

EDUARDO AKIRA YONEKURA

**SEGMENTAÇÃO DE ENVELOPES POSTAIS
BRASILEIROS POR ABORDAGEM
MORFOLÓGICA UTILIZANDO MATRIZ DE
CO-OCORRÊNCIA E HISTOGRAMA 2D**

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

CURITIBA

2002

EDUARDO AKIRA YONEKURA

**SEGMENTAÇÃO DE ENVELOPES POSTAIS
BRASILEIROS POR ABORDAGEM
MORFOLÓGICA UTILIZANDO MATRIZ DE CO-
OCORRÊNCIA E HISTOGRAMA 2D**

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

Área de Concentração: *Ciência da Imagem*

Orientador: Prof. Dr. Jacques Facon

CURITIBA

2002

Yonekura, Eduardo Akira

Segmentação de Envelopes Postais Brasileiros por Abordagem Morfológica utilizando Matriz de Co-Ocorrência e Histograma 2D. Curitiba, 2002. 127p.

Dissertação(Mestrado)– Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Informática Aplicada.

1. Segmentação. 2. Envelopes Postais Brasileiros 3. Morfologia Matemática 4. Matriz de Co-Ocorrência e Histograma 2D. I. Pontifícia Universidade Católica do Paraná. Centro de Ciências Exatas e de Tecnologia. Programa de Pós-Graduação em Informática Aplicada II-t

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.

Aos meus pais e à minha irmã.

Agradecimentos

Ao orientador e amigo Prof. Dr. Jacques Facon pela amizade, excelente orientação, compreensão e apoio nos momentos difíceis, meus sinceros agradecimentos.

Aos meus pais e à minha irmã que sempre me apoiaram. Obrigado por me escutarem sempre.

Aos meus grandes amigos David Menoti Gomes, pela ajuda no Visual C++ sem o qual a implementação não teria sido possível e Horacio Andrés Legal Ayala, pelos conselhos úteis durante a escrita da presente dissertação.

À Cristiane Bastos Rocha Ferreira e Gerson Bochnia pela ajuda imprescindível no MatLab.

À Joesny Fagner de Oliveira pela ajuda na preparação da base de envelopes postais brasileiros.

Aos Professores, Dr. Bráulio Coelho Ávila, Dr. Carlos Alberto Maziero e Dr. Robert Carlisle Burnett pelas cartas de recomendação.

Aos Professores, Dr. Edouard Lethelier, Dr. Jacques Facon, Dr. Júlio César Nievola e Dr. Roberto Max Protil pelos ensinamentos durante a fase de créditos.

Aos Professores, Dr. Alceu de Souza Britto Júnior, Dr^a. Cinthia Obladen de Almendra Freitas e Dr. Edson José Rodrigues Justino pelo apoio.

Aos Professores, Dr. Alceu de Souza Britto Júnior, Dr. Díbio Leandro Borges e Dr. João Marques de Carvalho por terem aceito o convite para participar da banca examinadora.

Enfim, a todos que participaram direta ou indiretamente dessa dissertação.

Sumário

Agradecimentos	vii
Sumário	ix
Lista de Figuras	xv
Lista de Tabelas	xxi
Resumo	xxiii
Abstract	xxv

Capítulo 1

Introdução	1
1.1. Desafio	2
1.2. Motivação	3
1.3. Proposta	3
1.4. Contribuição	3
1.5. Organização	3

Capítulo 2

Estado da Arte	5
2.1. Técnicas de Segmentação	6
2.1.1. Segmentação por Limiarização	7
2.1.2. Segmentação por Cor	9
2.1.3. Segmentação por Textura	10
2.1.4. Segmentação por Contorno	12
2.1.5. Segmentação por <i>Clustering</i>	13
2.2 Segmentação de Imagens de Envelopes Postais	13
2.2.1 Método de [WOL97]	14

2.2.2 Método de [YU97]	15
2.2.3 Método de [WAN88]	17
2.2.4 Método de [DOW90]	18
2.2.5 Método de [YEH87]	19
2.2.6 Método de [JAI95]	19
2.2.7 Método de [XUE99]	21
2.3 Conclusão	22
Capítulo 3	23
Histograma 2D e Matriz de Co-Ocorrência	
3.1 Entropia	23
3.1.1 Entropia local e global	24
3.1.2 Entropia condicional	25
3.1.3 Exemplos de Segmentação usando a Entropia	26
3.1.4 Limiarização por Entropia de [PUN80]	26
3.1.5 Limiarização por Entropia de [KAP85]	27
3.2 Histograma 2D	28
3.2.1 Definição de Histograma 2D	28
3.2.2. Exemplo de Segmentação usando o Histograma 2D	29
3.2.3. Vantagens e Desvantagens do Histograma 2D	31
3.3 Matriz de Co-Ocorrência	32
3.3.1. Definição de [HAR92]	33
3.3.2. Definição de [PAL89]	34
3.3.3. Definição de [ZHA98]	35
3.3.4. Exemplo de Segmentação usando a Matriz de Co-	35
Ocorrência	
3.3.5. Vantagens e Desvantagens da Matriz de Co-Ocorrência	38
3.4. Conclusão	43
Capítulo 4	44
Reconstrução Morfológica em Níveis de Cinza e Segmentação Watershed	
4.1 Reconstrução Morfológica	44
4.1.1. Notações utilizadas	44
4.1.2. Reconstrução Binária	45

4.1.3. Reconstrução em níveis de cinza	45
4.1.4. Algoritmos de Reconstrução em níveis de cinza	47
4.1.5. Máximos regionais e reconstrução	49
4.2 Segmentação Watershed	55
4.2.1. Algoritmos de Segmentação Watershed	57
4.2.2. Filas Ordenadas	57
4.2.3. Watershed por Filas Ordenadas	59
4.2.4. Algoritmo de Watershed por Filas Ordenadas	60
4.3. Conclusão	64
Capítulo 5	65
Base de Envelopes Postais Brasileiros	
5.1 Regras recomendadas pela Empresa Brasileira de Correios e Telegráfos (ECT)	65
5.2. Características da base de imagens de envelopes postais	66
Capítulo 6	68
Análise e Comparação de Resultados Obtidos	
6.1. Ambiente de desenvolvimento utilizado	68
6.2. Metodologia com Matriz de Co-Ocorrência	68
6.2.1. Estratégia de avaliação de resultados e dificuldades encontradas	72
6.2.2. Resultados obtidos com matriz de co-ocorrência (Junção horizontal e vertical)	73
6.2.3. Resultados obtidos com matriz de co-ocorrência (Junção das 4 direções)	75
6.2.4. Análise comparativa dos resultados obtidos com matriz de co-ocorrência	77
6.2.5. Vantagens e desvantagens da abordagem proposta com matriz de co-ocorrência	79
6.3. Análise de Resultados com Histograma 2D	80
6.3.1. Metodologia com Histograma 2D	80
6.3.2. Resultados obtidos com histograma 2D	84
6.3.3. Análise quantitativa e qualitativa dos resultados obtidos com histograma 2D	86

6.3.4. Vantagens e desvantagens da abordagem proposta com histograma 2D	86
6.4 Análise comparativa entre Histograma 2D e Matriz de Co-Ocorrência	87
6.4.1 Metodologia de Avaliação por competição	87
6.4.2. Resultados obtidos com as imagens escolhidas utilizando histograma 2D	89
6.4.3. Comparação de resultados obtidos com Histograma 2D e Matriz de Co-Ocorrência	95
6.4.4. Comparação da abordagem proposta (histograma 2D e matriz de co-ocorrência) X Estado da Arte (Capítulo 2)	97
6.5. Conclusão	98
Conclusão	99
Referências Bibliográficas	101
Bibliografia Complementar	106
Apêndice A	107
Morfologia Matemática em níveis de cinza	
A.1. Operadores Morfológicos Básicos em Níveis de Cinza	107
A.1.1. Erosão	107
A.1.2. Dilatação	108
A.1.3. Abertura	110
A.1.4. Fechamento	111
A.1.5. TopHat por Abertura	112
A.1.6. TopHat por Fechamento	113
A.1.7. Gradiente Morfológico	114
A.1.8. Gradiente Morfológico por erosão em níveis de cinza	115
A.1.9. Gradiente Morfológico por dilatação em níveis de cinza	116
A.1.10. Gradiente Morfológico por dilatação e erosão em níveis de cinza	117
Apêndice B	118
Resultados Adicionais	
B.1. Resultados adicionais por matriz de co-ocorrência (Junção horizontal e vertical)	119

B.2. Resultados adicionais por matriz de co-ocorrência (Junção 4 direções)	122
B.3. Resultados adicionais por histograma 2D	125

Lista de Figuras

Figura 1.1	Exemplo de cheque bancário	2
Figura 1.2	Exemplo de envelope postal brasileiro	2
Figura 2.1	Diferentes tipos de textura	11
Figura 3.1	Quadrantes de um histograma 2D gerado pelo vetor 2D de limiarização (S, T)	30
Figura 3.2	Imagem Original e Filtrada para o processamento do histograma 2D	31
Figura 3.3	Visualização 3D do histograma 2D das imagens da Figura 3.2	32
Figura 3.4	Imagem de dimensões 4x4 e níveis de cinza correspondentes	33
Figura 3.5	Resultados da matriz de co-ocorrência para diversos ângulos e distância 1 para o exemplo de imagem da Figura 3.4	34
Figura 3.6	Quadrantes de uma matriz de co-ocorrência	36
Figura 3.7	Imagem necessária para o processamento da matriz de co-ocorrência	39
Figura 3.8	Visualização 3D da matriz de co-ocorrência da imagem da Figura 3.7 (direção: horizontal, distância entre pixels: 1)	40
Figura 3.9	Visualização 3D da matriz de co-ocorrência da imagem da Figura 3.7 (direção: vertical, distância entre pixels: 1)	41
Figura 3.10	Visualização 3D da matriz de co-ocorrência da imagem da Figura 3.7 (direção: junção das direções horizontal e vertical, distância entre pixels: 1)	42

Figura 3.11	Visualização 3D da matriz de co-ocorrência da imagem da Figura 3.7 (direção: junção das 4 direções, distância entre pixels: 1)	43
Figura 4.1	Vizinhança 4	44
Figura 4.2	Vizinhança 6	44
Figura 4.3	Vizinhança 8 (N_G)	44
Figura 4.4	Porção de uma vizinhança 4	45
Figura 4.5	Porção de uma vizinhança 6	45
Figura 4.6	Porção de uma vizinhança 8	45
Figura 4.7	Sentido raster (N_G^+)	45
Figura 4.8	Sentido anti- raster (N_G^-)	45
Figura 4.9	Decomposição de limiares de uma imagem em níveis de cinza	47
Figura 4.10	Princípio da reconstrução em níveis de cinza da máscara I pelo marcador J	51
Figura 4.11	Exemplo de reconstrução em níveis de cinza	54
Figura 4.12	Exemplo de aplicação da reconstrução em níveis de cinza	54
Figura 4.13	Enchimento e determinação das bacias de retenção e das linhas watersheds	55
Figura 4.14	Uso da informação do gradiente na segmentação Watershed	56
Figura 4.15	Estrutura de uma fila simples	58
Figura 4.16	Estrutura de uma Fila Ordenada: (a) a fila ordenada; (b) extração de um elemento de mais alta prioridade; (c) redução do número de filas simples; (d) tratamento do elemento de mais alta prioridade quando a fila correspondente tiver sido suprimida.	59
Figura 4.17	Exemplos de Segmentação Watershed sobre uma imagem de gradiente	61
Figura 4.18	Segmentação Watershed por filas ordenadas	63
Figura 5.1	Exemplos de imagens da base de envelopes postais	67

	brasileiros	
Figura 6.1b)	Número de bacias segmentadas sem pós-processamento com a abordagem proposta de matriz de co-ocorrência para a imagem a)	70
Figura 6.2	Diagrama com a estrutura da abordagem proposta para a segmentação de envelopes postais brasileiros	71
Figura 6.3	Exemplo de segmentação boa	73
Figura 6.4	Exemplo de segmentação ruim	74
Figura 6.5	Exemplo de segmentação boa	75
Figura 6.6	Exemplo de segmentação ruim	76
Figura 6.7	Tempos de processamento com matriz de co-ocorrência	78
Figura 6.8	Tempos de processamento médio com matriz de co-ocorrência	78
Figura 6.9	Reconstrução da imagem I pela imagem I-1	81
c) Figura 6.10-	Número de bacias segmentadas sem pós-processamento com a abordagem proposta de histograma 2D para as imagens a) - b)	82
Figura 6.11	Diagrama com a estrutura da abordagem proposta para a segmentação de envelopes postais brasileiros	83
Figura 6.12	Exemplo de segmentação boa	84
Figura 6.13	Exemplo de segmentação ruim	85
Figura 6.14	Tempos de processamento com histograma 2D	86
Figura 6.15	Imagens escolhidas para a metodologia de avaliação por competição	88
Figura 6.16	Classes segmentadas com abordagem por matriz de co-ocorrência e por histograma 2D respectivamente da imagem 6.17-a)	89
Figura 6.17	Classes segmentadas com abordagem por matriz de co-ocorrência e por histograma 2D respectivamente da imagem 6.26-a)	91

Figura 6.18	Classes segmentadas com abordagem por matriz de co-ocorrência e por histograma 2D respectivamente da imagem 6.27-a)	92
Figura 6.19	Classes segmentadas com abordagem por matriz de co-ocorrência e por histograma 2D respectivamente da imagem 6.28-a)	93
Figura 6.20	Classes segmentadas com abordagem por matriz de co-ocorrência e por histograma 2D respectivamente da imagem 6.29-a)	94
Figura 6.21	Tempos de processamento com as duas técnicas	96
Figura 6.22	Tempos de processamento médio com as duas técnicas	96
Figura A.1	Exemplo de Erosão	108
Figura A.2	Exemplo de Dilatação	109
Figura A.3	Exemplo de Abertura	110
Figura A.4	Exemplo de Fechamento	111
Figura A.5	Exemplo teórico da transformação Tophat por Abertura	112
Figura A.6	Exemplo de TopHat por Abertura	113
Figura A.7	Exemplo teórico da transformação Tophat por Fechamento	114
Figura A.8	Exemplo de TopHat por Fechamento	114
Figura A.9	Exemplo de Gradiente por erosão	116
Figura A.10	Exemplo de Gradiente por dilatação	116
Figura A.11	Exemplo de Gradiente por dilatação–erosão	117
Figura B.1	Exemplo de segmentação boa, porém com a última classe contendo bordas adicionais geradas por defeitos de aquisição da imagem	119
Figura B.2	Exemplo de segmentação média	120
Figura B.3	Exemplo de segmentação ruim onde as classes foram superiores a 3	121

Figura B.4	Exemplo de segmentação boa, porém com a última classe contendo bordas adicionais geradas por defeitos de aquisição da imagem	122
Figura B.5	Exemplo de segmentação média	123
Figura B.6	Exemplo de segmentação ruim onde as classes foram superiores a 3	124
Figura B.7	Exemplo de segmentação boa, porém com a última classe contendo bordas adicionais geradas por defeitos de aquisição da imagem	125
Figura B.8	Exemplo de segmentação média	126
Figura B.9	Exemplo de segmentação ruim onde as classes foram superiores a 3	127

Lista de Tabelas

Tabela 6.1	Classificação dos resultados obtidos com junção horizontal e vertical	77
Tabela 6.2	Classificação dos resultados obtidos com junção das 4 direções	77
Tabela 6.3	Classificação dos resultados obtidos com histograma 2D	86
Tabela 6.4	Comparação da abordagem proposta (histograma 2D e matriz de co-ocorrência) X Estado da Arte (Capítulo 2)	97

Resumo

O presente trabalho apresenta um novo método de segmentação de envelopes postais brasileiros por abordagem composta combinando os conceitos de morfologia matemática, matriz de co-ocorrência e histograma 2D. Técnicas de segmentação morfológica ajudam a interpretar as informações geradas pela matriz de co-ocorrência e pelo histograma 2D com o objetivo de extrair o conteúdo dos envelopes postais brasileiros. Pouco conhecimento *a priori* é exigido sobre os mesmos.

No histograma 2D, a grande dificuldade reside na escolha adequada da 2ª imagem (imagem filtrada). Já na matriz de co-ocorrência, o maior problema é a escolha do sentido de varredura bem como da distância entre os *pixels*. As vantagens desta abordagem são descritas e ilustradas com testes efetuados em uma base composta de 80 imagens de envelopes postais brasileiros totalmente diferentes onde as regras de posicionamento parte manuscrita, carimbos e selos não são respeitados. Contudo, os resultados são encorajadores, se comparados com outros métodos presentes na literatura com relação aos resultados e tempo de processamento. O mérito da abordagem proposta é a combinação de três conceitos (morfologia matemática, matriz de co-ocorrência e histograma 2D) que leva em consideração as vantagens de cada um.

Palavras-Chave: Segmentação, Morfologia Matemática, Matriz de Co-Ocorrência, Histograma 2D.

Abstract

This work presents a new segmentation method by compound approach combining the concepts of mathematical morphology, co-occurrence matrix and two-dimensional histogram. Morphological segmentation techniques assist to interpret the information generated by the co-occurrence matrix and two-dimensional histogram with the objective to extract the contents of the brazilian postal envelopes. Very little a priori knowledge is required over the images.

In the two-dimensional histogram, the great difficulty is the correct choice of the second image (filtered image). In the co-occurrence matrix, the problem is the choice of the image scanning as well as the distance between the pixels. The advantages of this approach are described and illustrated with tests carried out in a database compound by 80 different brazilian postal envelopes where there is no fixed position for the handwritten address block, postmarks and stamps. However, the results are encouraging, if compared to other methods presents in the literature regarding the results and time processing. The merit of the proposed method is the combination of three approaches (mathematical morphology, co-occurrence matrix and two-dimensional histogram) which takes into consideration the advantages of each one.

Keywords: Segmentation, Mathematical morphology, co-occurrence matrix, two-dimensional histogram.

Capítulo 1

Introdução

O processamento de imagens de documentos tem recebido considerável atenção na área de processamento digital de imagens. De fato, devido à sua importância em diferentes contextos a serem analisados, tais como, revistas, jornais, cheques bancários, envelopes postais, etc, uma crescente necessidade de automatizar processos manuais se faz presente, aliada aos avanços tecnológicos que estão tornando os computadores pessoais mais baratos, confiáveis e rápidos.

Diversas são as aplicações relacionadas ao processamento de documentos. Como exemplo, pode-se citar os bancos que necessitam automatizar processos tais como, compensação e verificação de assinaturas de cheques bancários; as bibliotecas, que necessitam manter o controle de perda de material impresso, degradação, etc; e os correios, que manipulam milhares de envelopes postais diariamente.

Contudo, é importante salientar que na maior parte dos documentos, devido às particularidades próprias de cada contexto, e também à falta de padronização dos mesmos, uma única estratégia que seja capaz de processar com eficiência todos os tipos de documentos citados torna-se impossível. Na Figura 1.1, um exemplo de cheque bancário preenchido é mostrado. É possível perceber que existem traços manuscritos que sobrepõem o logotipo e a parte manuscrita dificultando a segmentação de ambos [JES99]. Já na Figura 1.2, percebe-se que a parte manuscrita sobrepõe os selos e carimbos. É possível perceber também que os carimbos não aparecem em nenhuma posição específica.



Figura 1.1 Exemplo de cheque bancário

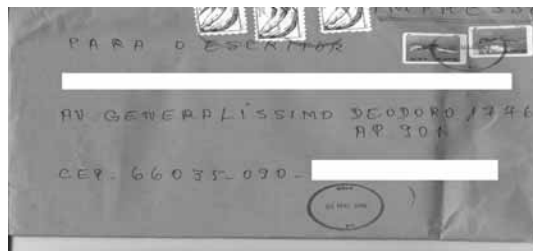


Figura 1.2 Exemplo de envelope postal brasileiro

Na presente pesquisa, é proposto um método capaz de segmentar envelopes postais brasileiros em classes que os compõem, a saber: parte manuscrita com carimbos, selos e fundo. Na sequência, serão descritos em detalhes o desafio, a motivação, a proposta e as contribuições deste trabalho.

1.1. Desafio

Os desafios para a presente pesquisa são:

- Desenvolver uma metodologia de segmentação, apesar da existência de poucos estudos sobre a segmentação de envelopes postais, que se restringem principalmente à localização do bloco-endereço. Na literatura, encontram-se diversos estudos de envelopes postais com fundo apresentando textura simples, gerando técnicas de segmentação por textura (Capítulo 2), nem sempre adequados para a grande variedade e complexidade de envelopes postais brasileiros;

O objetivo de possibilitar uma segmentação eficiente de envelopes postais brasileiros com pouco uso de conhecimento *a priori* a respeito dos mesmos. Devido à variabilidade e complexidade acima citadas, tal meta torna-se um grande desafio.

1.2. Motivação

Tem-se como motivação para o desenvolvimento da presente pesquisa:

- O processamento de imagens de envelopes postais, que representa uma crescente necessidade, pois milhões deles são manipulados diariamente no mundo inteiro;
- A extração automática das classes anteriormente citadas, o que até agora foi pouco explorado;
- A possibilidade de mesclar abordagens que até agora não foram combinadas, a saber, Histograma 2D, Matriz de Co-Ocorrência e Morfologia Matemática.

1.3. Proposta

- A proposta da presente pesquisa é segmentar os envelopes postais brasileiros nas classes que os constituem, com pouco uso de conhecimento a priori e empregando uma metodologia de segmentação baseadas em Histograma 2D, Matriz de Co-Ocorrência e Morfologia Matemática.

1.4. Contribuição

As principais contribuições deste trabalho são:

- Uma proposta de segmentação de envelopes postais brasileiros baseadas no Histograma 2D e Matriz de Co-Ocorrência juntamente com métodos morfológicos;
- A possibilidade de extensão do método proposto para outros contextos, como por exemplo, imagens de texturas e imagens médicas;
- Uma contribuição para pesquisas futuras na área de segmentação de envelopes postais brasileiros.

1.5. Organização

Essa dissertação está dividida em 6 Capítulos descritos da seguinte maneira:

O Capítulo 1 apresentou uma descrição sucinta dos desafios, motivação, proposta, contribuição e organização bem como da importância do estudo da segmentação de imagens de documentos, particularmente de envelopes postais brasileiros;

O Capítulo 2 trata das técnicas de segmentação de imagens em geral e mais especificamente, de imagens de envelopes postais;

O Capítulo 3 descreve detalhadamente o Histograma 2D e a Matriz de Co-Ocorrência. Suas propriedades, características e principalmente, suas importâncias na presente pesquisa;

O Capítulo 4 descreve a morfologia matemática em níveis de cinza. As operações utilizadas na presente pesquisa, tais como, a reconstrução em níveis de cinza e a ferramenta de segmentação Watershed são detalhadas;

O Capítulo 5 apresenta a base de imagens de envelopes postais. Uma descrição a respeito da variabilidade e complexidade da mesma é detalhada neste capítulo;

O Capítulo 6 descreve a metodologia e análise de resultados empregando as abordagens por Matriz de Co-Ocorrência e Histograma 2D bem como uma avaliação por competição entre as abordagens empregadas;

O Capítulo 7, por fim, apresenta as conclusões e propostas de trabalhos futuros na segmentação de envelopes postais brasileiros.

Capítulo 2

Estado da Arte

O processo de segmentação consiste em decompor uma imagem em diferentes áreas de interesse. Contudo, por ser extremamente dependente do contexto a ser analisado, a segmentação pode ser considerada como um dos processos mais difíceis na área de Processamento Digital de Imagens [GON92].

De acordo com [HAR87], as seguintes propriedades são desejáveis para um bom processo de segmentação:

- Regiões de uma imagem segmentada devem ser uniformes e homogêneas com relação à alguma característica como, por exemplo, o nível de cinza, cor ou textura;
- Regiões adjacentes devem ter valores significativamente diferentes com relação à característica segmentada;
- O interior das regiões deve ser simples e sem falhas (buracos).

Diversas são as aplicações na área de segmentação de imagens, tais como: imagens médicas [GON98] [SCH94], imagens de cheques bancários [JES99] [MAR00], localização de placas de veículos [BUS01] entre outras. Em todas essas aplicações, o desempenho final da metodologia está diretamente relacionado com a eficiência do processo de segmentação.

A imagem digital I é uma função discreta bidimensional $f(x, y)$ digitalizada pela coordenada espacial e pela magnitude do valor da característica com dimensões M e N , respectivamente, altura e largura. Uma imagem digital pode ser vista como uma matriz bidimensional, cujos índices linha e coluna, i e j respectivamente, identificam um ponto (*pixel*) na imagem, e o valor do elemento da matriz correspondente identifica o nível de intensidade da característica.

É importante salientar que a imagem, que é a “matéria prima” da segmentação, pode ser encontrada de variados tipos:

- Imagem de intensidade de luz: representa a variação da intensidade da luz na cena;
- Imagem de profundidade: é o mapeamento da informação de profundidade em diferentes pontos da cena;
- Imagem de ressonância magnética nuclear: representa a variação de intensidade de ondas de rádio geradas por sistemas biológicos, quando expostas aos pulsos de frequência de rádio;
- Imagem de temperatura: infravermelho;
- Imagem de ultra-sonografia: representa a reflexão de ondas.

Na literatura, encontram-se várias técnicas de segmentação, mas não existe um método simples que seja bom para todos os tipos de imagens, e nem todos os métodos são igualmente bons para um tipo particular de imagem. As metodologias baseiam-se na textura [TUC98], cor [BRI96], detecção de bordas [CAN86], técnicas de limiarização [MAT99], entre outros.

Na próxima seção serão apresentadas algumas técnicas de segmentação bem como pesquisas na literatura.

2.1. Técnicas de Segmentação

Alguns biólogos mostraram que o sistema de visão do ser humano baseia-se na noção de região e na noção de contorno:

- Uma região de uma imagem é um conjunto de pontos "ligados" onde, de qualquer ponto da região pode-se chegar a qualquer outro ponto por um caminho inteiramente contido nessa região. As regiões que se desejam detectar, em geral, são regiões "homogêneas", ou seja, apresentam alguma propriedade local aproximadamente constante. Comumente, esta propriedade é a continuidade no nível de cinza.
- Um contorno é uma mudança brusca do nível de cinza entre duas regiões relativamente homogêneas. Ele pode aparecer como uma sequência de pontos, uma linha, um segmento, uma curva ou uma forte variação do nível de cinza médio.

Baseadas nessas duas noções, pode-se dizer que existem basicamente três abordagens de segmentação:

- Segmentação por região: A detecção de regiões em uma imagem pode ser realizada de duas maneiras, por fusão de *pixels* ou por particionamento;

- Segmentação por contorno: Uma das técnicas de detecção de bordas mais utilizadas consiste em usar um operador de derivação local, tais como o gradiente (derivada de primeira ordem) e o laplaciano (derivada de segunda ordem), para realçar as bordas. Esses operadores também são muito utilizados na pesquisa de contornos. Após a etapa de detecção de bordas, é necessária a identificação dos objetos a partir de seus contornos, através do processamento da imagem de bordas. A localização de fronteiras envolve o processo de perseguição de bordas (*edge tracking*);
- Segmentação por textura: a textura é nada mais que uma região que possui propriedades específicas, seja de periodicidade, de repetição de forma aleatória ou não, dependendo da textura [TAM78].

A seguir serão apresentadas algumas metodologias de segmentação por região, por contorno, por textura ou ainda híbridas.

2.1.1. Segmentação por Limiarização

A limiarização é uma das principais técnicas de segmentação no processamento digital de imagens. A limiarização converte uma imagem originalmente em níveis de cinza para uma imagem binária.

A limiarização pode ser classificada como:

- Limiarização Global: é utilizada para isolar os objetos de interesse do fundo em que se situam. Utiliza o histograma como critério de análise, onde, no caso ideal, o mesmo é bem definido como tendo dois picos e um vale profundo, representando os objetos e o fundo.
- Limiarização Adaptativa: seleciona um limiar individual para cada *pixel* baseado no alcance de intensidade estimado em sua vizinhança local e, desta forma, permite uma limiarização mais eficiente para imagens cujos vales não estão bem definidos. A grande dificuldade reside no fato do tamanho da janela a ser utilizada, pois requer o conhecimento prévio da imagem [RIT96].
- Limiarização multimodal ou Multilimiarização: é utilizada quando se tem vários objetos que se distinguem do fundo. A presença dos mesmos faz com que o histograma de distribuição de níveis de cinza seja multimodal. O limiar é encontrado pela localização dos vales que separam os objetos. A limiarização multimodal pode ser então entendida como sendo uma extensão dos métodos de limiarização bimodal, com a diferença que possibilita a segmentação em múltiplas classes [MAR00].

A importância da segmentação por limiarização é tão elevada, que existem na literatura centenas de algoritmos quanto à mesma [MAT99]. Na presente pesquisa será detalhado o método de limiarização global clássico de [OTS79]. Os métodos de limiarização baseados na entropia serão detalhados no Capítulo 3.

[OTS79]: este método baseia-se na análise de discriminantes. A operação de limiarização é obtida pelo particionamento dos *pixels* de uma imagem em duas classes $C_0 = \{0, 1, 2, \dots, t\}$ e $C_1 = \{t+1, t+2, \dots, l-1\}$, isto é, objetos e fundo. Pode-se definir as características das classes C_0 e C_1 em termos de distribuições de Gauss (distribuição normal), seja a partir de suas médias $\mu_0(t)$, $\mu_1(t)$, das variâncias intra-classes $\sigma_w^2(t)$, inter-classes $\sigma_B^2(t)$ e da variância total σ_T^2 .

Essas médias e variâncias podem ser calculadas da seguinte maneira:

$$\mu_0(t) = \frac{\mu_t(t)}{\omega_0(t)} \quad (2.1)$$

$$\mu_1(t) = \frac{\mu_t - \mu_t(t)}{1 - \omega_0(t)} \quad (2.2)$$

$$\sigma_B^2 = \omega_0(t)\omega_1(t)[\mu_0(t) - \mu_1(t)]^2 \quad (2.3)$$

$$\sigma_w^2(t) = \sigma_t^2 - \sigma_B^2 \quad (2.4)$$

$$\sigma_T^2 = \sum_{g=0}^{255} (g - \mu_t)^2 p_g \quad (2.5)$$

em que:

$$\omega_0(t) = \sum_{g=0}^t p_g \quad (2.6)$$

$$\omega_1(t) = \sum_{g=t+1}^{255} p_g = 1 - \omega_0(t) \quad (2.7)$$

$$\mu_t(t) = \sum_{i=0}^t gp_g \quad (2.8)$$

$$\mu_T = \sum_{g=0}^{255} gp_g \quad (2.9)$$

$$p_g = \frac{h_g}{N} \quad (2.10)$$

em que:

N é o número de *pixels* e $h(g)$ denota o valor do histograma para cada nível de cinza g .

2.1.2. Segmentação por Cor

O uso da cor no processamento digital de imagens é motivado principalmente por dois fatores [GON92]):

- A motivação pelo uso da cor é que o olho humano é capaz de distinguir centenas de cores, mesmo sob as mais severas condições;
- A cor é um descritor poderoso que freqüentemente simplifica a extração de características de uma imagem.

Uma das mais importantes contribuições da cor verifica-se em imagens médicas, onde a mesma é utilizada como auxiliar na visualização de imagens microscópicas. A utilização de corantes possibilita o destaque de regiões de interesse caracterizadas por um determinado tipo de substância, objeto ou microorganismo presente nas amostras microscópicas [BRI96].

Contudo, para aplicar a cor no processamento digital de imagens, é necessário que se tenha conhecimento dos principais modelos de representação da mesma. De uma forma geral, os modelos fornecem uma padronização para a cor, permitindo a sua quantificação independente dos observadores. Diversos são os modelos para a representação da cor encontrados na literatura. Entretanto, as três principais abordagens são:

- RGB: consiste na utilização de um espaço tridimensional onde cada cor é obtida a partir das cores primárias vermelha, verde e azul, respectivamente baixa, média e alta freqüências do espectro visual pelo olho humano.

- HSI: H é a matiz (descreve a cor pura: vermelha, azul, amarela); S a saturação da cor (grau de pureza); I é a componente que representa a intensidade de luz (brilho) da cor.
- CMY: é formado pelas cores ciano, magenta e amarelo respectivamente. Está associado ao modelo RGB e é utilizado por muitos dispositivos que depositam pigmentos coloridos em papel, tais como impressoras e copiadoras.

Como exemplo de segmentação empregando a cor, pode citar-se [BRI96] que propôs um algoritmo para segmentação por cor. O algoritmo caracteriza-se pelo uso do sistema de representação da cor HSI e toma por base propriedades estatísticas da imagem, através da transformada de *Karhunen-Loève*, que melhor distribui a informação para a devida segmentação. Este algoritmo é aplicado em um processo de auxílio a diagnóstico médico, baseado na quantificação de mucina a partir de imagens de biópsia retal realizadas em ratos da linhagem *Sprague-Dawley*. Basicamente, o seu objetivo é a extração de informações de uma cena, tomando por base o atributo cor.

Do ponto de vista computacional, o processamento da informação cor implica em considerável aumento dos custos, além de precisar de um espaço de armazenamento três vezes maior. Na maior parte das aplicações, como meio de ganho de performance, utiliza-se a transformação de imagens originalmente coloridas para níveis de cinza. Mas, dependendo do contexto a ser analisado, nem sempre o contraste da imagem monocromática é suficiente para se obter bons resultados no processamento digital de imagens.

2.1.3. Segmentação por Textura

Como descrito anteriormente, a textura pode ser definida como sendo padrões que são repetidos periodicamente [TAM78]. Suas aplicações são diversas, como por exemplo, análise de imagens médicas, processamento de documentos e sensoriamento remoto, dentre outros.

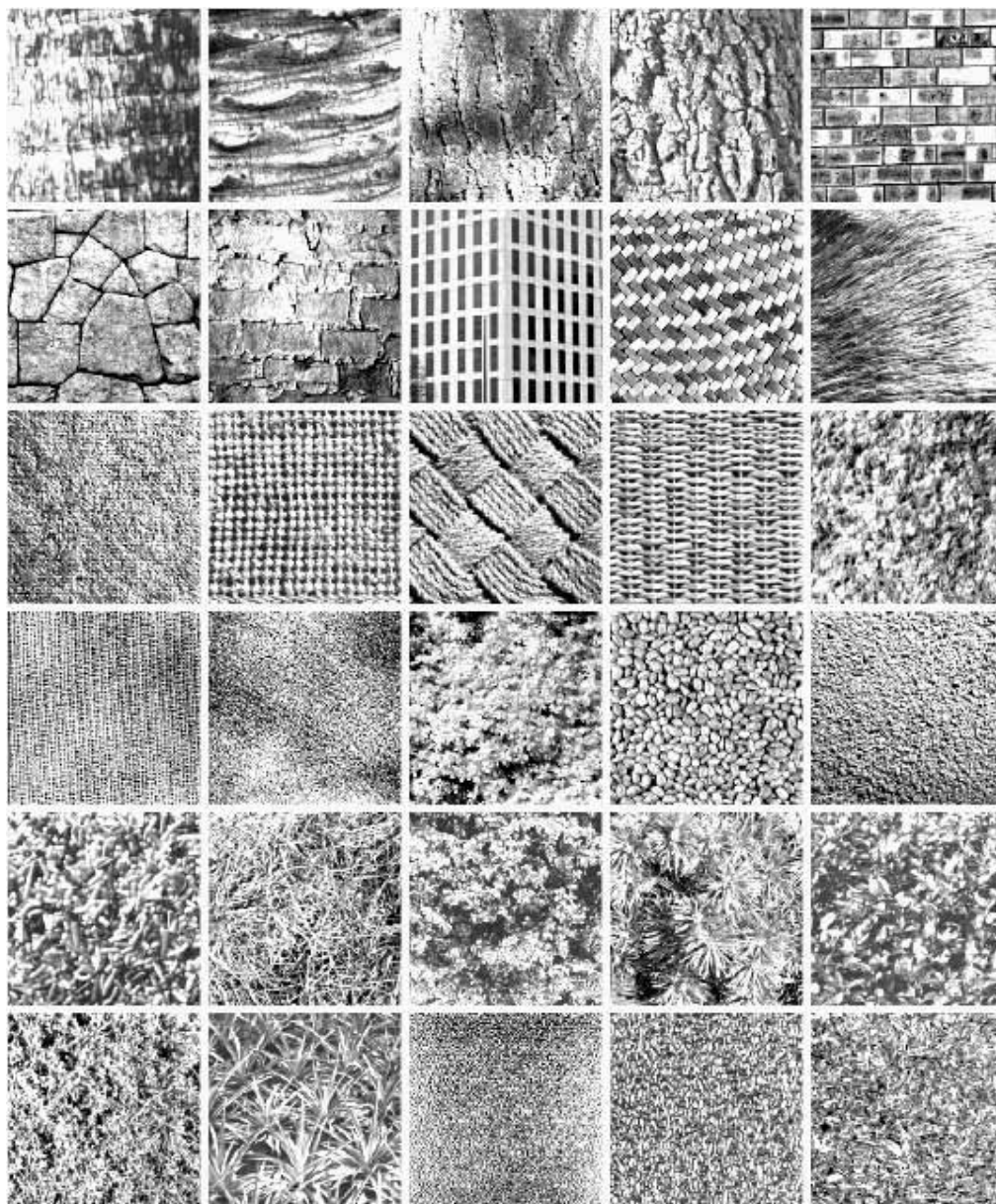


Figura 2.1 Diferentes tipos de textura

Algumas propriedades seguem a noção de textura:

- A textura de um único ponto é indefinida. Portanto, é uma propriedade contextual e sua definição envolve níveis de cinza em uma dada vizinhança;
- A textura envolve a distribuição espacial dos níveis de cinza. Portanto, a Matriz de Co-Ocorrência, que será detalhada no Capítulo 3, é essencial e amplamente utilizada na análise de textura.

Na literatura, existem diversos métodos para a análise de textura, dentre os quais podem ser classificados [FAC01]:

- Métodos estatísticos: são importantes na análise de textura graças à desejável propriedade de envolver a distribuição espacial dos níveis de cinza. A Matriz de Co-Ocorrência, que será descrita detalhadamente no Capítulo 3, se constitui na principal abordagem dos métodos estatísticos;
- Métodos estruturais: 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.

Pode-se citar vários trabalhos relacionados com a segmentação de texturas. Na área médica, [LAN78] extraíram as estatísticas de primeira e segunda ordem (ver Capítulo 3) para diferenciar células do sangue. [CHE89] utilizaram fractais para classificar imagens de ultrassom de fígados. Na área de sensoriamento remoto, [RIG90] utilizaram a Matriz de Co-Ocorrência para extrair características de textura com o objetivo de segmentar imagens SAR.

2.1.4. Segmentação por Contorno

Uma borda é, por definição, uma mudança brusca do nível de cinza entre duas regiões relativamente homogêneas. Ela pode aparecer como uma seqüência de pontos, uma linha, um segmento, uma curva ou uma forte variação do nível de cinza médio. Basicamente, a idéia predominante nas técnicas de detecção de bordas é o cálculo de um operador de derivada local, tais como os gradientes de *Sobel* e *Prewitt* (derivada primeira) e o *Laplaciano* (derivada segunda) [GON92] [FAC01].

De acordo com [CAN86] e [DER87], um detector de bordas eficiente deve possuir as seguintes propriedades:

- Detecção: o operador deve fornecer uma resposta na vizinhança da borda;
- Localização: a borda deve ser localizada com precisão;
- Unicidade: uma borda deve provocar uma única resposta do operador de extração.

É importante salientar que os operadores de detecção de bordas de [CAN86] e [DER87] são atualmente considerados como os melhores em [FAC01]:

- Detectar bordas verdadeiras e não gerar falsos contornos;
- Gerar bordas de um *pixel* de largura;
- Avaliar de forma objetiva os resultados obtidos.

2.1.5. Segmentação por *Clustering*

As técnicas de segmentação por *clustering* rotulam regiões de uma imagem através do particionamento de um conjunto de características em grupos compactos e bem definidos no espaço de características. São inúmeras as técnicas de *clustering*. Nesta seção, apresentaremos 2 das técnicas mais conhecidas:

Fuzzy c-means: A técnica de clustering por Fuzzy c-means minimiza a função objetivo J_m com relação ao grau fuzzy $\mu_{i,j}$ e ao centro do cluster V_i . A função objetivo é definida como:

$$J_m = \sum_{i=1}^c \sum_{j=1}^n (\mu_{i,j})^m d^2(X_j, V_i)$$

em que:

$$d^2(X_j, V_i) = (X_j, V_i)^T A(X_j, V_i)$$

A pode ser qualquer positivo definido em uma matriz $p \times p$, onde p é a dimensão do vetor de características X_j , c é o número de clusters e n o número de pontos.

Erro quadrático: Este método de *clustering* usa o critério de erro quadrático para um conjunto de partições no espaço de características. Após uma inicialização de k clusters seguida pela determinação da medida do erro quadrático mínimo entre os dados em que k é o limite superior dos clusters, o processo de particionamento se encerra quando não há mudança na rotulação dos clusters.

2.2. Segmentação de Imagens de Envelopes Postais

A automação postal representa um segmento da economia de extrema importância, devido à quantidade de envelopes postais manipulados diariamente. Com o avanço tecnológico de sistemas de reconhecimento ótico de caracteres (*OCR-Optical Character Recognition*), de fato, tornou-se possível o reconhecimento de envelopes postais que possuem o bloco-endereço pré-impresso (*machine-printed*) e, em alguns casos, manuscrito (*handwritten*). Contudo, a localização rápida e eficiente da posição e orientação do bloco-endereço, sobretudo em casos complexos, onde o envelope postal contém, além do bloco-endereço manuscrito, selos, carimbos e fundo com textura complexa, ainda não apresenta uma solução definitiva. Os métodos de localização do bloco-endereço baseiam-se em técnicas de limiarização, inteligência artificial (sistemas especialistas), características geométricas ou análise textural por filtragem.

2.2.1. Método de [WOL97]

[WOL97] propuseram um método rápido de localização do bloco- endereço de imagens de envelopes postais. Primeiramente, a imagem é dividida em pequenos blocos retangulares $B_{l,m}$ e a magnitude do gradiente médio $\bar{G}_{l,m}$ é calculada para cada bloco. Esta etapa é necessária para a classificação de blocos homogêneos e heterogêneos. Considere $|B_{l,m}|$ ser o número de *pixels* do bloco $B_{l,m}$, Então, $\bar{G}_{l,m}$ pode ser definida da seguinte maneira:

$$\bar{G}_{l,m} = \frac{1}{|B_{l,m}|} \sum_{(j,k) \in B_{l,m}} g_{j,k} \quad (2.11)$$

em que:

$g_{j,k}$ denota a magnitude do gradiente na posição (j,k)

Um limiar θ é definido através da magnitude do gradiente médio da imagem ponderada por uma constante c :

$$\theta = c * \left(\frac{1}{w * h} \sum_i^w \sum_j^h g_{i,j} \right) \quad (2.12)$$

em que:

c denota um valor constante

$g_{j,k}$ denota a magnitude do gradiente na posição (j,k)

w e h denotam o tamanho da imagem reduzida

Um bloco $B_{l,m}$ é marcado como heterogêneo se a magnitude do gradiente médio $\bar{G}_{l,m}$ do bloco correspondente é maior que o limiar θ . Para envelopes que contém somente o bloco-endereço, selos iguais e remetente, a detecção de blocos heterogêneos revela-se eficiente. Se o envelope contiver logotipos e diversos selos diferentes, além de carimbos, o método falha porque o limiar θ avaliado tende a aumentar justamente devido as regiões com alto contraste, uma vez que a magnitude do gradiente médio dos blocos $B_{l,m}$ correspondentes ao bloco-endereço é muito baixa. Uma melhoria necessária é então introduzida ao se considerar dois limiares θ_l e θ_r ao invés de um único limiar θ , correspondentes à metade esquerda e à direita respectivamente do envelope.

$$\theta_l = c * \frac{1}{num_l} \sum_{(i,j) \in L} g_{i,j} \quad (2.13)$$

$$\theta_r = c * \frac{1}{num_r} \sum_{(i,j) \in R} g_{i,j} \quad (2.14)$$

Para detectar os candidatos a bloco-endereço, os blocos significativos (homogêneos e heterogêneos) são fundidos a áreas conectadas de forma arbitrária. Isto é feito através da rotulação dos blocos encontrados. Na primeira iteração, um único rótulo de valor 1 é atribuído a todos os blocos, começando da esquerda superior até a direita inferior. Blocos à esquerda e acima que não possuem rótulo são marcados com um novo rótulo. Se um dos vizinhos já está rotulado, este rótulo é então atribuído ao bloco em questão. Se os rótulos de ambos os vizinhos diferem, o bloco atual é rotulado com o rótulo do vizinho à esquerda. Este método foi testado para 2000 imagens de envelopes, com uma taxa de acerto de 98%. Contudo, o sistema se mostra ineficiente em casos onde o contraste do bloco-endereço é pobre, ou se o carimbo está misturado com o bloco-endereço. Também não está claro em quantos blocos a imagem deve ser dividida.

2.2.2. Método de [YU97]

Nesse método, os autores trabalharam sobre imagens de envelopes plastificados e de jornais. Devido a grande variabilidade, [YU97] propuseram uma classificação dos envelopes postais em simples ou complexo, de acordo com o número de classes presentes no envelope postal tais como selos, carimbos, logotipos, ruídos, a presença ou não de um fundo complexo e embalagem plástica. Portanto, o número de classes presentes é fator determinante para a classificação dos mesmos.

Os envelopes postais ainda foram classificados em pré-impressos e manuscritos. Para os envelopes pré-impressos, algumas regras foram observadas:

- satisfaz a regra do alinhamento à esquerda;
- o bloco-endereço geralmente está encoberto por uma embalagem plástica, e apresenta ainda alguma inclinação, dificultando assim sua localização.

O método proposto por [YU97] pode ser dividido em duas fases distintas. Na primeira, as imagens coloridas (formato RGB) são transformadas em imagens binárias através do canal vermelho. O método de binarização empregado é uma adaptação da técnica de [OTS79]. De fato, no caso de envelopes postais, segundo os autores, o desequilíbrio entre as variâncias dos níveis de cinza do fundo e do primeiro plano não permitem à técnica de Otsu de fornecer um valor de binarização adequado.

Para contornar esta deficiência, [YU97] generalizaram o critério de seleção do limiar k^* do método de Otsu como sendo $k^* = \operatorname{argmax}_k \xi^\alpha(k)$ onde :

$$\xi^\alpha(K) = \omega_0(K)\omega_1(K) |\mu_0(K) - \mu_1(K)|^\alpha \quad (2.15)$$

em que:

$\mu_0(K)$, $\mu_1(K)$ e $\omega_0(K)$, $\omega_1(K)$ representam respectivamente as médias e as variâncias dos níveis de cinza das duas classes em função do limiar k .

O caso $\alpha=2$ corresponde ao método tradicional de Otsu. Os autores selecionaram $\alpha=0.75$ como sendo o valor ideal para os testes. A segunda fase da abordagem de localização do bloco endereço de envelopes postais de [YU97] consiste em:

- extrair, a partir da imagem binária todos os componentes conexos, pelo método de grafos de blocos adjacentes (BAG);
- avaliar a complexidade dos envelopes a partir do número M de componentes conexos. No caso onde $M < 350$, o envelope é considerado simples, e no caso contrário complexo;
- preservar, supondo que o bloco endereço é impresso em um rótulo de cor clara, os componentes conexos com a metade de seus *pixels* vizinhos de cor clara. O critério usado é $\min[r(x,y), g(x,y), b(x,y)] > 160$, onde r , g e b representam os canais R,G e B;
- modelar o bloco endereço caracterizando, nas direções horizontais e verticais, os blocos de linhas de texto que podem ser identificados como sendo blocos endereços. Para isto, para um bloco de linhas de texto candidato, critérios de alinhamento de texto são usados assim que o número de *pixels* escuros antes e depois do bloco, e que a densidade de *pixels* escuros no bloco;
- definir um critério de confiança para cada bloco de linhas de texto classificado como sendo um bloco endereço. São conservadas somente os blocos de linhas de texto candidatos com critérios de confiança mais elevados.

O sistema foi testado em uma base de 109 imagens de envelopes do tipo carta, revista e jornal que apresentavam endereços manuscritos, rótulos inclinados e envelopes plastificados com pouco contraste. A vantagem do sistema proposto é o tratamento de imagens de envelopes complexos, porém, novamente, como desvantagens, pode-se citar o uso de muitas heurísticas, além da complexidade e tempo de processamento elevados.

2.2.3. Método de [WAN88]

Neste método, [WAN88] propõem um sistema de localização automático do bloco-endereço (*ABLS- Address Block Location Subsystem*) utilizando sistemas especialistas. O sistema ABLs consiste de 6 etapas principais:

1. Base de informações estatísticas de envelopes postais: contém informações estatísticas das características geométricas dos rótulos dos endereços. Esta base de dados foi gerada após uma análise criteriosa da base de imagens;
2. Caixa de ferramentas: contém ferramentas de processamento de imagens para limiarização, segmentação, rotulação de componentes conectados, discriminação entre texto impresso e manuscrito, divisão e junção de blocos, além de análise de layout e características dos blocos de texto;
3. Máquina de inferência baseada em regras: atua como um interpretador de regras associado com cada ferramenta;
4. Sistema de controle: seleciona as ferramentas apropriadas da caixa de ferramentas. A utilidade de cada ferramenta no contexto atual é estimada e a que apresenta a maior utilidade é selecionada. O sistema de controle também combina novas evidências;
5. Dados de controle: usados pelo sistema de controle para obter informações sobre as interdependências entre ferramentas e critérios para aceitar um bloco como sendo o bloco-endereço;
6. *Blackboard*: contém os atributos geométricos dos blocos obtidos após o processamento de diferentes tipos de imagens.

O sistema ABLs pode usar uma ou mais ferramentas de processamento de imagens para processar os envelopes postais. As evidências acumuladas através dessas ferramentas são integradas ao *blackboard*. Cada envelope postal contém um valor de confiança que é uma medida do grau (probabilidade *a posteriori*) em que um bloco em particular seja o bloco-endereço. Baseado na evidência disponível, um score é atribuído a cada candidato a bloco-endereço. Finalmente, o candidato com o mais alto score é selecionado como sendo do bloco-endereço.

Uma base contendo mais de 3000 imagens foi utilizada para desenvolver as regras associadas com as ferramentas. Experimentos em 174 envelopes postais mostraram que o método localizou o bloco-endereço em 81% dos casos. Em 4% dos casos, o bloco-endereço

foi corretamente localizado, mas o mesmo não continha informações suficientes para o reconhecimento. O método falhou em 10% dos casos e nos 5% das imagens restantes, o método foi incapaz de decidir qual bloco era o bloco-endereço. A vantagem do método proposto é sua capacidade de processar imagens multi-espectrais (níveis de cinza, coloridas e infravermelhas). Porém, a desvantagem reside no elevado número de imagens (3000) necessárias para desenvolver as regras necessárias para o funcionamento do mesmo.

2.2.4. Método de [DOW90]

[DOW90] propõem um método de localização do bloco-endereço que utiliza características geométricas. Primeiramente, a imagem de entrada em níveis de cinza é limiarizada. Então, para cada *pixel* na imagem limiarizada, seus vizinhos na horizontal são analisados para identificar regiões de alta densidade de *pixels* pretos. O valor da vizinhança é escolhido baseado no espaçamento entre palavras, tal que as palavras em cada linha do endereço estejam conectadas.

A imagem limiarizada é reduzida de 2048 x 512 para 256 x 64 *pixels* para diminuir o tempo de processamento. Para cada bloco adjacente de 8 x 8 na imagem limiarizada, o *pixel* correspondente na imagem reduzida é preto se o bloco de entrada contém mais de 30 *pixels* pretos. Esta redução na resolução é necessária para remover pequenas regiões de alta densidade de *pixels* pretos que não fazem parte do bloco-endereço. Os blocos que contém *pixels* pretos são então agrupados para que formem regiões separadas. Então, as seguintes características são extraídas para cada uma das regiões:

- Coordenadas de cada região;
- Orientação de cada região;
- Densidade de *pixels* pretos de cada região;
- Densidade de *pixels* pretos da imagem limiarizada.

Estas características são utilizadas para identificar regiões que são candidatos a fazer parte do bloco-endereço. Portanto, para cada *pixel* destas regiões, seus vizinhos na vertical são analisados para juntar linhas nas regiões. Finalmente, o bloco que se encontra mais perto de um ponto que é definido como sendo a metade da largura da imagem de entrada e 1/3 da altura é considerado como sendo o bloco-endereço.

[DOW90] testaram o método proposto em vários envelopes postais. Para envelopes pré-impressos, houve um acerto de 98%.

2.2.5. Método de [YEH87]

Assim como [DOW90], [YEH87] também utilizam características geométricas para a localização do bloco-endereço. Este método aplica o filtro do Laplaciano, seguido por uma análise criteriosa de componentes conectados para identificar cada conjunto individual de 8 *pixels*. Isto produz um número muito grande de componentes, muitos deles espúrios. Um processo de junção é utilizado para eliminar os componentes conectados indesejáveis. Este processo junta componentes adjacentes se o contraste relativo entre os mesmos for baixo. Para cada um dos componentes remanescentes, as seguintes características são extraídas:

1. Posição;
2. Dimensões;
3. Nível de cinza médio;
4. Espessura;
5. Comparação com o fundo.

Esta última etapa determina a probabilidade de que o componente seja mais escuro que o fundo. Os componentes conectados são agrupados baseados nas similaridades de suas características geométricas. Uma árvore de proximidade, que é baseada no algoritmo de agrupamento hierárquico (*hierarchical clustering algorithm*), é construída, e representa a similaridade entre componentes conectados. Portanto, cada cluster deve pertencer a uma das seguintes categorias: destinatário, remetente e selos. Funções ponderadas correspondentes às diferentes categorias são utilizadas para atribuir scores a cada cluster, baseados em suas posições. Por exemplo, quanto mais próximo um cluster estiver do canto superior direito, maior será o seu score relativo à categoria selos. Finalmente, o cluster com o mais alto score relativo a categoria destinatário é selecionado como o bloco-endereço do destinatário. A vantagem desse método é sua simplicidade. Porém, a principal desvantagem reside no fato de que é necessário um conhecimento a priori dos envelopes postais, o que é quase impossível devido a variabilidade dos mesmos.

2.2.6. Método de [JAI95]

[JAI95] propõem um método cujo objetivo é identificar regiões que consistem de texto, não-texto (espaços em branco e regiões cuja variação de intensidade são baixas), e bordas entre as regiões com texto e sem texto.

Para tanto, [JAI95] utilizam filtros de Gabor bidimensionais, uma vez que os mesmos apresentam propriedades de localização tanto no domínio espacial quanto freqüencial e

portanto são ideais para problemas de segmentação por textura. Um filtro de Gabor bidimensional pode ser visto como um plano sinusoidal de uma frequência em particular e orientação, modulados por um envelope Gaussiano. No domínio espacial, um filtro de Gabor bidimensional é dado por:

$$h(x, y) = \exp\left\{-\frac{1}{2}\left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right]\right\} \cos(2\pi\mu_0 x) \quad (2.16)$$

em que:

σ_x e σ_y são os desvios padrões do envelope Gaussiano nas direções x e y .

μ_0 é a frequência da onda no plano sinusoidal na direção x .

Um esquema de representação de textura utilizando filtros de Gabor consiste de duas etapas principais [JAI95]:

1. Filtrar a imagem de entrada através de n filtros de Gabor para obter n imagens filtradas;
2. Extrair as características das imagens que consistem no cálculo da “energia local” estimada através de janelas Gaussianas de tamanho apropriado centradas em cada *pixel* em cada uma das imagens filtradas. Portanto, cada *pixel* em uma imagem filtrada é representado como um ponto em um espaço de características n -dimensional.

Regiões com textura homogênea podem ser encontradas tanto no modo supervisionado quanto no modo não-supervisionado. Para uma imagem com N colunas, onde N é potência de 2, os filtros de Gabor com frequências $1\sqrt{2}$, $2\sqrt{2}$, $4\sqrt{2}$, ..., $(N/4)\sqrt{2}$ ciclos/imagem são selecionados. Para cada frequência radial μ_0 , 4 filtros, com orientações $\theta = 0^\circ, 45^\circ, 90^\circ$ e 135° são utilizados. Portanto para uma imagem com $N = 256$, um total de 28 filtros, ($n = 4 \times 7$) seriam utilizados.

Baseados em experimentos iniciais com imagens de documentos, [JAI95] escolheram somente 8 filtros correspondentes às 2 frequências mais altas. A justificativa reside no fato de que estas 2 frequências identificam as regiões correspondentes às texturas de mais alta frequência que são típicas das regiões que contém texto. No caso dos envelopes postais, [JAI95] utilizaram o modo não-supervisionado para a localização do bloco-endereço.

O algoritmo que descreve o método proposto por [JAI95] pode ser sumarizado da seguinte maneira:

1. Obter uma segmentação em 3 clusters da imagem de entrada;

2. Identificar os componentes conectados pertencentes às regiões com texto;
3. Eliminar componentes conectados muito pequenos ($< 30 \text{ pixels}$) e muito altos ($> 1500 \text{ pixels}$);
4. Para cada componente conectado, efetuar uma análise de componentes principais (ACP) para calcular os eixos principais e secundários dos componentes conectados. Construir o menor retângulo (*bounding box*) para cada componente conectado;
5. Construir um histograma dos eixos principais. Selecionar a direção correspondente à moda do histograma como o eixo principal comum. O eixo secundário é ortogonal ao eixo principal;
6. Projetar as coordenadas dos *pixels* de cada componente conectado no eixo secundário para agrupar linhas de texto. Para cada grupo de linhas, projetar as coordenadas dos *pixels* no eixo principal para identificar grupos de palavras;
7. Para cada grupo de palavras, construir um retângulo (*bounding box*) alinhado com o sistema de coordenadas definidos pelos eixos principal e secundário calculados no passo 5. Cada retângulo é portanto, um candidato a bloco-endereço.

[JAI95] propõem o uso de heurísticas que podem variar de simples análise das posições dos candidatos a bloco-endereço a outras mais complexas como análise de layout para decidir de forma correta o bloco-endereço. Os testes foram realizados em envelopes de tamanho padrão que não apresentavam nenhuma outra mensagem de texto que o bloco endereço. As vantagens do método são que, nenhuma técnica de binarização é necessária para localizar os componentes conexos e o método é invariante à inclinação dos envelopes postais. Porém, segundo os autores, a principal desvantagem é a complexidade do método de *clustering*.

2.2.7. Método de [XUE (1999)]

[XUE (1999)] propõem um método de localização do bloco-endereço que utiliza características geométricas. Para tal, uma base composta de 500 imagens de envelopes postais chineses com textura simples (fundo branco e regular) é utilizado. [XUE (1999)] dividem este problema em 3 sub-problemas: extração de blocos-texto, linhas-bloco e localização do bloco-endereço baseada em linhas-texto. Para a extração de blocos-texto, os componentes conectados são classificados em 4 categorias: ruído, texto, gráfico e imagem. Primeiramente, se os componentes conectados satisfazem as condições de que a largura e altura do retângulo circunscrito do componente conectado são menores do que algum limiar definido previamente, então o mesmo é classificado como ruído. Feito isso, algumas regras são usadas

para encontrar blocos correspondentes a gráficos. Se os componentes conectados satisfazem certas regras, então este é rotulado como sendo correspondente a gráfico. Por exemplo, certas regras propostas são definidas da seguinte maneira:

- Se ($\text{largura}(\text{cc})/\text{altura}(\text{cc}) > T_w$) então cc is a gráfico;
- Se ($\text{altura}(\text{cc})/\text{largura}(\text{cc}) > T_v$) então cc is a gráfico;
- Se ($\text{pixnum}(\text{cc})/\text{area}(\text{cc}) > T_d$) então cc is a gráfico.

em que:

$\text{largura}(\text{cc})$, $\text{altura}(\text{cc})$, $\text{area}(\text{cc})$ representam a largura, altura e área do retângulo circunscrito correspondente ao componente conectado cc, respectivamente; $\text{pixnum}(\text{cc})$ denota o número de pixels pretos do componente conectado cc; T_w , T_v e T_d são limiares pré-definidos.

Finalmente, somente blocos correspondentes às imagens e blocos-texto são analisados. [XUE (1999)] extraíram duas características dos componentes conectados: uma é a área do retângulo circunscrito e a outra é o número de componentes conectados. Linhas-texto podem ser então geradas através da fusão de blocos-texto. Os autores propõem um método *split-and-merge* baseado em várias heurísticas, como por exemplo, a distância vertical entre os blocos-texto. A localização do bloco-endereço é efetuada através de uma análise criteriosa das linhas-texto. A linha-texto com o mais alto *score* é selecionada como sendo o bloco-endereço. O sistema foi capaz de localizar corretamente o bloco-endereço em 81.7% dos casos.

2.3. Conclusão

Neste capítulo foram apresentadas as principais técnicas de segmentação presentes na literatura, tais como, a segmentação por limiarização, cor, textura, contorno e *clustering* bem como uma descrição detalhada dos principais métodos de localização do bloco-endereço.

Capítulo 3

Histograma 2D e Matriz de Co-Ocorrência

Neste capítulo serão descritos em detalhes a entropia, o histograma 2D e a matriz de co-ocorrência. Suas aplicações no processamento digital de imagens, exemplos de segmentação por limiarização utilizando entropia, histograma 2D e matriz de co-ocorrência, estatísticas de primeira e segunda ordem bem como suas vantagens e desvantagens também são assuntos deste capítulo.

3.1. Entropia

[SHA49] define a entropia de um sistema como uma função de probabilidade de ocorrência de diferentes estados de um sistema. Se um sistema tem n diferentes estados com probabilidade de ocorrência p_i , $i = 1, 2, \dots, n$, $\sum_{i=1}^n p_i = 1$, então o ganho de informação da ocorrência do evento i é definido como: $\Delta I = -\log_2 p_i$. O valor esperado de ganho de informação é definido como a entropia de um sistema. Portanto, a entropia H do sistema é:

$$H = -\sum_{i=1}^n p_i \log_2 p_i \quad (3.1)$$

Baseado no conceito de entropia de Shannon, diversos autores definiram a entropia para uma imagem [PUN80] [KAP85].

Considere F_i ser a freqüência do nível de cinza i em uma imagem de dimensões $M \times N$ com L níveis de cinza. Então $\sum_{i=0}^{L-1} F_i = M \times N = I$. Seguindo a definição de entropia de Shannon, [PUN80] definiu a entropia de uma imagem (histograma) da seguinte forma:

$$H = -\sum_{i=0}^{L-1} p_i \log_2 p_i \quad (3.2)$$

em que:

$$p_i = F_i / I \quad (3.3)$$

3.1.1. Entropia local e global

Sabe-se que em uma imagem as intensidades dos *pixels* (níveis de cinza) não são independentes umas das outras. Esta dependência pode ser analisada considerando as seqüências de *pixels* para estimar a entropia. Para tanto, Shannon demonstrou o seguinte teorema:

Teorema: Considere $p(s_i)$ ser a probabilidade de uma seqüência de níveis de cinza s_i de comprimento q . Esta seqüência é definida como uma permutação de q níveis de cinza. Portanto, pode-se definir:

$$H^{(q)} = -\frac{1}{q} \sum_i p(s_i) \log_2 p(s_i), \quad (3.4)$$

para todas as seqüências de níveis de cinza de comprimento q .

Portanto $H^{(q)}$ é uma função decrescente monotônica de (q) e $\lim_{q \rightarrow \infty} H^{(q)} = H$, ou seja, a entropia da imagem. Para diferentes valores de q , obtemos então várias ordens de entropia.

Caso 1 (Estatística de primeira ordem): Se $q = 1$, ou seja, comprimento unitário. Tem-se portanto:

$$H^{(1)} = -\sum_{i=0}^{L-1} p_i \log_2 p_i \quad (3.5)$$

em que:

p_i é a probabilidade de ocorrência do nível de cinza i .

Todavia, diferentes imagens com histogramas idênticos resultariam em um mesmo valor $H^{(1)}$. As definições utilizadas por [PUN80] e [KAP85] de fato pertencem a este caso.

Caso 2 (Estatística de segunda ordem): $q = 2$, ou seja, comprimento dois. Tem-se então de acordo com a equação (3.4):

$$H^{(2)} = -\frac{1}{2} \sum_{i=0}^{L-1} p(s_i) \log_2 p(s_i) \quad (3.6)$$

em que:

s_i é uma seqüência de níveis de cinza de comprimento dois, ou seja,

$$H^{(2)} = -\frac{1}{2} \sum_i \sum_j p_{ij} \log_2 p_{ij} \quad (3.7)$$

em que:

p_{ij} é a probabilidade de co-ocorrência dos níveis de cinza i e j .

Portanto, $H^{(2)}$ pode ser obtida através de uma matriz de co-ocorrência [PAL89] que será descrita detalhadamente na seção 3.3. $H^{(2)}$ ainda apresenta a vantagem desejável de levar em consideração a distribuição espacial dos *pixels*. Então, se duas imagens apresentam histogramas idênticos porém com distribuições espaciais diferentes resultarão em entropias diferentes, ou melhor, valores $H^{(2)}$ diferentes. Expressões para entropias de ordem ($q > 2$) podem ser deduzidas da mesma maneira. $H^{(i)}$, $i \geq 2$ pode ser chamada "entropia local" de ordem i de uma imagem.

3.1.2. Entropia condicional

Suponha que uma imagem tenha duas "porções" distintas, o objeto X e o fundo Y . Suponha também que os objetos consistem dos níveis de cinza $\{x_i\}$ e o fundo contém os níveis de cinza $\{y_i\}$. A entropia condicional do objeto X dado o fundo Y , pode ser definido como:

$$H(X/Y) = -\sum_{x_i \in X} \sum_{y_i \in Y} p(x_i/y_i) \log_2 p(x_i/y_i) \quad (3.8)$$

Similarmente, a entropia condicional do fundo Y dado o objeto X é definido como:

$$H(Y/X) = -\sum_{y_i \in Y} \sum_{x_i \in X} p(y_i/x_i) \log_2 p(y_i/x_i) \quad (3.9)$$

O *pixel* y_j , em geral, pode ser o vizinho de ordem m do *pixel* x_i , ou seja, y_j pode ser o m -ésimo *pixel* depois de x_i . Uma vez que a estimação de tal probabilidade é muito difícil, [PAL89] impuseram a restrição de que x_i e y_j sejam *pixels* adjacentes. Portanto, as equações (3.8) e (3.9) podem ser reescritas da seguinte maneira:

$$H(X/Y) = -\sum_{x_i \in X} \sum_{(x_i, y_i) \text{ adjacentes}} p(x_i/y_i) \log_2 p(x_i/y_i) \quad (3.10)$$

$$H(Y/X) = - \sum_{y_i \in Y} \sum_{x_i \in X} p(y_i/x_i) \log_2 p(y_i/x_i) \quad (3.11)$$

A entropia condicional de uma imagem pode então ser definida como:

$$H^{(c)} = (H(X/Y) + H(Y/X))/2 \quad (3.12)$$

em que:

X e Y representam o objeto e o fundo de uma imagem, respectivamente.

3.1.3. Exemplos de Segmentação usando a Entropia

A seguir serão apresentadas abordagens de segmentação empregando ao mesmo tempo histograma 1D e entropia. Vale lembrar que um histograma 1D reúne informações sobre a ocorrência dos L níveis de cinza em uma imagem. Para isto é computado, para cada nível de cinza, o número de *pixels* nos quais ele ocorre.

3.1.4. Limiarização por Entropia de [PUN80]

Considere s como sendo o limiar que divide a imagem em objeto e fundo. Considere também N_B e N_W ser o número de *pixels* pretos e brancos da imagem. Então a probabilidade a posteriori de um *pixel* preto é $P_B = N_B/N$ e de um *pixel* branco é $P_W = N_W/N$. Então, a entropia a posteriori da imagem é:

$$\begin{aligned} H'_L(s) &= -P_B \log_2 P_B - P_W \log_2 P_W \\ &= -P_S \log_2 P_S - (1-P_S) \log_2 (1-P_S) \end{aligned} \quad (3.13)$$

em que:

$$P_S = \sum_{i=0}^s p_i = P_B \text{ e } P_W = 1 - P_S \quad (3.14)$$

[PUN80] maximizou uma função $g(s)$ de $H'_L(s)$, onde:

$$g(s) = \frac{H_B^s \log_2 P_S}{H_L \log_2 [\max(p_0, p_1, \dots, p_s)]} + \frac{(H_L - H_B^s) \log_2 (1-P_S)}{H_L \log_2 [\max(p_{s+1}, p_1, \dots, p_{L-1})]} \quad (3.15)$$

em que:

$$H_L = - \sum_{i=0}^{L-1} p_i \log_2 p_i \quad (3.16)$$

e

$$H_B^{(s)} = -\sum_{i=0}^s p_i \log_2 p_i \quad (3.17)$$

O valor de s que maximiza $g(s)$ pode ser então definida como o limiar para a classificação de objeto e fundo.

3.1.4. Limiarização por Entropia de [KAP85]

Os autores acima citados também fazem o uso da definição de entropia de Shannon, porém de um ponto de vista diferente. Ao invés de considerarem somente uma distribuição de probabilidade para a imagem, é levada em consideração duas distribuições de probabilidade, uma para o objeto e outra para o fundo. A soma das entropias do objeto e do fundo são então maximizados. Se s é o limiar encontrado, então a distribuição de probabilidade dos níveis de cinza escuros é:

$$\frac{p_0}{P_s}, \frac{p_1}{P_s}, \dots, \frac{p_s}{P_s},$$

e a distribuição de probabilidade dos níveis de cinza claros é:

$$\frac{p_{s+1}}{1-P_s}, \frac{p_{s+2}}{1-P_s}, \dots, \frac{p_{L-1}}{1-P_s},$$

A entropia da parte preta da imagem (objeto) da imagem é:

$$H_B^{(s)} = -\sum_{i=0}^s \frac{p_i}{P_s} \log_2 \left(\frac{p_i}{P_s} \right) \quad (3.18)$$

e da parte branca:

$$H_W^{(s)} = -\sum_{i=s+1}^{L-1} \frac{p_i}{1-P_s} \log_2 \left(\frac{p_i}{1-P_s} \right) \quad (3.19)$$

A entropia total da imagem é então definida como:

$$H_T^{(s)} = H_B^{(s)} + H_W^{(s)} \quad (3.20)$$

Portanto, o valor de s que maximiza $H_T^{(s)}$ fornece o limiar para a classificação do fundo e objeto.

Os dois métodos descritos acima assumem que uma imagem é representada somente pelo seu histograma. Então, é possível que diferentes imagens que apresentem histogramas idênticos resultem em valores de entropia idênticos independentemente das diferentes distribuições espaciais dos pixels, o que com certeza, pode prejudicar muito a etapa da segmentação [PAL89].

3.2. Histograma 2D

Como explicado anteriormente, um histograma 1D reúne informações sobre a ocorrência dos L níveis de cinza em uma imagem. Para isto, é computado para cada nível de cinza, o número de pixels correspondentes naquela imagem. Porém, com um histograma 1D, perde-se a noção de distribuição espacial dos níveis de cinza.

Pode se dizer que, com o histograma 1D, muita informação contida na imagem é perdida. Se a mesma fosse usada, com certeza resultaria em uma segmentação mais eficiente [BRI92] [ABU89].

3.2.1. Definição de Histograma 2D

Para recuperar a noção de distribuição espacial, foi definido o histograma 2D. Para tal, [ABU89] emprega duas imagens, a imagem original $f(x, y)$ e uma nova imagem $g(x, y)$ obtida pela aplicação do filtro da Média, como segue:

$$g(x, y) = \frac{1}{n^2} \sum_{i=-n/2}^{n/2} \sum_{j=-n/2}^{n/2} f(x+i, y+j) \quad (3.21)$$

em que:

$n \leq N$ representa o tamanho da máscara usada no filtro da Média (em geral $n = 3$ ou 5).

Considere agora a variável r_{ij} representando o número de ocorrências do par (i, j) com $i = f(x, y)$ e $j = g(x, y)$ sendo os níveis de cinza do pixel (x, y) na imagem original e na imagem da Média, respectivamente. Criou-se assim o histograma 2D que é a relação do número de ocorrências de todos os pares (i, j) , onde $0 \leq i, j \leq L$. Portanto, temos que $0 \leq r_{ij} \leq MN$.

A partir do histograma 2D, é possível definir a probabilidade p_{ij} do par (i, j) por:

$$p_{ij} = r_{ij} / MN \quad (3.22)$$

Por ser uma probabilidade, p_{ij} verifica que:

$$\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_{ij} = 1 \quad \text{para } i, j = 0, \dots, L-1; \quad (3.23)$$

3.2.2. Exemplo de Segmentação usando o Histograma 2D

Na literatura, percebe-se que o histograma 2D é bastante usado em várias abordagens de segmentação, principalmente para a segmentação por limiarização. Das abordagens existentes, destaca-se a abordagem de [ABU89] que será apresentada a seguir.

Na hipótese da imagem $f(x, y)$ ser limiarizada a partir de seu histograma 1D, e se $T \in G$ representa o seu limiar, então a imagem limiarizada $f_T(x, y)$ é uma imagem binária definida da seguinte maneira:

$$f_T(x, y) = \begin{cases} b_0, & \text{se } f(x, y) < T \\ b_1, & \text{se } f(x, y) \geq T \end{cases} \quad (3.24)$$

em que:

$$1 \leq b_0, T, b_1 \leq L$$

Considere agora que deseja-se limiarizar a imagem original $f(x, y)$ a partir de seu histograma 2D, representado pelo par $[f(x, y), g(x, y)]$, onde $g(x, y)$ representa a imagem processada pelo filtro da Média. Neste caso, um vetor 2D (S, T) , $1 \leq S, T \leq L$, representa o vetor limiar e então a imagem limiarizada $f_{S,T}(x, y)$ é uma imagem binária definida da seguinte maneira:

$$f_{S,T}(x, y) = \begin{cases} b_0 & \text{se } f(x, y) < S \text{ e } g(x, y) < T \\ b_1 & \text{se } f(x, y) \geq S \text{ e } g(x, y) \geq T \end{cases} \quad (3.25)$$

em que:

$$1 \leq b_0, S, T, b_1 \leq L$$

Como mostrado na Figura 3.1, existem L^2 elementos e o vetor 2D de limiarização (S, T) divide o histograma em quatro quadrantes. Teoricamente, os pixels no interior dos objetos ou do fundo devem contribuir principalmente perto da diagonal por causa da homogeneidade. Portanto, para os pixels dentro dos objetos ou do fundo, o nível de cinza de um pixel e sua média são similares. Já para os pixels na vizinhança das bordas entre objetos e fundo, o nível de cinza de um pixel difere bastante de sua média [ABU89].

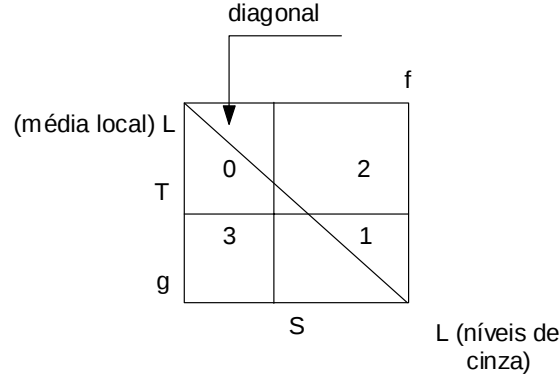


Figura 3.1 Quadrantes de um histograma 2D gerado pelo vetor 2D de limiarização (S, T)

Portanto, os quadrantes 0 e 1 contêm as distribuições dos objetos e do fundo, enquanto os quadrantes 2 e 3 contêm as distribuições dos pixels perto das bordas e o ruído.

Isto gera uma superfície composta de dois picos e um vale. Os picos correspondem aos objetos e o fundo [ABU89]. Eles podem ser separados através da escolha de um vetor (S, T) que maximize a entropia dos dois grupos (objetos e fundo). O método é descrito da seguinte maneira:

Se o limiar (na realidade, o vetor (S, T)) está localizado no par (s, t), as entropias dos dois grupos são definidas como:

$$H_0(s, t) = \log P_{st} + H_{st} / P_{st} \quad (3.26)$$

$$H_1(s, t) = \log(1 - P_{st}) + (H_{LL} - H_{st}) / (1 - P_{st}) \quad (3.27)$$

com:

$$P_{st} = \sum_{i=1}^s \sum_{j=1}^t p_{ij} \quad (3.28)$$

$$H_{st} = - \sum_{i=1}^s \sum_{j=1}^t p_{ij} \log p_{ij} \quad (3.29)$$

$$H_{LL} = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_{ij} \log p_{ij} \quad (3.30)$$

O vetor (S, T) selecionado tem que satisfazer o seguinte critério:

$$H(S, T) = \max \left(\min_{s=1, \dots, L; t=1, \dots, L} \{H_0(s, t), H_1(s, t)\} \right) \quad (3.31)$$

É obvio que, se comparado com a entropia 1D, a entropia 2D exige um tempo de processamento muito maior.

3.2.3. Vantagens e Desvantagens do Histograma 2D

O Histograma 2D é amplamente utilizado na segmentação por limiarização [ABU89] [BRI92] [CHE94] [LI97] [GON98]. A principal vantagem da utilização do Histograma 2D é o “resgate” da noção de distribuição espacial dos pixels através da ocorrência do par $[f(x,y), g(x,y)]$, que era perdido com o Histograma 1D. Porém, a grande desvantagem reside na escolha da 2ª imagem (imagem filtrada), que na verdade se trata de uma heurística. A escolha da mesma como sendo o filtro da média pode gerar novos valores que não existem na imagem original, o que pode ser inconveniente dependendo do contexto a ser analisado. Isso não acontece com operadores morfológicos básicos (erosão e dilatação), o que pode ser de suma importância. Na Figura 3.3 é mostrada a visualização do histograma 2D para as imagens da Figura 3.2.



a) Imagem Original

b) Filtro da Média (5x5)

Figura 3.2 Imagem Original e Filtrada para o processamento do histograma 2D

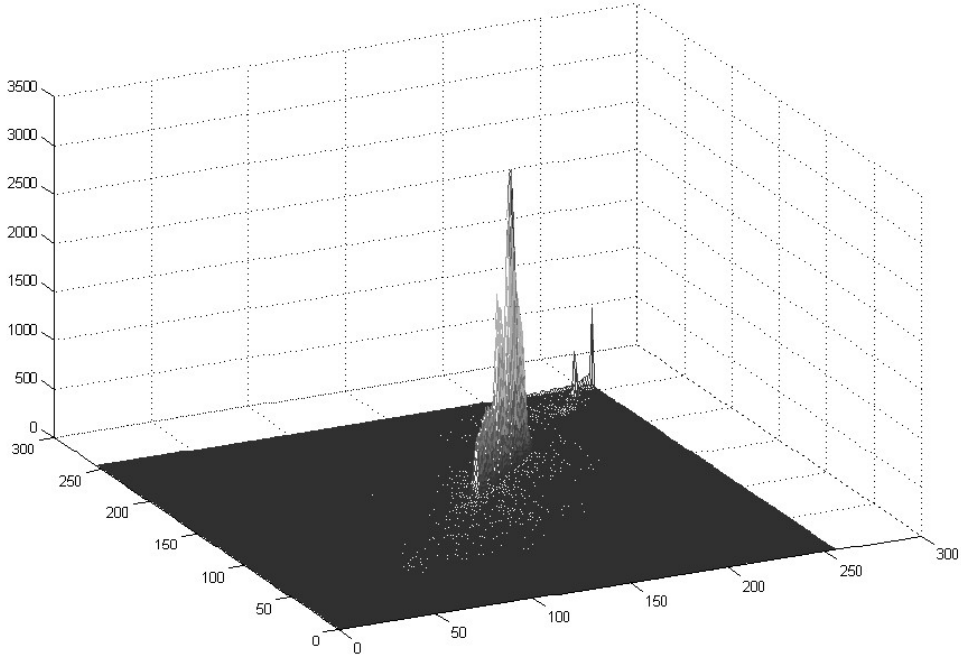


Figura 3.3 Visualização 3D do histograma 2D das imagens da Figura 3.2

3.3. Matriz de Co-Ocorrência

A matriz de co-ocorrência de uma imagem é uma matriz $T = [t_{ij}(v, h)]_{L \times L}$ de dimensão $(L \times L)$ que contém informação a respeito das transições de intensidade entre pixels adjacentes [PAL89]. Em outros termos, $t_{ij}(v, h)$ representa o número de vezes que o par de níveis de cinza (i, j) ocorre para dois pixels (k, l) e (m, n) distantes de h e v , respectivamente na horizontal e na vertical.

Pode-se formalizar $t_{ij}(v, h)$ da seguinte maneira:

$$t_{ij}(v, h) = \sum_{l=0}^{M-v} \sum_{k=0}^{N-h} \delta \quad (3.32)$$

em que:

$$\delta = 1 \text{ se } f(k, l) = i \text{ e } f(m, n) = j \text{ para} \quad (3.33)$$

$$|k - m| = h \text{ e } |l - n| = v$$

$\delta = 0$, caso contrário.

em que:

h e v representam a distância entre os dois pontos respectivamente na horizontal e na vertical.

3.3.1. Definição de [HAR92]

[HAR92] define a matriz de co-ocorrência como sendo simétrica e como função do relacionamento angular e da distância entre pixels. Abaixo, segue um exemplo de uma imagem de dimensões 4x4 com 4 níveis de cinza (0 a 3):

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Figura 3.4 Imagem de dimensões 4x4 e níveis de cinza correspondentes

Pode-se formalizar a matriz de co-ocorrência de acordo com diversos relacionamentos angulares bem como a distância entre os pixels:

$$C(i, j, d, 0^\circ) = \# \{[(k, l), (m, n)] \mid k - m = 0, |l - n| = d, I(k, l) = i, I(m, n) = j\}$$

$$C(i, j, d, 90^\circ) = \# \{[(k, l), (m, n)] \mid |k - m| = d, l - n = 0, I(k, l) = i, I(m, n) = j\}$$

$$C(i, j, d, 45^\circ) = \# \{[(k, l), (m, n)] \mid (k - m = d, l - n = -d) \text{ ou } (k - m = -d, l - n = d), I(k, l) = i, I(m, n) = j\}$$

$$C(i, j, d, 135^\circ) = \# \{[(k, l), (m, n)] \mid (k - m = d, l - n = d) \text{ ou } (k - m = -d, l - n = -d), I(k, l) = i, I(m, n) = j\}$$

em que:

o símbolo # denota o número de elementos do conjunto

0° = Horizontal, 90° = Vertical, 45° = Diagonal superior e 135° = Diagonal inferior

d é a distância entre os pixels

C denota as frequências de ocorrência do par (i, j) da matriz de co-ocorrência

$$\begin{array}{cc}
0^\circ \begin{pmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} & 45^\circ \begin{pmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\
90^\circ \begin{pmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix} & 135^\circ \begin{pmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix}
\end{array}$$

Figura 3.5 Resultados da matriz de co-ocorrência para diversos ângulos e distância 1 para o exemplo de imagem da Figura 3.4

3.3.2. Definição de [PAL89]

[PAL89] definiram a matriz de co-ocorrência como sendo assimétrica considerando somente as transições horizontal à direita, vertical para baixo, diagonal superior (sentido baixo-cima) e diagonal inferior (sentido cima-baixo). De acordo com [PAL89] a representação de uma matriz de co-ocorrência simétrica anteriormente definida por [HAR92] que leva em consideração as quatro direções, a saber: horizontal (à direita e à esquerda), vertical (para cima e para baixo), diagonais superior (sentidos baixo-cima e cima-baixo) e inferior (sentidos cima-baixo e baixo-cima) necessitam de um tempo maior em termos de processamento sem, contudo, trazer grandes benefícios. Portanto, $t_{ij}(v, h)$ (nas direções horizontal e vertical) é definido como segue:

$$t_{ij}(v, h) = \sum_{l=0}^{M-v} \sum_{k=0}^{N-h} \delta \quad (3.34)$$

em que:

$$\begin{array}{l}
\delta = 1 \text{ se} \\
\left\{ \begin{array}{l} f(l, k) = i \text{ e } f(l, k+1) = j \\ \textbf{ou} \\ f(l, k) = i \text{ e } f(l+1, k) = j \end{array} \right. \\
\delta = 0, \text{ caso contrário.}
\end{array} \quad (3.35)$$

3.3.3. Definição de [ZHA98]

[ZHA98] propôs uma modificação da matriz de co-ocorrência através da junção de várias direções, a saber: junção das direções horizontal e vertical, e junção das 4 direções (horizontal, vertical e diagonais). Portanto, a equação 3.34 sofre modificações considerando a junção horizontal e vertical:

$$\delta = 1 \text{ se, e somente se,} \quad (3.36)$$

$$\left\{ \begin{array}{l} f(l, k) = i \text{ e } f(l, k + 1) = j \\ f(l, k) = i \text{ e } f(l + 1, k) = j \end{array} \right.$$

$\delta = 0$, caso contrário.

No caso da junção das 4 direções (horizontal, vertical e diagonais) tem-se:

$$\delta = 1 \text{ se, e somente se,} \quad (3.37)$$

$$\left\{ \begin{array}{l} f(l, k) = i \text{ e } f(l, k + 1) = j \\ f(l, k) = i \text{ e } f(l + 1, k) = j \\ f(l, k) = i \text{ e } f(l + 1, k + 1) = j \\ f(l, k + 1) = i \text{ e } f(l + 1, k) = j \end{array} \right.$$

$\delta = 0$, caso contrário.

3.3.4. Exemplo de Segmentação usando a Matriz de Co-Ocorrência

Na literatura, percebe-se que a matriz de co-ocorrência é bastante usada em várias abordagens de segmentação, principalmente para a segmentação por limiarização e por textura. Das abordagens existentes, destaca-se a abordagem de limiarização de [PAL89] que será apresentada a seguir.

A probabilidade de co-ocorrência p_{ij} dos níveis de cinza i e j podem ser definidos como:

$$p_{ij} = \frac{t_{ij}}{(\sum_i \sum_j t_{ij})} \quad (3.38)$$

em que:

$$0 \leq p_{ij} \leq 1$$

Se s , $0 \leq s \leq L-1$, é um limiar, então s particiona a matriz de co-ocorrência em quatro quadrantes, denominadas A, B, C e D:

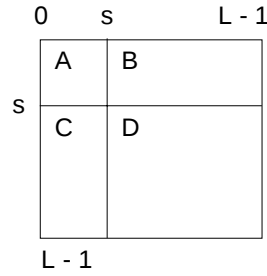


Figura 3.6 Quadrantes de uma matriz de co-ocorrência

Os quatro quadrantes (A, B, C, D) podem ser divididos em dois tipos: o quadrante D representa os objetos, cujos pixels em níveis de cinza se encontram acima do limiar s e o quadrante A representa o fundo, cujos pixels em níveis de cinza se encontram abaixo do limiar s . Os quadrantes B e C representam as transições ao longo das fronteiras entre fundo e objetos.

Portanto, pode-se definir as probabilidades com relação a cada quadrante [PAL89]:

$$P_A = \sum_{i=0}^s \sum_{j=0}^s p_{ij} = \sum_{i=0}^s \sum_{j=0}^s \left\{ \frac{t_{ij}}{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} t_{ij}} \right\} \quad (3.39)$$

$$P_B = \sum_{i=0}^s \sum_{j=s+1}^{L-1} p_{ij} = \sum_{i=0}^s \sum_{j=s+1}^{L-1} \left\{ \frac{t_{ij}}{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} t_{ij}} \right\} \quad (3.40)$$

$$P_C = \sum_{i=s+1}^{L-1} \sum_{j=0}^s p_{ij} = \sum_{i=s+1}^{L-1} \sum_{j=0}^s \left\{ \frac{t_{ij}}{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} t_{ij}} \right\} \quad (3.41)$$

$$P_D = \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} p_{ij} = \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} \left\{ \frac{t_{ij}}{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} t_{ij}} \right\} \quad (3.42)$$

Normalizando as probabilidades dentro de cada quadrante individual, tal que a soma das probabilidades de cada quadrante seja igual a um, obtemos então as seguintes probabilidades para quadrantes diferentes:

$$\begin{aligned} p_{ij}^A = \frac{p_{ij}}{P_A} &= \frac{\frac{t_{ij}}{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} t_{ij}}}{\sum_{i=0}^s \sum_{j=0}^s \left\{ \frac{t_{ij}}{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} t_{ij}} \right\}} \\ &= \frac{t_{ij}}{\sum_{i=0}^s \sum_{j=0}^s t_{ij}} \end{aligned} \quad (3.43)$$

para $0 \leq i \leq s$,

$$0 \leq j \leq s$$

Similarmente,

$$p_{ij}^B = \frac{p_{ij}}{P_B} = \frac{t_{ij}}{\sum_{i=0}^s \sum_{j=s+1}^{L-1} t_{ij}} \quad (3.44)$$

para $0 \leq i \leq s$,

$$s+1 \leq j \leq L-1$$

$$p_{ij}^C = \frac{p_{ij}}{P_C} = \frac{t_{ij}}{\sum_{i=s+1}^{L-1} \sum_{j=0}^s t_{ij}} \quad (3.45)$$

para $s+1 \leq i \leq L-1$,

$$s+1 \leq j \leq L-1$$

e

$$p_{ij}^D = \frac{p_{ij}}{P_D} = \frac{t_{ij}}{\sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} t_{ij}} \quad (3.46)$$

para $s+1 \leq i \leq L-1$,

$$0 \leq j \leq s$$

Agora com a ajuda das equações (3.7) e (3.44), a entropia local de segunda-ordem do objeto pode ser definida como:

$$H_A^{(2)}(s) = -\frac{1}{2} \sum_{i=0}^s \sum_{j=0}^s p_{ij}^A \log_2 p_{ij}^A \quad (3.47)$$

Similarmente, a entropia de segunda ordem do fundo pode ser definida como:

$$H_D^{(2)}(s) = -\frac{1}{2} \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} p_{ij}^D \log_2 p_{ij}^D \quad (3.48)$$

Portanto, a entropia total de segunda ordem pode ser escrita como:

$$H_T^{(2)}(s) = H_A^{(2)}(s) + H_D^{(2)}(s) \quad (3.49)$$

O nível de cinza correspondente ao valor máximo de $H_T^{(2)}(s)$ fornece portanto o limiar para a classificação objeto-fundo.

3.3.5. Vantagens e Desvantagens da Matriz de Co-Ocorrência

A matriz de co-ocorrência, assim como o histograma 2D e, ao contrário do histograma 1D, possui a desejável característica de “resgatar” a noção da distribuição dos pixels através da ocorrência do par $[f(k,l)=i, f(m,n)=j]$. A possibilidade de se analisar seqüências de pixels também é de extrema importância. Suas aplicações no processamento digital de imagens são diversas, principalmente relacionadas a segmentação por limiarização e por textura (Capítulo 2). Contudo, a grande desvantagem da matriz de co-ocorrência reside no fato da escolha do sentido de varredura e da distância entre os pixels. Neste caso, é necessário um conhecimento a priori da imagem a ser analisada, o que nem sempre é possível. Nas Figuras 3.8 e 3.9 são mostradas as visualizações da matriz de ocorrência (sentido horizontal e vertical com distância entre pixels de 1) para a imagem da Figura 3.7 enquanto que nas Figuras 3.10 e 3.11 tem-se as visualizações da junção das direções horizontal e vertical e junção das 4 direções, propostas por [ZHA98], respectivamente.



a) Imagem original

Figura 3.7 Imagem necessária para o processamento da matriz de co-ocorrência

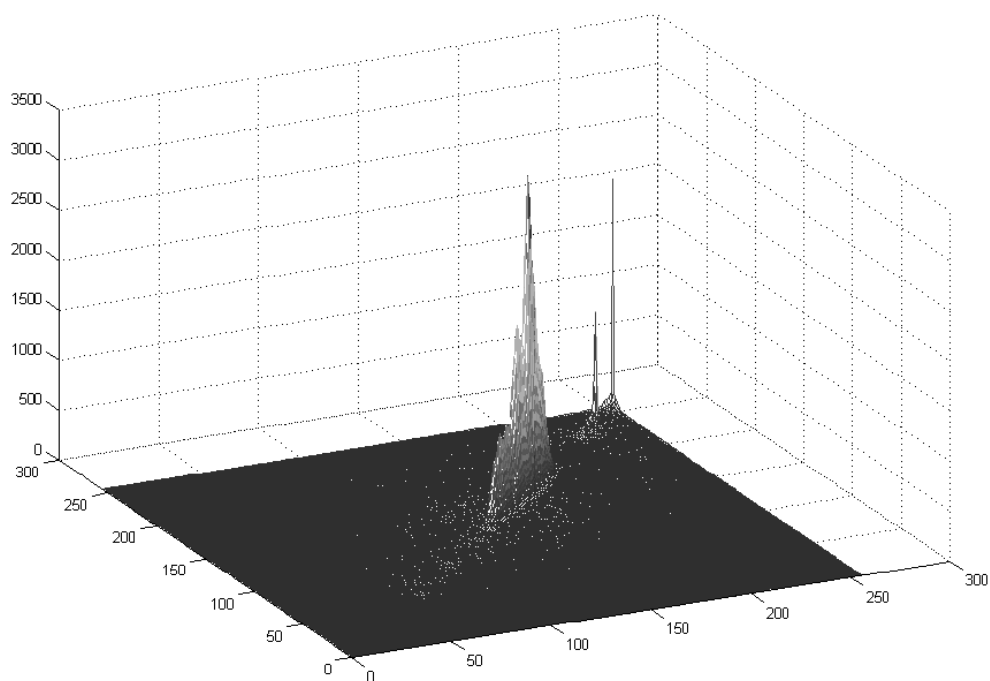


Figura 3.8 Visualização 3D da matriz de co-ocorrência da imagem da Figura 3.7
(direção: horizontal, distância entre pixels: 1)

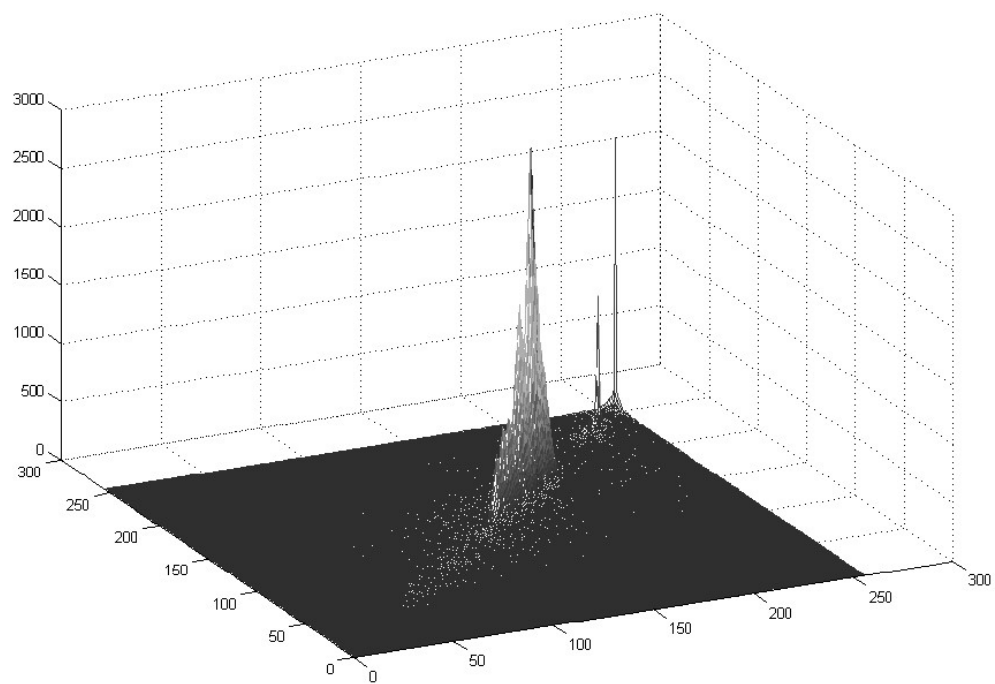


Figura 3.9 Visualização 3D da matriz de co-ocorrência da imagem da Figura 3.7
(direção: vertical, distância entre pixels: 1)

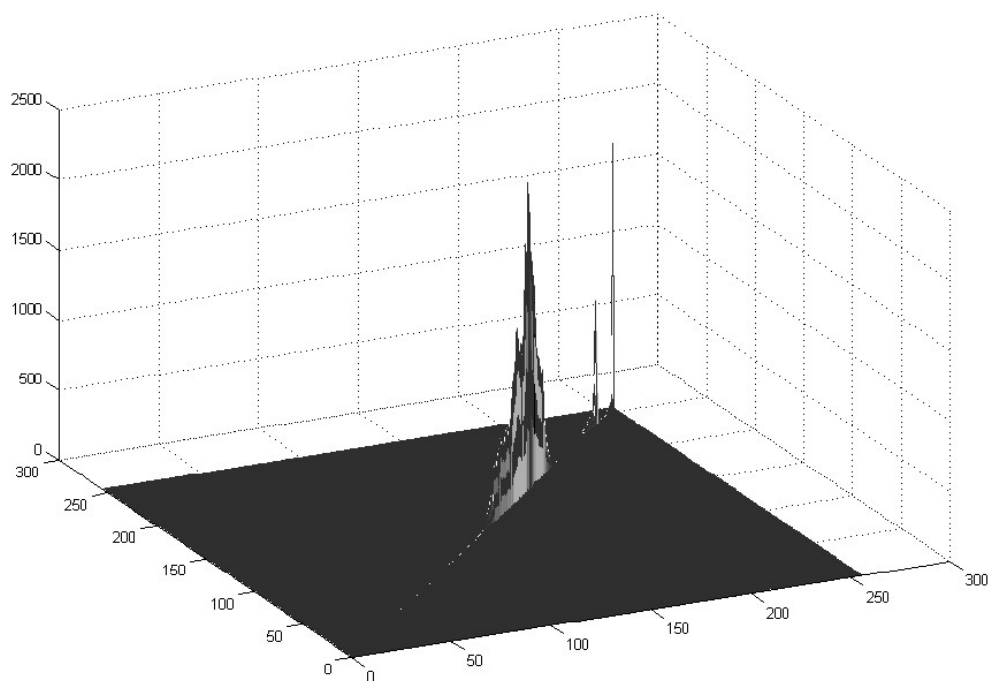


Figura 3.10 Visualização 3D da matriz de co-ocorrência da imagem da Figura 3.7 (direção: junção das direções horizontal e vertical, distância entre pixels: 1)

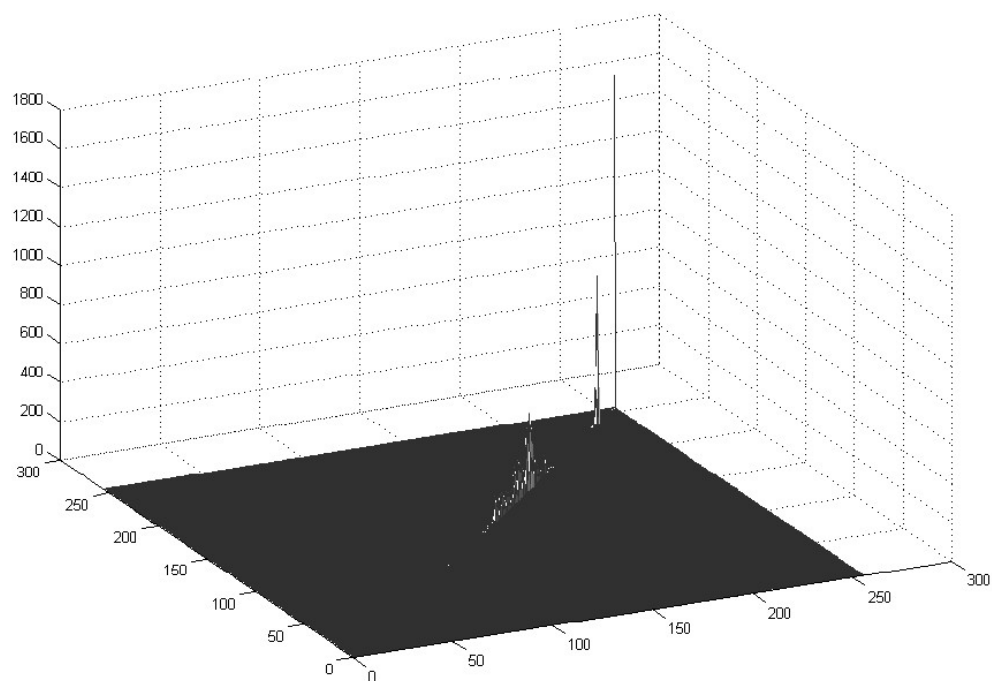


Figura 3.11 Visualização 3D da matriz de co-ocorrência da imagem da Figura 3.7
(direção: junção das 4 direções, distância entre pixels: 1)

3.4. Conclusão

Neste capítulo foram apresentadas as noções de entropia, matriz de co-ocorrência e histograma 2D. Suas vantagens e desvantagens, bem como suas principais aplicações no processamento digital de imagens e nas abordagens propostas foram introduzidas.

Capítulo 4

Reconstrução Morfológica em Níveis de Cinza e Segmentação Watershed¹

4.1. Reconstrução morfológica

A reconstrução é uma ferramenta extremamente útil proporcionada pela morfologia matemática. É bem conhecida no caso binário, onde ela simplesmente extrai os componentes conectados de uma imagem que estão “marcados” por outra imagem. Contudo, a reconstrução pode também ser definida para imagens em níveis de cinza. Suas aplicações são diversas, tais como, filtragem, segmentação e extração de características.

4.1.1. Notações utilizadas

O grid discreto $G \subset \mathbb{Z}^2 \times \mathbb{Z}^2$ fornece os relacionamentos de vizinhança entre pixels: p é vizinho de q se, e somente se $(p, q) \in G$. Podemos considerar grids quadrados, em uma vizinhança 4 ou vizinhança 8, ou um grid hexagonal, neste caso, com vizinhança 6.

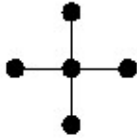


Figura 4.1 Vizinhança 4

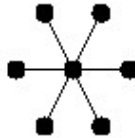


Figura 4.2 Vizinhança 6

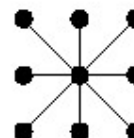


Figura 4.3 Vizinhança 8

(N_G)

¹ Para uma descrição completa dos operadores morfológicos básicos em níveis de cinza, ver Apêndice A

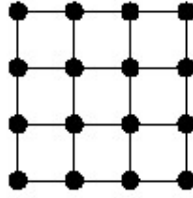


Figura 4.4 Exemplo de uma
vizinhança 4

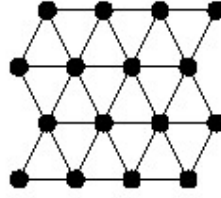


Figura 4.5 Exemplo de uma
vizinhança 6

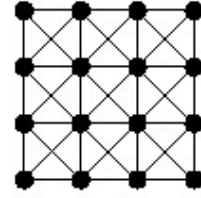


Figura 4.6 Exemplo de uma
vizinhança 8

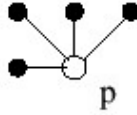


Figura 4.7 Sentido raster
(N_G^+)

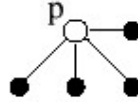


Figura 4.8 Sentido anti-
raster (N_G^-)

4.1.2. Reconstrução Binária

Definição em termos de componentes conectados

Considere I e J serem duas imagens definidas no mesmo domínio discreto D tal que $J \subseteq I$. Em termos de mapeamento, isto significa que: $\forall p \in D, J(p)=1 \Rightarrow I(p)=1$. J é denominada de imagem marcadora, ou simplesmente marcador, e I é denominada de imagem máscara, ou simplesmente máscara.

A reconstrução $\rho_I(J)$ da máscara I pelo marcador J é a união dos componentes conectados de I que contém pelo menos um pixel de J :

$$\rho_I(J) = \bigcup_{J \cap I_k \neq \emptyset} I_k \quad (4.1)$$

Uma discussão completa a respeito de suas definições e aplicações são detalhadas em [FAC96].

4.1.3. Reconstrução em níveis de cinza

Definição usando superposição de limiares:

É sabido que, pelo menos no caso discreto, qualquer transformação crescente definida para imagens binárias pode ser estendida para imagens em níveis de cinza. Por crescente, entende-se que uma transformação ψ tal que:

$$Y \subseteq X \Rightarrow \psi(Y) \subseteq \psi(X) \quad (4.2)$$

$$\forall X, Y \subset Z^2$$

Com o objetivo de estender tal transformação ψ para imagens em níveis de cinza I com valores $\{0, 1, \dots, N-1\}$, é suficiente considerar as sucessivas limiarizações $T_k(I)$ de I , para $k = 0$ a $N - 1$:

$$T_k(I) = \{p \in D_I \mid I(p) \geq k\} \quad (4.3)$$

Estes são ditos para constituir a decomposição dos limiares de I . Esses conjuntos satisfazem a seguinte relação de inclusão:

$$T_k(I) \subseteq T_{k-1}(I), \quad (4.4)$$

$$\forall k \in [1, N - 1]$$

Quando da aplicação da operação crescente ψ para cada um desses conjuntos, as relações de inclusão são preservadas. Então, podemos estender ψ para imagens em níveis de cinza da seguinte maneira:

$$\forall p \in D_I, \psi(I)(p) = \max\{k \in [0, N - 1] \mid p \in \psi(T_k(I))\} \quad (4.5)$$

Portanto, seguindo o princípio da superposição de limiares, definimos a reconstrução em níveis de cinza como segue:

Considere I e J serem duas imagens em níveis de cinza definidas no mesmo domínio, cujos valores no conjunto discreto $\{0, 1, \dots, N - 1\}$ e tal que $J \leq I$ (i.e., para cada pixel $p \in D_I, J(p) \leq I(p)$). A reconstrução em níveis de cinza $\rho_I(J)$ de I em J é dada por:

$$\forall p \in D_I,$$

$$\rho_I(J)(p) = \max\{k \in [0, N - 1] \mid p \in \rho_{T_k(I)}(T_k(J))\} \quad (4.6)$$

Assim como a reconstrução binária extrai os componentes conectados da máscara que estão marcadas, a reconstrução em níveis de cinza extrai os picos da máscara que estão marcadas pela imagem de marcadores.

Contudo, a definição anterior não provê nenhum método computacional interessante para determinar a reconstrução em níveis de cinza em imagens digitais. De fato, mesmo se um algoritmo de reconstrução binária completamente otimizado for utilizado, é preciso aplicar 256 vezes para determinar a reconstrução em níveis de cinza!

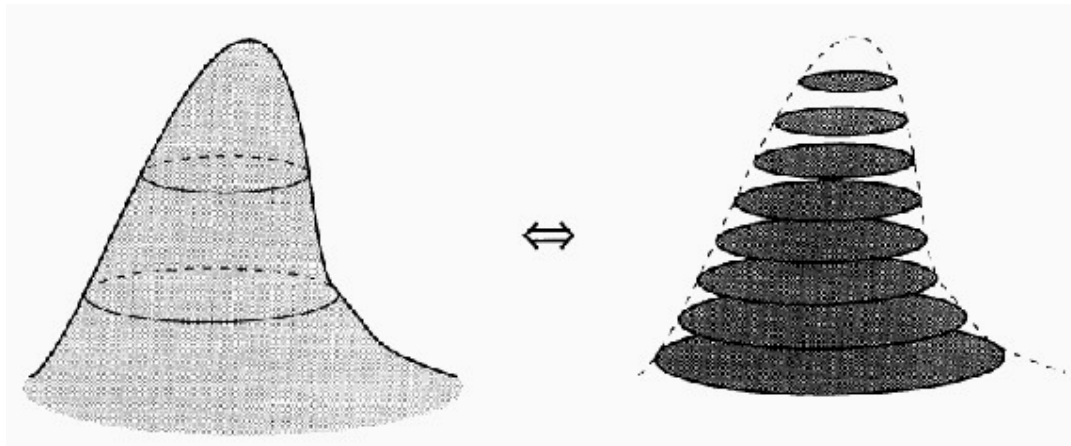


Figura 4.9 Decomposição de limiares de uma imagem em níveis de cinza

4.1.4. Algoritmos de Reconstrução em níveis de cinza

Nesta seção, um histórico dos algoritmos de reconstrução em níveis de cinza será detalhado até o algoritmo mais eficiente em termos de memória e tempo de execução implementado na presente pesquisa que é o algoritmo híbrido rápido.

Algoritmo paralelo

O algoritmo correspondente faz parte de um conjunto de métodos clássicos denominados paralelos. Basicamente, são iteradas diversas dilatações seguidas pela determinação do ponto mínimo até que a estabilidade seja alcançada:

Algoritmo: reconstrução paralela

- I : Imagem máscara
- J : Imagem marcadora, definida no domínio $D_I, J \leq I$. A reconstrução é determinada diretamente em J .
- Aloque uma nova imagem K também definida no domínio D_I .
- Repita até que a estabilidade seja alcançada (i.e., até que não haja mais modificações nos níveis de cinza dos pixels da imagem):

Etapa da Dilatação: para todo pixel q

faça

$K(p) \leftarrow \max\{J(q), q \in N_G(p) \cup \{p\}\}$

fim-faça

Etapas da determinação do ponto de mínimo

Para todo pixel $p \in D_I$

$J(p) \leftarrow \min(K(p), I(p))$

fim

Em cada uma das etapas anteriores, os pixels da imagem podem ser varridos em qualquer ordem, tornando possível a implementação do algoritmo em uma máquina paralela. Contudo, são exigidas muitas iterações, às vezes centenas, impossibilitando sua implementação em computadores convencionais.

Algoritmo seqüencial

Em uma tentativa de reduzir o número de varreduras exigidas para a obtenção de uma imagem em níveis de cinza reconstruída, algoritmos seqüenciais e recursivos foram propostos. De uma forma geral, eles se baseiam nos seguintes princípios:

Os pixels da imagem são varridos em uma ordem pré-definida, geralmente raster ou anti-raster.

O novo valor do pixel atual, determinado pelos valores dos pixels de sua vizinhança, é escrito na mesma imagem, portanto sendo levados em consideração quando da determinação dos valores dos novos pixels.

Aqui, diferentemente dos algoritmos paralelos, o sentido de varredura é essencial. Este tipo de algoritmo foi primeiramente introduzido para o cálculo de funções distância e então estendido para um número de transformações morfológicas. Tanto a reconstrução binária quanto em níveis de cinza podem ser obtidos seqüencialmente utilizando o seguinte algoritmo, onde a informação é primeiramente propagada horizontalmente da esquerda para a direita e verticalmente de cima para baixo (raster) e então horizontalmente da direita para a esquerda e verticalmente de baixo para cima (anti-raster).

Algoritmo: reconstrução seqüencial

- I : imagem máscara (binária ou em níveis de cinza)
- J : imagem marcadora definida no domínio D_I , $J \leq I$

A reconstrução é diretamente determinada em J .

- Repita até que seja alcançada a estabilidade

Varrer D_I no sentido raster

Considere p como sendo o pixel atual

$$J(p) \leftarrow (\max\{J(q), q \in N_G^+(p) \cup \{p\}\}) \wedge I(p)$$

Varrer D_I no sentido anti-raster

Considere p como sendo o pixel atual

$$J(p) \leftarrow (\max\{J(q), q \in N_G^-(p) \cup \{p\}\}) \wedge I(p)$$

Este algoritmo geralmente exige poucas iterações (tipicamente uma dúzia) até que a estabilidade seja alcançada, e é portanto muito mais eficiente que o algoritmo paralelo apresentado na seção anterior. Contudo, é importante salientar que, em um dado momento de execução, uma iteração é necessária para modificar o valor de poucos pixels (ou, no pior caso, somente um, uma vez que a estabilidade ainda não foi alcançada). Na próxima seção, será detalhado um novo método que leva em consideração somente os pixels que podem ser modificados.

4.1.5. Máximos regionais e reconstrução

De acordo com a seção anterior, uma etapa além no projeto de implementação da reconstrução seria levar em consideração somente os pixels cujo valor poderia ser mudado. Para tanto, uma primeira varredura é necessária para detectar os pixels que são os iniciadores do processo e que estão tipicamente localizados nas fronteiras dos objetos ou regiões de interesse. Portanto, partindo desses pixels, a informação é propagada somente nas partes relevantes da imagem. Um dos métodos mais eficientes é tratar a imagem em estudo como grafos e realizar buscas por largura (“*breadth-first scanings*”) destes grafos partindo de pixels estrategicamente localizados. As buscas por largura (*breadth-first scanings*) envolvidas são implementadas usando filas de pixels, i.e., *First-In - First-Out* (FIFO): os pixels que são primeiramente colocados na fila são os primeiros a serem extraídos. Em outras palavras, cada novo pixel incluído na fila é colocado de um lado enquanto que o pixel a ser removido é obtido do outro lado. Na prática, uma fila é simplesmente um array de ponteiros para pixels, em que as seguintes operações podem ser efetuadas:

- `fifo_add(p)`: coloca o (ponteiro para) pixel p na fila.
- `fifo_first()`: retorna o (ponteiro para) pixel que está no começo da fila, e o remove.
- `fifo_empty()`: retorna verdadeiro se a fila está vazia e falso caso contrário.

No caso binário, é extremamente fácil implementar a reconstrução usando o algoritmo FIFO: é suficiente inicializar a fila inserindo-a os pixels de bordas da imagem marcadora. Então os valores desses pixels são propagados nos componentes conectados relevantes da imagem máscara. O algoritmo correspondente trabalha da seguinte forma:

Algoritmo de reconstrução binária usando uma fila de pixels

- I : imagem máscara (binária)
- J : imagem marcadora (binária), definida no domínio $D_I, J \subseteq I$. A reconstrução é determinada diretamente em J .
- Inicialização da fila com pixels de borda da imagem marcadora: para todo pixel $p \in D_I$:

Se $J(p) = 1$ e $\exists q \in N_G(p), J(q) = 0$ e $I(p) = 1$

$\text{fifo_add}()$

Propagação: enquanto $\text{fifo_empty}() = \text{false}$

$p \leftarrow \text{fifo_first}()$

para todo $q \in N_G(p)$ (vizinho de p)

se $J(q) = 0$ e $I(q) = 1$

$J(q) \leftarrow 1$

$\text{fifo_add}(q)$

Este algoritmo é extremamente eficiente, uma vez que após a inicialização da fila, somente os pixels relevantes são considerados. O tempo de execução desse algoritmo é de $\frac{1}{4}$ de segundo em uma estação Sun IPC Workstation, para imagens de 256×256 . A extensão desse algoritmo para imagens em níveis de cinza não é imediata: é necessário considerar os máximos regionais da imagem marcadora. Considere $R(I)$ como sendo a seguinte imagem:

$$R(I)(p) = \begin{cases} I(p), & \text{se } p \text{ pertence a um máximo, } \forall p \in D_I \\ 0, & \text{caso contrário} \end{cases}$$

Proposição 1: Considere I e J como sendo duas imagens em níveis de cinza tal que $J \leq I$. Então:

$$\rho_I(J) = \rho_I(R(J)) \quad (4.7)$$

Prova: É suficiente provar que para todo limiar $h \in \{0, 1, \dots, N-1\}$

$$\rho_{T_h(I)}(T_h(J)) = \rho_{T_h(I)}(T_h(R(J))) \quad (4.8)$$

$R(J) \leq J$ implica que $T_h(R(J)) \subseteq T_h(J)$. Como a reconstrução binária é uma operação crescente, temos que:

$$\rho_{T_h(I)}(T_h(J)) \supseteq \rho_{T_h(I)}(T_h(R(J))) \quad (4.9)$$

Similarmente, considere C como sendo um componente conectado de $T_h(J)$. Considere também que $h_{\max} = \max\{J(q), q \in C\}$ e $C_{\max}$ é o conjunto dos pixels de C com valor $C_{\max}$. C' é um componente conectado de $C_{\max}$. C' é obviamente um máximo regional de J de altitude $h_{\max}$. Portanto, por definição, $\forall p \in C', R(J)(p) = h_{\max}$. Uma vez que $h \leq h_{\max}$, isto implica em: $\forall p \in C', T_h(R(J))(p) = 1$. Portanto, $C \cap T_h(R(J)) \neq \emptyset$. Isto é verdadeiro para todo componente conectado C de $T_h(J)$:

$$\rho_{T_h(I)}(T_h(J)) \subseteq \rho_{T_h(I)}(T_h(R(J))) \quad (4.10)$$

que completa a prova.

Na prática, a proposição acima significa que somente os máximos regionais da imagem marcadora J precisam ser levados em consideração para o cálculo de $\rho_I(J)$. O algoritmo descrito abaixo se aproveita desse fato propagando os valores dos máximos regionais de J usando uma fila de pixels.

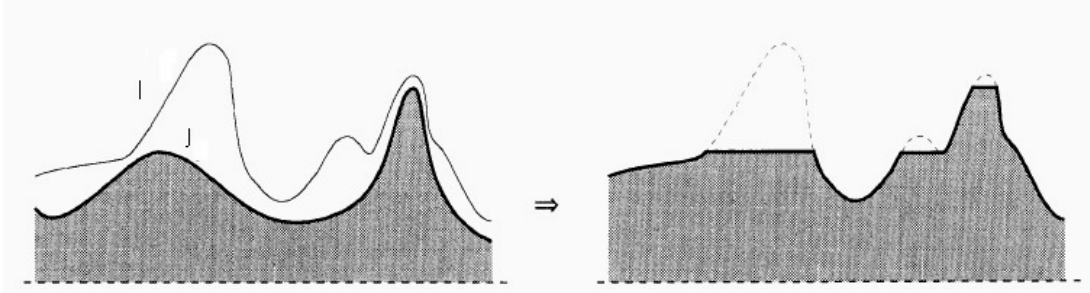


Figura 4.10. Princípio da reconstrução em níveis de cinza da máscara I pelo marcador J

Algoritmo usando uma fila de pixels

- I : imagem máscara (em níveis de cinza)
- J : imagem marcadora (em níveis de cinza), definida no domínio $D_I, J \leq I$. A reconstrução é determinada diretamente em J .
- Calcular os máximos regionais de J : $J \leftarrow R(J)$;
- Inicialização da fila com as fronteiras dos máximos:

```

Para todo pixel  $p \in D_I$ 
Se  $J(p) \neq 0$  e  $\exists q \in N_G(p), J(q) = 0$ 
fifo_add()
Propagação: Enquanto fifo_empty() = false
 $p \leftarrow \text{fifo\_first}()$ 
para todo pixel  $q \in N_G(p)$ 
se  $q$  é menor que  $p$  e se é necessário propagá-lo:
se  $J(q) \leq J(p)$  e  $I(q) \neq J(q)$ 
 $J(q) \leftarrow \min\{J(p), I(q)\}$ 
fifo_add(q)

```

O algoritmo acima constitui uma grande melhoria com relação ao algoritmo seqüencial apresentado na seção anterior. Podemos perceber uma desejável melhoria também no tempo de execução que alcança em torno de 2.5 segundos para imagens de 256x256 em uma estação Sun SparcStation.

Algoritmo híbrido rápido

Embora muito mais rápido do que as técnicas previamente propostas, o algoritmo acima tem o seu desempenho prejudicado pela determinação inicial dos máximos regionais da imagem marcadora. Além disso, algumas regiões da imagem podem ser “varridas” mais de uma vez durante a etapa de varredura em largura. Isto é verdadeiro quando dois máximos regionais de J com diferentes elevações estão próximos um do outro. Por outro lado, o algoritmo de reconstrução em níveis de cinza seqüencial não apresente esta desvantagem, porém, após duas varreduras, são exigidas diversas varreduras adicionais em que somente alguns pixels são modificados.

Na verdade, estes dois algoritmos apresentam vantagens e desvantagens, e esta é a principal motivação para o desenvolvimento do algoritmo híbrido rápido: a idéia é começar com as duas varreduras do algoritmo seqüencial. Durante a segunda varredura (anti-raster), todo pixel p tal que seu valor atual ainda poderia ser propagado durante a próxima varredura (raster), i.e., tal que:

$$\exists q \in N_G^-(p), J(q) \text{ e } J(q) < I(q)$$

é colocado na fila. O último passo do algoritmo é exatamente igual ao passo da propagação em largura do algoritmo FIFO anteriormente proposto. Contudo, o número de pixels a serem considerados durante esta etapa é consideravelmente menor do que o anterior. O algoritmo é descrito abaixo:

Algoritmo de reconstrução em níveis de cinza híbrido rápido

- I : imagem máscara (binária ou em níveis de cinza)
- J : imagem marcadora, definida no domínio $D_I, J \leq I$. A reconstrução é determinada diretamente em J .
- Varrer D_I no sentido raster:

Considere p ser o pixel atual

$$J(p) \leftarrow (\max\{J(q), q \in N_G^+(p) \cup \{p\}\}) \wedge I(p)$$

Varrer D_I no sentido anti- raster:

Considere p ser o pixel atual

$$J(p) \leftarrow (\max\{J(q), q \in N_G^-(p) \cup \{p\}\}) \wedge I(p)$$

Se existe um pixel p tal que

$$J(q) < J(p) \text{ e } J(q) < I(q)$$

fifo_add(p)

Passo da propagação: Enquanto fifo_empty() = false

$$p \leftarrow \text{fifo_first}()$$

Para todo pixel $q \in N_G(p)$:

Se $J(q) < J(p)$ e $I(q) \neq J(q)$

$$J(q) \leftarrow \min\{J(p), I(q)\}$$

fifo_add(q)

Conforme citado antes, este é o algoritmo mais eficiente tanto em tempo de processamento quanto em exigência de memória. Para tanto, uma execução típica deste algoritmo leva menos de um segundo em uma *Sun IPC Workstation* para uma imagem de tamanho 256x256. Na Figura 4.11 segue-se um exemplo do algoritmo acima:

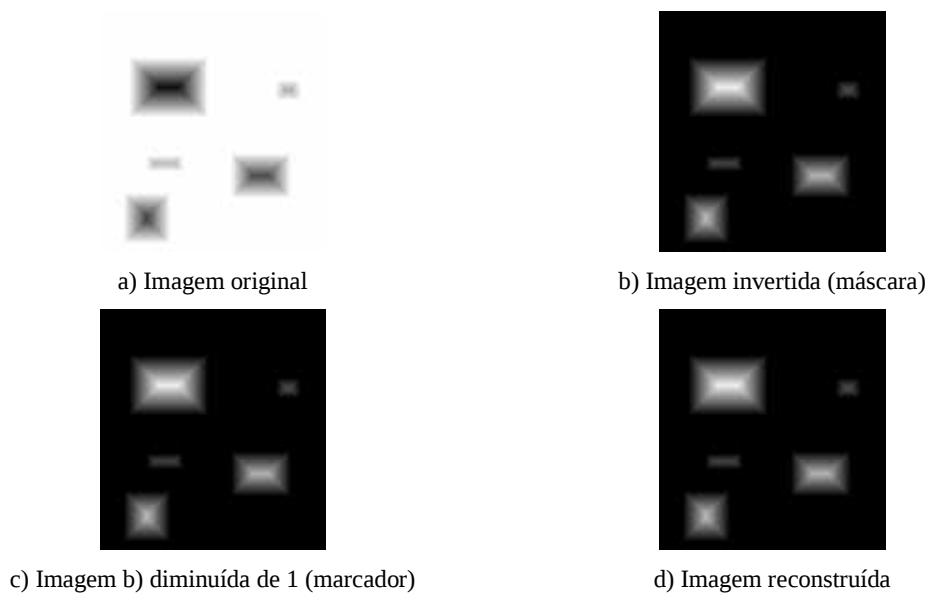


Figura 4.11 Exemplo de reconstrução em níveis de cinza

Pode parecer que a imagem reconstruída (Figura 4.11-d)) não apresenta nenhuma diferença com relação a imagem da Figura 4.11-b). Contudo, os picos (as partes mais claras) foram extraídos. Isso é de suma importância para obter-se a imagem de marcadores necessários para a segmentação Watershed que será detalhada na próxima seção. Na Figura 4.12, segue-se um exemplo de tal aplicação:

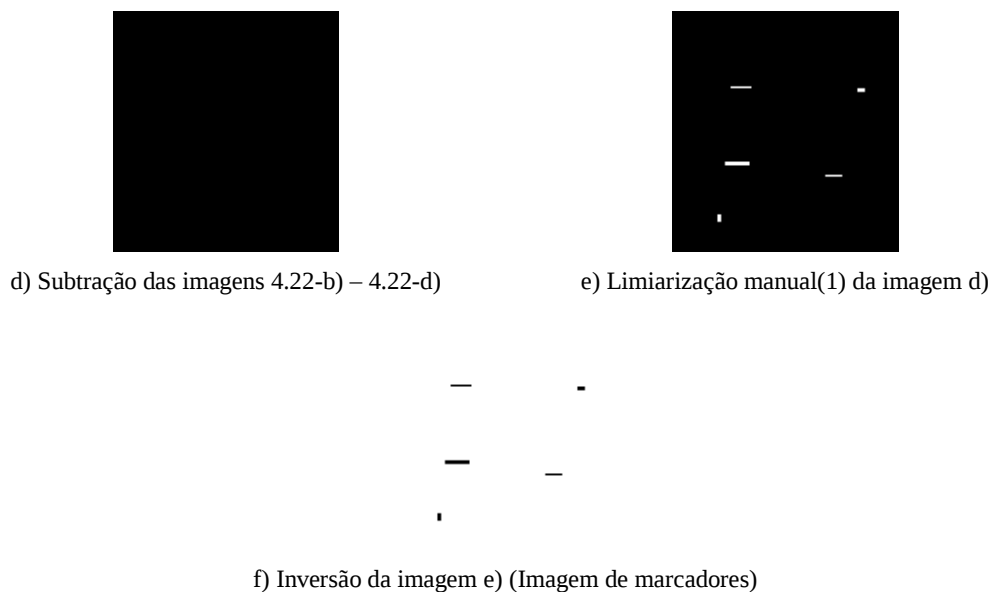


Figura 4.12 Exemplo de aplicação da reconstrução em níveis de cinza

4.2. Segmentação Watershed

Uma imagem contém, de forma geral, diversas informações que podem ser consideradas, em função do valor médio no nível de cinza da região em estudo, como picos e/ou vales [FAC96].

Considere a imagem f como uma superfície topográfica e defina as bacias de retenção de f e as linhas Watershed, através de um processo de enchimento. Imagine que perfuremos cada mínimo $M_i(f)$ da superfície topográfica S , e que afundemos esta superfície em um lago com uma velocidade vertical constante. A água vai entrando através dos mínimos $M_i(f)$ da superfície S . Durante o processo de enchimento, podem ocorrer o encontro de águas de diferentes mínimos. É justamente isso que queremos evitar através da construção de diques nos pontos da superfície S onde haveriam enchimentos. Estes diques, portanto, definem as linhas Watershed da função f . Eles separam as várias bacias de retenção $CB_i(f)$, cada um contendo um, e somente um único mínimo $M_i(f)$.

Vale ressaltar que a segmentação Watershed pode ser inteiramente definida em termos de operações de morfologia matemática, contudo, as implementações mais eficientes usam pilhas de pixels para efetuar o processo de enchimento e detecção das bacias de retenção [MAR00].

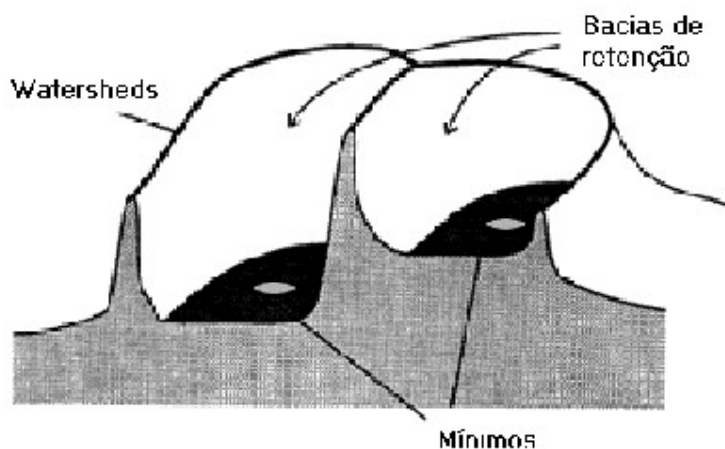
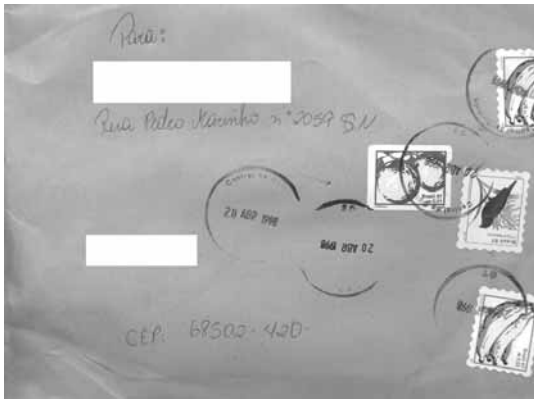


Figura 4.13 Enchimento e determinação das bacias de retenção e das linhas watersheds

A abordagem de segmentação por Watershed constitui um método morfológico de extração de linhas de crista entre regiões. Portanto, pela própria definição das linhas Watersheds, pode-se relacionar essa abordagem com a segmentação por gradiente. Realmente,

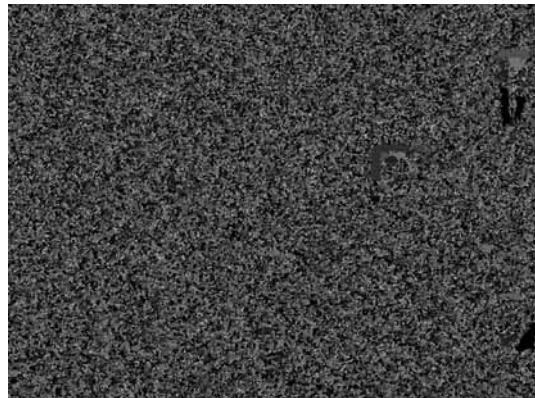
os contornos de uma imagem em níveis de cinza podem ser visualizados como regiões onde, nas bordas, os níveis de cinza mostram uma rápida variação, isto é, regiões de máximo gradiente. De fato, é muito interessante e pertinente a aplicação da segmentação Watershed empregando imagens do gradiente. A Figura 4.14 mostra a possibilidade e as limitações da segmentação Watershed empregando diretamente a informação do gradiente. A imagem original consiste de um envelope postal brasileiro complexo. É portanto natural esperar, do resultado da segmentação Watershed, bacias correspondendo às classes procuradas (carimbos, parte manuscrita, fundo e selos). Contudo, na Figura 4.14-b), devido ao excesso de bordas e ruído geradas pela imagem do gradiente, aparecem bacias em número muito maior que o número de classes presentes na imagem. A imagem das bacias rotuladas (Figura 4.14-c)) atesta do número exagerado de bacias segmentadas.



a) Imagem original



b) Imagem do gradiente



c) Bacias segmentadas

Figura 4.14 Uso da informação do gradiente na segmentação Watershed

A segmentação Watershed sempre gera contornos fechados e constitui uma abordagem geral de detecção de contornos. Contudo, como foi ilustrado na Figura 4.24, o Watershed

muito raramente é usado diretamente sobre o gradiente de imagens sem resultar em excesso de segmentação gerando imagens segmentadas em muitas bacias, muitas delas não relevantes. Esse problema é devido principalmente aos ruídos contidos nos dados: ruídos da imagem original resultam em ruído no gradiente morfológico, gerando muitos mínimos regionais. Isso se traduz diretamente em uma grande quantidade de bacias de retenção (hipersegmentação). Na literatura, existem várias técnicas que permitem vencer a hipersegmentação sendo que uma dessas técnicas é a modificação da homotopia. Entretanto, nenhuma delas é totalmente satisfatória devido à dificuldade de incorporar conhecimentos específicos sobre o conjunto de imagens em estudo [FAC96].

4.2.1. Algoritmos de Segmentação Watershed

Existem dois grandes grupos de segmentação Watershed, o de técnicas de simulação de enchimento e o de técnicas de busca direta das cristas de Watershed. Cada um desses grupos pode ser dividido em três subgrupos: algoritmos paralelos, seqüenciais e ordenados [FAC96].

Um algoritmo é dito paralelo se os pixels vizinhos do pixel a ser processado tomam os valores na imagem original. O algoritmo é dito paralelo porque o resultado não depende da ordem na qual os pixels são processados.

Num algoritmo seqüencial, o novo valor de um pixel serve para processar pixels vizinhos ainda não processados. Portanto, o resultado depende completamente do sentido de varredura. De forma geral, o sentido escolhido é o rastreamento direto ou inverso.

Um algoritmo ordenado é parecido com o seqüencial, exceto no sentido de varredura. O sentido de varredura é realizado de maneira que cada pixel é testado um única vez e processado no momento em que a sua vizinhança é completamente ou suficientemente definida.

4.2.2. Filas Ordenadas

Um método apresentado aqui será o algoritmo seqüencial do grupo de técnicas de simulação de enchimento, porém, empregando filas ordenadas (FIFO - *first in first out*). A principal vantagem é a facilidade de armazenamento dos pixels em qualquer ordem e de recuperação dos mesmos na ordem de enchimento das bacias de retenção, o que faz esta estrutura ser otimizada para a detecção de linhas de *watershed* [FAC96].

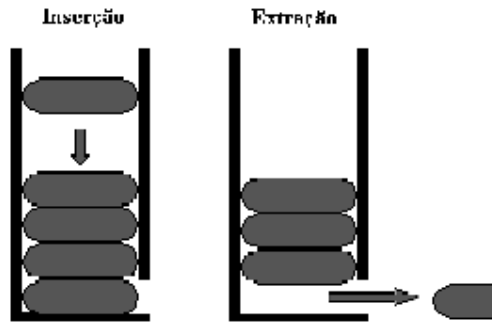


Figura 4.15 Estrutura de uma fila simples

Durante o enchimento de uma superfície topográfica, existem várias relações de ordem entre os pixels. A primeira delas é o fato de que um pixel P_1 é atingido pelo enchimento antes de outro pixel P_2 se $P_1 < P_2$, ou seja, se o nível de cinza do pixel P_1 for menor do que o nível de cinza do pixel P_2 no relevo. Uma segunda relação acontece com os pixels de um mesmo planalto de altura h . O enchimento desse planalto não se faz de forma instantânea mas de forma progressiva onde os primeiros pixels atingidos são os vizinhos de precedentes pixels atingidos. Portanto, se dois pixels P_1 e P_2 pertencem ao mesmo planalto de altura h , P_1 é atingido pelo enchimento antes do pixel P_2 se a distância geodésica entre esse planalto e o pixels de altura $h' < h$ for menor para P_1 que para P_2 .

A fila ordenada permite introduzir uma relação de ordem hierárquica entre os pixels. Pixels como clientes são "atendidos" na ordem de chegada em função de sua prioridade, que segue a lógica de uma enchente. Um pixel é atendido:

- Antes de todos os pixels de menor prioridade, mesmo se esses chegaram antes;
- Antes de todos os pixels de igual prioridade que chegaram depois;
- Depois de todos os pixels de maior prioridade, mesmo se esses chegaram depois.

Uma fila hierárquica pode ser vista como múltiplas filas. Na chegada de um novo pixel, esse é colocado no fim da fila em função de sua prioridade e é atendido depois dos pixels de igual prioridade que chegaram antes. Um único pixel pode ser atendido de cada vez. Quando uma fila é esvaziada, ela é eliminada. Se um pixel de maior prioridade chega após a supressão da fila de igual prioridade, ele é colocado na fila de maior prioridade. Essa fila ordenada é constituída de várias filas simples que podem ser enchidas pelo topo por pixels que tem a prioridade correspondente. O seu ordenamento faz-se da esquerda para a direita, sendo que à

esquerda estão os pixels armazenados nas filas simples de menor prioridade (maior nível de cinza) e que à direita estão os pixels armazenados nas filas simples de maior prioridade (menor nível de cinza). Na Figura 4.16 é exemplificada o esquema de múltiplas filas de prioridade, aqui representados por cilindros, onde os discos coloridos representam os pixels analisados e a fila de mais alta prioridade, neste caso, se encontra à extrema direita.

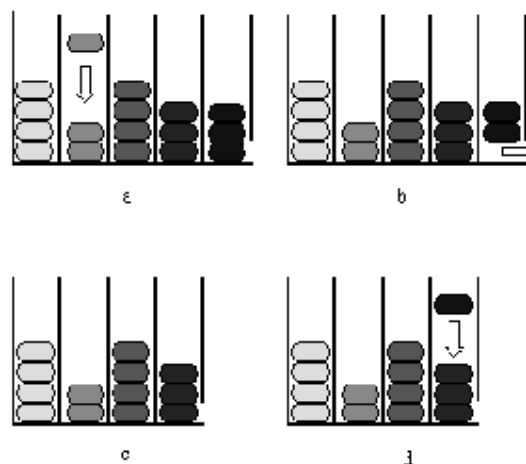


Figura 4.16 Estrutura de uma Fila Ordenada: (a) a fila ordenada; (b) extração de um elemento de mais alta prioridade; (c) redução do número de filas simples; (d) tratamento do elemento de mais alta prioridade quando a fila correspondente tiver sido suprimida.

No exemplo em questão, pode-se identificar cinco níveis de prioridade. Todos os cilindros possuem uma abertura no topo, o que significa que a qualquer momento é possível a introdução de um cliente em quaisquer das filas. Somente uma das filas, a de mais alta prioridade, tem a abertura na base, logo, a extração de um elemento da estrutura só poderá ocorrer nesta fila, para o cliente que tiver sido o primeiro a chegar. Se a tentativa de extrair um novo elemento falhar, a fila é suprimida e a fila com prioridade imediatamente abaixo é aberta para extração dos próximos elementos. Na Figura 4.16-d) podemos verificar a chegada de um elemento de maior prioridade cuja fila já tenha sido suprimida; neste caso o cliente é posto no final da fila corrente de mais alta prioridade.

4.2.3. Watershed por Filas Ordenadas

O princípio consiste em "encher" uma imagem em níveis de cinza S a partir de "poços" representados pela imagem dos marcadores M . Se os marcadores constituem os mínimos regionais da imagem S , então o resultado será linhas Watershed associadas ao relevo de S . Na

prática, isso acontece poucas vezes, devido ao ruído e a natureza das imagens tratadas. Novamente, alertamos que uma modificação da homotopia é necessária antes de efetuar o enchimento dos marcadores.

A vantagem de empregar filas ordenadas está na possibilidade de encher as bacias sem precisar reconstruir. Os marcadores são identificados por rótulos. Uma maneira de preencher as bacias refletindo a realidade da imagem consiste em definir os rótulos como sendo os valores dos níveis de cinza dos marcadores na imagem original. Por consequência, a bacia enchida terá o rótulo de seu marcador. Existem duas versões desse algoritmo: a primeira onde as bacias se tocam sem fronteiras ou diques entre elas, versão que será detalhada a seguir, e a segunda onde as fronteiras ou diques entre as bacias são geradas.

4.2.4. Algoritmo de Watershed por Filas Ordenadas

Para este algoritmo, a ordem de prioridade usada é a do nível de cinza. A fila ordenada tem um número de filas simples equivalente aos níveis de cinza que representam a imagem f . O objetivo desta etapa consiste em armazenar na fila ordenada todos os pixels de borda de cada mínimo regional da imagem rotulada de marcadores M . Esse armazenamento efetua-se usando como prioridade o nível de cinza destes pixels na imagem f .

Assim, o algoritmo pode ser sumarizado em:

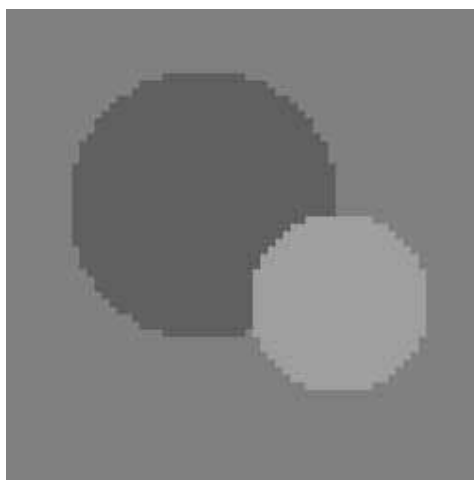
Repita até que a fila ordenada esteja vazia

{

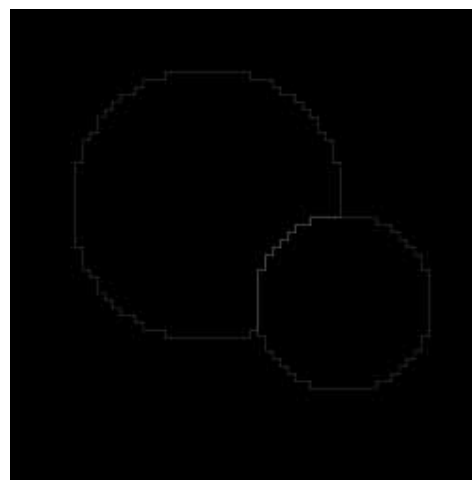
- um pixel p é extraído da fila ordenada;
- para cada vizinho p' de p não rotulado na imagem W o mesmo tratamento é aplicado:
- ao pixel p' é dado o mesmo rótulo que o ponto p na imagem W ;
- o pixel p' é armazenado na fila ordenada em função do seu nível de cinza na imagem f ;

}

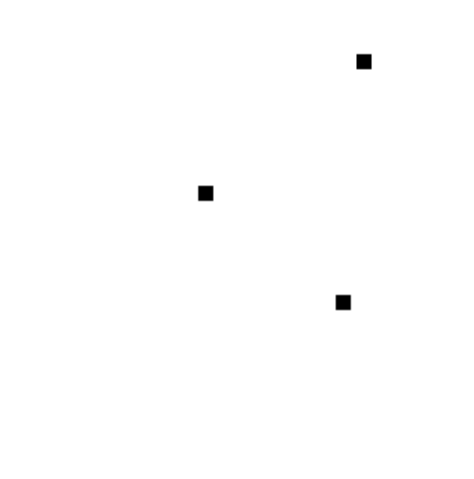
Na Figura 4.17, é apresentado um exemplo do algoritmo de segmentação Watershed empregando uma imagem de gradiente e duas imagens de marcadores rotulados.



a) Imagem Original



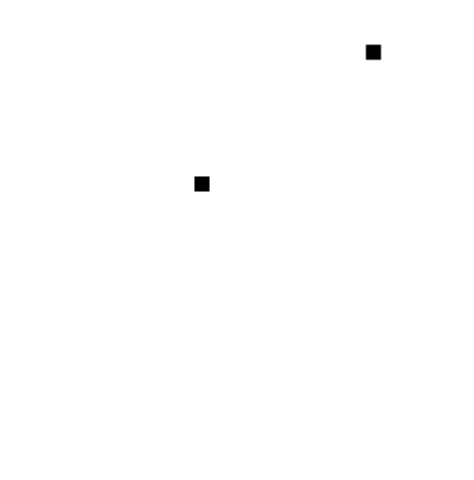
b) Gradiente por erosão com EE quadrado



c) Imagem com três marcadores rotulados



d) Imagem resultante com três bacias



a) Imagem com dois marcadores rotulados



a) Imagem resultante com duas bacias

Figura 4.17 Exemplos de Segmentação Watershed sobre uma imagem de gradiente

Para facilitar a compreensão, é usada uma imagem com um sinal unidimensional S representada por paralelepípedos brancos e a imagem em níveis de cinza dos marcadores W por paralelepípedos em hachuras. Esta imagem W em função do algoritmo representará no final a construção do *Watershed*.

Em uma primeira etapa, a imagem está representada por 6 níveis de cinza (de 0 até 5), a fila ordenada é composta por 6 filas simples. A imagem dos marcadores que representa um dos requisitos deste algoritmo são os pontos (b), (f), (g) e (K). Os pontos de borda desses marcadores são armazenados na fila ordenada com a prioridade que corresponde ao nível de cinza na imagem S : sejam os pontos (b), (g) e (k) com prioridade [1] e (f) com prioridade [2] (Figura 4.18-b).

Na próxima etapa, a fila mais prioritária [0] é vazia. Ela é, portanto eliminada e o primeiro ponto a ser tratado, (b) na fila [1], é extraída da fila. Os pontos (a) e (c) não tem rótulos em W e são rótulos com o mesmo valor que (b). (a) e (c) são armazenados na fila ordenada com as respectivas prioridades (segundo os níveis de cinza em S) [1] e [3]. (Figura 4.18-c)

O próximo ponto a sair da fila ordenada é (g). O seu vizinho (f) já está rotulado em W (3) enquanto o outro vizinho seu (h) não está rotulado em W . Logo, (h) recebe o rótulo de (g), o valor 3 e é armazenado na fila ordenada na fila de prioridade [5] (Figura 4.18-d).

Para os demais pontos:

O ponto (k) sai da fila ordenada e rotula os seus vizinhos (j) e (l) com o valor 5. (j) e (l) são armazenados respectivamente na filas de prioridade [3] e [1]. (Figura 4.18-e).

O ponto (a) não tem pontos sem rótulo e sai da fila ordenada.

O ponto (l) não tem pontos sem rótulo e sai da fila ordenada. A fila de prioridade [1] é vazia e é eliminada. (Figura 4.18-f).

O próximo ponto (f) sai da fila ordenada e rotula o seu único vizinho não rotulado (e) com o valor 3. (e) é armazenado na fila de prioridade [2] (Figura 4.18-g).

O próximo ponto (e) sai da fila ordenada e rotula o seu único vizinho não rotulado (d) com o valor 3. (d) é armazenado na fila de prioridade [3]. A fila de prioridade [2], vazia neste instante é portanto eliminada (Figura 4.18-h).

O ponto (c) não tem pontos sem rótulo e sai da fila ordenada.

O ponto seguinte (j) sai da fila ordenada e rotula o seu único vizinho não rotulado (i) com o valor 5. (i) é armazenado na fila de prioridade [3] (Figura 4.18- i).

O penúltimo ponto (d) não tem pontos sem rótulo e sai da fila ordenada. Neste instante, as filas de prioridade [3] e [4] são vazias e portanto eliminadas.

O último ponto (h) não tem pontos sem rótulo e sai da fila ordenada. Neste instante, a fila ordenada é completamente vazia e o processo de segmentação Watershed converge na solução da Figura 4.18-j.

Considerando a imagem S desse exemplo como sendo uma imagem do gradiente com cristas nos pixels (c), (d) e (h), os três marcadores sendo os “poços” das bacias, o algoritmo reconstituiu as três bacias com alturas correspondentes aos níveis de cinza dos marcadores.

É possível imaginar que caso uma bacia tivesse somente dois marcadores, uma bacia desapareceria, isto independentemente dos picos do gradiente. É importante salientar mais uma vez que esse algoritmo exige a presença da imagem dos marcadores [BEU93] [FAC96].

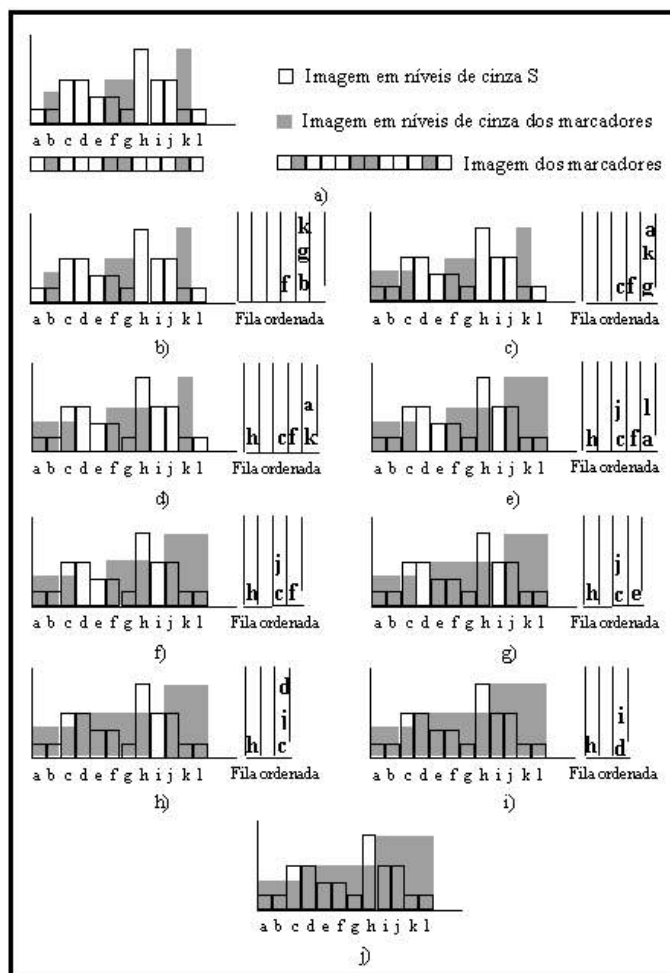


Figura 4.18 Segmentação Watershed por filas ordenadas

4.3. Conclusão

Neste capítulo foram apresentadas as duas ferramentas mais amplamente utilizadas na morfologia matemática: a reconstrução em níveis de cinza e a segmentação *Watershed*. Uma revisão dos principais algoritmos de reconstrução morfológica bem como sua aplicabilidade na segmentação *Watershed* foram analisados.

Capítulo 5

Base de Envelopes Postais Brasileiros

A definição da base de imagens é um fator de extrema importância para qualquer sistema de processamento digital de imagens. No caso dos envelopes postais brasileiros isso não é diferente. Neste capítulo, será descrita em detalhes as características da base de imagens utilizadas na presente pesquisa. Para tanto, um total de 80 imagens totalmente diferentes onde as regras de posicionamento com relação à parte manuscrita, carimbos e selos não são respeitadas foram selecionadas.

5.1. Regras recomendadas pela Empresa Brasileira de Correios e Telegráfos (ECT)

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;

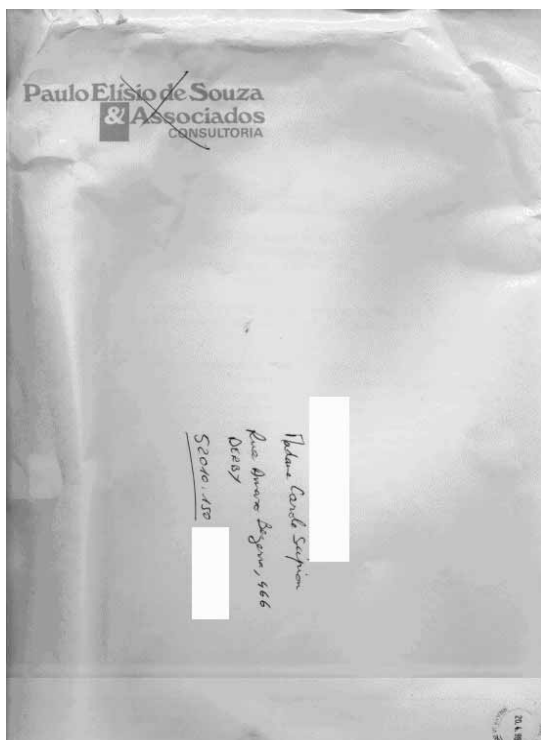
Porém, na maior parte dos casos, as mesmas não se verificam, inviabilizando portanto, uma técnica de segmentação baseada na posição das classes que compõem o envelope postal brasileiro. Na próxima seção, serão descritos os principais problemas que violam as regras da ECT observadas na base de imagens em questão.

5.2. Características da base de imagens de envelopes postais

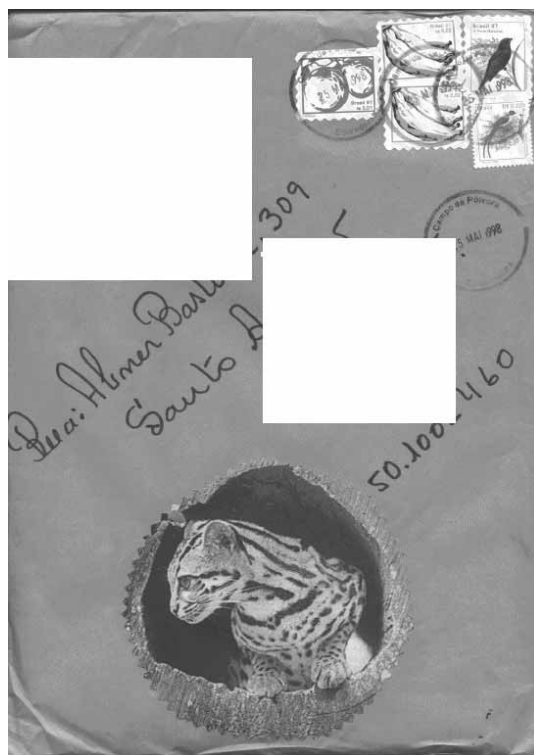
A base de imagens de envelopes postais não é padronizada e não segue as regras gerais da empresa de correios e telégrafos (ECT) (seção 5.1). Portanto, algumas características com relação a algumas imagens de envelopes postais foram observadas:

- Os selos e os carimbos não ocupam a parte superior à direita;
- A parte manuscrita se sobrepõe aos selos e aos carimbos;
- A presença de desenhos feitos à mão torna o fundo mais complexo e portanto mais difícil de ser segmentado;
- Problema de aquisição gerando bordas adicionais que não fazem parte dos envelopes postais.

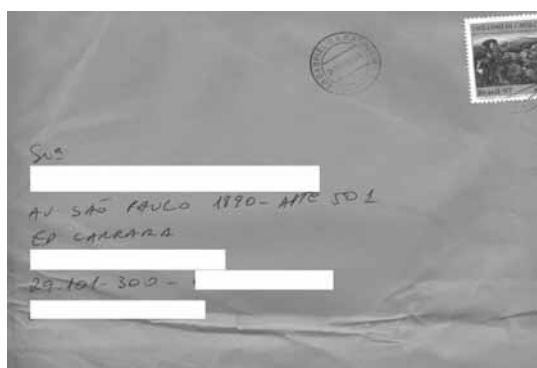
As características acima citadas tornam inviável a utilização de um único método de segmentação. Decidiu-se portanto, devido à complexidade e variabilidade da base de imagens postais brasileiros, o que inviabiliza a utilização de um único método, a aplicação da metodologia proposta somente das imagens da 1ª categoria, ou seja, manuscrito amassado.



a)



b)



c)



d)

Figura 5.1 Exemplos de imagens da base de envelopes postais brasileiros

Capítulo 6

Análise e Comparação de Resultados Obtidos

Neste capítulo, serão detalhadas as metodologias utilizadas com matriz de co-ocorrência e histograma 2D baseadas em uma segmentação supervisionada, bem como uma estratégia de avaliação por competição dos resultados obtidos e as dificuldades encontradas devido às características dos envelopes postais brasileiros.

6.1. Ambiente de desenvolvimento utilizado

O ambiente de desenvolvimento utilizado foi o Microsoft Visual C++ 5.0 com a biblioteca MFC (*Microsoft Foundation Classes*) e tecnologia COM (*Component Object Model*). O software desenvolvido foi uma extensão do FePI (Ferramentas para Processamento de Imagens), desenvolvida e projetada pelo Prof. Dr. Jacques Facon e por David Menoti Gomes.

6.2. Metodologia com Matriz de Co-Ocorrência

Aproveitando-se das características da escrita manuscrita e buscando também uma homogeneidade das outras classes procuradas, quais sejam, carimbos, fundo e selos, decidiu-se como sentido de varredura pela utilização das junções horizontal e vertical e das 4 direções. De acordo com [ZHA98], a utilização somente das varreduras horizontal ou vertical implica em que muitas informações importantes na vertical ou horizontal, respectivamente, são perdidas.

No caso dos envelopes postais brasileiros, 4 classes (bacias) são esperadas, a saber: carimbos, parte manuscrita, fundo e selos. Contudo, devido à complexidade dos mesmos e

principalmente a **similaridade de níveis de cinza** entre carimbos e parte manuscrita, **decidiu-se pela segmentação em 3 classes (bacias)**, que são:

Classe 1: carimbos com parte manuscrita;

Classe 2: fundo;

Classe 3: selos.

Considerando a matriz de co-ocorrência como uma superfície topográfica composta de vários picos, faz-se necessária a segmentação dos mesmos. Para tanto, a ferramenta de segmentação morfológica Watershed (Capítulo 4), é utilizada. No entanto, além de exigir a inversão dos picos para que se transformem em bacias, a mesma necessita da determinação de marcadores, que são obtidos pela reconstrução morfológica em níveis de cinza (Capítulo 4) [VIN93].

Contudo, é importante salientar que, devido à quantidade e complexidade dos picos gerados pela matriz de co-ocorrência, o número de marcadores obtidos é exageradamente alto, conforme mostrado na Figura 6.1 (em média 50). Faz-se necessário então uma etapa de pós-processamento para diminuir o número de picos e consequentemente, de marcadores. A solução proposta é a utilização de operadores morfológicos básicos em níveis de cinza (Apêndice A), quais sejam, a erosão e a dilatação. No caso específico da matriz de co-ocorrência, a erosão apresenta as seguintes propriedades:

- Eliminação de picos pequenos, que podem ser relevantes;
- Diminuição da altura de picos, reduzindo a importância dos mesmos, gerando perda de informação.

Já a dilatação se comporta da seguinte maneira:

- Junção de picos pequenos, sem perda de informação;
- Junção de picos pequenos com grandes, sem aumentar a importância dos mesmos;
- Não há modificação da altura dos picos.

Portanto, conforme as características acima citadas, neste contexto decidiu-se pela utilização de n dilatações até que, por reconstrução morfológica, seja possível a obtenção de 3 marcadores, correspondentes a 3 bacias segmentadas, conforme a Figura 6.2.

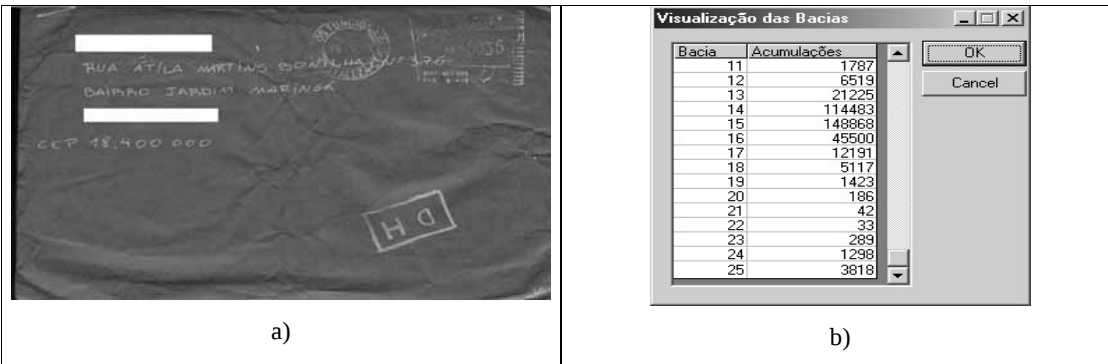


Figura 6.1b) Número de bacias segmentadas sem pós-processamento com a abordagem proposta de matriz de co-ocorrência para a imagem a)

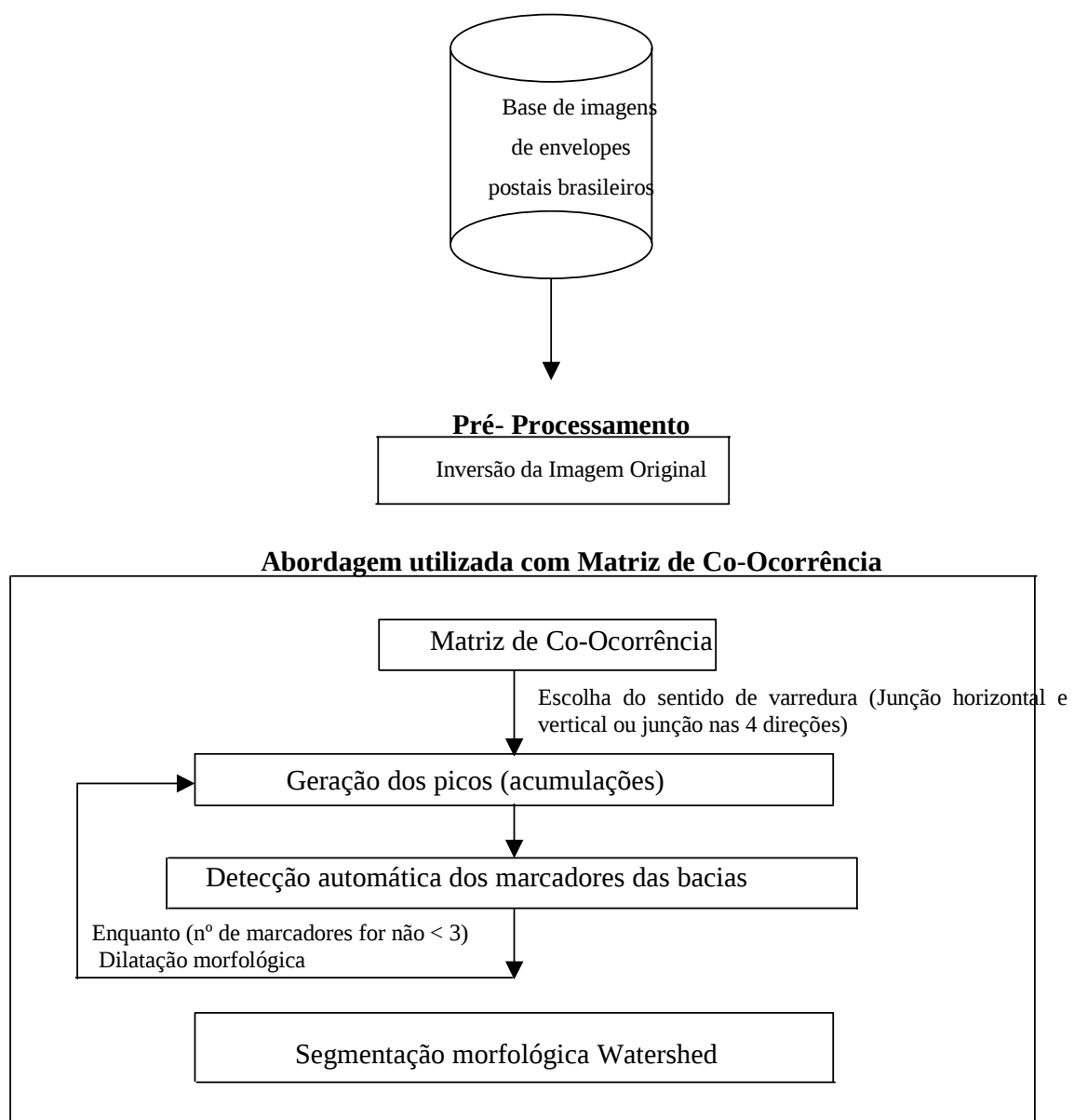


Figura 6.2 Diagrama com a estrutura da abordagem proposta para a segmentação de envelopes postais brasileiros

6.2.1. Estratégia de avaliação de resultados e dificuldades encontradas

Nesta seção serão detalhadas a estratégia de avaliação dos resultados obtidos na segmentação em 3 classes e também as dificuldades encontradas, face às características dos envelopes postais brasileiros. Para tanto, uma análise quantitativa (número de bacias) e qualitativa (análise visual das bacias segmentadas) dos resultados obtidos fazem-se necessárias, e são propostas da seguinte maneira:

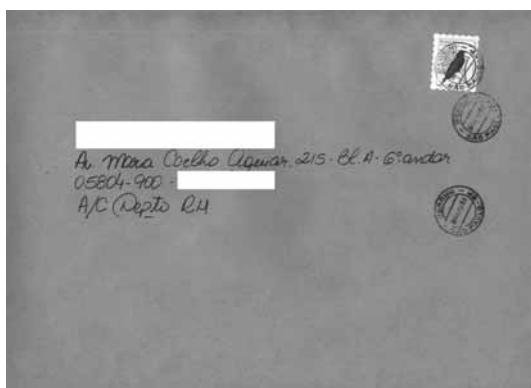
- Avaliação do número de imagens segmentadas em 3 bacias;
- Avaliação do número de imagens segmentadas em mais de 3 bacias;

Devido à complexidade dos resultados obtidos, propõe-se, no caso de 3 bacias, uma análise visual das mesmas, com atribuição de conceitos bom, médio e ruim para a classificação dos resultados:

- O conceito bom é atribuído ao caso em que foi possível a separação de carimbos e parte manuscrita, fundo e selos (Figura 6.3);
- O conceito médio refere-se ao caso em que a bacia correspondente a carimbos e parte manuscrita não foi completamente separada do fundo (geralmente a parte manuscrita apresenta falhas dificultando a etapa de segmentação);
- O conceito ruim refere-se à segmentação em 3 bacias, porém a bacia correspondente a carimbos e parte manuscrita (1ª bacia) não foi segmentada, aparecendo como parte da 2ª bacia correspondente ao fundo (Figura 6.4). Também considerou-se como ruim exemplos em que se obteve mais de 3 bacias.

É importante salientar também que, como descrito anteriormente no Capítulo 5 (Base de imagens de envelopes postais brasileiros), algumas imagens apresentam problemas de aquisição, gerando bordas adicionais que não fazem parte do envelope postal brasileiro, mas que são segmentadas com a classe correspondente aos selos (classe 3). Para facilitar a análise de resultados obtidos (principalmente da classe 1), todas as imagens aparecem no sentido horizontal.

6.2.2. Exemplos de resultados obtidos com matriz de co-ocorrência (Junção horizontal e vertical)



a) Imagem original



b) Imagem original invertida



c) Classe 1



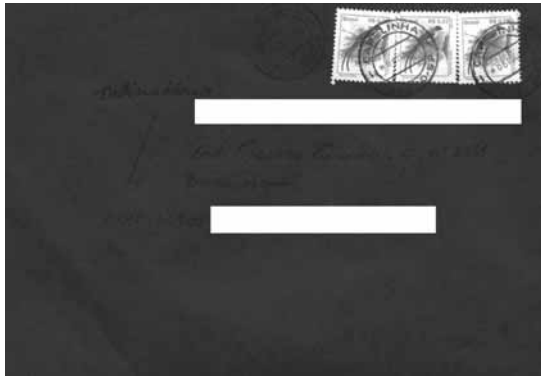
d) Classe 2



e) Classe 3

Figura 6.3 Exemplo de segmentação boa

Na Figura 6.3-a) é apresentada uma imagem de envelope postal com fundo relativamente complexo. Os carimbos aparecem aleatoriamente sobre o mesmo. Contudo, a segmentação obtida foi perfeita nas três classes esperadas (parte manuscrita com carimbos, fundo e selos).



a) Imagem original



b) Imagem original invertida



c) Classe 1



d) Classe 2

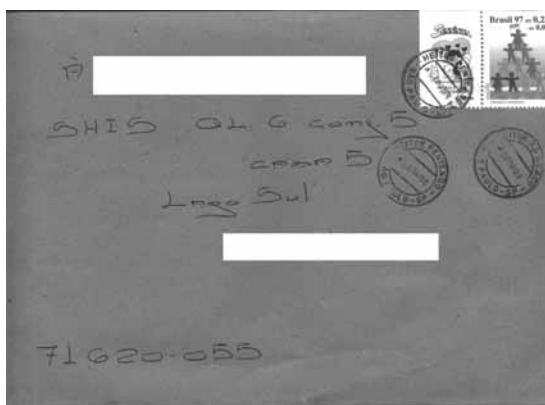


e) Classe 3

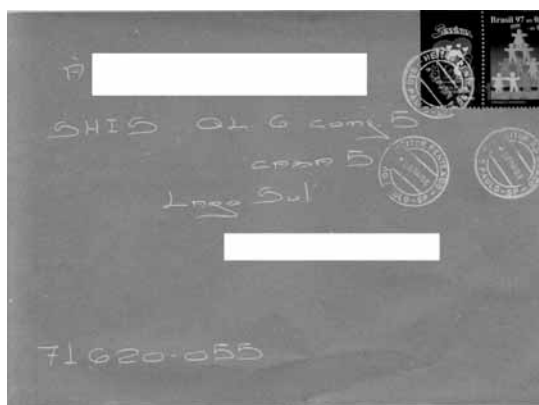
Figura 6.4 Exemplo de segmentação ruim

Na Figura 6.4-a), é possível perceber toda a complexidade do envelope postal em questão. É visível a falta de contraste entre a parte manuscrita e o fundo complexo. Tal complexidade pode ser verificada na obtenção dos resultados em que não foi possível segmentar a parte manuscrita.

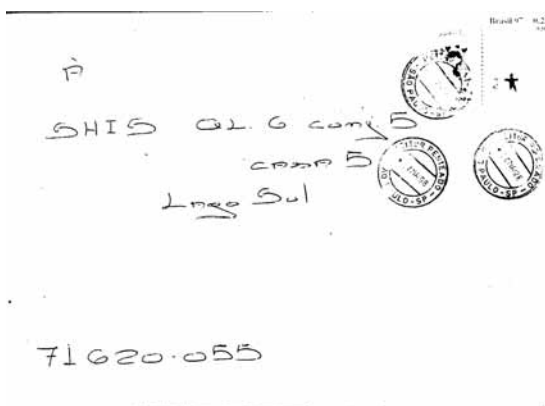
6.2.3. Exemplos Resultados obtidos com matriz de co-ocorrência (Junção das 4 direções)



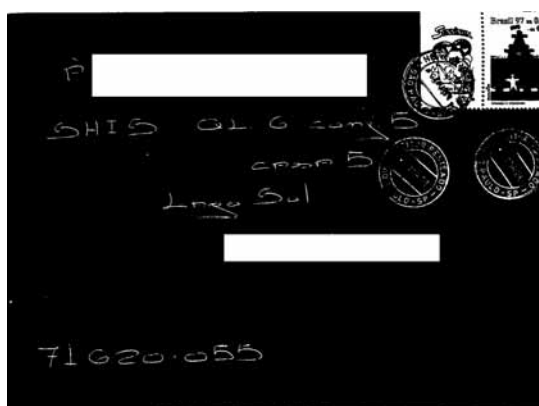
a) Imagem original



b) Imagem original invertida



c) Classe 1



d) Classe 2



e) Classe 3

Figura 6.5 Exemplo de segmentação boa

Apesar do envelope complexo conforme apresentado na Figura 6.5-a), em que a parte manuscrita e os carimbos aparecem de forma aleatória sobre um fundo complexo (amassado), obteve-se uma boa segmentação.

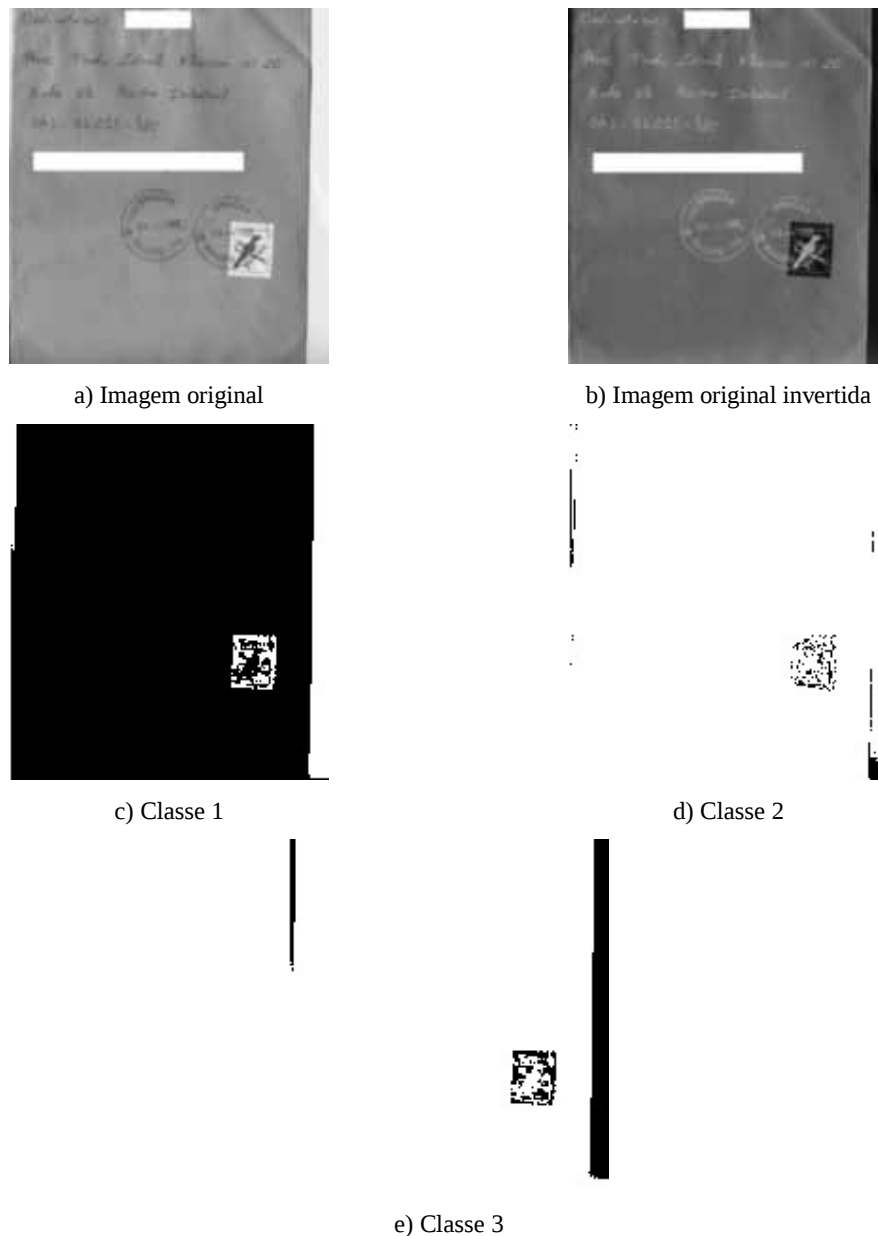


Figura 6.6 Exemplo de segmentação ruim

Na Figura 6.6-a), não existem regras quanto ao posicionamento da parte manuscrita, carimbos e selos. O fundo revela-se extremamente complexo (amassado e com bordas adicionais devido ao defeito de aquisição). Os resultados foram ruins, uma vez que a parte manuscrita não foi segmentada.

6.2.4. Análise comparativa dos resultados obtidos com matriz de co-ocorrência

Com relação à análise qualitativa, conforme as Tabelas 6.1 e 6.2, podemos afirmar que os resultados com matriz de co-ocorrência (junção vertical e horizontal) se revelam mais promissores do que os resultados obtidos pela junção das 4 direções. Isso se deve, principalmente ao fato de que, no 2º caso (junção das 4 direções), uma boa segmentação só é possível se a **parte manuscrita for muito bem definida e não apresentar falhas na escrita**, o que geralmente não ocorre na base de imagens de envelopes postais.

Junção Horizontal e Vertical		
Bom	Médio	Ruim
47 (58,75%)	16 (20%)	17 (21,25 %)

Tabela 6.1 Classificação dos resultados obtidos com junção horizontal e vertical

Junção das 4 direções		
Bom	Médio	Ruim
43 (53,75%)	15 (18,75%)	22 (27,5%)

Tabela 6.2 Classificação dos resultados obtidos com junção das 4 direções

Com relação aos tempos de processamento obtidos com matriz de co-ocorrência (junção das direções horizontal e vertical e junção das 4 direções), podemos afirmar que, além de apresentar melhores resultados qualitativos, a junção das direções horizontal e vertical também apresentou um menor tempo de processamento, conforme a Figura 6.7, bem como um menor tempo de processamento médio, conforme a Figura 6.8.

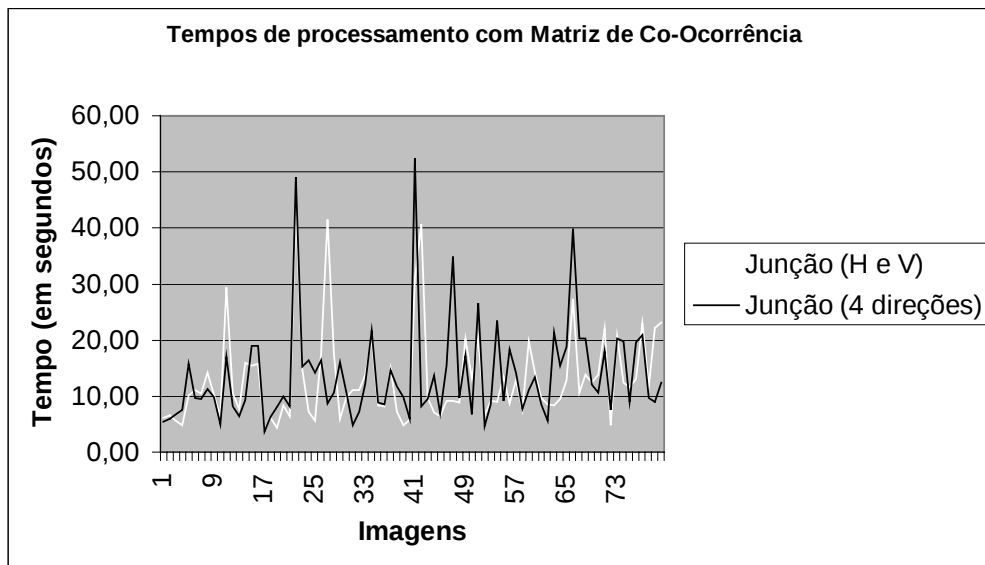


Figura 6.7 Tempos de processamento com matriz de co-ocorrência

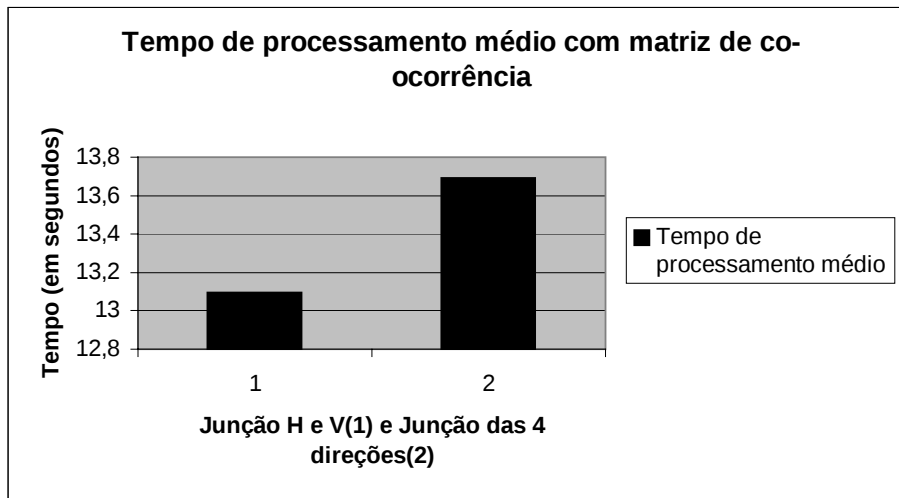


Figura 6.8 Tempos de processamento médio com matriz de co-ocorrência

6.2.5. Vantagens e desvantagens da abordagem proposta com matriz de co-ocorrência

Para a abordagem proposta com matriz de co-ocorrência é possível destacar as seguintes desvantagens:

- o método falha em imagens com pouco contraste entre a parte manuscrita e o fundo;
- o método não é ótimo na busca de 3 bacias, em alguns casos, gerando bacias espúrias.

Contudo, pode-se destacar as seguintes vantagens:

- não é necessário nenhum conhecimento a priori das imagens a serem processadas;
- não é necessário, exceto pela inversão da imagem original, nenhum pré-processamento das imagens;
- o método é invariante à inclinação do envelope e segmenta as classes independentemente das posições em que se encontram;
- o método é robusto para imagens ruidosas;
- o método pode ser estendido para outros contextos, como por exemplo, imagens de texturas.

6.3. Análise de Resultados com Histograma 2D

Nesta seção, serão detalhadas a metodologia utilizada com histograma 2D baseada em uma segmentação supervisionada, bem como uma estratégia de avaliação de resultados e as dificuldades encontradas devido as características dos envelopes postais brasileiros.

6.3.1. Metodologia com Histograma 2D

Conforme citado anteriormente, as imagens de envelopes postais brasileiros são constituídas de 4 classes, a saber: carimbos, parte manuscrita, fundo e selos. Portanto, na abordagem de segmentação dos mesmos, é de se esperar a obtenção dessas 4 classes. Contudo, devido à complexidade dos mesmos e principalmente à **similaridade de níveis de cinza** entre carimbos e parte manuscrita, **decidiu-se pela segmentação em 3 classes (bacias)**, que são:

Classe 1. carimbos com parte manuscrita;

Classe 2. fundo;

Classe 3. selos.

A grande incógnita da abordagem por histograma 2D (Capítulo 3) reside na escolha da 2ª imagem (imagem filtrada). Por ser extremamente utilizada na segmentação por limiarização (Capítulo 2), vários autores propuseram o filtro da Média [ABU89] [BRI92] [CHE94] [LI97] [GON98]. Porém, na segmentação de envelopes postais brasileiros, o filtro da Média tem a indesejável propriedade de gerar novos valores (níveis de cinza) que não existem na imagem original, fator este que, no cálculo do histograma 2D, pode-se gerar vários picos irrelevantes que não correspondem às classes reais.

Portanto, decidiu-se utilizar como 2ª imagem a imagem reconstruída, pois a mesma “encosta” na imagem original até o limite do marcador (nível de cinza), ou seja, somente os picos mais acentuados (máximos regionais) não serão totalmente reconstruídos. A reconstrução em níveis de cinza possui a propriedade de não alterar drasticamente a topologia da imagem (perfil de níveis de cinza) [VIN93], que é o que se pretende na presente abordagem. Após vários testes, chegou-se a conclusão, devido aos resultados, de que a imagem reconstruída seria obtida através da imagem *I-1* (Figura 6.9).

Então, o histograma 2D das duas imagens em questão (imagem original e imagem reconstruída) apresenta dois comportamentos distintos:

- as regiões não afetadas pela reconstrução encontram-se na diagonal principal do histograma 2D;
- as regiões afetadas pela reconstrução encontram-se fora da diagonal principal do histograma 2D.

Portanto, o histograma 2D é processado de tal maneira a somente considerar as regiões afetadas pela reconstrução. Estas regiões caracterizam-se por representar partes das classes a serem extraídas (parte manuscrita, carimbos, fundo e selos). Contudo, é importante salientar que, devido à quantidade e complexidade dos picos gerados pelo histograma 2D, o número de marcadores obtidos pelo processo de reconstrução morfológica é exageradamente alto, conforme mostrado na Figura 6.10 (em média 300). Faz-se necessário então, uma etapa de pós-processamento para que se diminuam o número de picos e consequentemente, de marcadores. A solução proposta é a utilização de operadores morfológicos básicos em níveis de cinza (Apêndice A), quais sejam, a erosão e a dilatação. No caso específico do histograma 2D, a erosão apresenta as seguintes propriedades:

- Eliminação de picos pequenos que podem ser relevantes;
- Diminuição da altura de picos, reduzindo a importância dos mesmos, gerando perda de informação.

Já a dilatação se comporta da seguinte maneira:

- Junção de picos pequenos, sem perda de informação;
- Junção de picos pequenos com grandes, sem aumentar a importância dos mesmos;
- Não há modificação da altura dos picos.

Portanto, conforme as características acima citadas, neste contexto decidiu-se pela utilização de n dilatações até que, por reconstrução morfológica, seja possível a obtenção de 3 marcadores, correspondentes as 3 bacias segmentadas pela ferramenta de segmentação morfológica Watershed, conforme a Figura 6.11.



Figura 6.9 Reconstrução da imagem I pela imagem I-1

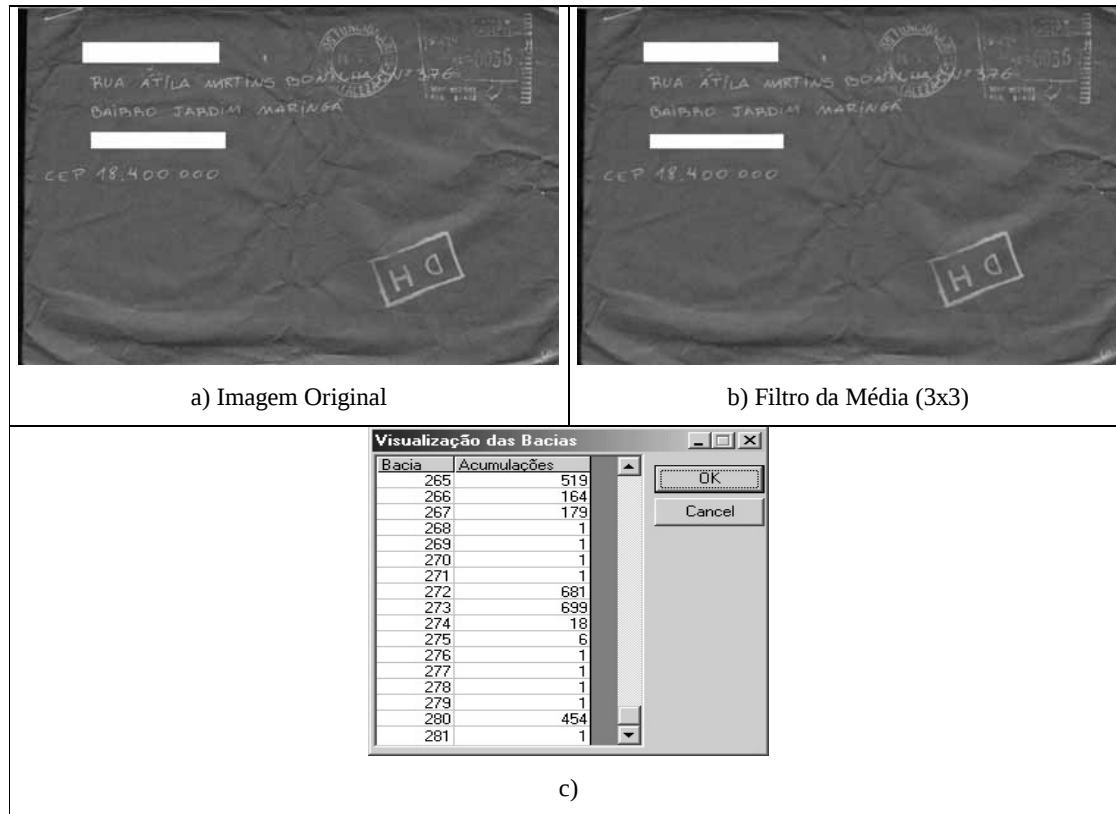


Figura 6.10. c) Número de bacias segmentadas sem pós-processamento com a abordagem proposta de histograma 2D para as imagens a) - b)

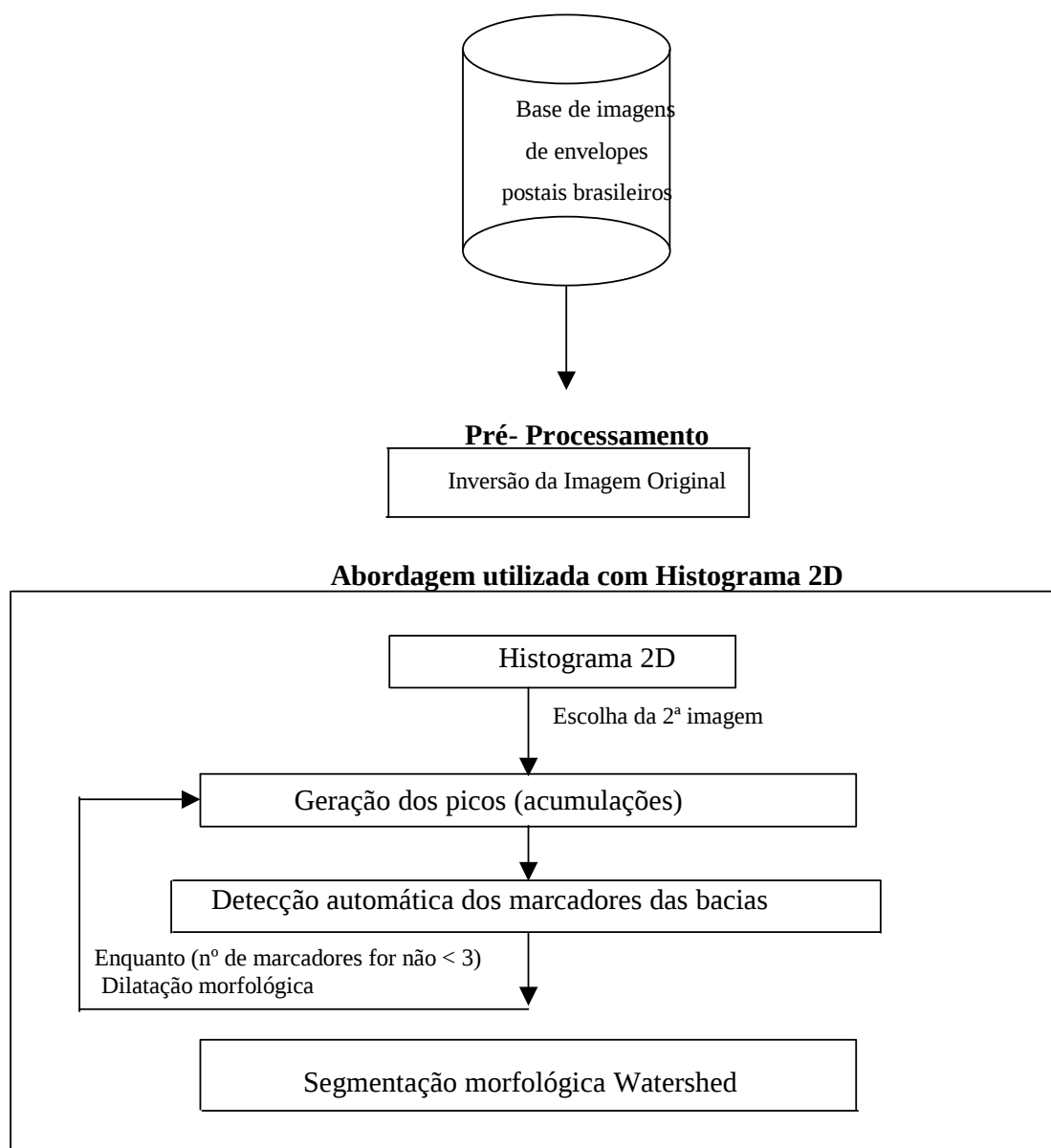
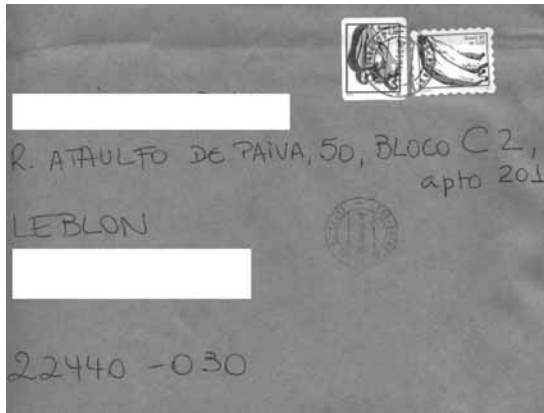
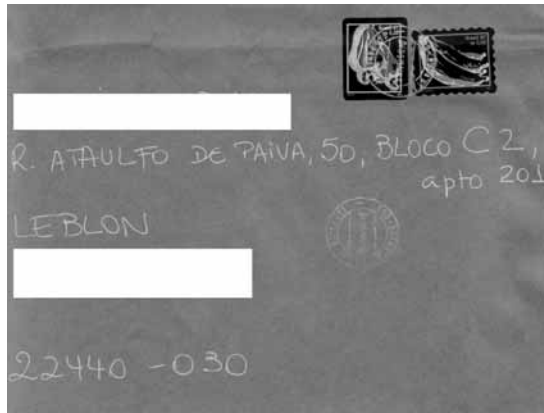


Figura 6.11 Diagrama com a estrutura da abordagem proposta para a segmentação de envelopes postais brasileiros

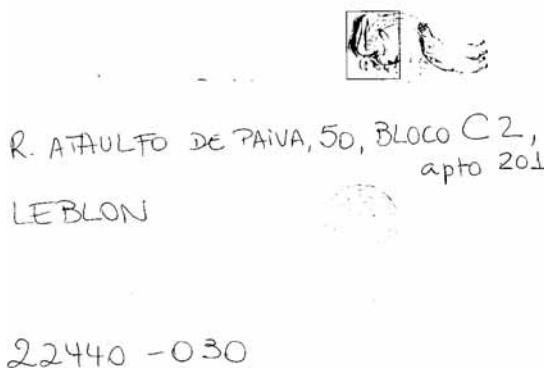
6.3.2. Resultados obtidos com histograma 2D



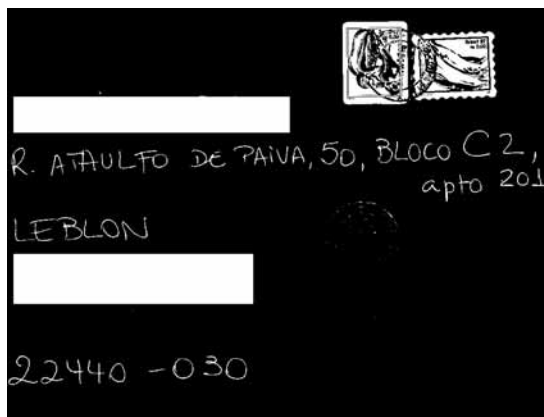
a) Imagem original



b) Imagem original invertida



c) Classe 1



d) Classe 2



e) Classe 3

Figura 6.12 Exemplo de segmentação boa

Na Figura 6.12-a), é apresentado um exemplo segmentação perfeita nas três classes esperadas, apesar da complexidade do envelope postal em questão, que caracteriza-se por apresentar um fundo complexo (amassado), e à falta de posicionamento dos carimbos.

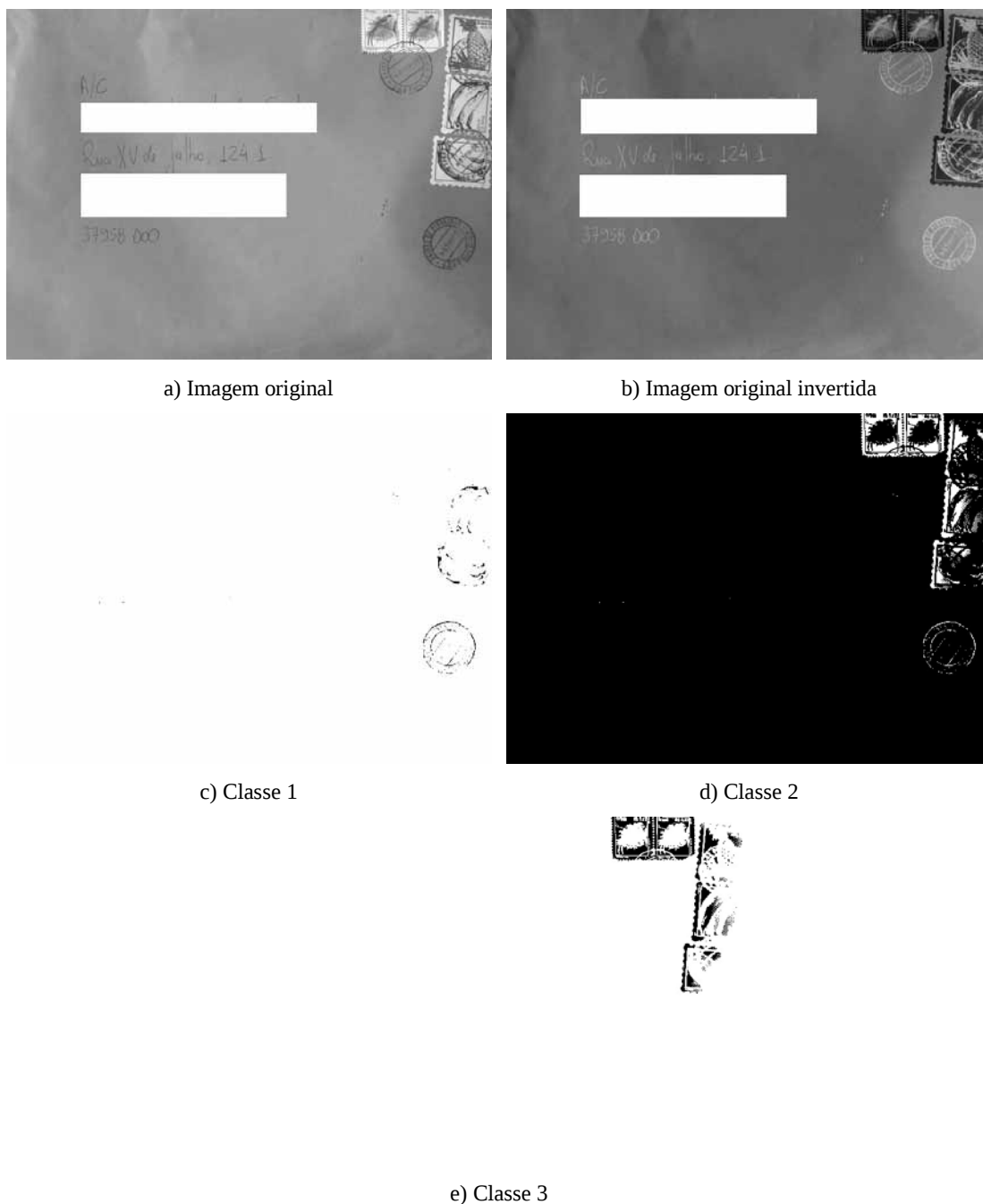


Figura 6.13 Exemplo de segmentação ruim

Na Figura 6.13-a), é possível perceber a falta de contraste entre a parte manuscrita e o fundo complexo (amassado) além do posicionamento incorreto e aleatório dos selos e carimbos. Obteve-se uma segmentação ruim, também devido à espessura da parte manuscrita.

6.3.3. Análise quantitativa e qualitativa dos resultados obtidos com histograma 2D

A Tabela 6.3 apresenta a análise quantitativa com relação aos resultados obtidos com a abordagem proposta utilizando histograma 2D. Os resultados revelam-se promissores considerando-se a porcentagem em que foi possível obter uma boa segmentação. É preciso salientar também a baixa porcentagem de resultados ruins (apenas 10%).

Histograma 2D		
Bom	Médio	Ruim
55 (68,75%)	17 (21,25%)	8 (10%)

Tabela 6.3 Classificação dos resultados obtidos com histograma 2D

A Figura 6.14 mostra o tempo de processamento obtido com histograma 2D para cada uma das imagens. Já o tempo médio de processamento ficou em torno de 14 segundos.

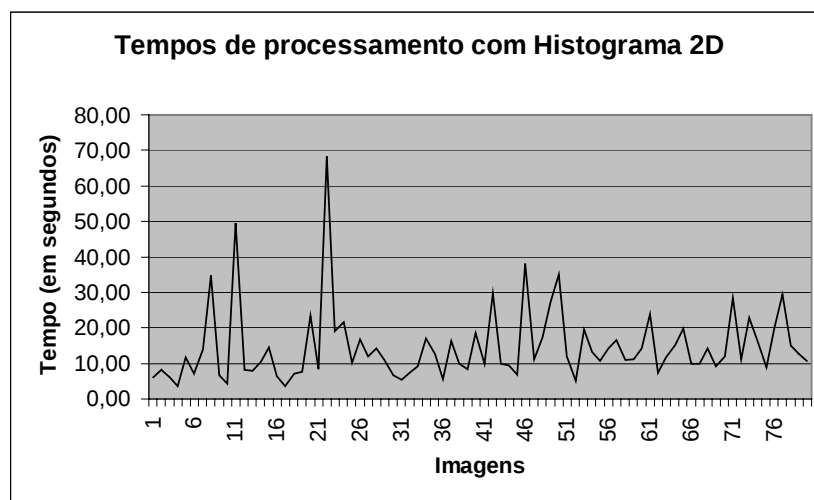


Figura 6.14 Tempos de processamento com histograma 2D

6.3.4. Vantagens e desvantagens da abordagem proposta com histograma 2D

Para a abordagem proposta com histograma 2D é possível destacar as seguintes desvantagens:

- o método não é robusto em imagens com pouco contraste entre a parte manuscrita e o fundo;
- o método não é ótimo na busca de 3 bacias, em alguns casos, gerando bacias espúrias.

Contudo, pode-se destacar as seguintes vantagens:

- pouco conhecimento a priori é necessário sobre as imagens a serem processadas;
- não é necessário, exceto pela inversão da imagem original, nenhum pré-processamento das imagens;
- o método é invariante à inclinação do envelope e segmenta as classes independentemente das posições em que se encontram;
- o método é robusto para imagens ruidosas;
- baixa porcentagem de resultados ruins, ao se considerar a complexidade da base de envelopes postais brasileiros.

6.4 Análise comparativa entre Histograma 2D e Matriz de Co-Ocorrência

O objetivo desta seção é o de comparar os resultados obtidos com a abordagem proposta utilizando histograma 2D com aqueles produzidos pela matriz de co-ocorrência. Também será efetuada uma comparação com os métodos de segmentação de envelopes postais presentes na literatura (Capítulo 2). Para tanto, vários parâmetros, tais como, base de imagens utilizadas, heurísticas, conhecimento a priori, invariância quanto à inclinação da parte manuscrita, complexidade e tempo de processamento serão analisados. A escolha da solução final baseia-se no resultado da competição entre as técnicas empregadas (histograma 2D e matriz de co-ocorrência).

6.4.1 Metodologia de Avaliação por competição

Nesta seção, é proposta uma metodologia de avaliação por competição entre as técnicas empregadas para a segmentação de envelopes postais brasileiros (histograma 2D e matriz de co-ocorrência). Para tal, foram escolhidas 5 imagens da base de envelopes postais brasileiros (Figura 6.15) que apresentam características particulares tais como, textura complexa, problemas de aquisição que geram bordas adicionais que não pertencem a nenhuma das classes, inclinação e sobreposição da parte manuscrita com os carimbos, e ainda casos ainda mais complexos (Figura 6.15e) em que uma região da parte manuscrita é sobreposta por um corretivo líquido. Para facilitar a comparação das classes segmentadas, nas Figuras 6.16 a 6.29, na primeira coluna tem-se os resultados obtidos com abordagem por matriz de co-

ocorrência, e na segunda coluna tem-se os resultados obtidos com abordagem por histograma 2D.

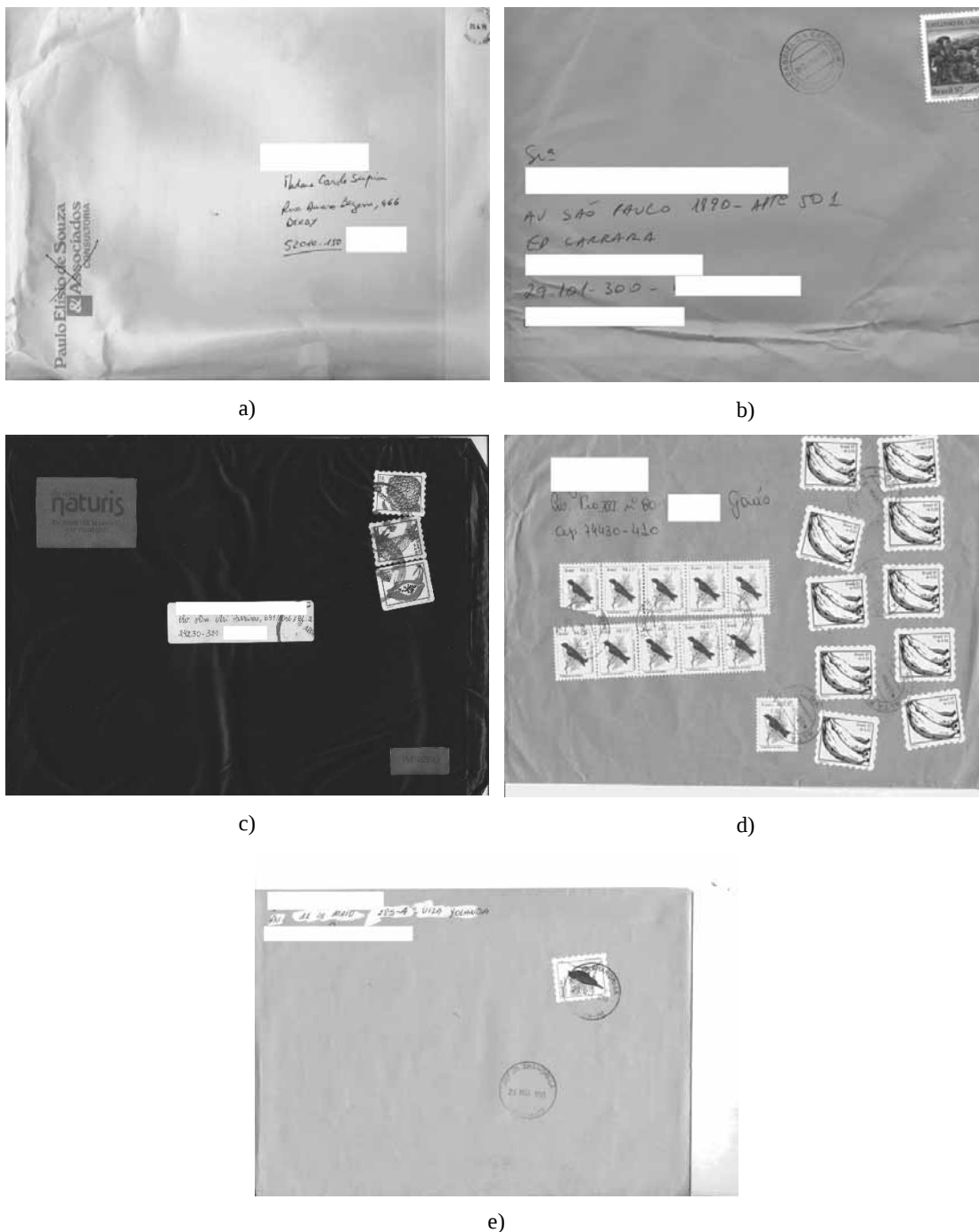
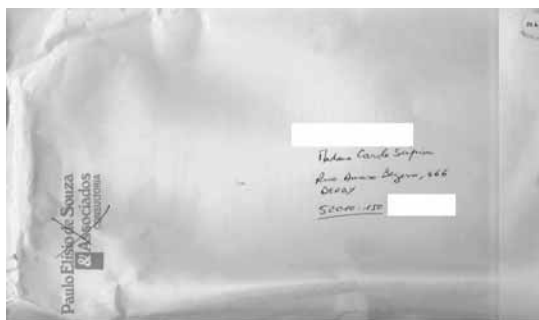


Figura 6.15 Imagens escolhidas para a metodologia de avaliação por competição

6.4.2. Resultados obtidos com as imagens escolhidas utilizando histograma 2D



a) Imagem Original



b) Classe 1



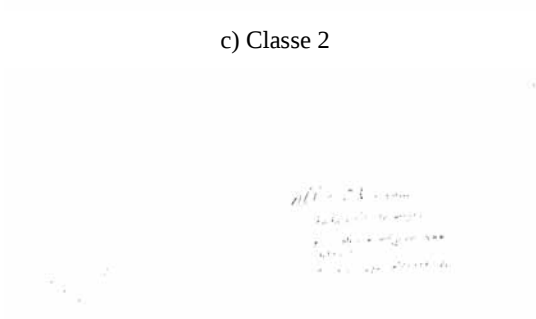
h) Classe 1



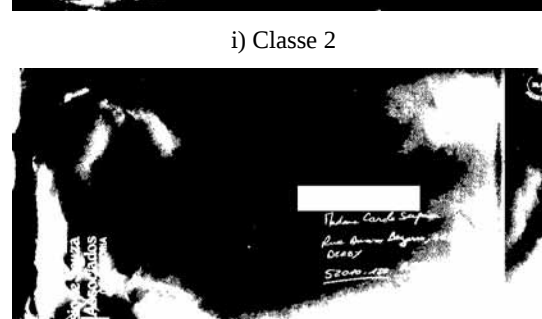
c) Classe 2



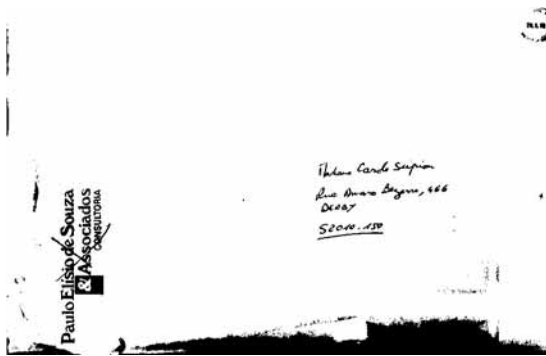
i) Classe 2



d) Classe 3



j) Classe 3



e) Classe 4



f) Classe 5



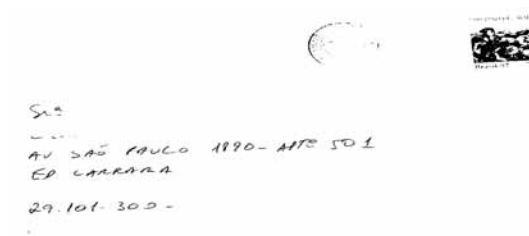
g) Classe 6

Figura 6.16 Classes segmentadas com abordagem por matriz de co-ocorrência e por histograma 2D respectivamente da imagem 6.16-a)

Na Figura 6.16-a), percebe-se toda a complexidade do envelopes postal. O fundo é extremamente complexo (amassado e com muita variação nos níveis de cinza). O envelope postal em questão não apresenta selos. Contudo, a abordagem por histograma 2D mostra-se eficiente, segmentando corretamente a parte manuscrita e o fundo nas duas classes subsequentes devido à ausência de selos.



a) Imagem original



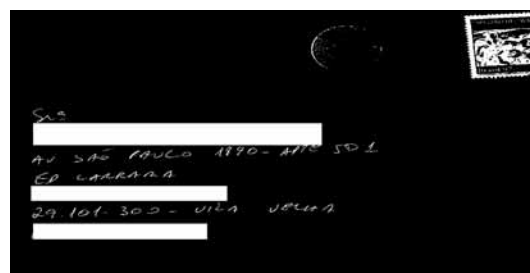
b) Classe 1



e) Classe 1



c) Classe 2



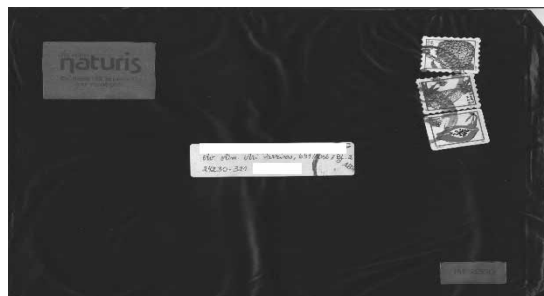
f) Classe 2

d) Classe 3

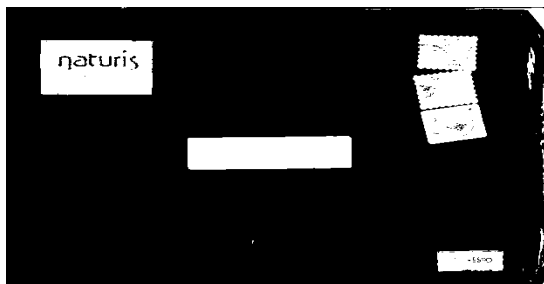
g) Classe 3

Figura 6.17 Classes segmentadas com abordagem por matriz de co-ocorrência e por histograma 2D respectivamente da imagem 6.17-a)

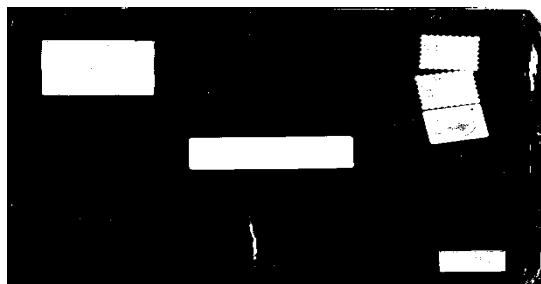
Na Figura 6.17-a), o envelope caracteriza-se por apresentar um fundo extremamente complexo (amassado) sobreposto sobre a parte manuscrita. Contudo, obteve-se uma segmentação perfeita nas três classes esperadas nas duas abordagens propostas.



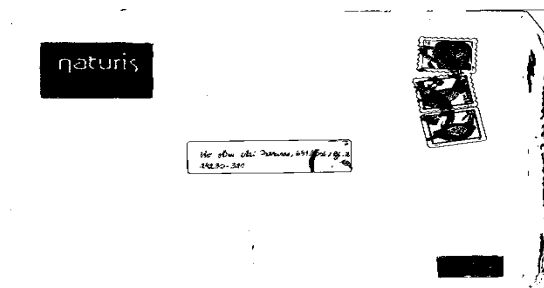
a) Imagem original



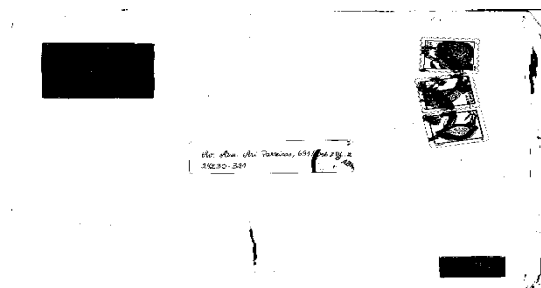
b) Classe 1



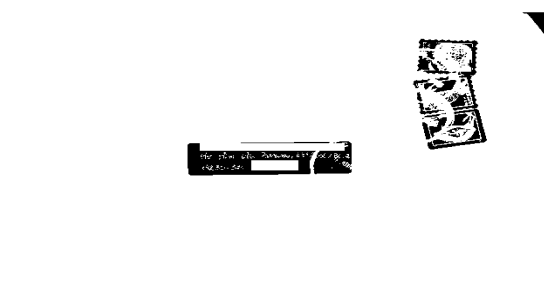
e) Classe 1



c) Classe 2



f) Classe 2



d) Classe 3



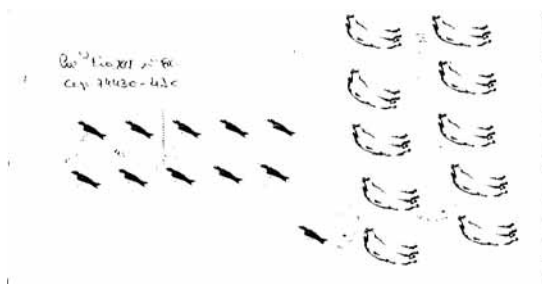
g) Classe 3

Figura 6.18 Classes segmentadas com abordagem por matriz de co-ocorrência e por histograma 2D respectivamente da imagem 6.18-a)

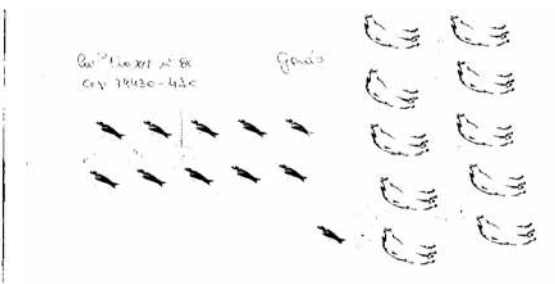
Na Figura 6.18-a) é apresentado um envelope postal particular, pois a parte manuscrita encontra-se sobreposta sobre um rótulo branco. Novamente, os resultados obtidos nas abordagens propostas revelam-se eficientes, uma vez que a parte manuscrita é segmentada com sucesso.



a) Imagem original



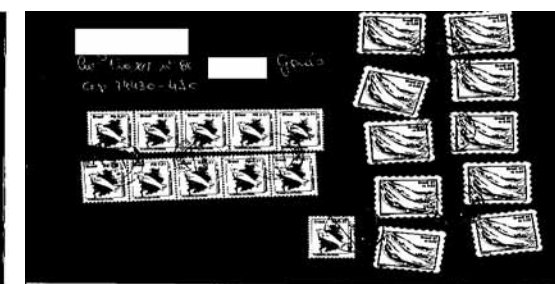
b) Classe 1



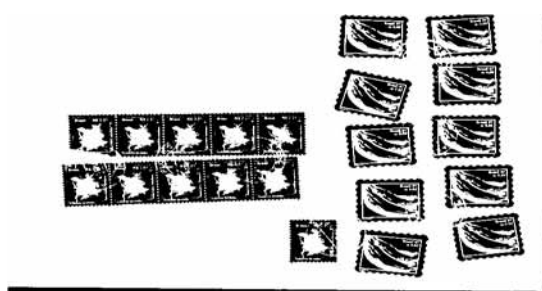
e) Classe 1



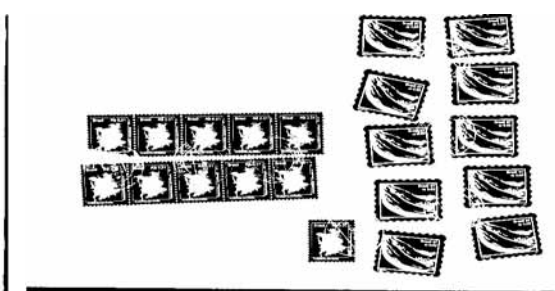
b) Classe 2



f) Classe 2



d) Classe 3



g) Classe 3

Figura 6.19 Classes segmentadas com abordagem por matriz de co-ocorrência e por histograma 2D respectivamente da imagem 6.19-a)

Na Figura 6.19-a), o envelope postal não respeita nenhum posicionamento quanto à parte manuscrita, carimbos e selos que aparecem exageradamente sobre o mesmo. O fundo caracteriza-se pela complexidade. Porém, a segmentação foi eficiente nas abordagens propostas, devido às dificuldades existentes.



a) Imagem original



b) Classe 1



e) Classe 1



c) Classe 2



f) Classe 2



d) Classe 3



g) Classe 3

Figura 6.20 Classes segmentadas com abordagem por matriz de co-ocorrência e por histograma 2D respectivamente da imagem 6.20-a)

Na Figura 6.20, o envelope postal caracteriza-se por apresentar uma região da parte manuscrita com um corretivo líquido. Os carimbos e selos não respeitam as regras de posicionamento, e o fundo complexo (amassado) apresenta bordas adicionais devido ao defeito de aquisição.

6.4.3. Comparação de resultados obtidos com Histograma 2D e Matriz de Co-Ocorrência

Conforme citado anteriormente, as imagens de envelopes postais brasileiros utilizadas neste capítulo caracterizam-se apresentam características particulares, tais como, textura complexa, problemas de aquisição que geram bordas espúrias que não pertencem a nenhuma das classes, inclinação e sobreposição da parte manuscrita com os carimbos, etc. Isto dificulta a obtenção de bons resultados. Nesta seção, uma análise qualitativa por competição em termos de resultados obtidos empregando as duas técnicas (histograma 2D e matriz de co-ocorrência) é proposta.

Pode-se dizer que, de uma forma geral, para as imagens apresentadas em questão (Figura 6.15), a abordagem de segmentação com histograma 2D apresentou melhores resultados. Isto se verifica para as imagens da Figura 6.15a (com relação ao número de bacias segmentadas) e 6.15e (segmentação da parte manuscrita prejudicada por um corretivo líquido). Houve uma similaridade de resultados obtidos para as imagens das Figuras 6.15b e 6.15c, e na Figura 6.15d a abordagem de segmentação com matriz de co-ocorrência apresentou melhores resultados. Contudo, é importante salientar que, se considerarmos os tempos de processamento obtidos para cada imagem e médios conforme as Figuras 6.21 e 6.22 respectivamente, percebe-se um melhor desempenho da abordagem de segmentação com matriz de co-ocorrência. Outro aspecto importante reside no fato de que, no caso do histograma 2D, após vários testes, chegou-se a uma escolha da 2ª imagem (imagem filtrada), o que tornou possível a obtenção de tais resultados, ao passo que com matriz de co-ocorrência as heurísticas reduzem-se somente à escolha do sentido de varredura (horizontal, vertical, diagonais, ou junção das direções combinadas entre si) e à distância entre pixels.

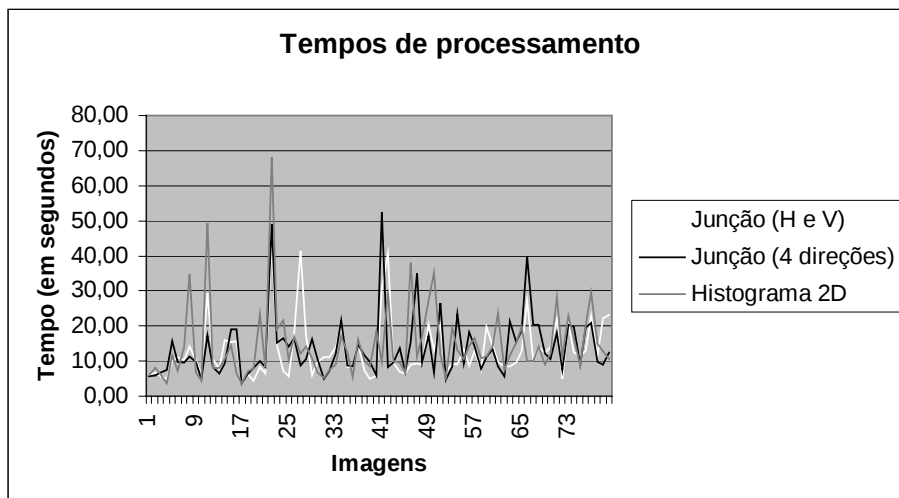


Figura 6.21 Tempos de processamento com as duas técnicas

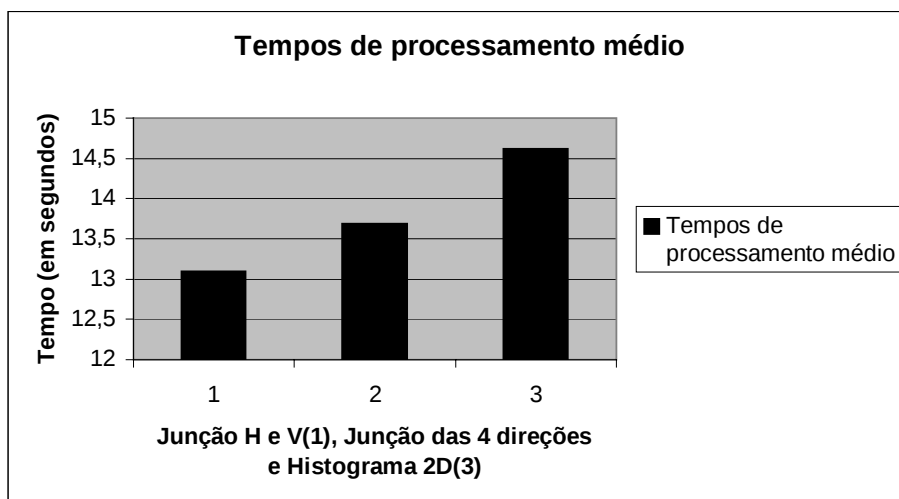


Figura 6.22 Tempos de processamento médio com as duas técnicas

6.4.4. Comparação da abordagem proposta (histograma 2D e matriz de co-ocorrência) X Estado da Arte (Capítulo 2)

Conforme citado anteriormente na literatura, os métodos de segmentação de envelopes postais empregam técnicas de limiarização [WOL97] [YU97], inteligência artificial (sistemas especialistas) [WAN88], características geométricas [DOW90] [YEH87] ou ainda, análise textural por filtragem [JAI95]. Nesta seção, uma comparação de diversos parâmetros, tais como, tipo de imagens utilizadas (complexa ou simples), heurísticas empregadas, conhecimento a priori, invariância quanto à inclinação da parte manuscrita, complexidade e tempo de processamento, entre as abordagens propostas neste trabalho e o estado da arte será efetuada:

	Abordagens propostas	Estado da Arte
Base de imagens utilizadas	Imagens de envelopes postais com textura complexa	Imagens de envelopes postais com textura complexa, com a vantagem de que o método de [YU97] é capaz de processar imagens de envelopes postais plastificados e de jornais
Heurísticas empregadas	Com histograma 2D, a heurística concentra-se na 2ª imagem (imagem filtrada), e com matriz de co-ocorrência, é preciso escolher o sentido de varredura e a distância entre pixels	Emprego de várias heurísticas, especialmente nos métodos de segmentação que empregam técnicas de limiarização [WOL97] [YU97] e inteligência artificial (sistemas especialistas) [WAN88]
Conhecimento a priori	Pouco uso de conhecimento a priori	Os métodos de segmentação que utilizam características geométricas baseiam-se no conhecimento a priori (inclinação da parte manuscrita e posição das

		classes que compõe o envelope postal)
Invariância à inclinação da parte manuscrita	Invariante à inclinação da parte manuscrita	Novamente os métodos de segmentação baseados em características geométricas são dependentes da inclinação da parte manuscrita
Complexidade	A abordagem proposta e os métodos de segmentação baseados em características geométricas caracterizam-se pela pouca complexidade	O método de [JAIN <i>et. al</i> (1995)], que utiliza filtros de Gabor bidimensionais destaca-se pela alta complexidade dentre os métodos pesquisados
Tempo de processamento	Destaca-se a abordagem proposta com matriz de co-ocorrência (junção das direções horizontal e vertical), com tempo de processamento médio de aproximadamente 13s	

Tabela 6.4 Comparação da abordagem proposta (histograma 2D e matriz de co-ocorrência) X Estado da Arte (Capítulo 2)

6.5 Conclusão

Neste capítulo, uma descrição detalhada dos resultados obtidos por matriz de co-ocorrência e histograma 2D, bem como uma avaliação por competição entre as duas abordagens foram propostas. Pode-se concluir, devido aos resultados obtidos, que a abordagem por histograma 2D mostrou-se mais eficiente. Apresentou-se também uma comparação das abordagens com os métodos de localização do bloco-endereço presentes no Capítulo 2.

Conclusão

Um dos grandes desafios no processamento digital de imagens consiste na busca de uma segmentação eficiente. Contudo, conforme citado anteriormente, não existe um único método ótimo de segmentação para todos os contextos. No caso em particular dos envelopes postais, os métodos pesquisados na literatura baseiam-se somente na localização do bloco-endereço sem se preocuparem com as demais classes que os compõem. As técnicas apresentadas (limiarização, inteligência artificial, características geométricas e análise textural por filtragem) nem sempre são adequadas para o problema em questão.

Para tanto propôs-se um método composto baseado em matriz de co-ocorrência e histograma 2D com morfologia matemática. O mérito da pesquisa reside justamente na combinação dessas abordagens aproveitando ao máximo as vantagens de cada uma com o objetivo de segmentar envelopes postais brasileiros complexos com pouco uso de conhecimento a priori a respeito dos mesmos.

Conforme já citado no capítulo 1, a principal contribuição deste trabalho é:

- Uma proposta de segmentação de envelopes postais brasileiros baseadas no Histograma 2D e Matriz de Co-Ocorrência juntamente com métodos morfológicos;
- Uma contribuição para pesquisas futuras na área de segmentação de envelopes postais brasileiros.

O mérito da pesquisa é a combinação de duas abordagens (matriz de co-ocorrência e histograma 2D) com a morfologia matemática aproveitando ao máximo as vantagens de cada uma com o objetivo de segmentar envelopes postais brasileiros complexos. Apesar dos resultados apresentados serem promissores, como trabalhos futuros na segmentação de envelopes postais brasileiros complexos, no caso específico da matriz de co-ocorrência, propõe-se ainda um estudo estatístico a respeito do mesmo [HAR92]. Já no histograma 2D, propõe-se um estudo detalhado sobre a escolha da 2ª imagem (imagem filtrada) que ainda permanece como uma incógnita.

Porém, é importante salientar que a abordagem de segmentação proposta não se restringe somente aos envelopes postais, podendo ser estendida para outros contextos, tais como:

- imagens de texturas;
- imagens médicas.

Referências Bibliográficas

- [ABU89] ABUTALEB, A.S. *Automatic Thresholding of Gray-Level Pictures using two-dimensional entropy*. CVGIP'89, 47, p. 22- 32.
- [BEU77] BEUCHER, S. *Random process simulation on the texture analyser*. Lecture notes in Biomathematics, 23, 1977.
- [BEU89] BEUCHER, S. *The Watershed Transformation applied to Image Segmentation*. Centre de Morphologie Mathématique, Ecole des Mines de Paris, 26 p.
- [BEU93] BEUCHER, S; MEYER, F. *The Morphological Approach to Segmentation: The Watershed Transformation*, Mathematical Morphology in Image Processing.
- [BRI92] BRINK, A.D. *Thresholding of Digital Images using two-dimensional entropies*. Pattern Recognition 25, p. 803- 808.
- [BRI96] BRITTO, A.S. *Segmentação por Cor utilizando a Transformada Karhunen-Loève*. Dissertação de Mestrado, Centro Federal de Educação Tecnológica do Paraná, Curitiba.
- [BUS01] BUSSADOR, A. *Localização Automática de Placas de Veículos em Fotos Digitais utilizando Abordagem Granulométrica*. Dissertação de Mestrado, Pontifícia Universidade Católica do Paraná, Curitiba.
- [CAN86] CANNY J. F. *A Computational Approach to Edge Detection*. IEEE Trans. Pattern Analysis Mach. Intell. 8, pp. 679-698.
- [CHE89] CHEN, C.C.; DAPONTE, J.S; FOX, M.D. *Fractal Feature Analysis and Classification in Medical Imaging*. IEEE Transactions on Medical Imaging, 8, p. 133-142.

- [CHE1994] CHEN, W.T.; WEN, C.H.; YANG, C.W. *A fast two-dimensional entropic thresholding algorithm*. Pattern Recognition 27, p. 885- 893.
- [DER87] DERICHE, R. *Optimal edge detection using recursive filtering*. International Journal of Computer Vision.
- [DOW90] DOWNTON, A.C.; LEEDHAM C.G. *Preprocessing and Presorting of Envelope Images for Automatic Sorting Using OCR*. Pattern Recognition, 23(4):347-362.
- [EVI96] EVIATAR H.; SOMORJAI R.L. *A Fast Simple Contour Algorithm for Biomedical Images*. Pattern Recognition Letters, 17, pp. 969-974.
- [FAC96] FACON, J. *Morfologia Matemática: Teoria e Exemplos*. Curitiba: PUC- PR. 304 p.
- [FAC98] FACON, J. *Processamento e Análise de Imagens*. Curitiba: PUC- PR, 170 p. (Apostila).
- [FAC01] FACON, J. *Metodologia de Avaliação de Abordagens de Segmentação de Imagens*. Curitiba: PUC- PR, 2001, 138p.
- [GON98] Gonçalves P.J.S., *Extração e Medição dos Contornos da Válvula Mitral Utilizando Modelos Activos de Contornos*. Dissertação de Mestrado em Engenharia Mecânica – Perfil de Sistemas, 1998.
- [GON97] GONG, J.; LI, L.; CHEN, W. *Fast Recursive algorithms for two- dimensional Thresholding*. Pattern Recognition 31, p. 295- 300.
- [GON92] GONZALEZ, R; WOODS, R.E. *Digital Image Processing*. Addison-Wesley.
- [HAR87] HARALICK R.M; STERNBERG, S.R; ZHUANG X. *Image Analysis using Mathematical Morphology*. IEEE Transactions On Pattern Analysis and Machine Intelligence, 4, p. 532-550.

- [HAR92] HARALICK R.M; SHAPIRO L. *Computer and Robot Vision*, Addison-Wesley.
- [JAI95] JAIN A.; BHATTACHARJEE S. *Address Block Location on Envelopes Using Gabor Filters*. Technical Report, Dept. of Computer Science, Michigan State University.
- [JES01] JESUS, A. *Uma Abordagem Morfológica para a Segmentação de Logotipos em Cheques Bancários Brasileiros sem conhecimento a priori*. Dissertação de Mestrado, Pontifícia Universidade Católica do Paraná, Curitiba.
- [KAP85] Kapur N.J.; Sahoo P.K.; Wong A. K. C. *A New Method for Gray-Level Picture Thresholding*, Computer Vision, Graphics and Image Processing, vol.29, pp. 273-285.
- [KAS87] Kass M.; Witkin A.; Terzopoulos D. *Snakes: Active Contour Models*. First Int. Conf. on Comp. Vision, London, pp. 259-268.
- [LAN78] LANDEWEERD, G.H; GELSEMA, E.S. *The Use of Nuclear Texture Parameters in the Automatic Analysis of Leukocytes*. Pattern Recognition, 10, p. 57-61.
- [LAI95] LAI, K.F.; CHIN R.T. *Deformable Contours: Modelling and Extraction*. IEEE Trans. Pattern Anal. Machine Intell., vol. 17, n°11, p.1084-1090.
- [LIE93] LIE, W.N. *An efficient threshold-evaluation algorithm for image segmentation based on spatial gray-level co-occurrences*. Signal Processing 33, p. 121- 126.
- [MAR00] MARTINS, P.R. *Segmentação de Histogramas Multimodais: Aplicação em Cheques Bancários Brasileiros*. Dissertação de Mestrado, Pontifícia Universidade Católica do Paraná, Curitiba.
- [MAT99] MATTANA M.F. *Avaliação da Qualidade da Segmentação por Binarização de Cheques Bancários a partir do Processo de Reconhecimento*. Dissertação de Mestrado, Pontifícia Universidade Católica do Paraná, Curitiba.

- [OTS79] OTSU N. *A Threshold Selection Method from Gray-Level Histograms*, IEEE Trans. Sys. Man and Cybernetics, v. SMC9, n° 1, pp. 62-66.
- [PAL89] PAL, N.R.; PAL, S.K. *Entropic Thresholding*. Signal Processing, 16, p. 97- 108.
- [PUN81] PUN, T. *Entropic Thresholding: a new approach*. Comp. Graphics and Image Proc., 16, p. 210-239.
- [RAN95] Ranganath, S. *Contour Extration from Cardiac MRI Studies Using Snakes*. IEEE Trans. Medical Imaging, vol. 14, n°2, p. 328-338,
- [RIG90] RIGNOT, E; KWOK, R. *Extraction of Textural features in SAR images: Statistical Model and Sensitivity*. Proceedings of International Geoscience and Remote Sensing Symposium, p. 1979-1982.
- [RIT96] RITTER, G.X.; WILSON, J. *Handbook of Computer Vision Algorithms in Image Algebra*. North Carolina : CRC Press. p. 125-141.
- [SHAN49] SHANNON, C.E.; WEAVER, W. *The Mathematical theory of communication*. The University of Illinois Press, Urbana.
- [SCH94] SCHULZE, M. *Biomedical Image Processing with Morphology-Based Nonlinear Filters*. Tese de Doutorado, The University of Texas at Austin
- [SCH67] SCHWARTZ, L. *Mathematics for Physical Sciences*. Hermann, Paris.
- [SWE95] SWETS, D.L.; PUNCH, B. *Genetic Algorithms for Object Localization in a Complex Scene*. Technical Report, Dept. of Computer Science, Michigan State University.
- [TAM78] TAMURA, H.S; YAMAWAKI, Y. *Textural Features corresponding to Visual Perception*. IEEE Transactions on Systems, Man, and Cybernetics, SMC-8, p. 460-473

- [TUC98] TUCERYAN, M; JAIN, A.J. *Texture Analysis*. The Handbook of Pattern Recognition and Computer Vision (2^a edição), World Scientific Publishing Co., p. 207-248.
- [YU97] YU, B.; JAIN, A.K., MOHIUDDIN, M. *Address Block Location on Complex Mail Pieces*. Proc. Fourth Intl. Conf. on Document Analysis and Recognition (ICDAR), p. 897-901.
- [YEH87] YEH P.S.; ANTOY S.; LITCHER A.; ROSENFELD A. *Address Location on Envelopes*. Pattern Recognition, 20(2): p.213-227.
- [WAN88] WANG C.H.; P.W. PALUMBO; SRIHARI S.N. *Object Recognition in Visually Complex Environments: An Architecture for Locating Address Blocks on Mail Pieces*. Proc. Ninth Intl. Conf. on Pattern Recognition, 365-367.
- [WOL97] WOLF, H.N; SCHMIDT, W. *Fast Address Block Location on Handwritten and Machine-Printed Mail-piece Images*. Proc. Fourth Intl. Conf. on Document Analysis and Recognition (ICDAR), p. 753-757.
- [ZHA98] ZHANG Y. *Detection of Urban Housing Development using MultisensorSatellite Data and Maps*. International Archives of Photogrammetry and Remote Sensing, vol. 32, 292-299.

Bibliografia Complementar

- [BOL80] BOLDRINI, J.L; COSTA, S.I.R; FIGUEIREDO, V.L.; WETZLER, H.G. *Álgebra Linear*. Departamento de Matemática da Universidade Estadual de Campinas (UNICAMP).
- [DAV95] DAVIS, S.R. *C++ para leigos*. Berkeley.
- [DOW98] DOWNING, D; CLARK, J. *Estatística Aplicada*. Editora Saraiva.
- [HOR97] HORTON, I. *Beginning MFC Programming*. Wrox.
- [KAI99] KAIN, E. *The MFC Answer Book*. Addison-Wesley.
- [JAI92] JAIN A.; BHATTACHARJEE S; CHEN, Y. *On Texture in document images*. Computer vision & Pattern Recognition. p. 677-680.
- [LEE89] LEE, J; GLICKMAN, F. *Applying intelligent robotics and vision to mail processing*. International Symposium on Intelligent Control. p. 724-729.
- [MAR98] MARSHALL, D. *ActiveX/OLE Programming*. R&D Books.
- [SCH92] SCHALKOFF, R. *Pattern Recognition: Statistical, Structural and Neural approaches*. John Wiley & Sons.
- [STR93] STROUSTRUP, B; ELLIS, M. *C++: Manual de Referência Comentado*. 5ª ed, Editora Campus.
- [STR00] STROUSTRUP, B. *A Linguagem de Programação C++*. 3ª ed, Editora Bookman.

Apêndice A

Morfologia Matemática em Níveis de Cinza

Na morfologia em níveis de cinza, diferentemente da morfologia binária, é necessário conhecer o nível de cinza do pixel mais escuro MIN e o nível de cinza do pixel mais claro MAX, na vizinhança de cada pixel ou em uma parte da vizinhança da imagem original. Uma operação morfológica em níveis de cinza é completamente definida através do tamanho e forma da vizinhança (elemento estruturante), das regiões de pesquisa de MIN e MAX e do algoritmo.

A.1. Operadores Morfológicos Básicos em Níveis de Cinza

A.1.1. Erosão

A erosão de um sinal f por um elemento estruturante g é a seguinte:

$$\varepsilon^g(f(x)) = \text{Min}\{f(x+y) - g(y) : y \in D[g]\} \quad (\text{A.1})$$

onde a erosão de f por g consiste em verificar se o elemento centrado em x está abaixo da função f . Min equivale a mínimo, x é o ponto a ser processado na imagem original, y são os pontos envolvidos pelo elemento estruturante e $D[g]$ é o domínio do elemento estruturante.

Uma imagem, de forma geral, apresenta um fundo que pode ser ou não uniforme e sobrepor padrões mais claros e/ou escuros. Portanto, uma imagem pode ser comparada a um relevo topográfico onde padrões claros são picos e escuros são vales. Através desta analogia, padrões de comportamento podem ser estabelecidos para a erosão, como exemplificado na Figura A.1.

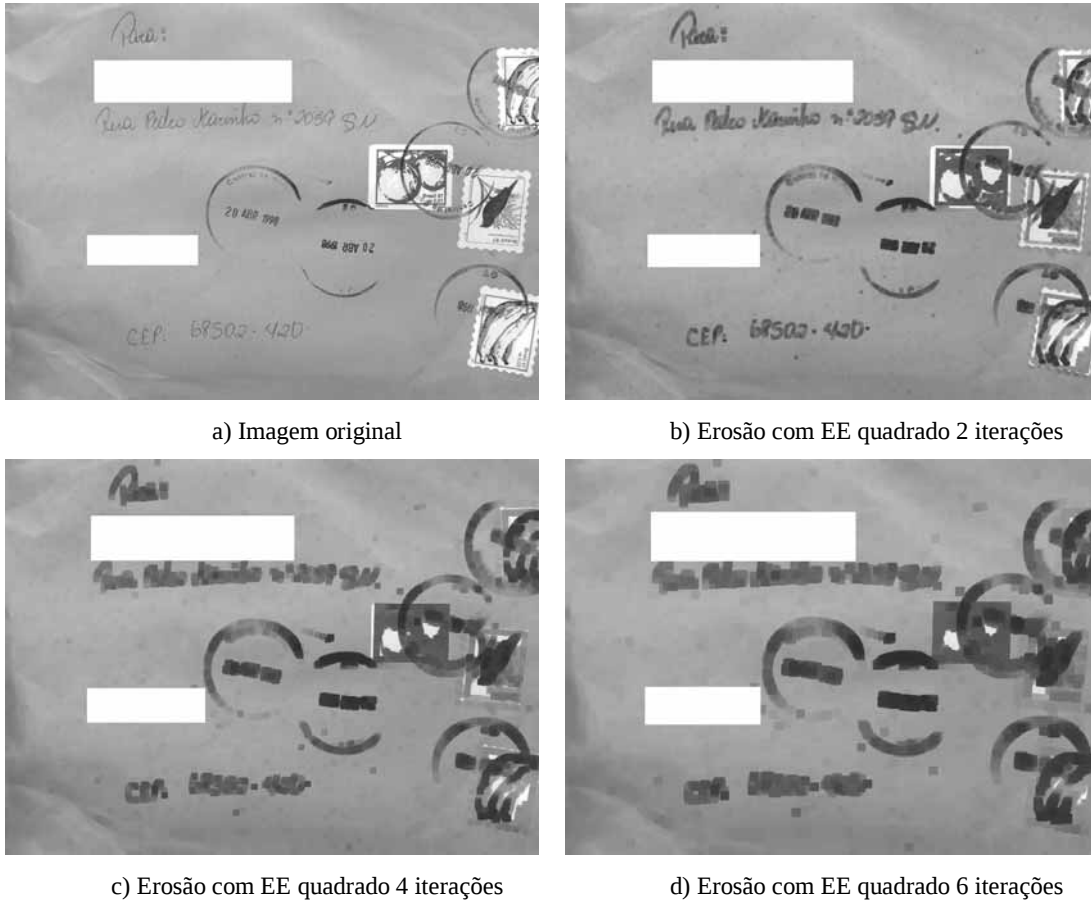


Figura A.1 Exemplo de Erosão

Os efeitos da erosão em níveis de cinza são:

- escurecer a imagem;
- alargar e engordar os vales (padrões escuros);
- conectar vales próximos;
- reduzir e às vezes eliminar picos (padrões claros);
- separar picos próximos.

A.1.2. Dilatação

A dilatação de um sinal f por um elemento estruturante g é a seguinte:

$$\delta^g(f(x)) = \text{Max}\{f(x-y) + g(y) : y \in D[g]\} \quad (\text{A.2})$$

onde a dilatação de f por g consiste em verificar se o elemento estruturante centrado em x está acima da função f . Max equivale a máximo, x é o ponto a ser processado na imagem

original, y são os pontos envolvidos pelo elemento estruturante e $D[g]$ é o domínio do elemento estruturante.

Considerando a analogia da imagem em níveis de cinza com o relevo topográfico onde os padrões claros são picos e os escuros são vales, é possível determinar padrões de comportamento para a dilatação, como vemos na Figura A.2.

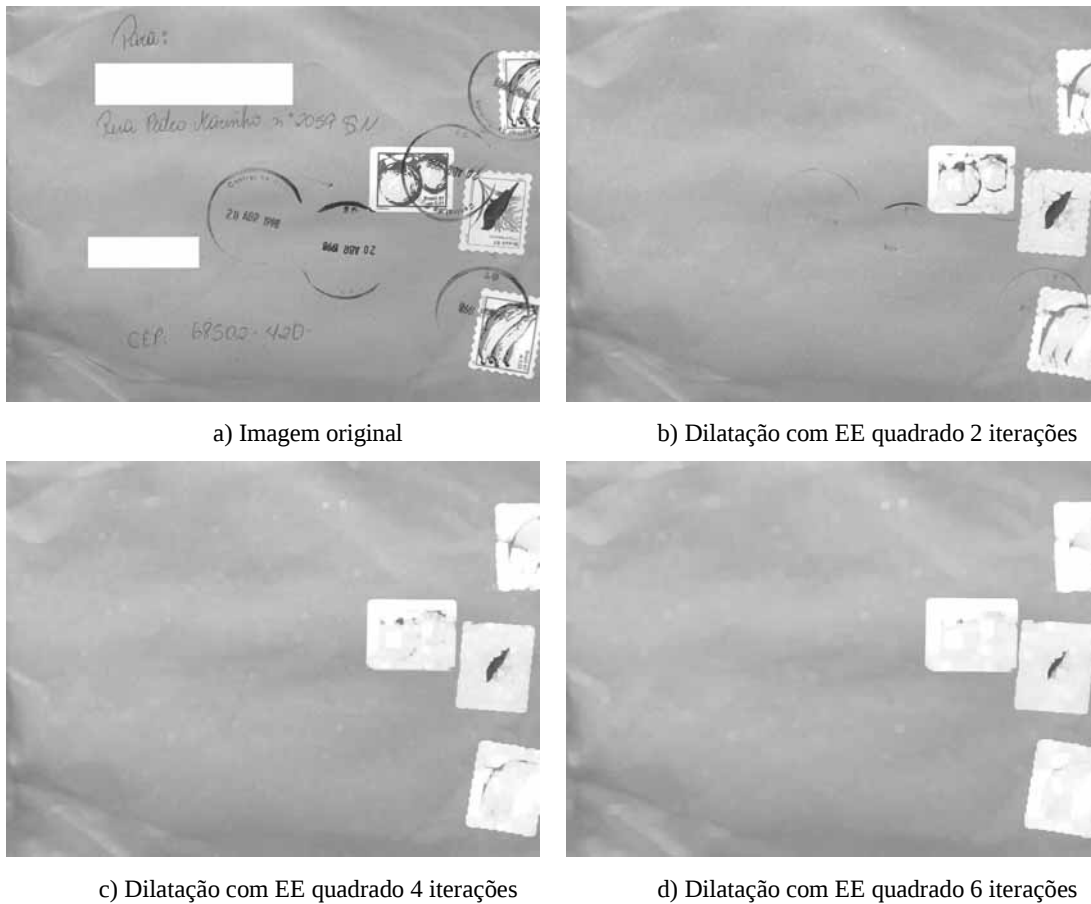


Figura A.2 Exemplo de Dilatação

Os efeitos da dilatação em níveis de cinza são:

- clarear a imagem;
- alargar e engordar os picos (padrões claros);
- conectar picos próximos;
- reduzir e às vezes eliminar vales (padrões escuros);
- separar vales.

A.1.3. Abertura

A abertura consiste em erodir um conjunto f por g e depois dilatar esse conjunto erodido pelo mesmo elemento estruturante. Como em morfologia binária, a operação morfológica de abertura em níveis de cinza é definida, inicialmente como:

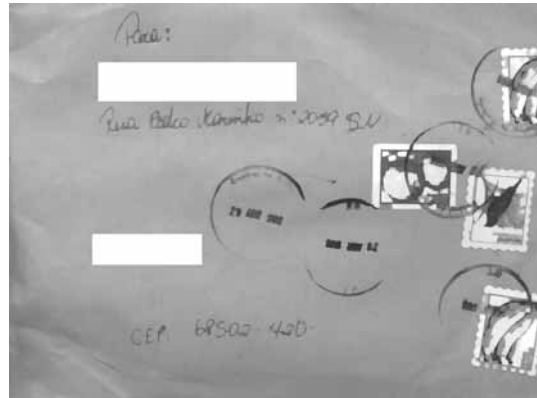
$$abe^g(f) = dil^g(ero^{\tilde{g}}(f)) \quad (A.3)$$

Seguindo a analogia da imagem em níveis de cinza com o relevo topográfico, o comportamento da abertura é:

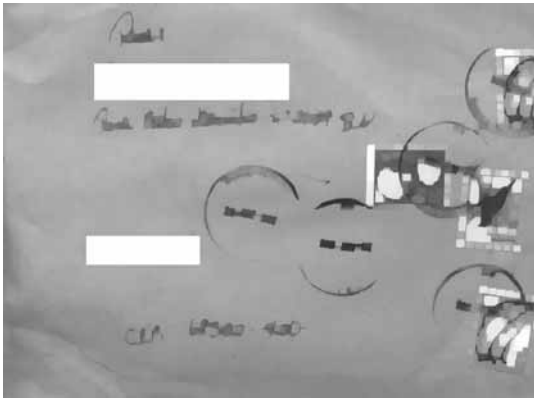
- separa picos próximos;
- elimina os picos inferiores em tamanho ao elemento estruturante;
- conserva vales afastados;
- emenda vales próximos;
- as entidades restantes após abertura ficam quase idênticas;
- a imagem aberta é mais regular que a imagem original;
- a imagem aberta é menos rica em detalhes que a imagem original.



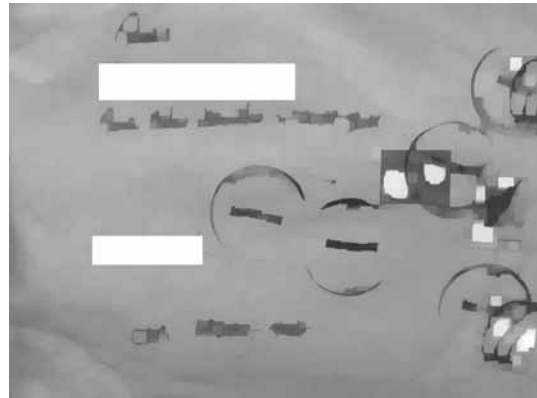
a) Imagem original



b) Abertura com EE quadrado 2 iterações



c) Abertura com EE quadrado 4 iterações



d) Abertura com EE quadrado 6 iterações

Figura A.3 Exemplo de Abertura

A.1.4. Fechamento

O fechamento de um sinal f por um elemento estruturante g é:

$$fec^g(f) = ero^g(dil^{\tilde{g}}(f)) \quad (A.4)$$

Explorando a propriedade de dualidade entre o fechamento e a abertura em níveis de cinza, podemos escrever:

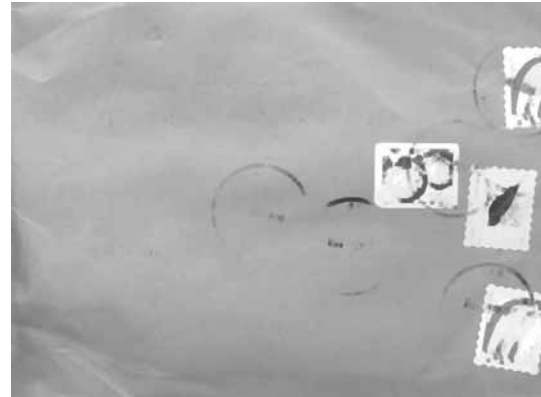
$$fec^g(f) = -abe^g(-f) \quad (A.5)$$

Seguindo a analogia da imagem em níveis de cinza como o relevo topográfico, o comportamento do fechamento em níveis de cinza:

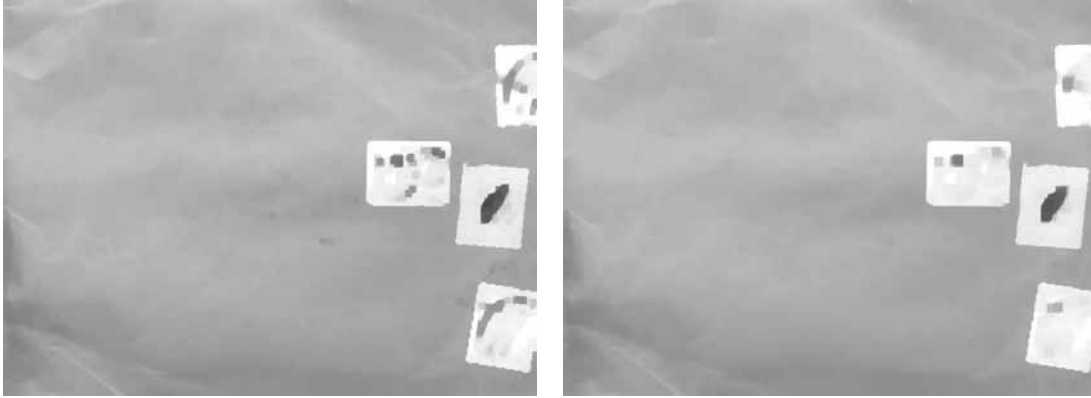
- separa vales próximos;
- elimina vales inferiores em tamanho ao elemento estruturante;
- conserva picos afastados;
- emenda picos próximos;
- as entidades restantes após o fechamento ficam quase idênticas;
- a imagem fechada é mais regular que a imagem original;
- a imagem fechada é menos rica em detalhes que a imagem original.



a) Imagem Original



b) Fechamento com EE quadrado 2 iterações



c) Fechamento com EE quadrado 4 iterações

d) Fechamento com EE quadrado 6 iterações

Figura A.4 Exemplo de Fechamento

A.1.5. TopHat por Abertura

A técnica de tophat por abertura é utilizada para a detecção de picos em uma imagem. O tophat por abertura $tophat_{abe}^g$ de um sinal f pelo elemento estruturante g é:

$$tophat_{abe}^g(f) = f - abe^g(f) \quad (A.6)$$

Como a abertura é um processo anti- extensivo, o seu resultado fica abaixo do sinal original, exemplificado na Figura A.5.

Com o uso de um elemento estruturante adequado, o processo de abertura permite a eliminação dos picos. Fazer a diferença, entre os sinais original e resultado da abertura, permite eliminar a heterogeneidade do fundo e consequentemente ressaltar a informação dos picos da imagem.

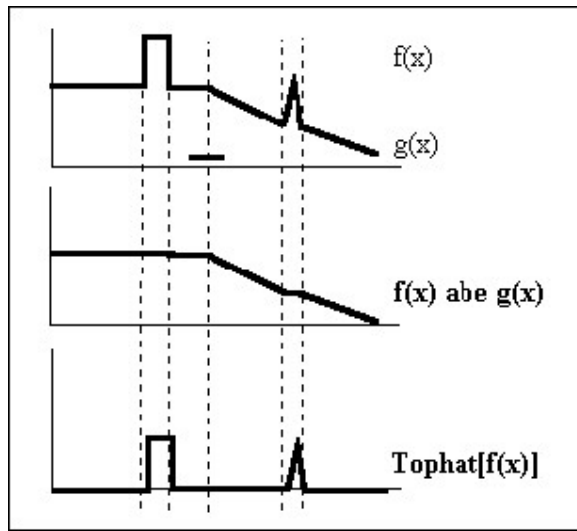
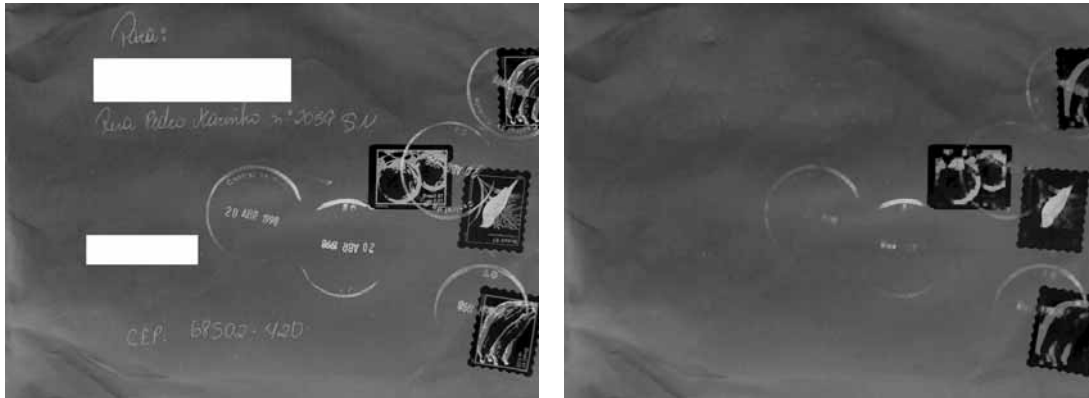
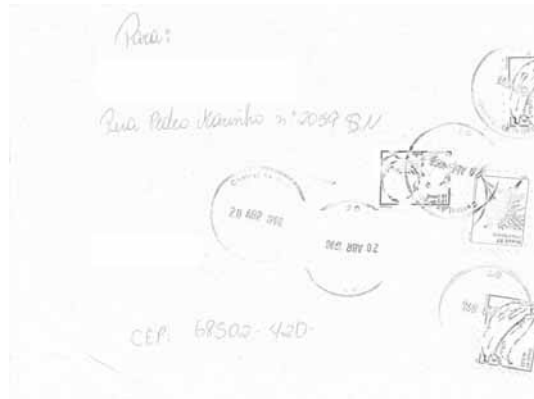


Figura A.5 Exemplo teórico da transformação Tophat por Abertura



a) Imagem original invertida

b) Abertura com EE quadrado 2 iterações



c) Tophat por Abertura invertido

Figura A.6 Exemplo de TopHat por Abertura

A.1.6. TopHat por Fechamento

De forma análoga, a técnica de tophat por fechamento é utilizada para a detecção dos vales em uma imagem. O tophat por fechamento $tophat_{fec}^g$ de um sinal f pelo elemento estruturante g é:

$$tophat_{fec}^g(f) = fec^g(f) - f \quad (A.7)$$

Como o fechamento é um processo extensivo, o seu resultado da transformação é sempre positivo, como vemos na Figura A.7.

Com o uso de um elemento estruturante adequado, o processo de fechamento permite a eliminação dos vales. Fazer a diferença, entre os sinais original e resultado fechado, permite eliminar a heterogeneidade e consequentemente ressaltar a informação dos vales da imagem.

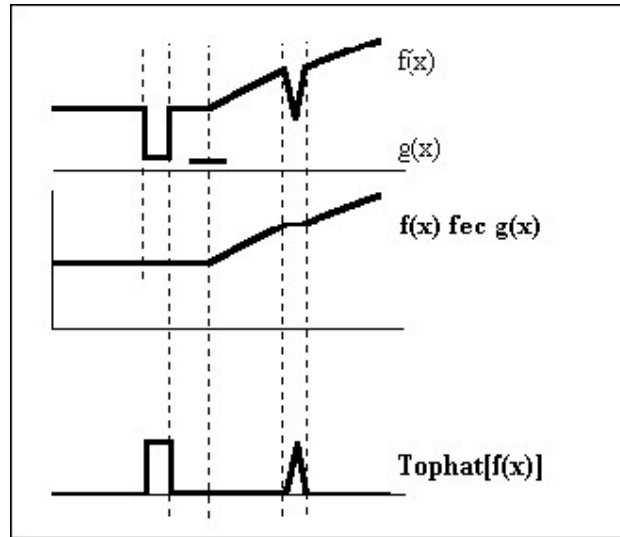


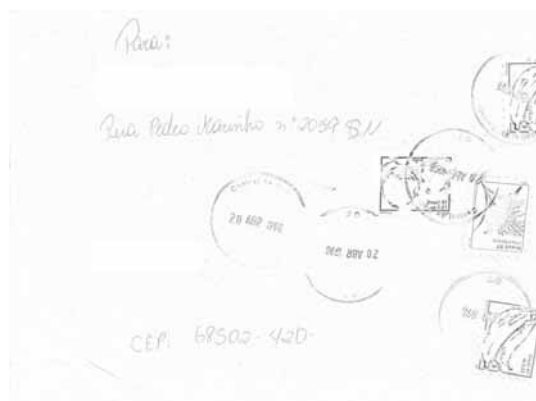
Figura A.7 Exemplo teórico da transformação Tophat por Fechamento



a) Imagem Original



b) Fechamento com EE quadrado 2 iterações



b) TopHat por Fechamento invertido

Figura A.8 Exemplo de TopHat por Fechamento

A.1.7. Gradiente Morfológico

A informação do gradiente é muito usada no processamento de imagens para detectar bordas. De forma geral, a informação gradiente é uma versão digital da informação diferencial. Supondo o sinal f diferenciável no seu domínio, pode-se mostrar [SCH67] que a derivada de f segundo a coordenada h na direção α na borda B é:

$$\frac{\partial f}{\partial h}(x) = \left(\frac{\partial f}{\partial h}\right)(x) + \sigma \cos \theta - \alpha \delta_x(B) \quad (\text{A.8})$$

onde a função $(\partial f / \partial h)$ representa a derivada do termo contínuo, σ o salto no ponto $x \in B$, o ângulo θ representa a reta perpendicular a B no ponto x orientada no sentido positivo do salto, e $\delta_x(B)$ é o pulso de Dirac no ponto x .

A definição do gradiente de f no ponto x é o vetor $(\partial f / \partial x)$, $(\partial f / \partial y)$ associado com a função diferencial df :

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \quad (\text{A.9})$$

A.1.8. Gradiente Morfológico por erosão em níveis de cinza

[BEU77] propôs uma avaliação do gradiente morfológico a partir de erosões e dilatações em níveis de cinza efetuadas com disco de raio unitário:

$$\text{Gradiente}(x) = \lim_{\lambda \rightarrow 0} \frac{(f \text{ dil } \lambda g) - (f \text{ ero } \lambda g)}{2\lambda} \quad (\text{A.10})$$

Na morfologia matemática, existem várias implementações digitais do gradiente e três delas serão aqui apresentadas.

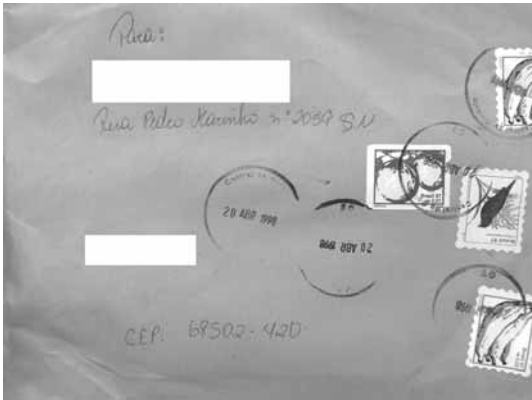
O primeiro tipo de gradiente morfológico $grad_{ero}^g$ de uma função $f(x)$ consiste em usar a operação de erosão em níveis de cinza.

O gradiente morfológico $grad_{ero}^g$ de uma função f por um elemento estruturante g a partir da erosão é:

$$grad_{ero}^g(f) = f - ero^g(f) \quad (\text{A.11})$$

Como nos processos anteriores, a obtenção de um resultado interessante depende da escolha do elemento estruturante.

O processo de Gradiente por erosão em níveis de cinza detecta bordas nas posições dos níveis de cinza mais elevados das bordas.



a) Imagem Original



b) Gradiente por erosão com EE quadrado invertido

Figura A.9 Exemplo de Gradiente por erosão

A.1.9. Gradiente Morfológico por dilatação em níveis de cinza

O segundo tipo de gradiente morfológico $grad_{dil}$ de uma função $f(x)$ consiste em usar a operação de dilatação em níveis de cinza.

O gradiente morfológico $grad_{dil}^g$ de uma função f pelo elemento estruturante g a partir da dilatação é:

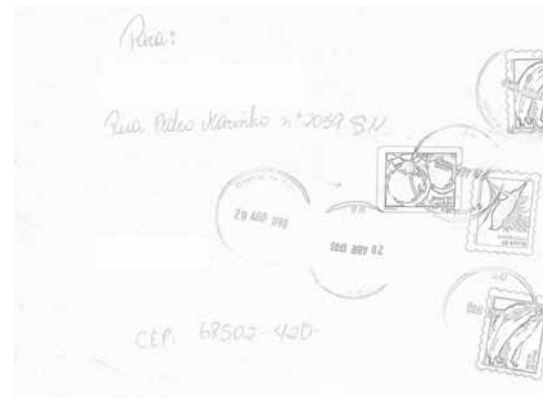
$$grad_{dil}^g(f) = dil^g(f) - f \quad (A.12)$$

Novamente, a obtenção de um resultado interessante depende da escolha do elemento estruturante.

O Gradiente por dilatação em níveis de cinza detecta bordas nas posições dos níveis de cinza mais baixos das bordas.



a) Imagem Original



b) Gradiente por dilatação com EE quadrado invertido

Figura A.10 Exemplo de Gradiente por dilatação

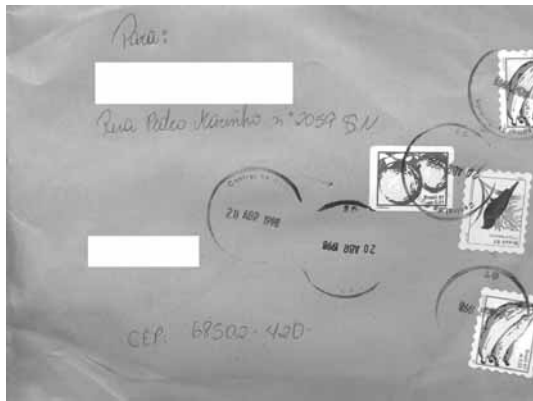
A.1.10. Gradiente Morfológico por dilatação e erosão em níveis de cinza

O terceiro tipo de gradiente morfológico $grad_{erodil}^g$ de uma função $f(x)$ consiste em usar simultaneamente as operações de erosão e dilatação em níveis de cinza.

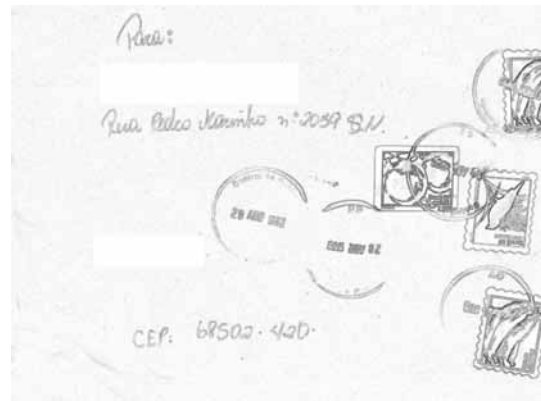
O gradiente morfológico $grad_{erodil}^g$ de uma função f pelo elemento estruturante g a partir da dilatação e da erosão é:

$$grad_{erodil}^g(f) = dil^g(f) - ero(f) \quad (A.13)$$

O gradiente por dilatação e erosão em níveis de cinza agrupa os resultados dos gradientes por erosão em níveis de cinza e por dilatação em níveis de cinza. Portanto, pode-se observar que os operadores de gradiente $grad_{ero}^g$ e $grad_{dil}^g$ fornecem contornos mais finos que o operador de gradiente $grad_{erodil}^g$.



a) Imagem Original



b) Gradiente por dilatação-erosão com EE quadrado invertido

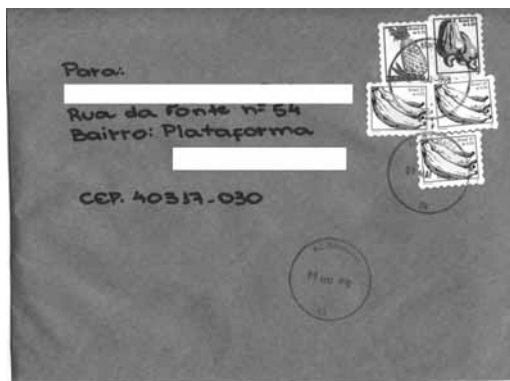
Figura A.11 Exemplo de Gradiente por dilatação-erosão

Apêndice B

Resultados Adicionais

Neste apêndice serão apresentados alguns resultados adicionais com as abordagens propostas por matriz de co-ocorrência e histograma 2D. Os resultados são complementares aos apresentados no Capítulo 6.

B.1. Resultados Adicionais por Matriz de Co-Ocorrência (Junção horizontal e vertical)



a) Imagem original



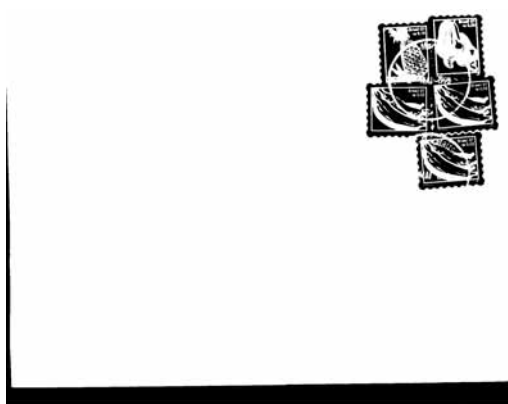
b) Imagem original invertida



c) Classe 1

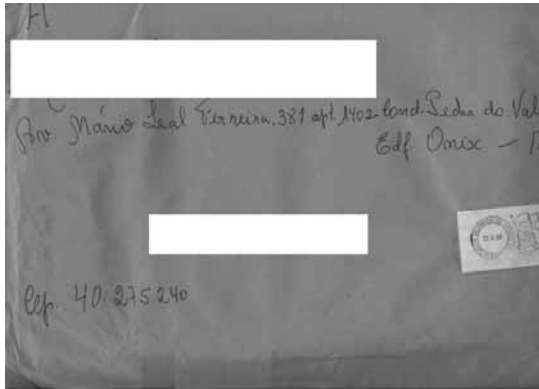


d) Classe 2

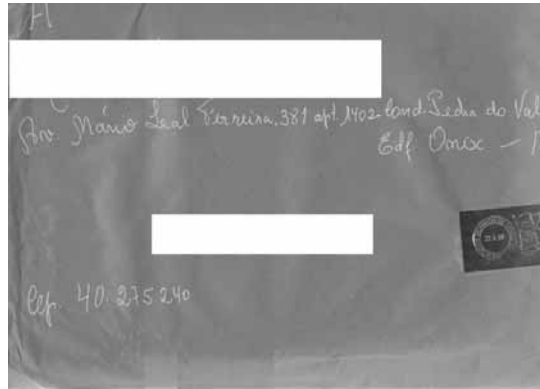


e) Classe 3

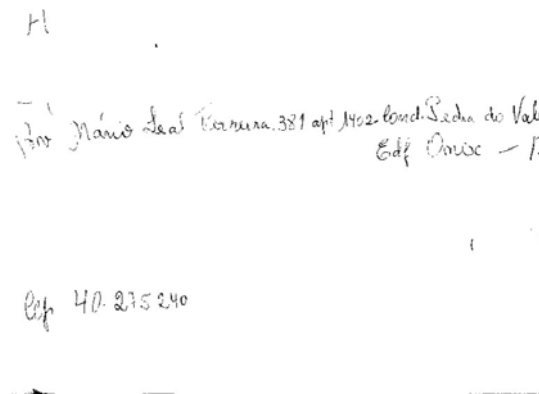
Figura B.1 Exemplo de segmentação boa, porém com a última classe contendo bordas adicionais geradas por defeitos de aquisição da imagem



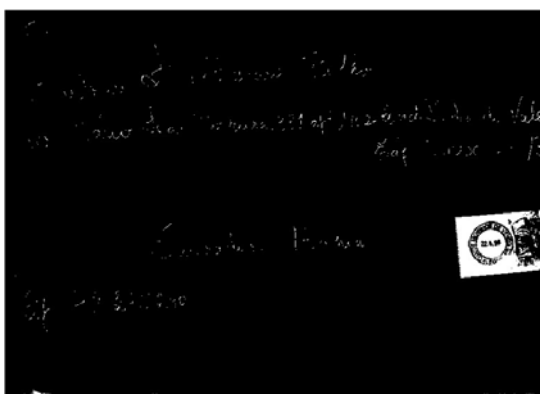
a) Imagem Original



b) Imagem Original invertida



c) Classe 1

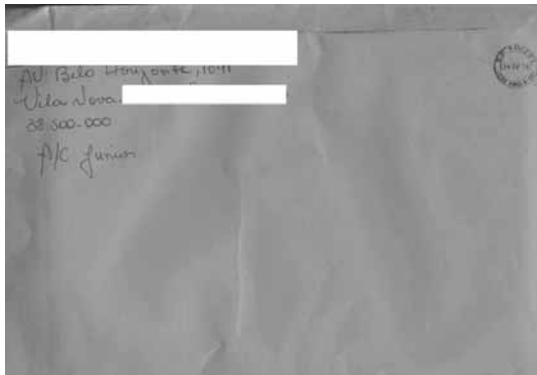


d) Classe 2

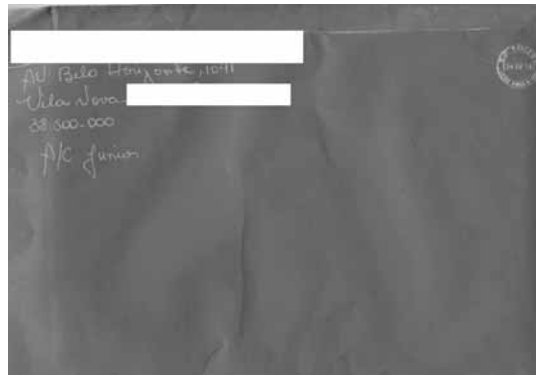


e) Classe 3

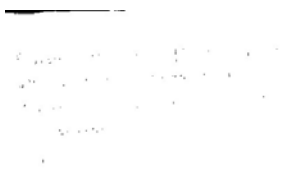
Figura B.2 Exemplo de segmentação média



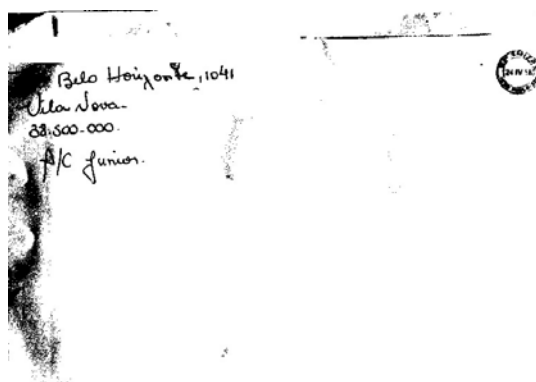
a) Imagem original



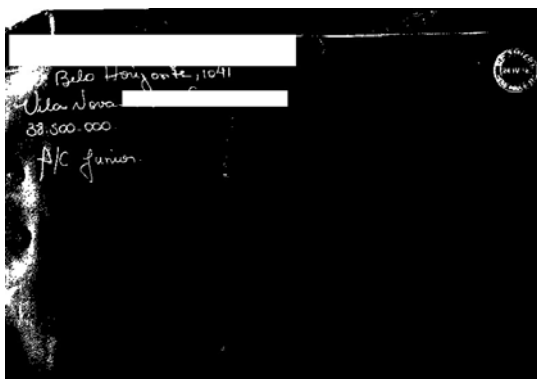
b) Imagem original invertida



c) Classe 1



d) Classe 2



e) Classe 3

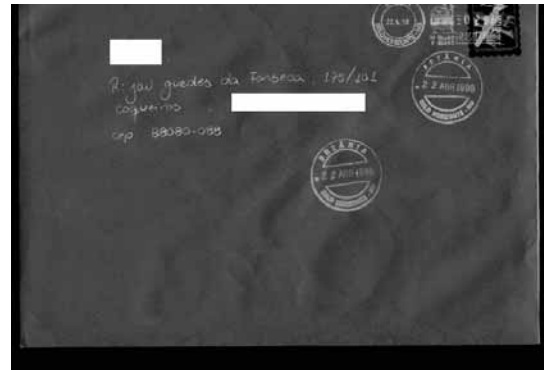
f) Classe 4

Figura B.3 Exemplo de segmentação ruim onde as classes foram superiores a 3

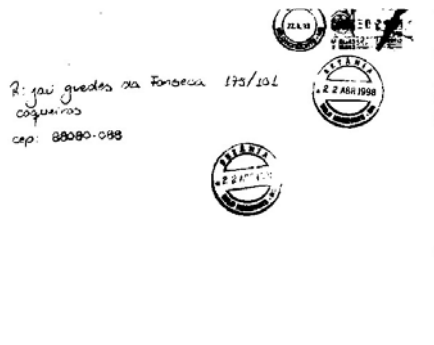
B.2. Resultados Adicionais por Matriz de Co-Ocorrência (Junção 4 direções)



a) Imagem original



b) Imagem original invertida



c) Classe 1



d) Classe 2



e) Classe 3

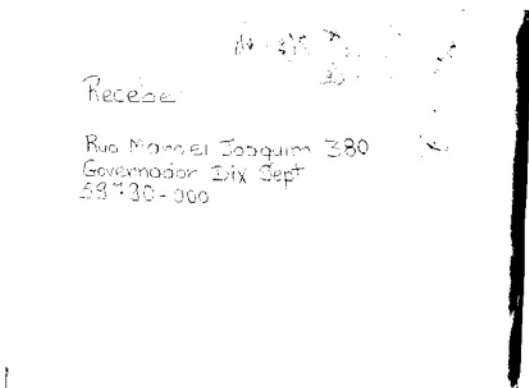
Figura B.4 Exemplo de segmentação boa, porém com a última classe contendo bordas adicionais geradas por defeitos de aquisição da imagem



a) Imagem original



b) Imagem original invertida



c) Classe 1



d) Classe 2



e) Classe 3

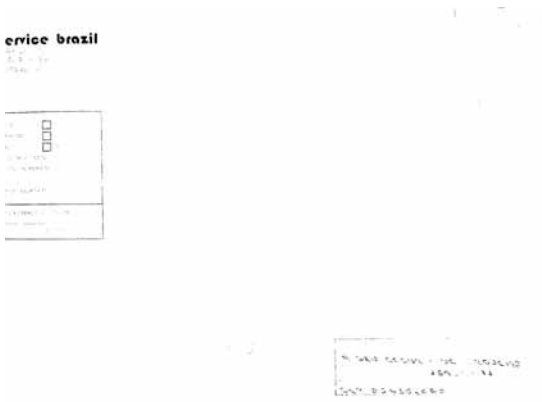
Figura B.5 Exemplo de segmentação média



a) Imagem original



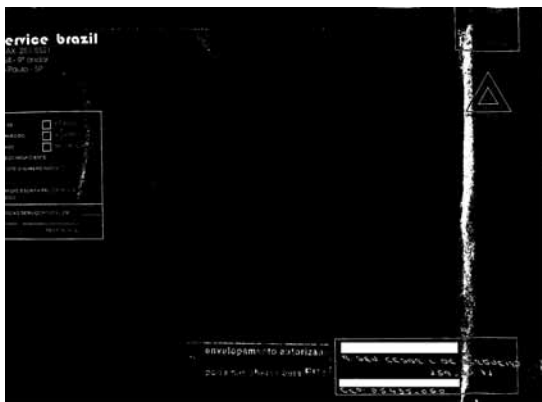
b) Imagem original invertida



c) Classe 1



d) Classe 2



d) Classe 3



e) Classe 4

Figura B.6 Exemplo de segmentação ruim onde as classes foram superiores a 3

B.3. Resultados Adicionais por Histograma 2D

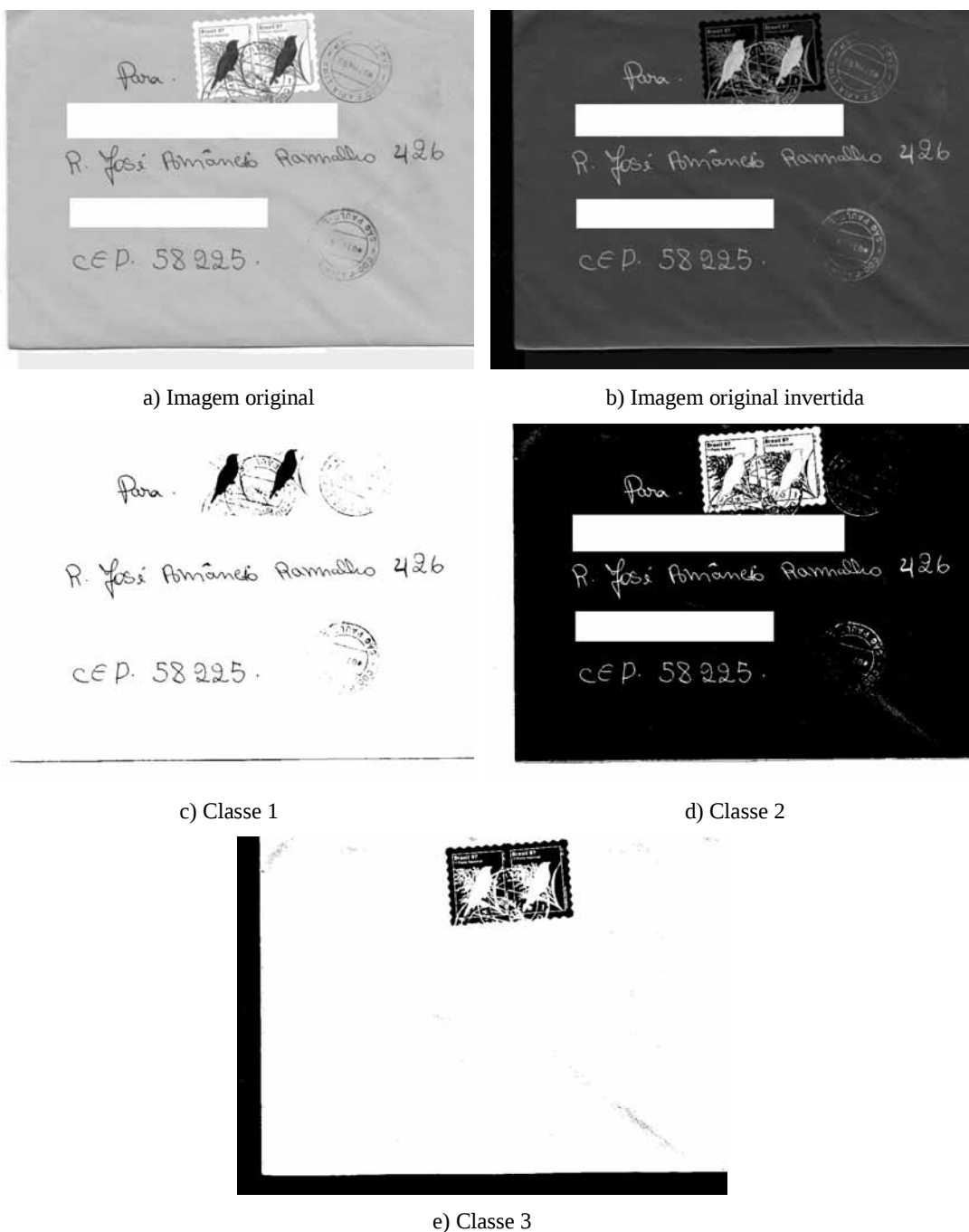


Figura B.7 Exemplo de segmentação boa, porém com a última classe contendo bordas adicionais geradas por defeitos de aquisição da imagem



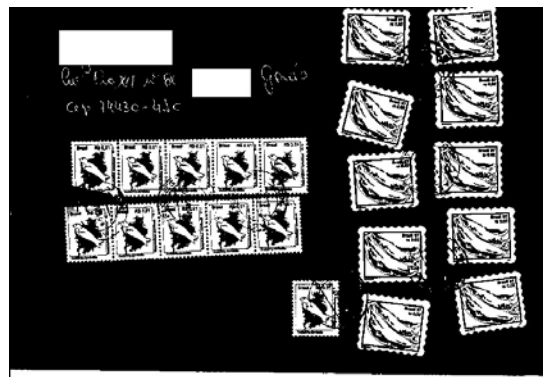
a) Imagem original



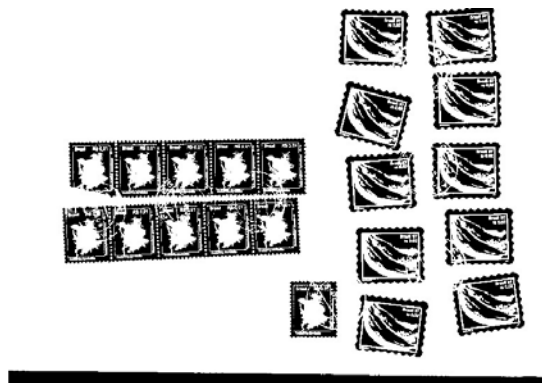
b) Imagem original invertida



c) Classe 1

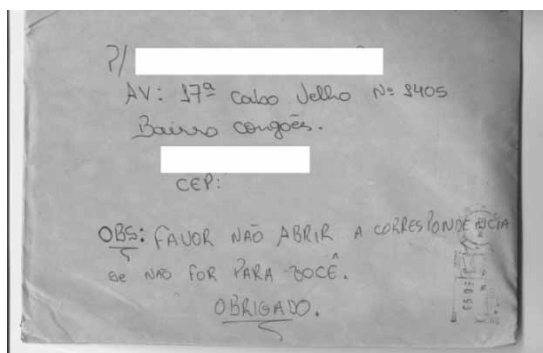


d) Classe 2

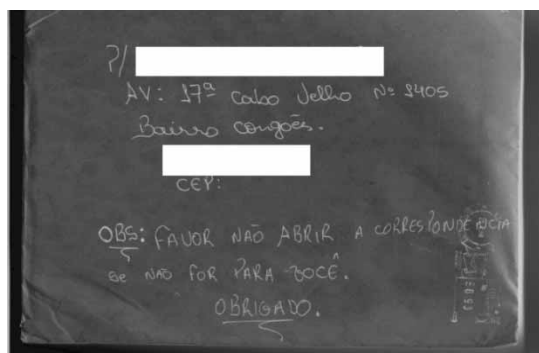


e) Classe 3

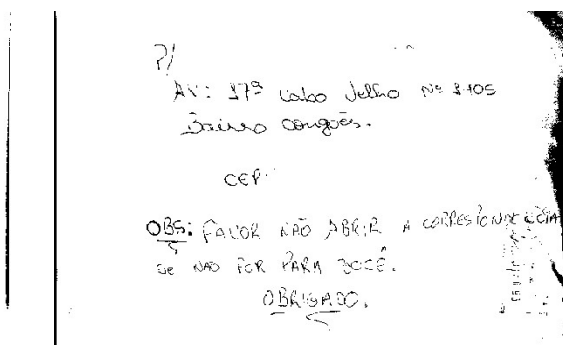
Figura B.8 Exemplo de segmentação média



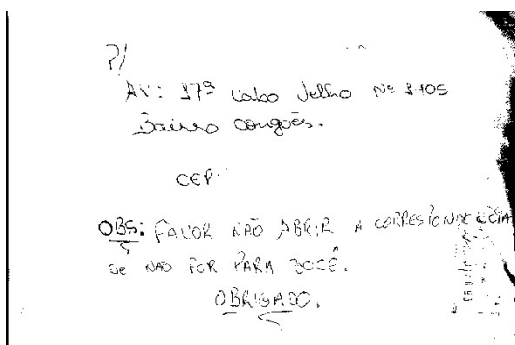
a) Imagem original



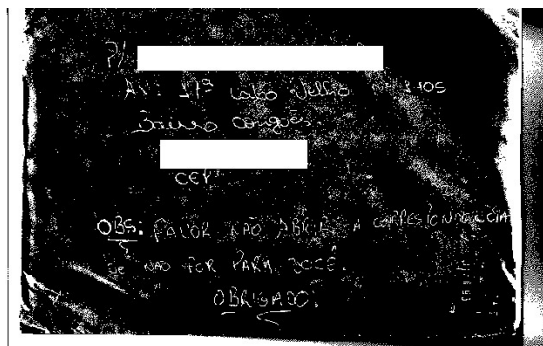
b) Imagem original invertida



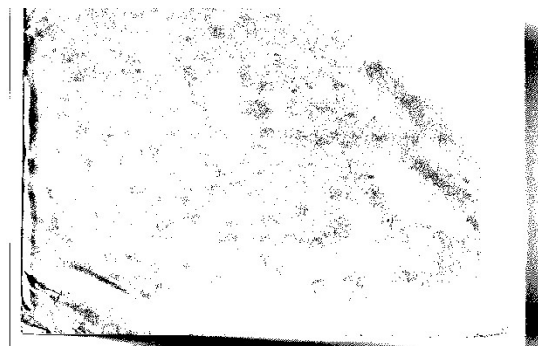
c) Classe 1



d) Classe 2



e) Classe 3



f) Classe 4



g) Classe 5

Figura B.9 Exemplo de segmentação ruim onde as classes foram superiores a 3