
### Estudo e Implementação de Estruturas de Otimização do Processo de Aprendizagem

Apresentam-se neste Apêndice, considerações sobre o desempenho do algoritmo e estudos realizados para a otimização do tempo de processamento das etapas de aprendizagem e segmentação.

Inicialmente, a estrutura de armazenamento da matriz de decisão foi projetada como sendo uma matriz que contém o conjunto de vetores de característica agrupados pela ordem crescente do nível de cinza de cada pixel. Portanto, o método de busca dentro da matriz de decisão é realizada de forma seqüencial. Foram estudadas e implementadas soluções baseadas em árvores binárias, em listas indexadas e por fim em estruturas paralelas. O trabalho foi realizado com a ajuda do aluno bolsista do PIBIC/CNPq Luciano Bastian Machado, sendo que a íntegra do Relatório Técnico encontra-se em [MAC04].

#### Árvores binárias

As árvores binárias são um conjunto finito de um ou mais nós onde existe um nó especial denominado raiz. Os demais nós estão organizados em 2 conjuntos disjuntos (Esquerda, Direita), denominados sub-árvores da raiz, e ainda existe uma relação hierárquica entre os dados (**Figura B.1**).



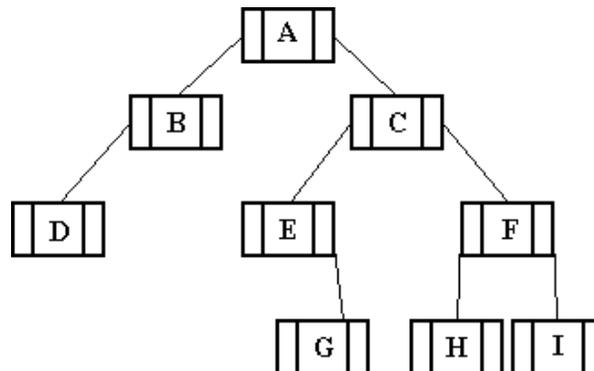
**Figura B.1** – Árvore binária

### Definição da Estrutura de Dados

```

Tipo nó
{
    <tipo de dado> info;
    nó esquerda;
    nó direita;
}
  
```

Essa estrutura permite que cada nó possua uma informação (um número inteiro por exemplo) e dois nós subsequentes com esta mesma estrutura.



**Figura B.2** – Estrutura dos dados da árvore binária

### Busca em árvores binárias

O sistema de busca em uma árvore binária fica eficiente quando se coloca algum tipo de classificação na estrutura da árvore. Por exemplo:

< Número >

<Menor>      <Maior>

Isso significa que todos os números da sub-árvore esquerda de um nó são menores que o número do nó. E todos os números da sub-árvore direita de um nó são maiores que o número do nó. Então a busca funcionaria assim:

O número procurado esta no nó atual?

- Sim. → Fim da Busca

- Não.

→ O número procurado é maior ou menor que o nó atual?

- Menor → Busque na sub-árvore esquerda.

- Maior → Busque na sub-árvore direita.

Porém as árvores binárias de pesquisa têm como principal desvantagem a dependência da estrutura da árvore em relação à ordem de inserção dos elementos. O que tem como consequência de poder afetar o tempo necessário para recuperar um item armazenado. Por exemplo, se inserir os elementos já em ordem, a estrutura da árvore será igual à de uma lista encadeada, e o tempo médio de recuperação de uma informação da árvore aumentará.

Para corrigir essa deficiência foi montado um controle de balanceamento da estrutura da árvore proposta em 1962 por Adelson-Velskii e Landis [BER98]. Daí o nome árvore AVL.

## Estrutura Indexada

A estrutura de uma busca indexada é extremamente simples. É constituída de um índice e da informação que deseja ser guardada. Para um melhor desempenho os índices foram montados em um vetor fixo de 0 a 255 e cada uma destas células aponta para uma lista encadeada de informações.

Índice	0		1		2 ...		255	
Lista	Info 1	↓						
Encadeada	Info 2	↓						
(Dados)	Info 3	↓						
	Info n	↓						

**Figura B.3** – Modelo de uma estrutura indexada

### Definição da Estrutura de Dados

```
Tipo celula
{
    <tipo de dado> info;
    celula proxima;
}
celula *Vetor [256];
```

Assim cada tom de cinza aponta para as suas informações e cada tom de cinza pode ser acessado diretamente.

### Acessando Informações do Vetor de Índices

No sistema de índice não existe busca, a tonalidade de cinza pode ser acessada diretamente da memória da seguinte forma:

```
Vetor[n]->info;
```

Onde  $n$  seria a tonalidade de cinza que esta sendo procurada.

Como cada índice do vetor aponta para uma lista encadeada uma busca é feita dentro da lista encadeada, mas esta é feita de forma seqüencial.

### Estrutura paralela

Foi também desenvolvida uma estrutura de processamento em paralelo que divide a altura da imagem em  $n$  pedaços, onde  $n$  é o número de computadores existentes em um *cluster*.

### *Cluster*

O cluster utilizado foi um conjunto de 9 computadores com o sistema operacional Linux Red-Hat, interligados em rede LAN 100 Mbps e com a pasta *home* do usuário cluster mapeada para um único computador (Master). O computador Master é responsável por mandar os outros nós (computadores) executarem o programa. Cada computador possui um

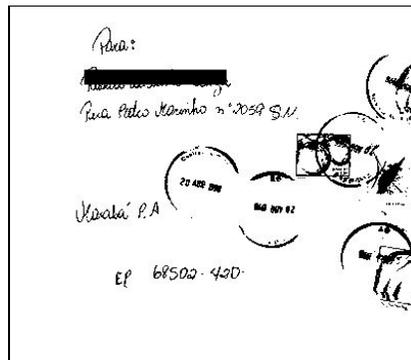


segmentadas e os tempos de processamento exigidos no caso de imagens de envelopes postais manuscritas.

A **Figura B.6** exemplifica os resultados obtidos da segmentação do bloco endereço, dos selos e carimbos em imagens de envelopes postais manuscritas pelas três abordagens. A primeira etapa da abordagem consiste no processo de aprendizagem. Neste processo, duas imagens são necessárias, a imagem original (**Figura B.5-(a)**) e a solução ideal da segmentação. No nosso caso a imagem (b) da **Figura B.5** representa essa solução ideal (gerada manualmente por um especialista) da segmentação do bloco endereço da imagem original (a). Ambas foram usadas no módulo de Aprendizagem para definir a matriz de decisão.



(a) Imagem Original



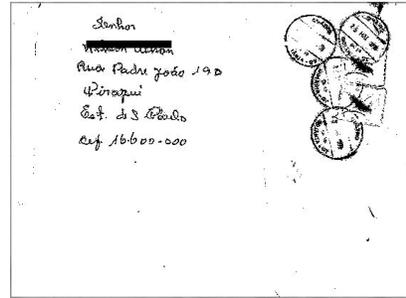
(b) Segmentação Ideal

**Figura B.5.** Imagens usadas na aprendizagem: (a) Imagem Original; (b) Segmentação Ideal.

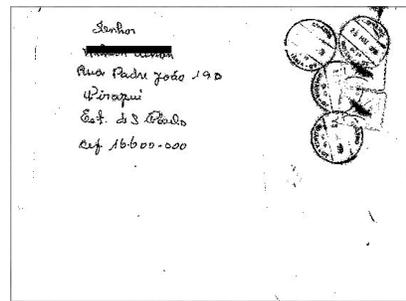
Em seguida, a segmentação de novas imagens faz-se acionando o módulo de segmentação. **Figura B.6** agrupa os resultados da segmentação por aprendizagem da nova imagem pelo processo inicialmente proposto na Tese, pela abordagem por Árvore Binária, por Estrutura Indexada e por Estrutura paralela.



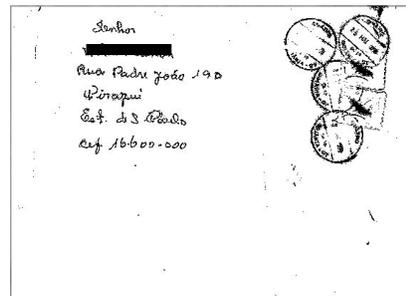
(a) Nova Imagem Submetida



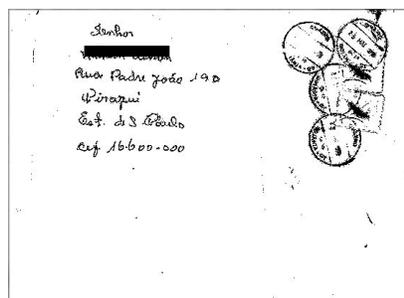
(b) Segmentação obtida pelo algoritmo original



(c) Segmentação obtida pela abordagem por árvore binária



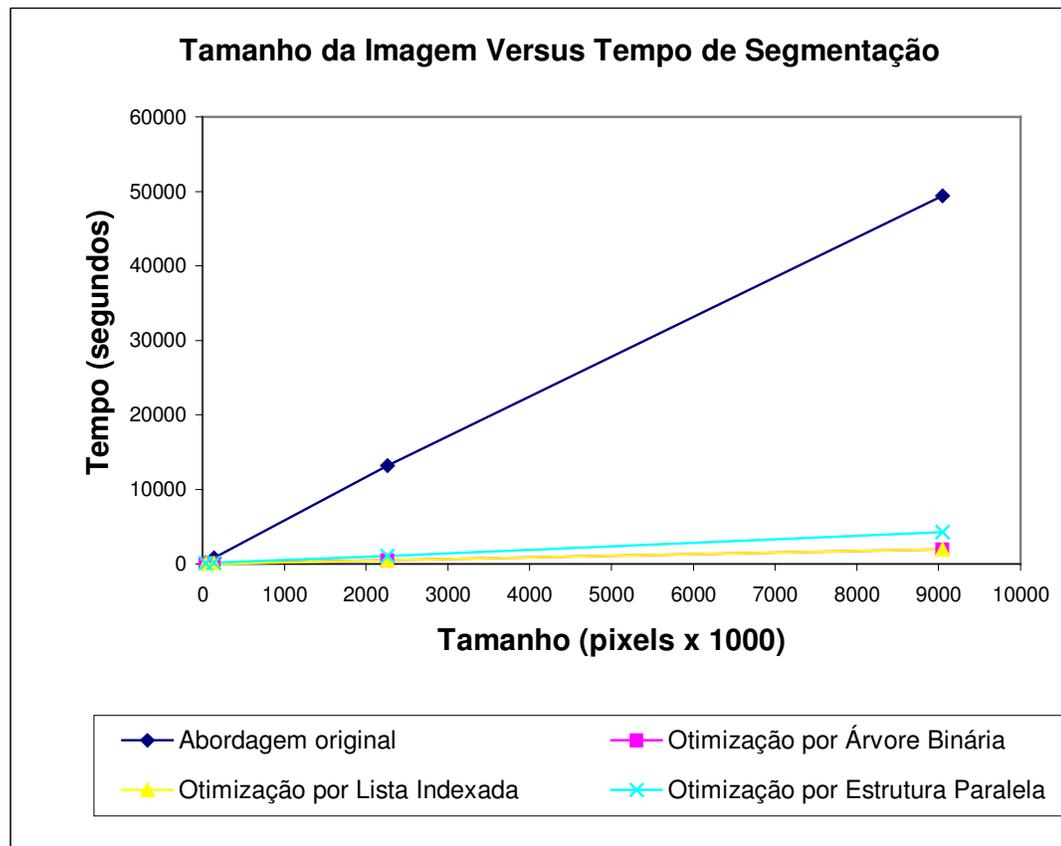
(c) Segmentação obtida pela abordagem por Estrutura Indexada



(e) Segmentação obtida por Estrutura paralela

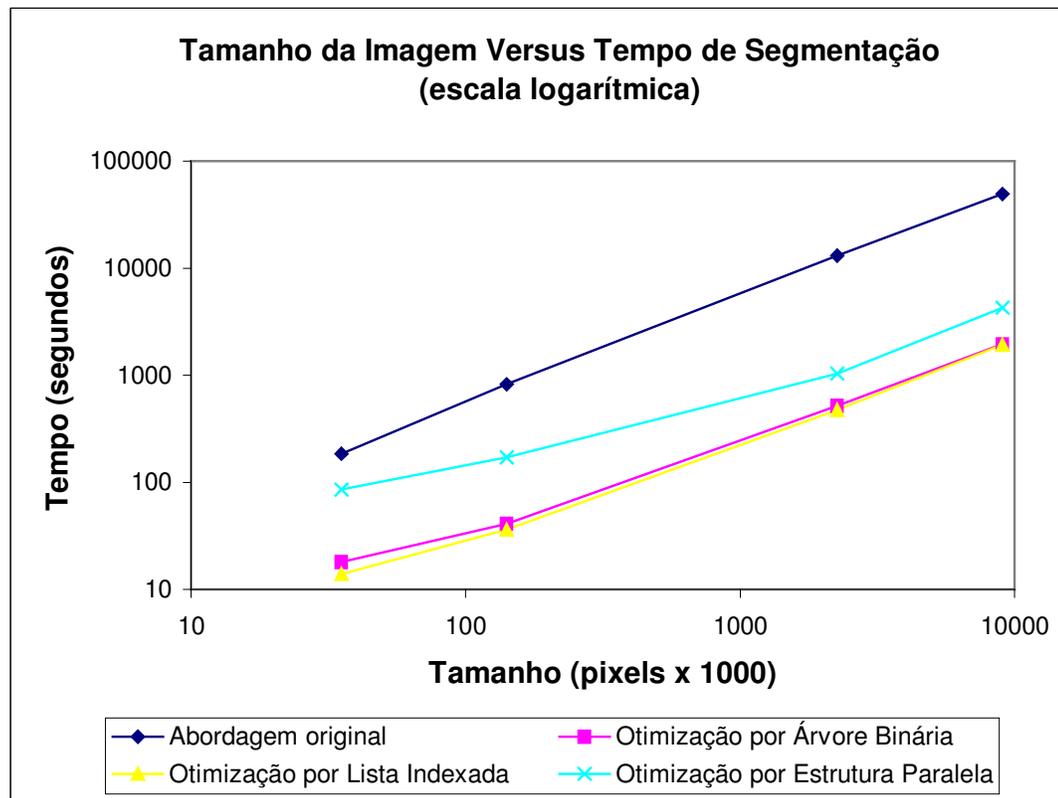
**Figura B.6.** Resultados obtidos no caso da segmentação de imagens de envelopes postais manuscritos. Os nomes dos destinatários foram escondidos propositalmente.

Na **Figura B.6**, é possível perceber que os resultados obtidos pelas três abordagens propostas são idênticos ao resultado obtido pelo processo inicialmente proposto (comprovadamente verificados por software). Isto demonstra que as novas abordagens implementadas não modificaram o resultado da segmentação inicialmente proposta.



**Figura B.7.** Tamanho da Imagem Versus Tempo de Segmentação no caso da segmentação de imagens de envelopes postais manuscritos.

A **Figura B.7** agrupa as medições de tempo de processamento versus o tamanho das Imagens. A **Figura B.8** apresenta os mesmos resultados porém, apresentados em escala logarítmica para facilitar a visualização. É possível perceber que as três novas abordagens propostas são realmente muito mais rápidas que a abordagem original. As abordagens por Árvore Binária e por Estrutura Indexada foram as mais rápidas (na ordem de 25 vezes mais rápidas para imagens maiores). A abordagem por estrutura paralela não se mostrou mais rápida (na ordem de 12 vezes mais rápidas para imagens maiores) que as outras duas abordagens devido ao uso de computadores mais antigos e menos velozes que os usados nas outras duas abordagens.



**Figura B.8.** Tamanho da Imagem Versus Tempo de Segmentação (escala logarítmica) no caso da segmentação de imagens de envelopes postais manuscritos.

## Conclusão

A completa semelhança entre os resultados obtidos pelas três abordagens propostas e o resultado obtido pelo processo inicialmente proposto demonstra claramente a melhoria do desempenho das abordagens implementadas.

Os resultados obtidos demonstraram que as abordagens por Árvore Binária, por Estrutura Indexada e por Estrutura Paralela podem significativamente acelerar o processo de segmentação por aprendizagem sem por tanto degradar a qualidade da mesma.

## Apêndice C

### Publicações Decorrentes da Pesquisa

Os artigos que foram produzidos a partir das pesquisas realizadas durante a elaboração desta tese estão listados a seguir.

[LAa03] LEGAL-AYALA, H. A. & FACON, J. *Image Segmentation by Learning Approach*. Proceedings of ICDAR 2003 - International Conference on Document Analysis and Recognition, Edinburgh, Scotland, v2 pp. 819-823, 2003.

[LAB03] LEGAL-AYALA, H. A. & FACON, J. *Comparative Results of Image Segmentation By Learning in Different Types of Images*. Proceedings of ICISP'2003 - IAPR International Conference on Image and Signal Processing, Agadir, Morocco, v2 pp 381-388, 2003.

[LAc03] LEGAL-AYALA, H. A. & FACON, J. *Segmentation Approach by Learning: Different Image Applications*. Proceedings of ICIAP 2003 - IAPR 12th International Conference on IMAGE ANALYSIS AND PROCESSING, Mantova, Italy, pp. 600-604, 2003.

[LAa04] LEGAL-AYALA, H. A. & FACON, J. *Unified Approach for Document Segmentation*. 5<sup>th</sup> Iberoamerican Meeting on Optics (RIO) and 8<sup>th</sup> Latin American Meeting on Optics, Lasers and Their Applications (OPTILAS), Proceedings SPIE Vol. 5622, pp. 1277-1282, 2004.

[Lab04] LEGAL-AYALA, H. A. & FACON, J. *Automatic Segmentation of Brain MRI Through Learning by Example*. In: Proceedings of 11<sup>th</sup> International Conference on Image Processing (ICIP), IEEE Signal Processing Society, Singapore, 24-27 October, 2004.