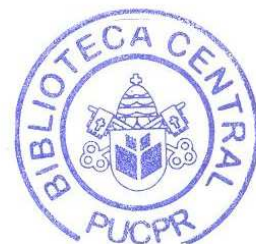


PAULO MARTINELLI



LOCALIZAÇÃO DO BLOCO ENDEREÇO EM IMAGENS DE ENVELOPES POSTAIS

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: *Metodologia e Técnicas de Computação.*

Orientador: Prof. Dr. Flávio Bortolozzi.

Co-orientador: Prof. Dr. Alceu de Souza Britto Jr

CURITIBA

2002



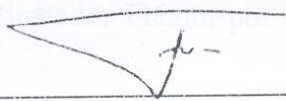


ATA DA SESSÃO PÚBLICA DE DEFESA DE DISSERTAÇÃO DE MESTRADO
DO PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA APLICADA
DA PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ

DEFESA DE DISSERTAÇÃO Nº 054

Aos 25 dias do mês de março de 2002 realizou-se a sessão pública de defesa da dissertação “Localização do Bloco Endereço em Imagens de Envelopes Postais”, apresentada por Paulo Martinelli como requisito parcial para a obtenção do título de Mestre em Informática Aplicada, perante uma Banca Examinadora composta pelos seguintes membros:

Prof. Dr. Flávio Bortolozzi
PUCPR (Presidente)




APROVADO
parecer (aprov/ reprov.)

Prof. Dr. Alceu de Souza Brito Jr.
PUCPR



APROVADO

Prof. Dr. Gerson Linck Bichinho
PUCPR



APROVADO

Prof. Dr. Guilherme Vilar
UNICAP



APROVADO

Conforme as normas regimentais do PPGIA e da PUCPR, o trabalho apresentado foi considerado APROVADO (aprovado/reprovado), segundo avaliação da maioria dos membros desta Banca Examinadora. Este resultado está condicionado ao cumprimento integral das solicitações da Banca Examinadora, conforme registrado no Livro de Defesas do programa.

Prof. Dr. Carlos Maziero
Diretor do PPGIA PUCPR

25/06/02 

Data e assinatura, após homologação da defesa pelo Colegiado.

Agradecimentos

Quero esternar meus agradecimentos a todos os que de forma direta ou indireta contribuíram para chegar a termo deste trabalho.

Em especial, agradeço:

A Deus pela vida, saúde e coragem concedidas.

À minha esposa Liliam, pela força, carinho e grandes idéias nas horas mais difíceis.

Aos meus filhos Lyane e Paulo Cesar, pelo carinho dispensados e compreensão na minha ausência.

Ao meu grande amigo e orientador Flávio, por ter acreditado em mim e dispensado especial atenção.

Ao meu co-orientador Alceu pela coragem de lutar contra o tempo.

Ao amigo Élcio pela grande força e encorajamento para enfrentar a realidade do mestrado.

À Madian e Ana Carolina pelas ajudas nas traduções.

Ao Marlon pela grande ajuda na implementação do software.

Aos colegas que não mediram esforços em dispensar colaborações.

Escolho sempre assuntos acima das minhas forças.

F. DOSTOIEVSKI

Sumário

Agradecimentos	iii
Sumário	iv
Lista de Figuras	vii
Lista de Tabelas	ix
Lista de Siglas	x
Resumo	xi
Abstract	Xii
Capítulo 1	
Introdução	1
1.1. Proposta.. .. .	4
1.2. Motivação.....	6
1.3. Objetivos.....	7
1.4. Estrutura do Trabalho.....	7
Capítulo 2	
Estado da Arte	9
2.1. Histórico do Correio.....	9
2.2. Correio no Brasil.....	10
2.3. A Tecnologia a Serviço dos Correios.....	11
2.3.1 Automação.....	12
2.3.2 Modernização das Empresas de Correios e Telégrafos Brasileiros...	12
2.4. Metodologia Utilizada pelo ECT	13
2.5. Sistema de Processamento e Análise de Imagens.....	14
2.6. Análise de Layout Imagens.....	15
2.7. Segmentação de Imagens.....	16

Capítulo 3

Fundamentação Teórica

22

3.1. Segmentação por Limiarização.....	22
3.2. Método de Otsu.....	23
3.3. Projeções Horizontais e Verticais.....	24
3.4. Método RLSA.....	25
3.5. Rotulação de Blocos.....	25
3.6. Rotulação Funcional.....	26
3.7. Tophat Diferencial.....	27

Capítulo 4

Morfologia Matemática

28

4.1. Operadores Morfológicos Binários.....	30
4.1.1. Erosão Binária.....	30
4.1.2. Dilatação Binária.....	32
4.1.3. Abertura.....	33
4.1.4. Fechamento.....	35
4.1.5. Reconstrução.....	38
4.2. Operadores Morfológicos em Níveis de Cinza.....	38
4.2.1. Dilatação.....	38
4.2.2. Erosão.....	39
4.2.3. Abertura.....	40
4.2.4. Fechamento.....	41
4.2.5. Tophat por Abertura.....	42
4.2.6. Tophat por Fechamento.....	43

Capítulo 5

Método Proposto

45

5.1. Visão geral.....	45
5.2. Desenvolvimento.....	47
5.2.1. Pré-processamento.....	47
5.2.2. Segmentação dos Blocos de Componentes.....	52

5.2.3. Detecção dos Componentes Conexos.....	52
5.3. Imagem Final.....	53
5.3.1. Reconstrução Dual.....	54
Capítulo 6	
Resultados Experimentais	56
6.1. Base de Imagens Postais da PUCPR.....	56
6.1.1. Descrição da Base de Imagens da PUCPR.....	56
6.1.2. Definição da Base de Imagens.....	60
6.1.3. Base de Testes	60
6.2. Resultados Gerais.....	60
6.3. Análise dos Resultados.....	62
6.3.1. Detecção de Bordas.....	62
6.3.2. Tophat.....	65
Capítulo 7	
Conclusão.	69
7.1. Considerações Finais.....	69
7.2. Conclusões.....	69
7.3. Recomendações e Sugestões.....	70
Referências Bibliográficas	71

Lista de Figuras

Figura 1.1	Tipos de Envelopes Postais	05
Figura 1.2	Seqüência do Processo	06
Figura 4.1	Tipos de elementos Estruturantes	28
Figura 4.2	Exemplos de Iteração do Elemento Estruturante	29
Figura 4.3	Exemplo de Iteração entre o Elemento Estruturante e a Imagem	30
Figura 4.4	Exemplo do Deslocamento dos Pixels na Erosão.....	31
Figura 4.5	Exemplo do Deslocamento dos Pixels na Dilatação.....	32
Figura 4.6	Resultado da Erosão e Dilatação.....	33
Figura 4.7	Exemplo da Imagem X Conectada.....	34
Figura 4.8	Intersecção da Imagem X com X.....	34
Figura 4.9	Exemplo da Imagem X com X.....	35
Figura 4.10	Exemplo da Imagem X Desconectada	36
Figura 4.11	Exemplo da Adição de X com X	36
Figura 4.12	Exemplo da Subtração e Intersecção de X com X	37
Figura 4.13	Resultado da Abertura e Fechamento.....	37
Figura 4.14	Exemplo de Dilatação	39
Figura 4.15	Exemplo de Erosão	40
Figura 4.16	Exemplo de Abertura	41
Figura 4.17	Exemplo de Fechamento	42
Figura 4.18	Diferença entre a Imagem Original e o Tophat Abertura	42
Figura 4.19	Diferença entre a Imagem Original e o Tophat Fechamento	43
Figura 5.1	Etapas do Processamento	46
Figura 5.2	Imagem Original	47
Figura 5.3	Detecção de Bordas	47
Figura 5.4	Bordas Erodidas	48
Figura 5.5	Limiarização por Otsu	48
Figura 5.6	Fechamento Horizontal	49

Figura 5.7	Abertura Vertical	49
Figura 5.8	Imagem Original	49
Figura 5.9	Tophat por Fechamento	50
Figura 5.10	Tophat Fechamento Invertido	50
Figura 5.11	Limiarização por Otsu	51
Figura 5.12	Dilatação Quadrada	51
Figura 5.13	Fechamento Horizontal	51
Figura 5.14	Abertura Vertical	52
Figura 5.15	Imagem com Blocos Localizados	53
Figura 5.16	Imagem Reconstruída	54
Figura 6.1	Imagens Tipo FN 256 NC	57
Figura 6.2	Imagens Tipo SE 256 NC	58
Figura 6.3	Imagens Tipo FN 16 NC	59
Figura 6.4	Imagens Tipo SE 16 NC	59
Figura 6.5	Imagens do Critério Muito Bom por Detecção de Bordas	63
Figura 6.6	Imagens do Critério Bom por Detecção de Bordas	63
Figura 6.7	Imagens do Critério Regular por Detecção de Bordas	64
Figura 6.8	Imagens do Critério Ruim por Detecção de Bordas	65
Figura 6.9	Imagens do Critério Muito Bom por Tophat	65
Figura 6.10	Imagens do Critério Bom por Tophat	66
Figura 6.11	Imagens do Critério Regular por Tophat	67
Figura 6.12	Imagens do Critério Ruim por Tophat	67

Lista de Tabelas

Tabela 6.1	Características da Base FN 256 Níveis de Cinza	57
Tabela 6.2	Características da Base SE 256 Níveis de Cinza	58
Tabela 6.3	Características da Base FN 16 Tonalidades de Cinza	58
Tabela 6.4	Características da Base SE 16 Tonalidades de Cinza	59
Tabela 6.5	Resumo Estatístico	61
Tabela 6.6	Resumo acumulado dos critérios Muito Bom e Bom	61

Lista de Abreviaturas

BMP	Bitmap – Mapa de Bits
BROPI	Biblioteca de Rotinas de Processamento de Imagens
CENPARMI	Center of Pattern Analysis and Machine Intelligence
CEP	Código de Endereçamento Postal
CNPq	Centro Nacional de Pesquisa
ECT	Empresa de Correios e Telégrafos
FN	Formato Normal
PAEP	Processamento Automático de Endereços Postais
PUCPR	Pontifícia Universidade Católica do Paraná
SE	Semi-embaraçoso
SRTP	Service de Recherche Technique de La Poste
OCR	Optical Character Recognition
RLSA	Run Length Smoothing Algorithm
DPI	Dots Per Inch

Resumo

Este trabalho apresenta a proposta de um método para a localização do bloco-endereço em imagens de envelopes postais manuscrito. As imagens destes envelopes não seguem padrões universais de preenchimento quanto ao tamanho, forma e tipos de caracteres, e o estudo desenvolvido nesta dissertação viabiliza automatizar parte do processo de reconhecimento estas informações. Esta pesquisa é parte integrante do projeto PAEP (CNPq e PUCPR), que tem por finalidade a localização e o reconhecimento do bloco-endereço de imagens postais.

O método tem como entrada uma imagem de envelope postal, em tons de cinza, que após pré-processada, com o objetivo de realçar a informação de interesse, é binarizada. Finalmente, a imagem binária do envelope tem seus componentes conexos detectados e submetidos a um processo de filtragem, onde são eliminados possíveis ruídos (selos, carimbos, ...). Duas abordagens são avaliadas durante o pré-processamento: a) abordagem baseada na detecção de bordas; b) abordagem baseada em TopHat.

Na realização deste trabalho, foi necessário a criação de uma sub base composta de 120 imagens de envelopes postais brasileiros, com resolução de 200 dpi em tons de cinza. Para os testes de validação do método proposto, foram utilizadas 80 imagens postais, nas técnicas de detecção de bordas e tophat, as quais apresentaram resultados considerado satisfatório, sendo que, a detecção de bordas obteve um rendimento considerado como bom, e a técnica do tophat apresentou um pouco inferior para mesmas imagens, também considerado bom, como havia sido proposto anteriormente.

Resumo

Este trabalho apresenta a proposta de um método para a localização do bloco-endereço em imagens de envelopes postais manuscrito. As imagens destes envelopes não seguem padrões universais de preenchimento quanto ao tamanho, forma e tipos de caracteres, e o estudo desenvolvido nesta dissertação viabiliza automatizar parte do processo de reconhecimento estas informações. Esta pesquisa é parte integrante do projeto PAEP (CNPq e PUCPR), que tem por finalidade a localização e o reconhecimento do bloco-endereço de imagens postais.

O método tem como entrada uma imagem de envelope postal, em tons de cinza, que após pré-processada, com o objetivo de realçar a informação de interesse, é binarizada. Finalmente, a imagem binária do envelope tem seus componentes conexos detectados e submetidos a um processo de filtragem, onde são eliminados possíveis ruídos (selos, carimbos, ...). Duas abordagens são avaliadas durante o pré-processamento: a) abordagem baseada na detecção de bordas; b) abordagem baseada em TopHat.

Na realização deste trabalho, foi necessário a criação de uma sub base composta de 120 imagens de envelopes postais brasileiros, com resolução de 200 dpi em tons de cinza. Para os testes de validação do método proposto, foram utilizadas 80 imagens postais, nas técnicas de detecção de bordas e tophat, as quais apresentaram resultados considerado satisfatório, sendo que, a detecção de bordas obteve um rendimento considerado como bom, e a técnica do tophat apresentou um pouco inferior para mesmas imagens, também considerado bom, como havia sido proposto anteriormente.

Abstract

This work presents the proposal of a method to locate address-blocks of images on handwritten postal envelopes. The images on those envelopes do not follow the universal patterns of filling due to size, form and character types. The study developed in this dissertation makes possible to recognize those information as part of the process. This search is part of the PAEP project (CNPq and PUCPR), which main purpose is to locate e identify the address-block of postal images.

The method has an input an image of a postal envelope, in grey levels, which were binaryzied after processed, with the purpose of highlight the information. Finally, the envelope binary image has its connected components and submitted to a cleaning process, were possible dirty information (stamps, notes).

Two approaches are done during pre-processing time: a) approach based on border detection; b) approach based on TopHat.

During this work, it was necessary to create a sub base made of 120 images of Brazilian postal envelopes, with a resolution of 200 dpi in grey level. The tests to validate the proposed method, 80 postal images were used on detection tests of borders and TopHat, which presented satisfactory.

The border detections had was considered as a good result, but in the other hand the TopHat technique was a little bit lower for the same images, also considered as good ones, as it was proosed before.

Abstract

This work presents the proposal of a method to locate address-blocks of images on handwritten postal envelopes. The images on those envelopes do not follow the universal patterns of filling due to size, form and character types. The study developed in this dissertation makes possible to recognize those information as part of the process. This search is part of the PAEP project (CNPq and PUCPR), which main purpose is to locate e identify the address-block of postal images.

The method has an input an image of a postal envelope, in grey levels, which were binaryzied after processed, with the purpose of highlight the information. Finally, the envelope binary image has its connected components and submitted to a cleaning process, were possible dirty information (stamps, notes).

Two approaches are done during pre-processing time: a) approach based on border detection; b) approach based on TopHat.

During this work, it was necessary to create a sub base made of 120 images of Brazilian postal envelopes, with a resolution of 200 dpi in grey level. The tests to validate the proposed method, 80 postal images were used on detection tests of borders and TopHat, which presented satisfactory.

The border detections had was considered as a good result, but in the other hand the TopHat technique was a little bit lower for the same images, also considered as good ones, as it was proosed before.

Capítulo 1

Introdução

A análise e o reconhecimento de imagens de documentos tornaram-se um importante objeto de estudo, devido ao grande número de aplicações. Pesquisadores em diversos países têm buscado novas técnicas que permitam a leitura automática de cheques, endereços postais, formulários do censo, dentre outros. Contudo, na prática, desenvolver um sistema geral capaz de processar todos os tipos de documentos (*cartas, jornais, revistas, ofícios, etc.*) é quase impossível, uma vez que cada documento possui características próprias.

Desta forma, os sistemas normalmente são direcionados para um tipo específico de documento e visam à integração de técnicas gerais de processamento e análise de imagens no caso específico da análise automática de imagens de envelopes postais.

A localização e o reconhecimento dos endereços postais representa uma área de pesquisa complexa [Yacoubi, 1995]. Como o documento analisado pode ser impresso ou manuscrito, a localização automática das regiões de interesse no bloco-endereço é um grande desafio. A falta de regras universais na disposição das regiões de interesse no bloco-endereço dificulta ainda mais essa localização.

As complexidades no processo de interpretação de endereços manuscritos provêm da grande variabilidade da escrita, do grande tamanho de léxicos no momento do reconhecimento do CEP [Heute, 1994], nomes de cidade [Yacoubi, 1999] e nome de rua [Yacoubi, 1995], e enfim da exigência de taxa de erro inferior a 1% no caso de aplicações reais.

Atualmente, a leitura de endereços postais depende, essencialmente, do reconhecimento do CEP [Heute, 1994]. Poucos trabalhos de pesquisa abordam o problema do reconhecimento conjunto do CEP, nome de cidade e nome de rua, no intuito de aumentar as taxas de reconhecimento dos endereços e manter baixa a taxa de erro.

Nos últimos anos tem-se verificado aumento sensível nas áreas abrangidas por sistemas computacionais inteligentes. Dentre essas áreas destaca-se a gestão automática de documentos que implica captação, armazenamento, indexação, localização e reconhecimento de imagens digitalizadas de documentos. Muitos países têm investido nessa área. Exemplos típicos são França e Canadá, os quais, através dos laboratórios Service - SRTP e Center of Pattern Analysis and Machine Intelligence - CENPARMI, têm buscado soluções no que diz respeito ao processamento automático de envelopes postais, cheques bancários e outros formulários.

No Brasil este tipo de atividade está praticamente inexplorada. No que diz respeito ao reconhecimento de envelopes postais, mesmo após minuciosa pesquisa, não foi encontrado nenhuma documentação sobre os hábitos relativos à disposição das regiões de interesse na elaboração de endereços em envelopes, bem como, nenhum trabalho de reconhecimento automático de endereços postais brasileiros publicado.

O Projeto Integrado de Pesquisa, PAEP - Processamento Automático de Endereços Postais, registrado no CNPq sob número 520.324/96-0, iniciado em março de 2.001 e desenvolvido pelo grupo de pós-graduação do Laboratório Unificado de Ciência da Imagem - LUCI, do Programa de Pós-Graduação em Informática Aplicada da Pontifícia Universidade Católica do Paraná, veio de encontro à necessidade de lançar o País nessa importante área de pesquisa, uma vez que a análise e tratamento de documentos manuscritos e/ou impressos requer processos customizados. De maneira simples, uma aplicação desenvolvida para documentos escritos, por exemplo em inglês, francês ou russo, não pode ser utilizada para documentos escritos em português. Trata-se assim de uma área estrategicamente importante para o Brasil.

Tudo isso demonstra que um esforço para montar um grupo de excelência na área da triagem postal faz-se necessário. Na PUCPR, esse esforço já iniciou com a aquisição de uma base de dados dos Correios Brasileiros composta de 50.000 imagens de envelopes reais, com endereços impressos e manuscritos. Essa mesma base foi utilizada na implementação dos sistemas de tratamento automático de endereços postais da SIEMENS e ALCATEL. Na ocasião, um termo de compromisso foi assinado pelo representante da ECT e pelo Pró-Reitor de Pesquisa da PUCPR, estipulando as regras de manuseio dessa base pelo LUCI.

O presente trabalho integra o Projeto PAEP e prevê a pesquisa e desenvolvimento da parte de um sistema de **processamento e reconhecimento de endereços postais em**

envelopes fornecidos pela ECT do Brasil, usando os conceitos de Processamento de Imagens e de Reconhecimento de Padrões.

O objetivo geral do projeto PAEP consiste em consolidar e elevar a qualidade do grupo de pós-graduação LUCI, através de pesquisa na área de processamento e reconhecimento automático de documentos. Para esse fim, o projeto prevê o desenvolvimento de um sistema de processamento e reconhecimento de endereços postais em envelopes dos Correios Brasileiros. Os objetivos específicos do PAEP são:

Formar recursos humanos na área de análise e reconhecimento automático de documentos, tanto em nível de graduação quanto em nível de pós-graduação.

Elaborar uma metodologia para o tratamento automático de imagens reais de envelopes dos Correios Brasileiros.

Pesquisar, implementar e avaliar novas ferramentas para o tratamento inteligente e automático de envelopes dos Correios Brasileiros, as quais deverão permitir a execução da referida metodologia. O que se espera é desenvolver e integrar novos componentes de *softwares* especializados na localização automáticos das regiões de interesse nos blocos-endereços, bem como no reconhecimento de seus principais componentes: CEP, nome de rua e nome de cidade. **Neste contexto, o trabalho desenvolvido se propõe a colaborar na busca de ferramentas para a localização automática de blocos de endereços.**

Uma das metas do projeto PAEP é a proposta deste trabalho e visa a localização automática do bloco-endereço.

A localização do bloco-endereço depende muito da posição, qualidade e aspecto visual do envelope (*textura, cor etc.*). Antes disso, deverão ser desenvolvidos algoritmos de filtragem adaptados a envelopes dos Correios Brasileiros e aos aspectos de textura e de cor que permitam livrar-se de problemas de inclinação, rabiscos, dados subscritos, dados sublinhados, etc. Na literatura encontram-se muitas técnicas de detecção de inclinação que serão estudadas e testadas para essa realidade. A morfologia matemática que já demonstrou ser uma abordagem eficiente em tipos mais diversos de imagens, será pesquisada e testada no caso de envelopes. Essa abordagem será mesclada a técnicas de binarização para verificar a possibilidade de desenvolver processos mais rápidos.

De um modo geral, a dificuldade na localização automática do bloco-endereço em um contexto multiescritor, deve-se a não utilização de regras no momento de escrever o endereço postal, tendo como consequência que as informações pertinentes (*CEP, nome de cidade, nome de rua*) não estão colocadas sempre em uma mesma ordem no documento (*envelope*).

1.1. Proposta

Para a proposta deste trabalho, inicialmente será feito um pequeno histórico dos correios.

Na medida em que a humanidade foi evoluindo, a comunicação tornou-se uma necessidade crescente entre os mais variados grupos, sendo que de um lado ela se aperfeiçoava para facilitar o contato entre as pessoas, e do outro caminhava para um processo cada vez mais complexo, obrigando o homem a desenvolver, através da sua inventividade, sistemas de comunicação mais eficientes para concretizar e veicular suas idéias.

E foi assim que a humanidade prosseguiu sua trajetória evolutiva no caminho da comunicação e na busca de meios para realizá-la, cujo histórico se perde nas brumas do tempo desde os desenhos rupestres, antes mesmo do envio de mensagens por tambores ou fumaça. Esse processo desenvolveu-se durante a Idade Antiga com os caracteres cuneiformes em pranchetas de terracota da Mesopotâmia, os ideogramas nos papiros do Nilo, e nos pergaminhos (de Pérgamo) da Ásia Menor.

Com o fim da Idade Média e começo do Renascimento, acentua-se o desenvolvimento, incentivado pela efervescência do desenvolvimento científico da época, como por exemplo, o surgimento da imprensa de Guttemberg. Desse período até os dias atuais, vem ocorrendo um crescimento constante de comunicação, pois segundo [SANTOS 2000, p. 25], vive-se numa sociedade performatizada, digitalizada pela tecnociência, na qual jornal, rádio, televisão, telefone, *internet*, videoconferência, etc., codificam e manipulam conhecimento e informação, distribuindo-os instantaneamente a qualquer parte do planeta Terra e já dirigem seu olhar para o cosmos na comunicação com astronautas, satélites e sondas espaciais.

Convém ressaltar que um dos meios mais antigos de comunicação, a escrita através de carta (*basta lembrar as cartas de Platão, as epístolas de São Paulo Apóstolo, etc.*), apesar de todo o avanço tecnológico, continua sendo o mais utilizado pela maioria da população. Este instrumento não só resiste ao tempo mas também tem aumentada sua utilização, de tal forma que a organização das cidades e países requer um eficiente sistema de controle e transporte de correspondência, já que através dele é que acontecem as trocas de informações e a realização de negócios, seja de forma manuscrita, mecanizada ou eletrônica.

A vantagem de se manter o processo da forma manuscrita, objetivo deste trabalho, circunscreve-se hoje à simplicidade, à agilidade de acesso e à manipulação das informações

armazenadas de forma digital, sendo este, ao que parece, o principal motivo do crescimento substancial de interesse no campo de análise e reconhecimento de imagens de documentos.

Quanto à forma mecanizada ou eletrônica não há tanta preocupação com a imagem, a princípio, já apresenta o ordenamento das linhas do endereço postal, portanto, com possibilidade de menos ruídos.

Contudo, seja na primeira forma ou segunda, o objetivo é tornar o sistema de transporte postal mais eficiente.

É fato que, ainda, não há procedimentos que apresentem solução satisfatória para um problema muito freqüente na realidade brasileira: o preenchimento do envelope postal se dá fora dos padrões definidos pelos Correios, principalmente no endereçamento do destinatário, conforme pode-se observar na Figura 1.1.

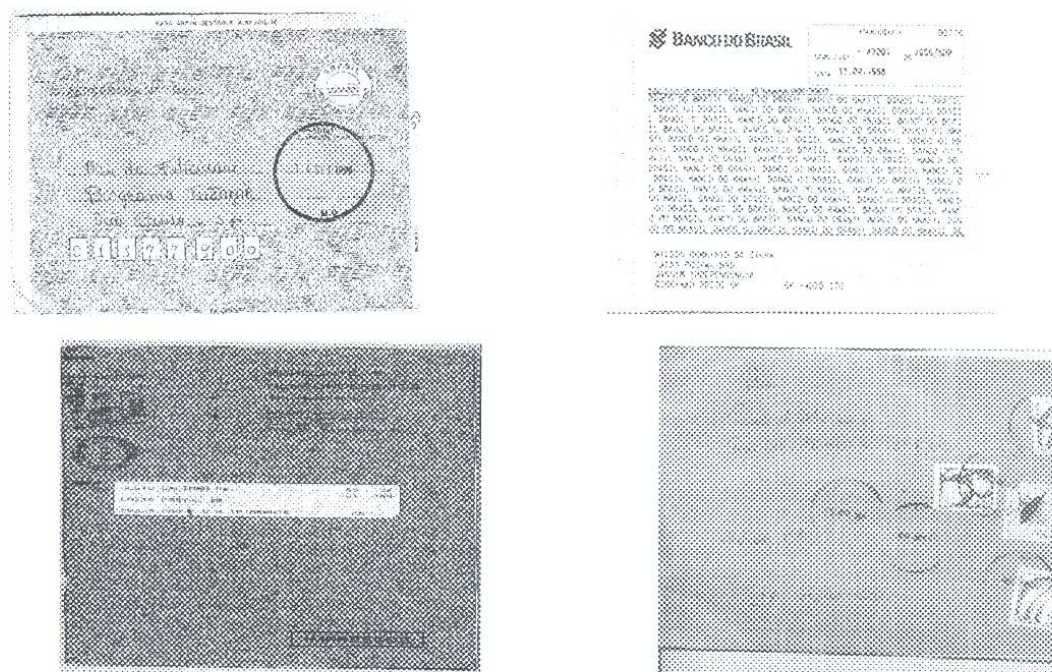


Figura 1.1: Tipos de envelopes postais da base de dados.

Supondo que fossem desenvolvidas campanhas educativas eficientes a respeito da importância da boa execução desta prática cotidiana, os resultados seriam percebidos em espaço de tempo muito longo, o que não é compatível com as necessidades advindas das transformações sócio-político-econômica-culturais a que a sociedade vem sendo submetida.

Diante desta problemática, **a proposta desta dissertação é a de pesquisar e propor um método para localização e extração do bloco-endereço em imagens de envelopes postais.** Na Figura 1.2 tem-se a seqüência do método proposto, onde é possível observar na

imagem original a presença de ruídos (*selos, carimbos, letras superpostas, carimbos superpostos em letras, entre outros*).

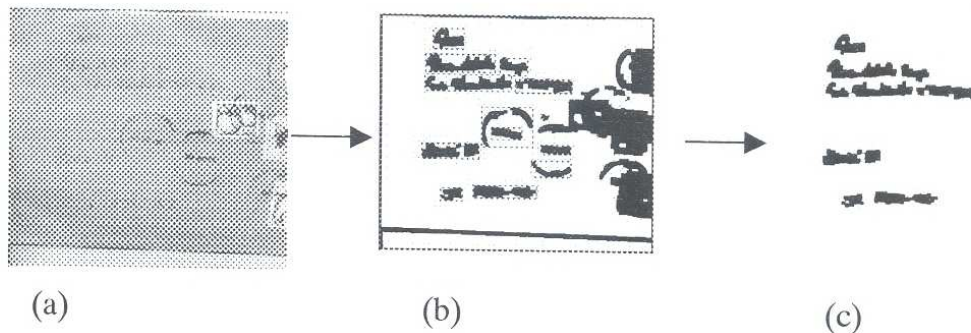


Figura 1.2: Seqüência do processo proposto. a) imagem original em níveis de cinza; b) imagem binarizada com localização dos blocos; c) imagem com extração dos blocos de endereço.

A localização eficiente dos blocos de endereços é indispensável na obtenção de uma imagem limpa destinada a um processamento posterior de reconhecimento do seu conteúdo, o qual já se considera indispensável à otimização das rotinas de transporte e distribuição de correspondências em envelopes postais, via Correios.

Com este propósito, pretende-se localizar os blocos de endereços através da utilização de técnicas de morfologia em níveis de cinza e binária. Tais imagens, quando segmentadas, identificariam as regiões onde estão localizados os caracteres das linhas do destinatário e endereço na imagem para, posteriormente, tratar cada uma delas em separado.

O desdobramento dessa pesquisa contribuirá para o reconhecimento de textos, através da identificação das regiões encontradas nas imagens digitais, submetidas a um processo de reconhecimento Optical Character Recognition - OCR.

1.2. Motivação

Tendo presente o caráter transitório da sociedade atual, denominada sociedade pós-industrial, gerado pelas mudanças sócio-político-econômico-culturais que vêm sendo percebidas a partir de 1960, destacando-se o desenvolvimento tecnológico e, em específico, da informática, a presente pesquisa traz, em si, o desafio de contribuir para a melhoria de um serviço indispensável à sociedade.

Um serviço postal é eficiente quando combina curto tempo de processo e um mínimo de erro na coleta, na classificação e na distribuição das correspondências. Sendo assim, a

triagem das correspondências e a agilidade da distribuição são processos vitais para o bom desempenho das organizações responsáveis por tais atividades. Neste sentido, os conhecimentos específicos da informática são de grande valia, pois os mesmos podem otimizar tais atividades, qualificando significativamente o atendimento social (*coleta, triagem, transporte, distribuição*).

Nesta perspectiva, a presente pesquisa deseja contribuir com o estudo e o desenvolvimento de algoritmos que possam compor um método capaz de agilizar a leitura de envelopes postais manuscritos. Tal contribuição será disponibilizada através do Projeto PAEP, o qual servirá de apoio para o desenvolvimento de algoritmos futuros que possibilitem o reconhecimento e a classificação dos diversos elementos do endereçamento postal.

Assim, pode-se perceber que a pesquisa não vem carregada somente de tecnologia pura e simples, mas busca atender também às questões sociais dentro do seu alcance.

1.3. Objetivos

O objetivo principal deste trabalho consiste na localização e extração do bloco-endereço em imagens de envelopes postais manuscritos dos Correios, através da segmentação de imagens em níveis cinza e binária, a fim de tornar sua leitura mais rápida e sua classificação mais eficiente.

Para que tal finalidade seja atingida se faz necessário:

- a) Eliminar ruídos e/ou componentes indesejáveis na imagem postal;
- b) Localização do bloco-endereço da imagem postal; e
- c) Desenvolver algoritmos que permitam a segmentação de blocos de endereços em imagens de envelopes postais, através do uso da morfologia em níveis de cinza e binária.

1.4. Estrutura do Trabalho

O presente trabalho compõe-se de 6 capítulos, dispostos na seguinte ordem: Introdução, Estado da Arte, Fundamentação Teórica, Morfologia Matemática, Método Proposto, Resultados Experimentais e Conclusão.

O Capítulo 1, Introdução, apresenta considerações iniciais acerca da problemática da automação postal, a proposta para resolver o problema da localização de endereços postais em

imagens de envelopes dos correios, os objetivos desta dissertação, a motivação e finalmente, a estrutura do trabalho.

No Capítulo 2, Estado da Arte, são tecidos comentários sobre a área de processamento e análise de imagens de documentos e algumas técnicas de segmentação de imagens, ou seja, como estão as pesquisas e experimentações a respeito de métodos e técnicas em níveis de cinza e binária empregada na caracterização de imagens de envelopes postais.

O Capítulo 3, Fundamentação Teórica, estão descritos os métodos e técnicas utilizadas na proposta, estabelecendo a base teórica para o processo, ou seja, criar um modelo para leitura das imagens, através de algoritmos de extração das características dos componentes da imagem.

No Capítulo 4, Morfologia Matemática, apresenta-se o conjunto de informações relacionadas à geometria e a topologia de um conjunto desconhecido. É, pois, a ferramenta que permite processar imagens com objetivos diferentes.

O Capítulo 5, Método Proposto, apresenta a metodologia e desenvolvimento da pesquisa através da descrição de cada fase das técnicas utilizadas, enfatizando a principal propriedade.

No Capítulo 6, Resultados Experimentais, é feito o estudo dos dados obtidos na experimentação realizada, apresentados estatisticamente, e demonstrando a eficiência ou não do método escolhido.

Finalmente no Capítulo 7, Conclusão, com base nos resultados obtidos, são apresentadas as possíveis conclusões juntamente com algumas recomendações e considerações de forma objetiva, acrescida de sugestões de novas pesquisas.

Partindo da visão geral da presente dissertação já se tornam perceptíveis os desafios que a mesma impõe. Na busca da superação dos mesmos, o primeiro passo é situar-se com relação ao conhecimento já construído pelos cientistas que se dedicam a este campo da Ciência. O Estado da Arte contém as indicações que foram julgadas mais pertinentes ao objeto de estudo em questão.

Capítulo 2

Estado da Arte

No decorrer deste capítulo são apresentados artigos na área de processamento e análise de imagem de documentos, enfocando soluções para diferentes problemas relacionados ao processamento e reconhecimento de imagens postais, os quais servem de apoio ao presente trabalho.

2.1. Histórico do Correio

O correio existe desde o início da Idade Moderna, quando as explorações marítimas dos séculos XV e XVI fizeram crescer as necessidades de algum tipo de comunicação, mesmo rudimentar, que permitisse a transmissão de notícias sobre as viagens e o envio de informações comerciais aos mercadores.

Em suas viagens em busca dos caminhos das Índias pelo cabo da Boa Esperança (África), os portugueses tripulantes dos navios que faziam a rota para o Oriente ou que voltavam de lá, adotaram as “pedras postais” – uma mistura de agências e caixas de correios, utilizadas naquela localidade, especialmente no bebedouro de Mossel Bay, uma cidade da atual África do Sul. Sob grandes pedras com os nomes dos navios e de seus comandantes, deixavam-se mensagens que deviam ser retiradas pelas tripulações dos navios que seguiam a rota oposta. Existem muitos exemplares de “pedras postais”, hoje conservadas no museu da Cidade do Cabo, com inscrições em português, espanhol, inglês e francês.

Os serviços de entrega eram feitos por mensageiros, primeiramente via fluvial através de canoas entre os povoados e cidades, e após a abertura de estradas entre as cidades, passou-se a utilizar cavalos e a galéria, uma espécie de carroça fechada puxada por cavalos.

Com a evolução dos meios de transporte, passou-se então a utilizar os meios mais rápidos disponíveis em cada época, como ferrovias, carros, caminhões, aviação, meios eletrônicos e *e-mail* até o momento.¹

2.2. O Correio no Brasil

O correio no Brasil teve o seu início em 1500, com a chegada de Pedro Álvares Cabral, quando surgiu a primeira correspondência oficial ligada ao País, escrita por Pero Vaz de Caminha e enviada ao Rei de Portugal, a qual relatava com notório entusiasmo o descobrimento de uma nova terra. Com este acontecimento, eternizado na história brasileira, estava sendo escrita a primeira página do surgimento dos correios no Brasil.

Os primórdios dos serviços postais no Brasil-Colônia reportam-se ao correio em Portugal e à sua atuação neste território. Durante os primeiros tempos da colonização do Brasil, os portugueses não dispunham de um sistema postal bem organizado, tendo, inclusive, que recorrer ao de nações vizinhas.

Entre o período Brasil-Colônia e o início do período da ECT (Empresa de Correios e Telégrafos) em 1969, houve vários processos e regulamentos estabelecendo administração terrestre e ultramarinha, os quais buscavam e promoviam o desenvolvimento dos serviços postais com a criação de Agências Postais, Correio Aéreo Nacional e Internacional, franquias e, inclusive, a subordinação ao Ministério das Comunicações por decreto em 1968. O processo de desenvolvimento aconteceu também no Período Imperial, de 1805 até 1888, no qual destaca-se a emissão de selos a partir de 1842. No Período Republicano de 1889 até 1928, destaca-se o interesse em organizar melhor os meios de comunicação, quando em 1911 é criada a União Postal Sul-americana. Em 1921, o transporte de malas diretas postais via aérea passa a ser aceito. Já em 1924, teve início o uso da máquina de franquear correspondências e em 1927 iniciava-se o transporte de correspondências via aérea regular entre a América do Sul e a Europa. Em 1929, foi criado o DCT (*Departamento de Correios e Telégrafos*), no qual o Código Postal Universal, elaborado por ocasião do IX Congresso Universal em Londres, viria a legislar e apresentar soluções para os problemas postais modernos, e dar início a uma nova era da história dos Correios².

¹ Fonte: Portugal em Linha – Artigos de Opinião – Belarmino D. Batista

² Fonte: Correios Brasileiros

No período de 1930 a 1969, houve alguns fatos nacionais que contribuíram para a melhora da administração e a qualidade nos serviços postais entre os quais se pode destacar a Revolução de 30, trazendo significativas alterações na estrutura político-administrativa do País, cujas conseqüências atingiram também o setor postal. Os Correios reagiram passando “a analisar não só sua estruturação, mas também a evolução de seu desempenho, seus meios e sua capacidade técnica de atender à necessidade de comunicação”³. No ano de 1931, o então presidente do Brasil, Getúlio Vargas, baixou o decreto pelo qual fundia a Direção-Geral dos Correios com a Repartição-Geral dos Telégrafos. Originava-se assim o Departamento de Correios e Telégrafos – o DCT. Pode-se citar, também, o surgimento da ECT (*Empresa dos Correios e Telégrafos*) trazendo uma nova postura dos poderes públicos com relação às comunicações e, particularmente, os serviços postais no desenvolvimento do País.

2.3. A Tecnologia a Serviço dos Correios

Em 1969, inicia-se o processo de desenvolvimento do Serviço Postal Brasileiro com a criação, em 20 de março, da Empresa Brasileira de Correios e Telégrafos.

Em 1971, é concluída a montagem do Centro de Triagem Mecanizada de São Paulo, padronizando os envelopes conforme recomendações da União Postal Universal e a edição do Guia Postal Brasileiro com o CEP - código de endereçamento postal, representado por 5 algarismos.

No ano seguinte, ou seja, em 1972, com vistas a levantar os diversos problemas postais, foi contratada a empresa SOMEPOST para fazer a relação dos setores de exploração, meios a serem utilizados, organização e racionalização de serviços e avaliação das necessidades de investimento, pois iniciavam-se aqui suas operações internacionais de correspondência agrupada, por meio de convênio com os Estados Unidos.

Durante os últimos anos, a área de análise e reconhecimento de documentos eletrônicos vem ganhando uma importância significativa no contexto internacional. No Brasil, as pesquisas, também, estão acontecendo em várias universidades federais e particulares.

O método utilizado no processamento digital de imagens, em envelopes postais, está direcionado para a segmentação e a binarização da imagem. Tal opção se faz para o tratamento da imagem, a fim de que se possa eliminar o fundo cinza e destacar a parte escrita e a parte gráfica (*carimbos e selos*).

³ Fonte: Correios Brasileiros

Utilizando-se do programa Bropi, faz-se a leitura da imagem em níveis de cinza disponível numa base digital do tipo Bitmap - BMP. A seguir aplica-se a técnica da detecção de bordas seguido de uma erosão da imagem na morfologia cinza; na seqüência é feita a binarização pelo método de Otsu, e aplicada às técnicas da morfologia binária: fechamento e abertura. Após isso, com os resultados obtidos da imagem, é feita a segmentação e a localização dos componentes; linhas dos blocos de endereço e outros componentes que fazem parte da imagem. Com a imagem segmentada e os blocos localizados, o próximo passo é a eliminação dos componentes indesejáveis, através de heurísticas pré-determinadas para limpar a imagem, disponibilizando-a com o mínimo de ruídos possível.

Desta forma, a dissertação busca agilizar o processo de localização dos blocos de endereço em imagens postais, através da automatização das técnicas que constituem a segmentação e binarização de imagens. Tal processo pode ser utilizado pelos Correios ou qualquer outra empresa que procura agilizar o tratamento de imagens postais ou similares.

Fazendo uma análise superficial de uma imagem de envelopes postais, encontram-se caracteres e imagens gráficas do tipo selos e carimbos. Para o presente trabalho foram pesquisadas técnicas que abordassem os problemas de localização e segmentação em imagens de envelopes postais.

A automação surgiu com o desenvolvimento da tecnociência que exigiu, em 1975, uma reestruturação.

2.3.1. Automação

O desenvolvimento da tecnociência exigiu, em 1975, uma reestruturação do Departamento Postal, através da qual foi implantada a Assessoria de Planos de Desenvolvimento, e o Serviço de Processamento de Dados, instalando-se, assim, os Centros de Triagem Automática, que possibilitaram maior rapidez no encaminhamento de objetos nas principais cidades do País, como São Paulo, Rio de Janeiro e Brasília.

2.3.2. Modernização das Empresas de Correios e Telégrafos Brasileiros

Em função da crescente demanda dos serviços postais e para que a ECT pudesse imprimir maior agilidade no tratamento dos objetos, permitindo, inclusive, a utilização de equipamentos eletrônicos, foi implantado o Código de Endereçamento Postal, atribuindo a todas as localidades brasileiras um código específico.

Em virtude das exigências dos usuários e do desenvolvimento tecnológico, tornou-se necessária uma readequação do sistema de CEP aos métodos de trabalho ainda mais modernos. Exemplificando melhor o problema, após chegar a seu destino, o CEP já não proporcionava uma racionalização para as atividades subseqüentes de tratamento das correspondências na fase preparação para entrega a seus respectivos destinatários. Desta forma, optou-se pela atribuição de um código individual a cada rua, acrescentando-se mais 3 dígitos ao CEP anterior.

Com esta alteração a ECT passou a oferecer a seus usuários um incremento na qualidade da prestação de seus serviços, principalmente nos aspectos de agilidade, rapidez e confiabilidade no encaminhamento e na entrega das correspondências. Sendo assim, com a crescente demanda dos serviços prestados pelos Correios, a empresa vem investindo na área e na melhoria de desempenho nos seus Centros de Triagem, aumentando a produtividade e a precisão dos serviços, assegurando maior agilidade, qualidade e segurança ao tratamento de objetos postais, inclusive o acesso do público à *internet*, constituindo mais um passo dos Correios com vistas à universalização dos serviços postais.

Em função dos grandes desafios dos mercados postais, nacional e internacional e do processo de grande efervescência reformista vivido pelos correios dos principais países do mundo, os Correios brasileiros vêm passando por um amplo processo de mudanças, desenvolvendo, dessa forma, ambiciosos programas de reestruturação. Seus balizadores são Programa de Qualidade Total e significativos investimentos em infra-estrutura, capacitação e treinamento, modernização tecnológica, atendimento ao cliente e lançamento de produtos e serviços.

Para fazer frente aos novos desafios, ao completar 31 anos, a ECT se prepara para assumir uma nova postura como empresa de economia mista, que passará a se denominar Correios do Brasil S.A. Esta reforma, proposta pelo anteprojeto da Lei Geral do Sistema Nacional de Correios, em tramitação no Congresso Nacional, propõe a abertura do mercado postal a operadores privados que, por sua vez, poderão competir pelos serviços ou explorar novos negócios em parceria com os Correios do Brasil.

2.4. Metodologia Utilizada pelo ECT

A base da metodologia proposta está centrada no endereçamento postal do destinatário em envelopes de cartas postais. Os sistemas atuais dos Correios já trabalham nesse sentido, mas por enquanto, o reconhecimento está sendo feito somente com o código postal CEP. A

proposta deste é preparar a base com os blocos do endereço do destinatário completo, para futuramente serem feitos os reconhecimentos das linhas com o nome, endereço, bairro, cidade, CEP ou outras informações que forem relevantes.

Na atualidade os Correios buscam as informações, em especial o CEP, nas etiquetas ou envelopes datilografados ou por qualquer outro meio mecânico impresso, os quais precisam ser paralelos à borda de maior dimensão do envelope. O aparecimento de qualquer linha, como por exemplo a moldura que delimita o campo do endereçamento do formulário das demais informações, provoca a rejeição da correspondência pelo equipamento de triagem eletrônica. As dimensões, também, devem ser ajustadas, evitando-se desta forma o deslocamento de CEP, impedindo a leitura do mesmo ou provocando a inclinação do envelope. Tal fato, além de causar a rejeição por parte do equipamento, dificulta a triagem manual.

Essa normatização serve tanto para envelopes manuscritos ou pré-impessos, segundo a qual as informações dos envelopes postais deveriam seguir o padrão de localização do endereço, selos e CEP, para que os equipamentos atuais o reconheçam a fim de agilizar o processo de triagem até a entrega das correspondências no seu destino com segurança.

No atual sistema, a presença de marcas de empresas, traços e desenhos presentes na imagem do envelope, são rejeitados.

No entanto, a modernização não pára e novas tecnologias são pesquisadas e implantadas em vários países desenvolvidos, como França, Canadá, EUA e outros países da Europa e Ásia, a fim de agilizar cada vez mais os processos postais internacionais.

No Brasil, em 2001, várias capitais implantaram novos equipamentos desenvolvidos com a tecnologia da empresa alemã Siemens, com o objetivo de classificar e agilizar o processamento postal de correspondências dos correios brasileiros⁴.

2.5. Sistema de Processamento e Análise de Imagens

O Processamento Digital de Imagens tem por objetivo a manipulação sistemática de imagens, com o intuito de gerar outra imagem que deve apresentar uma qualidade melhor para uma análise futura.

Um sistema de processamento e análise de imagens [Facon, 1993] pode ser constituído dos seguintes elementos:

⁴ Fonte: Correios Brasileiros

- Aquisição ou digitalização - A imagem capturada pelo sensor é transformada em uma imagem digital sobre a forma de uma matriz de valores discretos chamados pixels, abreviação de picture element, ou elemento de figura.
- Pré-processamento - Esta etapa permite corrigir um certo número de imperfeições e defeitos que aparecem na imagem decorrentes de problemas na aquisição, cujas causas podem ser as características físicas do sistema, as condições de iluminação, imperfeições no sistema de lentes da câmara, entre outras. O pré-processamento não é indispensável, mas na maioria dos casos é necessário.
- Segmentação – O objetivo principal da segmentação é subdividir uma imagem em suas partes constitutivas. Em uma imagem natural a segmentação é realizada pela detecção de descontinuidades e/ou regiões na imagem. Na maioria dos processos são baseados na pesquisa destas entidades, que são armazenadas sobre uma forma adequada, segmentos ou primitivas.
- Representação – O objetivo da representação é elaborar uma estrutura adequada, agrupando os resultados das etapas anteriores e o armazenamento dos diversos padrões que contêm o conhecimento a priori a respeito do documento. A representação é efetuada após a obtenção das primitivas, e às vezes, ao mesmo tempo. Algumas vezes a representação permite medir as propriedades das formas resultantes da segmentação.
- Interpretação/Classificação – Este é o processo mais “inteligente” em um sistema de visão por computador, pois representa o nível mais alto na abstração e permite obter a descrição do fenômeno representado na imagem, fazendo uso do conhecimento a priori dos casos estudados, e também utilizando os conhecimentos adquiridos nas fases anteriores do processo.

2.6. Análises de Layout de Imagens

A análise estrutural do *layout*, pode ser chamada também de análise física ou geométrica do *layout*, obtém a segmentação física entre os grupos de componentes de uma imagem. Dependendo do formato do documento, a segmentação pode ser realizada para palavras isoladas, linhas de texto, e estruturas de blocos (*grupos de linhas de texto, tais como parágrafos separados ou entradas de tabelas de conteúdo*). A análise estrutural pode ser realizada de modo *top-down* ou *bottom-up*. As técnicas que analisam o documento do geral para o detalhe são conhecidas como *top-down*, que se inicia encontrando os limites dos

grupos que existem na página e subdivide-se sucessivamente em grupos menores até que todos os blocos da página sejam encontrados. De maneira inversa ao *top-down*, a técnica *bottom-up* busca primeiramente os menores componentes do documento para o geral, onde os componentes conectados são fundidos em caracteres, depois palavras, linhas de texto e assim por diante, até que todos os blocos do documento sejam encontrados. O’Gorman et al. [1995] sugerem outro método, constituído da combinação das duas técnicas citadas, ou seja *top-down* e *bottom-up*.

A análise funcional do *layout*, também chamada análise sintática ou lógica do *layout*, utiliza informações de domínios dependentes, consistindo em regras de *layout* para uma página particular, a fim de realizar a rotulação das estruturas de blocos e de fornecer algumas indicações das funções deste bloco [O’Gorman, 1995].

2.7. Segmentação de Imagens

A segmentação de imagens é um processo que tipicamente particiona o domínio espacial de uma imagem em subconjuntos mutuamente exclusivos, chamados regiões, onde cada uma delas é uniforme e homogênea com respeito a algumas propriedades, como tom ou textura, e cujos valores diferem, em alguns aspectos e significados, das propriedades de cada região vizinha.

Yuan e Tan [2001] apresentam o método que detecta a extremidade das informações para extrair blocos do texto das imagens de documentos em escala de cinza e ainda, aponta as regiões em que há grande quantidade de ruídos nas imagens de jornais, revistas, envelopes postais e cartões, separando as regiões gráficas do texto. O algoritmo proposto traça os pontos em diferentes entidades e em grupos destas extremidades da região textual, usando a tecnologia de linha de aproximação, podendo futuramente categorizar e direcionar os blocos de textos. Finalmente, os blocos baseados nos componentes conectados podem ser introduzidos e fundidos nas dobras de regiões homogêneas, para obter a decomposição correta da página com eficiência computacional, reduzindo o tamanho do bloco por manipulação e segmentação de linhas ao invés de *pixels*. O método tem sido testado em um grupo considerado de imagens de jornais com múltiplos *layouts* de páginas, apresentando resultados promissores.

Em Pietikäinen e Okum [2001], o método apresentado é baseado nas extremidades de imagens complexas descobertas no texto de documentos, nas quais está embutido um colorido de fundo complexo. No referido artigo, propõe-se uma abordagem da textura baseada nas

informações das extremidades das imagens. Nas experiências por eles realizadas, o método desenvolveu-se satisfatoriamente para pequenos textos, ao passo que em textos maiores, obtiveram-se avaliações mais convincentes, ou seja, resultados de melhor qualidade. Tal método ultrapassa o problema da detecção do texto, podendo também ser usado para outras tarefas de interpretação de imagem. Os documentos complexos são crescentes e comuns nos dias de hoje, sendo necessária a descoberta de novos algoritmos para resolver problemas dessa natureza.

Segundo Jin e Tan [2001], propõe-se uma abordagem moderna que determina as posições em áreas de texto em imagens com fundo complexo, usando a decomposição de Wavelet e o pseudo movimento da imagem. Em tal método, a imagem, que sobressai, é traduzida, provê uma seqüência de mudanças da imagem – o pseudo movimento da imagem. Quando o pseudo movimento acontece sobre a imagem, os coeficientes de Wavelet deverão oscilar. Diante desta propriedade, pode-se localizar as áreas de texto da imagem. Sendo assim, o algoritmo de forma simples e eficiente, remove o fundo complexo a partir das imagens e áreas de texto com diferentes tipos de caracteres.

Em artigo apresentado por Xue et al. [1999], está a idéia fundamentada na localização do bloco de endereço do destinatário de envelopes manuscritos chineses, que possuem três características distintas: as linhas de textos escritos constituem a principal parte da imagem do envelope; não há uma margem entre o bloco de endereço do destinatário e o bloco de endereço do remetente; e as letras da caligrafia chinesa são compostas por complicados traços efetuados com liberdade. O método de segmentação da imagem do envelope em vários blocos de endereços de destinatários candidatos e a seleção do melhor bloco como destinatário ainda não é adequado para os envelopes manuscritos chineses. Este trabalho apresenta um novo método de localização do bloco de endereço do destinatário, baseado na extração de linhas de texto, no qual a imagem é segmentada em diversos blocos de acordo com as características dos envelopes manuscritos chineses, sendo dividido o problema de localização do bloco de endereço em subproblemas mais simples e resolvendo-os respectivamente. Apesar de resultados animadores, como por exemplo a redução do tempo de processamento e o índice de aproveitamento do processo, ainda não é satisfatório, porque, até o momento as máquinas automáticas de postagem reconhecem somente o código postal regional, ao passo que o bloco de endereço do destinatário inclui não só o código postal da região, mas também o endereço e o nome da região.

Bourgeois e Empotz [1998] escreveram sobre a segmentação de documentos compostos em níveis de cinza de periódicos e revistas; neles encontra-se a defesa da necessidade de um pré-tratamento dos níveis de cinza. Os referidos autores indicam um método de detecção das linhas de textos já utilizado com sucesso nos vídeos ou nas imagens naturais, o qual pode ser aplicado a outros documentos que contemplam os mesmos níveis de cinza. Em seguida, introduziram um método para rejeitar as falsas detecções e melhorar a separação entre o grafismo e as regiões que contêm o texto, técnica esta de binarização adaptada a cada linha que permite segmentar os caracteres e corrigir a inversão das cores. Levando em conta que periódicos e semanários são freqüentemente editados em cores, o que dificulta o reconhecimento dos caracteres por um sistema OCR, é necessário um método que possa avançar no sentido de minimizar essa dificuldade. Os referidos autores defendem que o método por eles elaborado pode, também, ser utilizado sobre documentos coloridos a partir da comparação das formas idênticas de caracteres e o seu agrupamento binário.

Cheng et al. [1997] propõem a segmentação que explora as características de textura locais e estrutura de imagens de documentos pesquisados em regiões, como texto, fundo, títulos e imagens. O método está baseado no uso de multiescala em estruturas bayesianas. Essas estruturas foram escolhidas porque permitem modelar com precisão as características de uma imagem e o contexto estrutural de cada região. Os parâmetros que descrevem as características das imagens típicas, são extraídas de base de dados de treinamentos de imagens, as quais são produzidas por escaneamento e segmentação manual de documentos típicos em seus componentes desejados. O algoritmo usado busca diminuir o tamanho da região mal interpretada da imagem o que resulta em maior eficiência do processo aperfeiçoamento da precisão da segmentação da mesma.

Bourgeois [1996] apresenta o método de localização de textos em uma imagem em níveis de cinza. O método foi desenvolvido para um sistema OCR destinado ao meio industrial, utilizando imagens em tons de cinza. A análise foi realizada com a ajuda de um filtro linear Run Length Smoothing Algorithm - RLSA e suas numerosas variantes, que colocam em evidência as zonas do texto. O algoritmo é utilizado para realizar a segmentação, efetuando a dilatação dos traços no sentido vertical e depois no sentido horizontal. Em contrapartida, tal algoritmo “não permite levar em conta as zonas do texto em qualquer orientação, pressupõe alinhamento do texto no eixo horizontal com grau de inclinação de mais ou menos 30°”. Constitui-se de um método simples e robusto que permite localizar e detectar, em tempo real, zonas de texto com a orientação de caracteres, por um cálculo direto sobre a

imagem. A vantagem está no “número reduzido de operações por *pixel* e no emprego de uma varredura única da imagem, o que permite efetuar o cálculo em tempo inferior à cadência de vídeo. Este método serve de base para o desenvolvimento de um sistema completo de reconhecimento de caracteres sobre imagens sem contraste para o setor industrial”.

Jain e Bhattacharjee [1992] apresentam o método denominado Segmentação de Textura Supervisionado, o que se baseia em textura para identificar blocos de endereço de destino em imagens de envelopes. As características de textura são computadas, usando filtros Gabor simétricos. Um classificador da camada de rede neural é usado para classificar *pixels* em categorias de texto e não texto, usando quatro características de textura. Também foi apresentada uma heurística simples para selecionar o bloco de endereço de destino correto, entre vários candidatos identificados. O método trabalha bem em várias imagens de envelope. O problema de automatização postal está ganhando importância rapidamente, sendo requerido um sistema de visão de máquina prático que possa esquadrihar cada parte do envelope de entrada e possa processar isto adequadamente. O problema de ler e entender o endereço de destino é controlado efetivamente através de um sistema de reconhecimento de caráter óptico (OCR). O aspecto mais difícil da automatização postal é localizar o bloco de endereço de destino com precisão. Isto pode ser atribuído às variações nas dimensões físicas das partes das cartas e o fato de que, na maioria delas, não há nenhuma posição predeterminada para o endereço de destino.

Jain et al. [1992] descrevem o método multicanal de filtragem de segmentação de textura no processamento de imagens de documentos, através do uso de filtros Gabor: separação de textos gráficos, localização de bloco de endereço e localização de barra de código, os quais vêm sendo utilizados para identificar regiões de texto ou barra de código de entrada em nível cinza de imagens de documentos. A classificação e a segmentação são mostradas em várias imagens escaneadas com um *scanner* de mesa plano, demonstrando a efetividade à segmentação e à aproximação de imagens de documentos, e a identificação das regiões homogêneas em uma dada imagem. Tanto o método supervisionado como não supervisionado têm sido usados para identificar regiões do texto. A maior vantagem do método de segmentação e classificação, através do uso dos filtros Gabor, é evitar a binarização da imagem.

O’Gorman [1992] apresentou um método chamado *Document Spectrum Analysis*, no qual todos os componentes conectados da imagem são primeiramente detectados, e a partir disso os blocos são gerados na técnica *bottom-Up* encontrando os vizinhos de cada

componente e agrupando os diferentes componentes. Este método de segmentação produz resultados muito bons para vários tipos de documentos, particularmente para páginas de jornais.

Haralick [1985] propôs as seguintes propriedades para um bom processo de segmentação:

- Regiões de uma imagem segmentada devem ser uniformes e homogêneas com respeito a alguma característica, como por exemplo, nível de cinza ou textura;
- Regiões adjacentes devem ter valores significativamente diferentes com respeito à característica segmentada; e
- Interior das regiões deve ser simples e sem falhas.

Entre as diferentes áreas de aplicação em que a segmentação atua, encontram-se 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; e
- Imagem de ultra-sonografia: representa a emissão, reflexão e recepção de ondas.

Baird [1990] propõe um método diferente, baseado na análise do fundo da página do documento antes dos objetos, para localizar os blocos de linhas. Esse método é conhecido pelo nome de partição do fundo da imagem. Ele consiste em delimitar os objetos pelo meio das regiões brancas modeladas por retângulos. Todos os retângulos brancos máximos são identificados e classificados, dos maiores para os menores. Os maiores retângulos são supostamente correspondentes às margens, espaços entre colunas ou entre parágrafos. Sua seleção permite realizar uma partição do fundo da imagem em blocos estruturados. A vantagem de considerar o fundo antes que os objetos resulta em maior independência frente à natureza do texto a se extrair (*o fundo é o delimitador mais comum em documentos de qualquer tipografia ou língua*). Um inconveniente desse estudo reside na escolha do tamanho mínimo necessário para um retângulo branco ser considerado como delimitador.

Considerando o conhecimento já construído a respeito de processos de localização dos componentes em imagens de envelopes postais, parte-se para a fundamentação teórica que

dará sustento ao processo de segmentação dos blocos de endereço, objeto de estudo desta dissertação.

Capítulo 3

Fundamentação Teórica

Neste capítulo buscar-se-á abordar os autores, teorias e seus aspectos mais significativo e relevantes para reforçar o entendimento a metodologia e método proposto a serem utilizados neste trabalho.

3.1. Segmentação por Limiarização

A limiarização é um método que se baseia no histograma da imagem, buscando encontrar ali “regiões bem definidas”, a fim de poder efetuar a divisão da imagem em objetos ou regiões. A continuidade dos níveis de cinza é a primitiva de maior valor na segmentação por região. Percebe-se, que a limiarização efetua a subdivisão da imagem em função das “regiões significativas” contidas no histograma da imagem [Facon, 1996].

Tomando-se uma imagem digital qualquer, onde $f(x,y)$ representa o nível de cinza de um ponto qualquer, e supondo a existência de apenas dois objetos, claramente definidos em função dos seus níveis de cinza, um fundo escuro sob um objeto claro, ao se observar o histograma dessa imagem, fica claro que ao selecionar um limiar L que esteja entre as suas regiões do histograma, que representam os objetos da imagem, pode-se extrair qualquer um desses objetos. Assim:

$$g(x,y) = \begin{cases} 1 & \text{se } f(x,y) > L \\ 0 & \text{se } f(x,y) \leq L \end{cases} \quad (3.1)$$

onde $g(x,y)$ representa o ponto limiarizado.

No entanto, o estudo torna-se mais interessante se for considerada uma imagem com dois tipos de objetos num fundo claro. Neste caso o histograma passaria a ser bimodal, e no caso anterior, multimodal. Para que seja possível a separação desses objetos, deve-se supor que:

$$g(x,y) = \begin{cases} \text{objeto 1 se } L1 < f(x,y) \leq L2 \\ \text{objeto 2 se } f(x,y) > L2 \\ \text{fundo se } f(x,y) \leq L1 \end{cases} \quad (3.2)$$

Pela multimodalidade do histograma, parece ser fácil perceber que esse tipo de limiarização apresenta um maior grau de complexidade, em função da dificuldade de se estabelecer múltiplos limiares que efetivamente isolem regiões de interesse.

3.2. Método de Otsu

Segundo Santos [1997], a aplicação da limiarização pelo método de Otsu em uma imagem que sofreu uma hiperbolização quadrática do histograma fornece resultados muito bons.

Este método de limiarização tem a vantagem de não utilizar nenhum conhecimento *a priori* da imagem, baseando-se apenas em seu histograma em níveis de cinza. A idéia é encontrar no histograma um limiar ótimo que separe os objetos da imagem através da construção de duas classes $C_0 = \{0,1,\dots,t\}$ e $C_1 = \{t+1,t+2,\dots,l-1\}$, a partir de um nível de cinza arbitrário, utilizando para isso a análise de discriminante.

Para encontrar o limiar ótimo pode-se minimizar a seguinte função (mais simples segundo Otsu) de critérios que se referem a t :

$$\eta = \sigma_{\beta}^2 / \sigma_T^2 \quad (3.3)$$

onde σ_T^2 é a variância total, independente do nível de cinza, e σ_{β}^2 é a variância interclasses. O limiar t ótimo será definido por:

$$t = \text{ArgMax} \sigma_{\beta}^2 \quad (3.4)$$

3.3. Projeções Horizontais e Verticais

Muitas diferenciações entre os blocos de texto podem ser feitas usando projeções. Por exemplo, geralmente uma projeção horizontal de um título irá indicar um tamanho de texto maior que o do corpo do texto. Ou o rodapé poderá ter um espaçamento menor entre a linha de texto do que o corpo do texto. Na prática não é necessário realizar um processamento em todos os *pixels* da imagem com resolução original quando se utiliza da técnica para analisar o *layout*. A imagem pode ser reduzida para um tamanho que melhore a eficiência e os resultados.

A desvantagem da técnica é a utilização somente em imagens binárias e sem inclinação. Outra desvantagem citada em Dias [1996] é o conhecimento a priori do *layout* esperado como critério de parada do algoritmo e ser unicamente utilizado em documentos com formato padronizado.

A proposta de Nagy [1991] consta do uso de cortes horizontais e verticais recursivos para dividir um documento em blocos. Nesse método o documento é recursivamente dividido usando a técnica de projeções horizontais e verticais. Essa projeção é calculada na horizontal e na vertical. Um valor zero indica a presença de um espaço em branco. Um corte horizontal ou vertical é, então, feito ao longo do espaço em branco, produzindo duas subimagens. Para cada uma dessas subimagens, a mesma operação é feita em outra direção. Deste modo, o processo é recursivamente aplicado para cada subimagem do documento até não existir mais a possibilidade de executar cortes. Os blocos de *pixels* pretos resultantes são considerados a representação dos blocos impressos indivisíveis no documento. Uma desvantagem deste método é sua sensibilidade à inclinação do documento.

Uma abordagem comum, utilizada em O’Gorman [1995] para segmentação de caracteres, é primeiro determinar as projeções horizontais e verticais de cada linha de texto em um documento. Por esta razão, este perfil terá picos pequenos entre as palavras, picos médios representando as palavras e picos grandes para partes ascendentes e descendentes das letras. O perfil real e sem ruído teria vales de altura zero entre os caracteres, indicando os locais de segmentação dos mesmos. Contudo, devido a ruídos, os tais caracteres podem estar unidos ou quebrados; também devido a fontes itálicas, pode não haver uma linha vertical separando-os claramente. Assim, a segmentação não pode ser realizada simplesmente identificando-se vales com valor zero entre os picos ao longo do perfil.

Segundo Normand [1995], um método baseado na extração do fundo do documento, oferece o benefício de ser adaptável para o contexto do documento e para a orientação dos

blocos de texto. Tal técnica envolve um elemento de estrutura isotrópica 2D usado para caracterizar uma cadeia de *pixels* brancos, onde este elemento é um disco aproximado por um octógono regular que pode ser gerado recursivamente. Outra vantagem do método proposto é que uma segmentação hierárquica pode ser derivada de uma imagem através de um padrão octogonal. Esta árvore permite realizar uma mancha isotrópica multi-escala, que leva a uma segmentação física dos blocos. A vantagem destes algoritmos é que podem ser aproveitados em qualquer problema de segmentação, onde os objetos são unidos de acordo com a organização de fundo do documento, encontrando blocos de texto com caracteres de diferentes tamanhos e vários tipos de grafos, independente da inclinação do documento.

3.4. Método RLSA – Run Length Smoothing Algorithm

Apresenta-se como um método interessante pela facilidade de implementação e obtenção de bons resultados em segmentação de imagens de envelopes postais ou textos, no qual os blocos de texto são compostos por muitos componentes fechados. Segundo Normand [1995], o princípio do RLSA é preencher as distâncias pequenas e unir as letras do mesmo bloco. Devem ser definidos um limiar de distância δ e uma direção ϕ . A imagem binária é analisada ao longo da direção ϕ escolhida, e a seqüência de *pixels* brancos adjacentes menores que o limiar δ é preenchida. O limiar poderia ser computado a partir de características estatísticas baseadas nas seqüências de pontos atuais na imagem inteira.

Outra maneira de se implementar este método é através de morfologia matemática. Em Oliveira [1995] é realizada a conversão dos *pixels* brancos em pretos nas seqüências horizontais e verticais, utilizando-se as operações morfológicas de fechamento, com os elementos estruturantes “linha horizontal” e “linha vertical”, respectivamente. O resultado obtido pela operação lógica E , entre as duas imagens, deve ser novamente fechada por um elemento estruturante “linha horizontal”. Deste resultado deve-se extrair as bordas, que representam os blocos da imagem.

3.5. Rotulação de Blocos

Segundo Saito [1999], a rotulação de uma imagem consiste em pegar ou demarcar uma determinada região de uma imagem pós-binarizada e rotular um *pixel* com rótulo L , de valor numérico, e os *pixels* vizinhos com as mesmas características, como cor, brilho, forma,

textura; isto quando acontece, determinam uma região. Antes de rotular uma imagem, é preciso segmentá-la, e a forma mais simples é binarizar.

No método citado em Le [1995], um bloco é definido como um componente conectado e é representado por uma coleção de seqüências de pontos pretos conectados na sua 8-vizinhança. A rotulação de blocos é similar a rotulação de componentes. Seu propósito é segmentar uma imagem binária rotulando seus blocos. A cada seqüência de pontos pretos é designado um número inteiro chamado rótulo e os rótulos de uma seqüência de pontos pretos conectados devem ser os mesmos.

O procedimento para a rotulação dos blocos identifica a seqüência de pontos pretos conectados para formar este bloco. O procedimento analisa a conectividade através das distâncias entre a linha corrente e a linha anterior de cada seqüência de pontos pretos. Se esta distância estiver entre o limiar δ para esta união, são designados rótulos para as seqüências destes pontos pretos da linha atual e/ou atualiza os rótulos das seqüências destes pontos pretos da linha anterior, definindo os blocos de uma imagem binária em termos de suas coordenadas.

3.6. Rotulação Funcional

Aborda-se aqui a implementação e a técnica de rotulação de imagens binárias a partir da técnica de *Labeling*. Esta técnica rotula um *pixel* e a partir deste verifica os vizinhos, rotulando os que têm as mesmas características. Posteriormente verificam-se os vizinhos dos vizinhos até esvaziar a lista, que é uma região da imagem. A seguir, recomeça-se o processo, fazendo a varredura de cima para baixo e da esquerda para direita, até encontrar um *pixel* sem rótulo.

A rotulação funcional é realizada e baseada em regras específicas para o tipo de imagem através do uso de características derivadas do mesmo. Essas características podem incluir as posições relativa e absoluta de um bloco em uma página, as posições relativa e absoluta de um campo dentro de um bloco, regras gerais de composição de espaços, tamanho da fonte e estilo do texto. As regras de formatação usadas para produzir a página são aquelas usadas no processo reverso de rotulação funcional [O’Gorman, 1995].

Devido à grande variação de formatos e aos problemas no processamento de imagens, tais como ruídos, detecção errônea de características e reconhecimento de texto imperfeito, a maior parte do trabalho, na rotulação funcional, está restrito a domínios particulares.

Jain, em [1998] usa uma abordagem baseada na extração dos componentes conectados para implementar a segmentação de imagens de documentos e identificação das suas regiões.

Um novo modelo de documento, o qual preserva informações geradas em *top-down*, é proposto baseado no fato de que um documento é logicamente representado por edições iterativas, armazenamento, recuperação, transferência e análise lógica. Tal método é aplicado para vários tipos de documentos e pode ajustar-se a quantias moderadas de inclinações e ruídos. Este algoritmo foi testado com 150 imagens (125 imagens com 300 dpi – banco de dados de documentos, 10 imagens com 100 dpi – texto japonês, e 15 imagens de artigos com 100 e 300 dpi). Os principais erros na segmentação dos documentos foram: tabelas sem conexão, títulos/legenda em uma linha simples, colunas com aberturas pequenas, parágrafos começando com caracter grande, títulos/cabeçalhos escassos, espaços largos e desenhos escassos ou muito claros.

3.7. Tophat Diferencial

Através de Gu [1998] é apresentado um sistema completo para extração de caracteres em imagens de cenas coloridas, utilizando um novo algoritmo de segmentação morfológico – Tophat diferencial (DTT). Este novo método pode lidar com problemas de segmentação mais difíceis do que os outros algoritmos conhecidos, especialmente com objetos nos quais existem, em um fundo complicado, tamanho irregular e direções que são afetadas por condições externas variantes. A abordagem proposta para detecção de caracteres em imagens de cenas é robusta e adaptável neste experimento. A segmentação é realizada a partir da imagem de entrada, onde é decomposta pelo DTT em uma série de subimagens, com diferentes tamanhos de objetos. O procedimento de segmentação começa com $r_1 B_{disk}$ e termina com $r_8 B_{disk}$, porque em uma investigação estatística o maior caracter encontrado em uma imagem de cena é menor do que 17 (*diâmetro do disco de $r_8 B_{disk}$*).

Neste capítulo tratou-se da fundamentação teórica necessária para a compreensão da morfologia matemática, que se destina ao estudo dos operadores morfológicos binários e níveis de cinza que serão estudados na etapa seguinte.

Capítulo 4

Morfologia Matemática

Morfologia vem das palavras gregas *morphê* (forma) e *logos* (ciência); portanto a morfologia é a ciência que trata das formas que a matéria pode tomar, em quaisquer áreas, como biologia, sociologia, gramática [Facon, 1996].

Na área de processamento de imagens, a morfologia matemática é uma ferramenta muito versátil porque ela permite processar imagens com objetivos de realce, de segmentação, de detecção de bordas, de esqueletização, de afinamento, de análise de formas, entre outras.

O princípio básico da morfologia matemática, em termos de imagem, é extrair informações relacionadas à geometria e à topologia de um conjunto desconhecido X . Este conjunto X é comparado, a partir de uma transformação, com um conjunto B completamente definido e conhecido (*forma, tamanho*), chamado de elemento estruturante. O resultado dessa transformação permite avaliar o conjunto X e extrair as informações de geometria e topologia. A figura 4.1 apresenta os tipos de elementos estruturantes binários básicos.

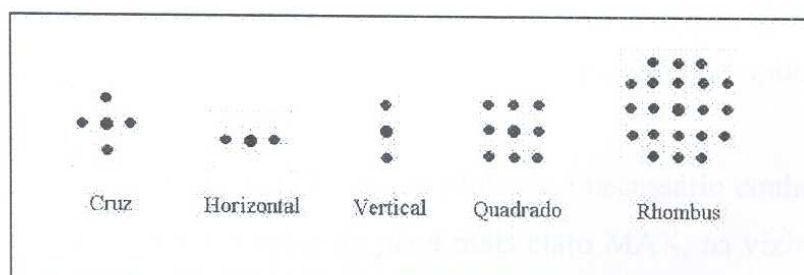


Figura 4.1: Tipos de elementos estruturantes básicos.

O formato e o tamanho do elemento estruturante possibilitam testar e quantificar de que maneira o elemento estruturante se comporta na imagem. A figura 4.2 ilustra uma imagem binária, um elemento estruturante e exemplos de interação.

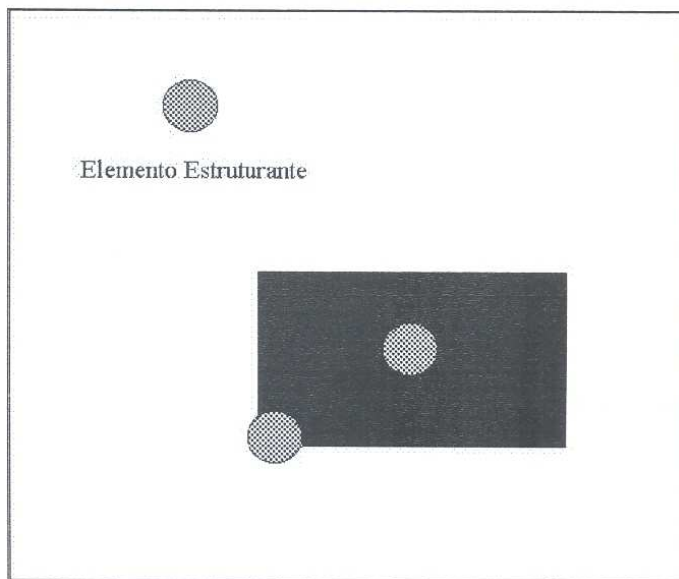


Figura 4.2: Exemplo de iteração do elemento estruturante em uma imagem.

Com o exemplo da figura 4.2, nota-se que, em algumas situações, o elemento estruturante encaixa-se à imagem e, em outras, não. Se o formato ou o tamanho desse elemento estruturante for mudado, outras respostas serão obtidas. Portanto, o tipo e a natureza da informação extraída depende necessariamente do tipo de elemento estruturante e do tipo de imagem estudada.

A morfologia matemática pode ser aplicada em imagem binária e em níveis de cinza:

- Morfologia binária: uma operação morfológica binária é completamente determinada através da vizinhança examinada ao redor do ponto central, da configuração dos pontos pretos (*pixel ativo*) e brancos (*pixel inativo ou neutro*) nessa vizinhança e do algoritmo. Este tipo de morfologia se aplica em imagens binárias.
- Morfologia em níveis de cinza: nessa morfologia é necessário conhecer o valor do *pixel* mais escuro MIN e o valor do *pixel* mais claro MAX, na vizinhança de cada *pixel* ou numa parte da vizinhança da imagem original. Uma operação morfológica em níveis de cinza é completamente determinada através do tamanho e forma da vizinhança, das regiões de pesquisa de MIN e MAX e do algoritmo. Este tipo de morfologia se aplica em imagens em níveis de cinza.

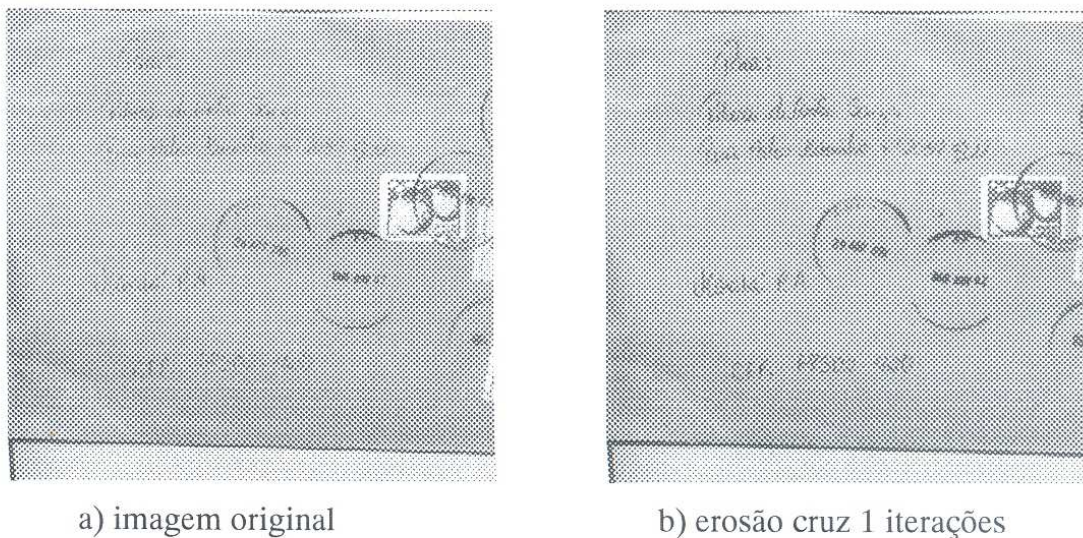


Figura 4.3: Exemplo de iteração entre o elemento estruturante e a imagem em níveis de cinza.

A morfologia matemática, ao contrário das técnicas apresentadas no capítulo 2, não requer cálculos matemáticos complexos; dispensa a fase de treinamento e não há necessidade de determinação de parâmetros, como o caso das redes neurais, contornos ativos e algoritmos genéticos, respectivamente.

Nas próximas seções serão apresentados alguns operadores morfológicos binários e em níveis de cinza. Estes operadores são definidos de maneiras diferentes por Banon [1994] e Barrera [1994]. Este trabalho adotou o formalismo de Minkowski [Serra, 1982] e as definições foram extraídas em Facon [1996] e Serra [1982].

4.1. Operadores Morfológicos Binários

Uma operação morfológica binária é completamente determinada através da vizinhança examinada ao redor do ponto central, da configuração dos pontos pretos (*pixel ativo*) e brancos (*pixels inativo ou neutros*) nessa vizinhança e do algoritmo.

Essa operação morfológica binária ocorre na aplicação das técnicas de erosão binária, dilatação binária, abertura e fechamento.

4.1.1. Erosão Binária

A erosão de um conjunto X pelo elemento estruturante B pode ser representada pela fórmula:

$$\varepsilon^B(X) = X \ominus \tilde{B} = \bigcup_{b \in \tilde{B}} X_b \quad (4.1)$$

onde \tilde{B} representa o transposto de B (o elemento transposto é obtido pela simetria do centro a origem dos eixos), \ominus representa a subtração de Minkowski [Facon, 1996].

Conforme a definição acima, o conjunto a ser erodido, X, é deslocado em função das posições permitidas pelo elemento estruturante B. Os deslocamentos são realizados em relação ao ponto central de \tilde{B} . Veja-se a figura 4.4.

A partir de agora um *pixel* inativo ou neutro é representado por “.”, e um *pixel* ativo é representado por “•”.

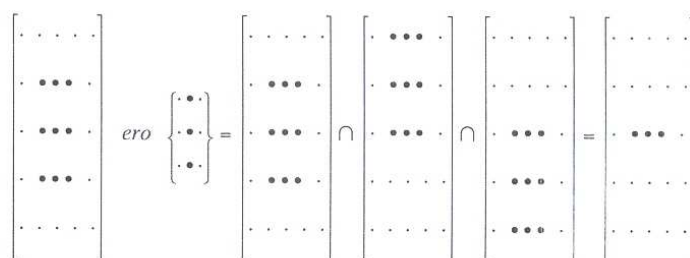
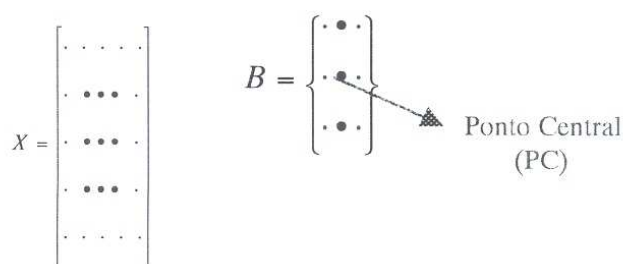


Figura 4.4: Exemplo dos deslocamentos dos pixels na erosão.

Cada movimento de X é realizado em relação ao ponto central (PC) de \tilde{B} . Neste caso, B é simétrico ($\tilde{B} = B$) e PC representa, portanto, a origem das translações. Verifica-se que a erosão diminuiu o conjunto X.

Os efeitos obtidos com a erosão binária são:

- Diminuição das partículas (*pixels* pretos),
- Eliminação de grãos (*pixels* pretos) de tamanho inferior ao tamanho do elemento estruturante,
- Aumento dos buracos (*pixels* brancos), e

- Separação de grãos próximos.

4.1.2. Dilatação Binária

A dilatação de um conjunto X pelo elemento estruturante B pode ser representada pela seguinte fórmula:

$$\delta^B(X) = X \oplus \tilde{B} = \bigcup_{b \in \tilde{B}} X_b \quad (4.2)$$

onde \tilde{B} representa o transposto de B , \oplus representa a adição de Minkowski [Facon, 1996].

Conforme a definição acima, o conjunto X a ser dilatado, é deslocado em função das posições permitidas pelo elemento estruturante B . Os deslocamentos são realizados em relação ao ponto central de \tilde{B} . Veja-se a figura 4.5.

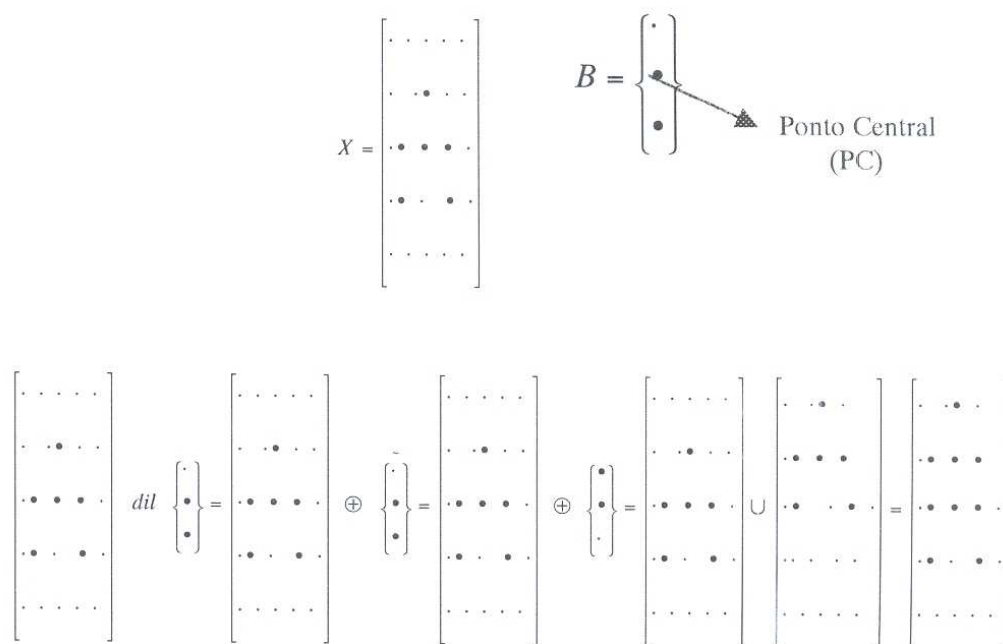
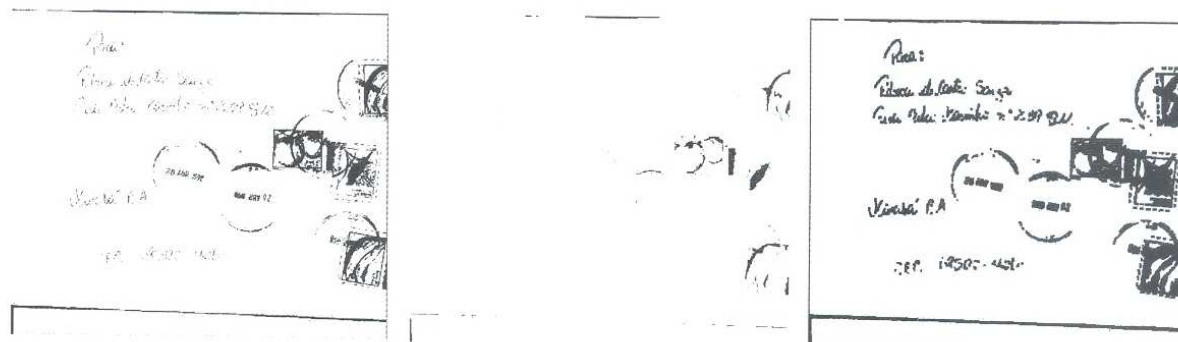


Figura 4.5: Exemplo dos deslocamentos dos pixels na dilatação.

X é movimentado em relação ao ponto central PC de \tilde{B} que representa, portanto, a origem das translações. PC faz parte dos pontos relevantes de \tilde{B} . Neste exemplo, o elemento estruturante não é simétrico; isto faz com que $B \neq \tilde{B}$. A dilatação é uma operação dual à erosão. Ela aumenta o conjunto X .

Os efeitos obtidos com a dilatação binária são:

- Expandir partículas,
- Preencher pequenos buracos, e
- Conectar grãos próximos.



a) imagem original limiarizada b) imagem erodida c) imagem dilatada

Figura 4.6: Resultado da erosão e dilatação com elemento estruturante quadrado com 2 iterações.

4.1.3. Abertura

Ao contrário da erosão e da dilatação, a abertura binária consiste em eliminar as partículas indesejáveis, sem modificar o tamanho das outras entidades. Intuitivamente, pode-se prever que esta operação consiste em erodir e, depois, dilatar o resultado da erosão.

A abertura de um conjunto X pelo elemento estruturante B pode ser representada pela seguinte fórmula:

$$\phi^B(X) = \delta^B(\varepsilon^{\tilde{B}}(X)) \quad (4.3)$$

A figura 4.7 apresenta uma imagem X com dois conjuntos conectados. A operação de abertura é aplicada para separar estes conjuntos.

$$X = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \quad B = \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\}$$

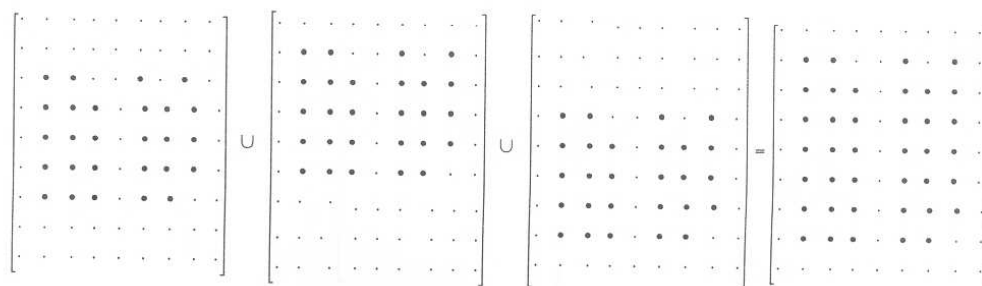


Figura 4.9: Exemplo da união da imagem X com X .

O resultado mostra os dois conjuntos de X desconectados e com tamanho e forma idênticos aos de X .

Os efeitos obtidos pela abertura binária são:

- Não devolver, de forma geral, o conjunto inicial,
- Nivelar os contornos pelo interior,
- Separar as partículas,
- Eliminar as pequenas partículas inferiores em tamanho, com relação ao elemento estruturante,
- As entidades restantes após a abertura ficam quase idênticas às originais,
- Conjunto aberto é mais regular que o conjunto inicial, e
- Conjunto aberto é menos rico em detalhes que o conjunto inicial.

4.1.4. Fechamento

O fechamento binário segue o mesmo raciocínio da abertura binária, apenas inverte a operação, ou seja, dilata e depois erode o resultado da dilatação.

O fechamento de um conjunto X pelo elemento estruturante B pode ser representado pela seguinte fórmula:

$$\gamma^B(X) = \varepsilon^B(\delta^{\tilde{B}}(X)) \quad (4.4)$$

A figura 4.10 apresenta uma imagem X com dois conjuntos desconectados. A operação de fechamento é aplicada para unir estes conjuntos.

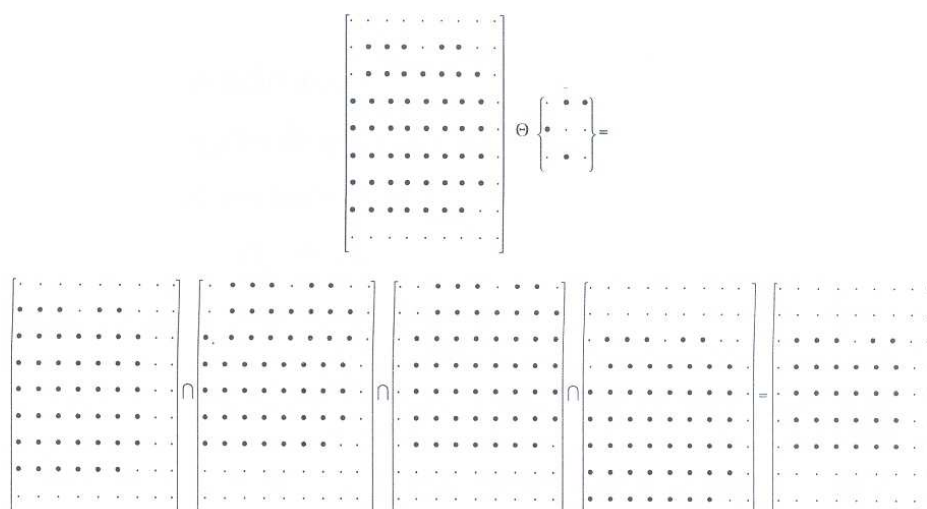
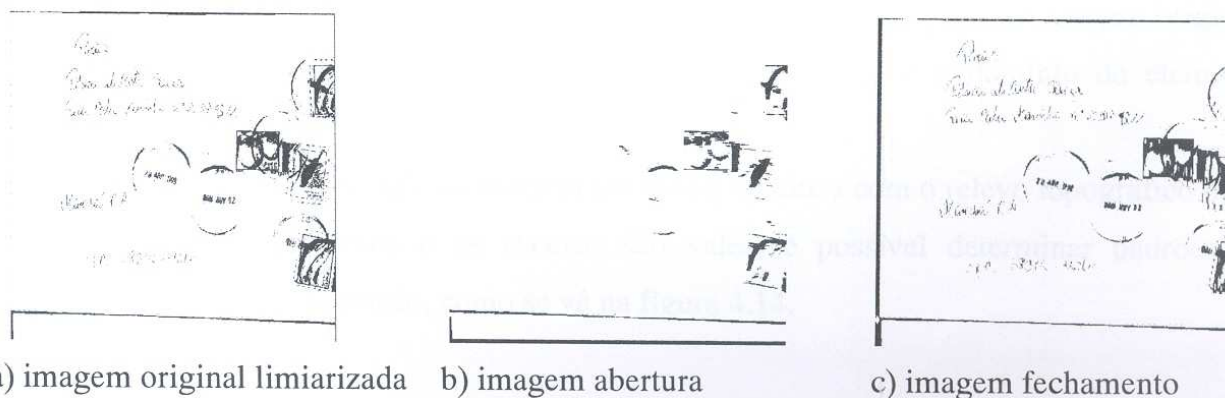


Figura 4.12: Exemplo da subtração e interseção de X com X.

O resultado final mostra os dois conjuntos de X conectados e com tamanho e forma idênticos ao conjunto X inicial.

Os efeitos obtidos pelo fechamento binário são:

- Suavizar as fronteiras pelo exterior,
- Preencher os buracos, inferiores em tamanho em relação ao elemento estruturante, no interior das partículas,
- Emendar partículas próximas,
- As entidades restantes após o fechamento ficam quase idênticas,
- Conjunto fechado é mais regular que o conjunto inicial, e
- Conjunto fechado é menos rico em detalhes que o conjunto inicial.



a) imagem original limiarizada b) imagem abertura c) imagem fechamento

Figura 4.13: Resultado da abertura com elemento estruturante vertical e fechamento com elemento estruturante horizontal com 4 iterações.

4.1.5. Reconstrução

O processo de reconstrução consiste em recuperar um conjunto S a partir de um conjunto Z , onde Z é um subconjunto de uma imagem binária X , e Z é uma amostra de S . Pode-se dizer que S é marcado por Z , portanto chama-se marcador.

A reconstrução binária $\rho_s(Z)$ de uma máscara (*conjunto finito*) S a partir do marcador Z ($Z \subset S$), usando o elemento B , é:

$$\rho_s(Z) = \lim_{n \rightarrow \infty} \delta_{cs}^B (\dots \delta_{cs}^B Z) \quad (4.5)$$

onde $\delta_{cs}^B(Z) = \delta^B(Z) \cap S$ representa a dilatação condicional relativamente a S Facon [1996].

A reconstrução permite eliminar os padrões indesejáveis sem deformar os desejáveis.

4.2. Operadores morfológicos em níveis de cinza

Nessa morfologia é necessário conhecer o valor do *pixel* mais escuro MIN e o valor do *pixel* mais claro MAX, na vizinhança de cada *pixel* ou numa parte da vizinhança da imagem original. Uma operação morfológica em níveis de cinza é completamente determinada através do tamanho e forma da vizinhança, das regiões de pesquisa de MIN e MAX e do algoritmo.

4.2.1. Dilatação

A dilatação de um sinal f por um elemento estruturante g pode ser representada pela seguinte equação:

$$\delta^g(f(x)) = \max\{f(y) + g(x - y) : y \in D[g]\} \quad (4.6)$$

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 [FACON, 1996].

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 se vê na figura 4.14.

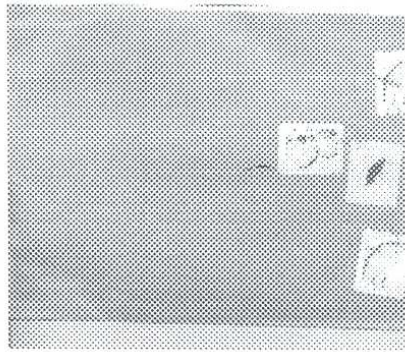


Figura 4.14: Exemplo de dilatação 2 iterações.

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*), e
- Separar vales.

4.2.2. Erosão

A erosão de um sinal f por um elemento estruturante g pode ser representada pela seguinte equação:

$$\varepsilon^g(f(x)) = \min\{f(y) - g(x - y) : y \in D[g]\} \quad (4.7)$$

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 [FACON, 1996].

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 4.15.

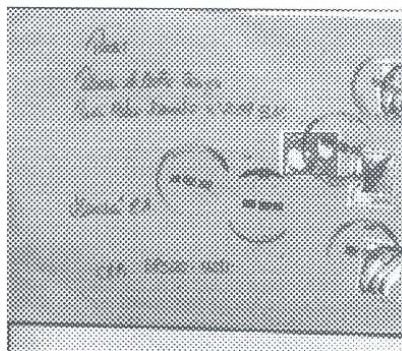


Figura 4.15: Exemplo de erosão 2 iterações.

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*), e
- Separar picos próximos.

4.2.3. 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 [Facon, 1996]:

$$abe^g(f) = dil^g(ero^g(f)) \quad (4.8)$$

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, e
- A imagem aberta é menos rica em detalhes que a imagem original.

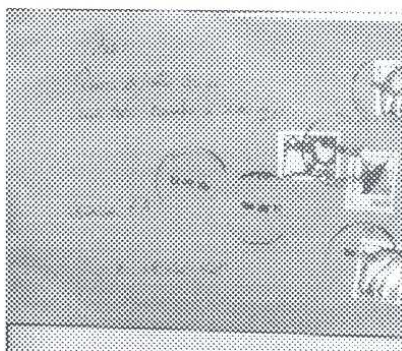


Figura 4.16: Exemplo de abertura 2 iterações.

4.2.4. Fechamento

O fechamento de um sinal f por um elemento estruturante g é definido pela fórmula [Facon, 1996]:

$$fec^g(f) = ero^g(dil^g(f)) \quad (4.9)$$

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

$$fec^g(f) = -abe^g(-f) \quad (4.10)$$

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, e
- A imagem fechada é menos rica em detalhes que a imagem original.

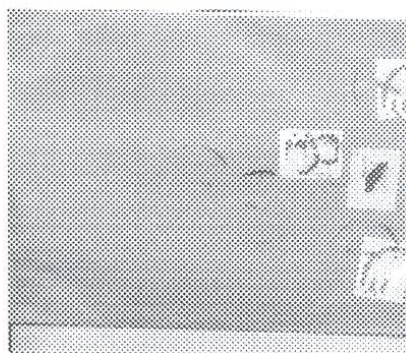


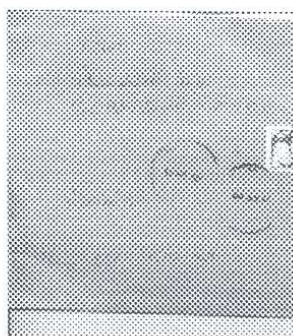
Figura 4.17: Exemplo de fechamento 2 iterações.

4.2.5. 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 pode ser representado pela fórmula:

$$tophat_{abe}^g(f) = f - abe^g(f) \quad (4.11)$$

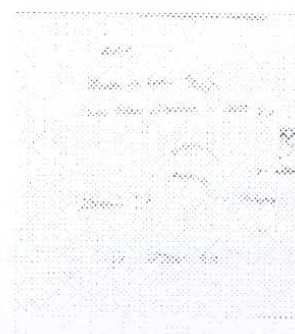
Como a abertura é um processo anti-extensivo, o seu resultado fica abaixo do sinal original. Com o uso de um elemento estruturante adequado, o processo de abertura permite a eliminação dos picos. Fazer a diferença, entre os sinais originais e o resultado da abertura, permite tirar o ruído e eliminar a falta de homogeneidade, ou seja, ressaltar a informação dos picos da imagem [FACON, 1996].



a) imagem original cinza



b) tophat abertura



c) tophat invertida

Figura 4.18: Diferença entre a imagem original cinza e a imagem tophat aberta com 4 iterações e elemento estruturante quadrado.

4.2.6. 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 pode ser representado pela fórmula:

$$tophat_{fec}^g(f) = fec^g(f) - f \quad (4.12)$$

Como o fechamento é um processo extensivo, o seu resultado da transformação é sempre positivo, como se vê na figura 4.19.

Com o uso de um elemento estruturante adequado, o processo de fechamento permite a eliminação dos vales. Fazer a diferença, entre os sinais originais e o resultado do fechado, permite tirar o ruído e eliminar a falta de homogeneidade, ou seja, ressaltar a informação dos vales da imagem [Facon, 1996].

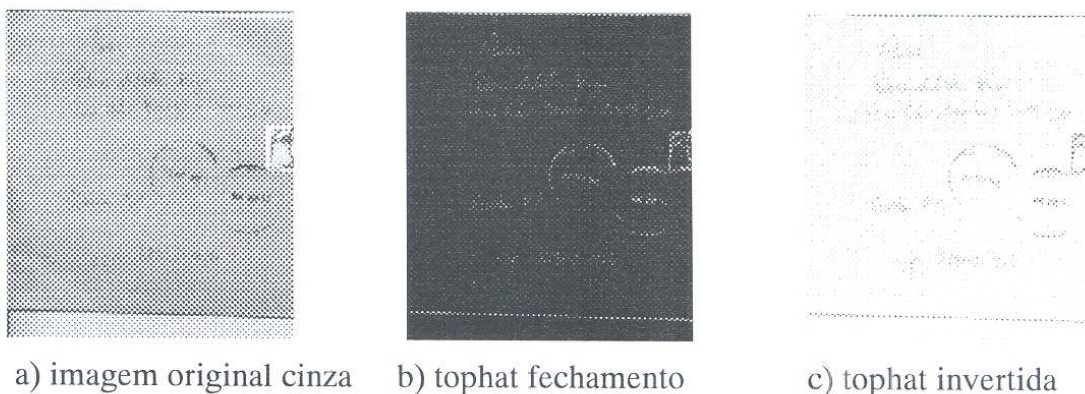


Figura 4.19: Diferença entre a imagem original cinza e a imagem tophat fechada com 4 iterações e elemento estruturante quadrado.

As técnicas de segmentação são processos utilizados na morfologia binária e níveis de cinza para corrigir imagens de suas fragmentações, ou seja, ruídos, elementos estes que não fazem parte da imagem, ou não se quer que o faça. Destaca-se aí o pré-processamento para melhorar a qualidade da imagem obtida quando escaneada, e a segmentação tem o objetivo de fornecer os blocos de todos os componentes que fazem parte da imagem, separando-os para localização e o futuro reconhecimento de cada bloco do endereço.

Neste capítulo foram abordadas as técnicas sobre processamento digital de imagens e morfologia matemática binária e em níveis de cinza, onde foram mostradas as operações que darão sustentação para as implementações necessárias neste trabalho. No capítulo seguinte será apresentada a metodologia aplicada para desenvolver o trabalho proposto, assim como o

processamento necessário para obter a segmentação das imagens dos envelopes dos Correios em níveis cinza.

Capítulo 5

Método Proposto

Este capítulo apresenta a metodologia utilizada no desenvolvimento do trabalho e a descrição de cada uma das etapas que o constituem. O problema a ser tratado consiste na identificação e ordenação dos processos necessários para a segmentação e localização, em uma imagem postal, dos elementos que a compõem.

5.1. Visão Geral

Percebe-se que as dificuldades existentes neste campo concentram-se na heterogeneidade das imagens e no acúmulo de ruídos ou elementos indesejáveis que interferem na interpretação das informações ali contidas. Em função disso, o processo de manipulação das imagens exige o desenvolvimento de uma automação mais confiável no que se refere ao reconhecimento das imagens. Em virtude do rápido desenvolvimento da tecnologia de informação e das necessidades por ele geradas, os sistemas computadorizados devem desempenhar o papel de ler e interpretar, de maneira eficiente, os documentos preparados. Pensando na eficiência do processo em estudo, limitou-se o trabalho a um tipo de imagem em tons de cinza, a fim de melhor comparar e selecionar as técnicas mais adequadas à implementação do processo.

Com o estudo desta base de imagens pretende-se ampliar a pesquisa na localização do bloco de endereço em envelopes postais, facilitando o acesso e a manipulação das imagens digitalizadas das correspondências dos Correios. Para tanto é necessário que se faça uma preparação das mesmas, tal que, futuramente permita além da localização, também a identificação dos blocos de endereço do destinatário, ou outras informações relevantes da imagem.

Para realizar essas tarefas o trabalho foi dividido em duas etapas:

- a primeira, a atualização quanto aos métodos e técnicas já propostos, bem como o suporte teórico necessário, os quais foram estudados e apresentados nos Capítulos 2, 3 e 4.
- a segunda fase consiste na criação de uma base de testes de imagens postais, a qual será submetida à experimentação através do método considerado mais adequado para a segmentação, na experimentação, foram utilizadas as rotinas de morfologia matemática disponíveis no *software*.

A plataforma de desenvolvimento empregada na implementação deste trabalho foi o Microsoft Visual C++ 6.0, e o *software* utilizado como base no desenvolvimento deste trabalho foi o BROPI, implementado no LUCI.

A figura 5.1 indica a seqüência das etapas que constituem o processo proposto, a partir da imagem escaneada do envelope postal original, designada aqui de imagem original, iniciando com o pré-processamento até obtermos a imagem do resultado final.

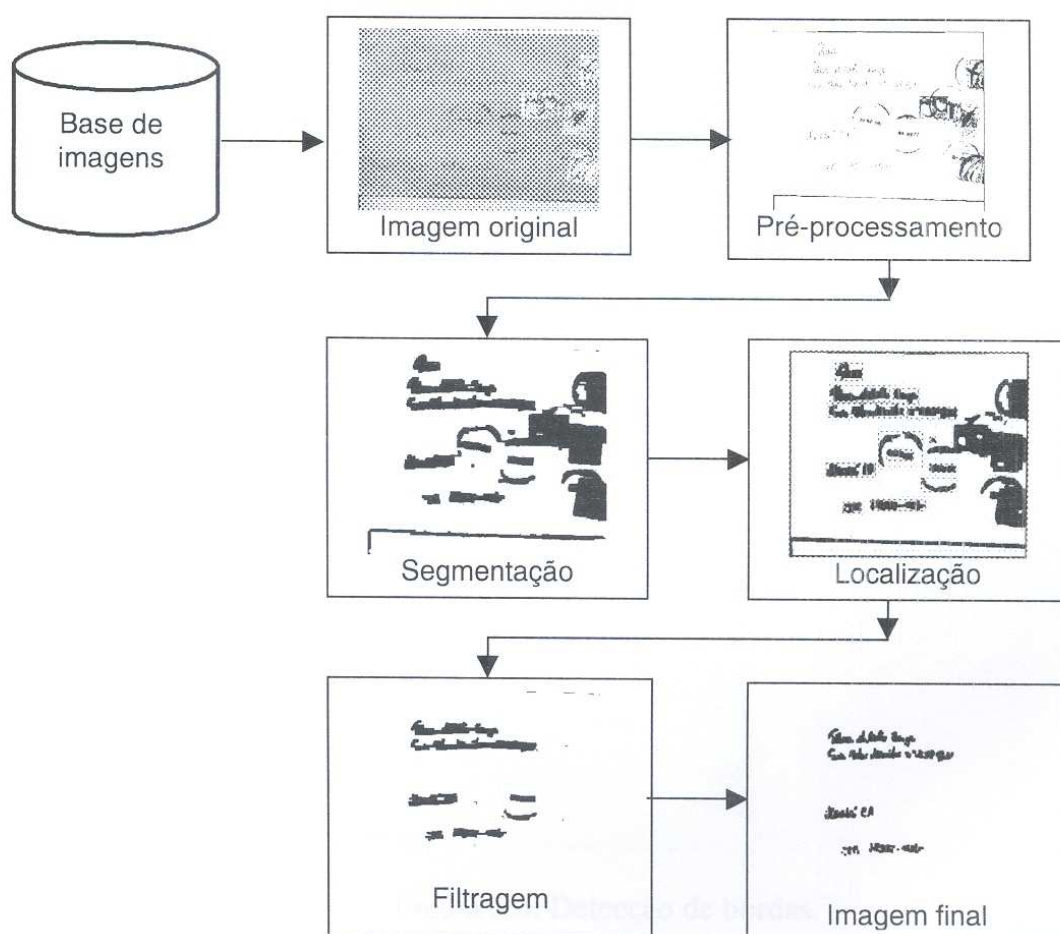


Figura 5.1: Etapas do processamento.

5.2. Desenvolvimento

Nesta etapa são detalhadas as fases que constituem o processamento necessário para se obter a localização dos blocos de endereço nas imagens digitalizadas de envelopes postais em níveis de cinza.

5.2.1. Pré-processamento

O pré-processamento foi realizado em duas abordagens distintas: abordagem baseada na detecção de bordas e abordagem baseada por tophat.

a) **na abordagem baseada na detecção de bordas**, a imagem original (figura 5.2) é obtida a partir do acesso ao arquivo da sub-base de testes criada para este trabalho.



Figura 5.2: Imagem original.

A aplicação da detecção de bordas na imagem original, destaca as bordas das regiões escuras, facilitando assim a marcação dos manuscritos, carimbos, selos e outros elementos com relação ao fundo (figura 5.3).

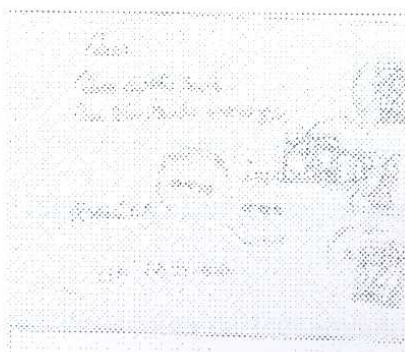


Figura 5.3: Detecção de bordas.

A erosão (figura 5.4) tem como objetivo desgastar, ou seja, neste caso borrar a imagem de tal forma que seja conservado um número de *pixels* suficiente para não alterar a natureza da imagem com elemento estruturante quadrado e sete iterações.

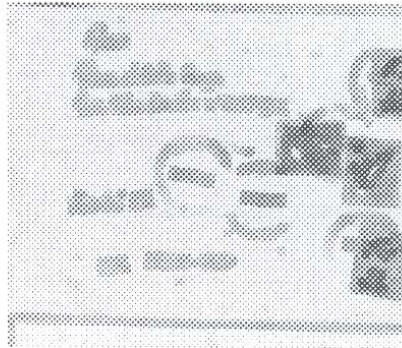


Figura 5.4: Bordas erodidas.

A técnica de limiarização por Otsu tem por objetivo encontrar uma imagem binária de boa qualidade para o processamento (figura 5.5). A vantagem deste método de limiarização está no fato de não utilizar nenhum conhecimento *a priori* da imagem. A utilização deste algoritmo limpa a imagem, deixando somente as partes de maior concentração binária em destaque, eliminando o fundo.

Convém lembrar que o processo de limiarização desgasta a imagem, o que, em determinadas circunstâncias, pode gerar perdas parciais ou totais de informações.

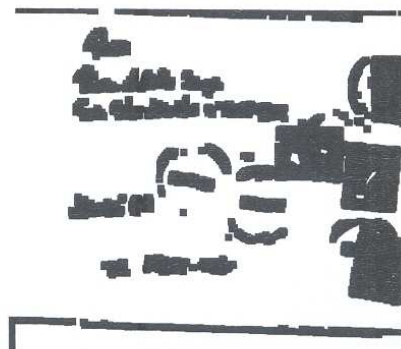


Figura 5.5: Limiarização por Otsu.

O objetivo da aplicação da técnica de fechamento em uma imagem, é a aproximação ou agrupamento das letras e/ou componentes próximos, formando assim, blocos com *pixels* de mesma intensidade, ou seja, borra a imagem (figura 5.6). Neste caso, utilizamos o elemento estruturante horizontal com cinco iterações.

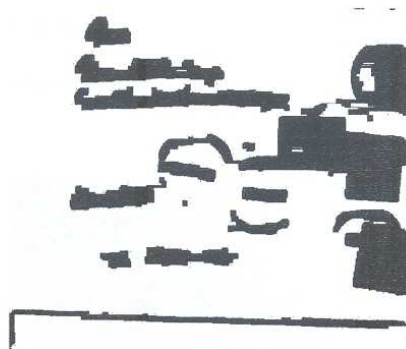


Figura 5.6: Fechamento horizontal, imagem borrada.

A aplicação da técnica de abertura com elemento estruturante vertical e três iterações, tem como objetivo a eliminação de ruídos na horizontal (figura 5.7), com o intuito de limpar mais a imagem e, em certas situações, dependendo do número de iterações realizadas, tem-se a separação de blocos que estavam conexos.

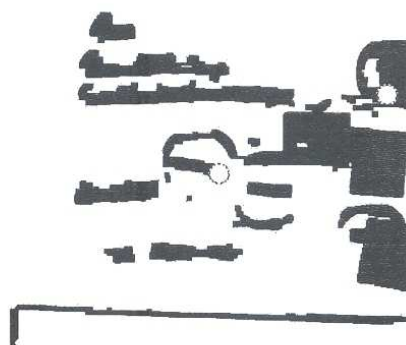


Figura 5.7: Abertura vertical e separação de blocos

b) na abordagem baseada por tophat, a imagem original (figura 5.8) é obtida a partir do acesso ao arquivo da sub-base de testes criada para este fim.

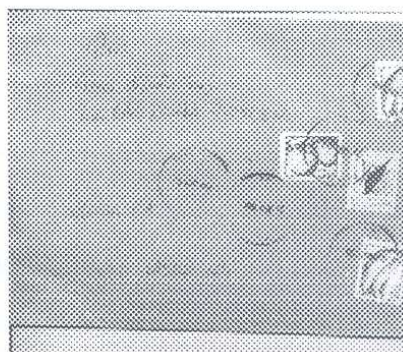


Figura 5.8: Imagem original.

Tophat por fechamento (figura 5.9) com elemento estruturante quadrado e três iterações tem como objetivo ressaltar os vales (*regiões escuras*) numa imagem, facilitando assim a marcação das partes escritas e os componentes que fazem parte da imagem postal.

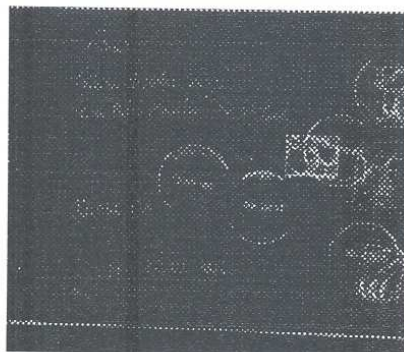


Figura 5.9: Tophat por fechamento.

O objetivo da inversão do tophat (figura 5.10) é deixar claro o fundo da imagem, convertendo os vales em picos e vice-versa, tendo em vista a melhor qualidade da binarização da imagem que será feita a seguir.

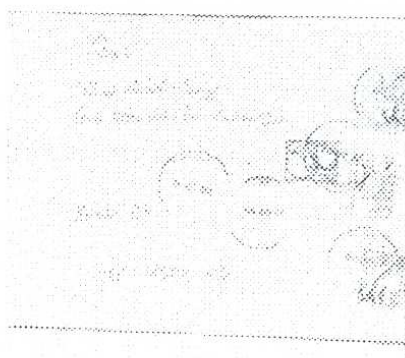


Figura 5.10: Tophat por fechamento invertido.

A técnica de limiarização por Otsu (figura 5.11) tem como objetivo encontrar uma imagem binária de boa qualidade para o seu processamento.

Entre as técnicas de limiarização conhecidas, como Entropia, Anisotropia, Kitler e Otsu, a última foi escolhida, cuja descrição encontra-se no Capítulo 3, pelo fato de que o seu algoritmo fornece uma imagem mais clara, com menor número de ruídos, *deixando somente* as partes mais escuras em destaque, eliminando assim, o fundo. Este método de limiarização tem a vantagem de não utilizar nenhum conhecimento *a priori* da imagem.

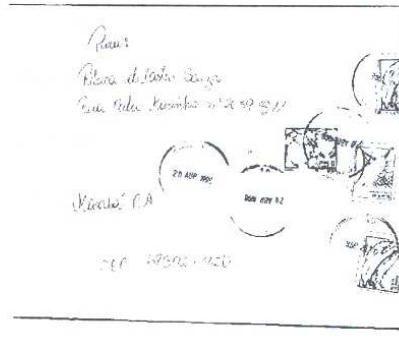


Figura 5.11: Limiarização por Otsu.

Aplicando a técnica da dilatação (figura 5.12) com elemento estruturante quadrado e três iterações tem como objetivo a dilatação das bordas, ou seja, borrar a imagem, procurando com isso agrupar *pixels* de mesma intensidade, formando blocos conexos dos componentes da imagem.

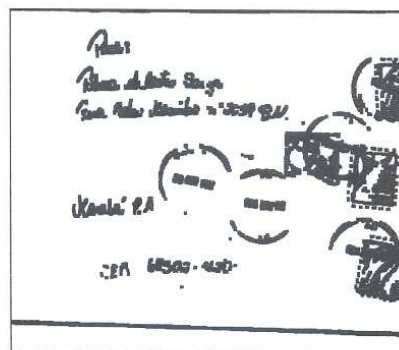


Figura 5.12: Dilatação quadrada.

O fechamento com elemento estruturante horizontal (figura 5.13) e vinte iterações tem o objetivo de eliminar ruídos através da dilatação e erosão, preenchendo os buracos que lhes são menores, o que melhora a qualidade da segmentação, tornando mais unidas as palavras e/ou componentes que fazem parte da imagem.

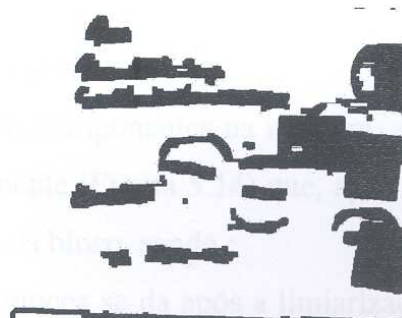


Figura 5.13: Fechamento horizontal.

O principal objetivo da abertura com elemento estruturante vertical (figura 5.14) e três iterações é a separação de blocos sem modificar-lhes o tamanho e a forma, separando as partículas e eliminando as de tamanho inferior ao elemento estruturante. O resultado é um conjunto mais regular e menos rico em detalhes que o conjunto inicial, ou seja, uma imagem mais limpa.

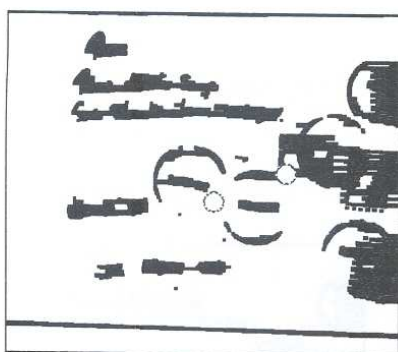


Figura 5.14: Abertura vertical.

5.2.2. Segmentação dos Blocos de Componentes

Esta etapa visa obter os blocos que indicam as linhas escritas contidas na imagem e seus componentes não escritos, como carimbos, selos e ruídos.

A segmentação da imagem é realizada através do algoritmo *Run Length Smoothing Algorithm* - RLSA composto das operações morfológicas descritas anteriormente.

Os elementos estruturantes utilizados nos processos morfológicos que compõem a segmentação da imagem possuem tamanho de 3x3 para o quadrado, 3x1 na horizontal e 1x3 na vertical. O número de iterações foi determinado empiricamente, observando-se os resultados obtidos com o processamento das imagens que compõem a base de testes.

Terminados estes processos morfológicos, passou-se para a localização dos componentes conexos da imagem resultante da operação.

5.2.3. Detecção dos Componentes Conexos

Para a localização dos componentes na imagem, aproveita-se a imagem resultante do processo realizado anteriormente (Figura 5.14) que, após realizado outros recursos, indicará a posição e a localização de cada bloco, sendo :

- A formação dos blocos se dá após a limiarização, onde é realizada uma busca por *pixels* binário cujo nível seja de mesma intensidade. Uma vez localizado um destes, todos os seus vizinhos que possuem um nível igual ou maior que o nível

correspondente a região de fundo da imagem, são considerados como parte do mesmo objeto;

- A demarcação de cada bloco acontece em função da homogeneização dos *pixels* nas regiões das imagens;
- A localização deste componente é determinada anotando-se a posição dos *pixels* mais a esquerda, a direita, acima e abaixo que estão ligados pela 8-vizinhança ao primeiro *pixel* encontrado.

A figura 5.15 mostra o resultado desse processo, onde os componentes na imagem aparecem conectados dentro de cada bloco.

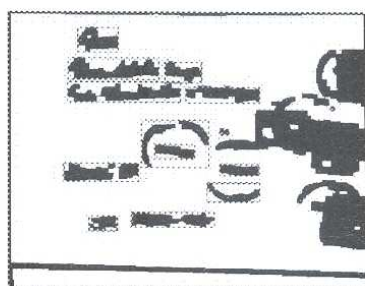


Figura 5.15: Imagem com blocos localizados.

5.3. Imagem Final

Terminada a etapa de segmentação e a contagem dos componentes, passa-se para a eliminação dos componentes não necessários ao resultado final da imagem. Busca-se deixar somente os blocos que devem fazer parte do endereço postal, conforme se descreve a seguir:

- No início de cada processo é criada uma imagem de contorno (*bordas*), do tamanho da imagem original, a qual será utilizada para o cálculo da diferença entre a imagem original e a imagem de bordas, afim de eliminar as bordas horizontal, vertical e os componentes que estão conexos a elas, ficando somente a parte mais interna da imagem, ou seja, o que não está tocando nas bordas;
- Numa segunda etapa, é realizado o processo da média aritmética das alturas e larguras dos blocos remanescentes. Utilizando-se da média das alturas, é calculado o desvio padrão das mesmas, o qual será usado para eliminar os blocos que estiverem fora do intervalo da média mais duas vezes o desvio padrão e média menos duas vezes o desvio;

- Na terceira etapa, é feita uma abertura com elemento estruturante quadrado com uma iteração, com objetivo de limpar pequenos ruídos que por ventura tenham ficado após a realização dos processos, e
- No último passo, temos a imagem resultante dos processos anteriores que serve como marcador para o processo de reconstrução binária, utilizando a imagem original limiarizada como máscara, obtendo-se assim a imagem final, conforme figura 5.16.

A aplicação destas regras no conjunto de componentes localizados na imagem, resulta na manutenção dos mesmos, que provavelmente se encontram na região do bloco-endereço.



Figura 5.16: Imagem final.

5.3.1. Reconstrução Dual

Utilizando a imagem de marcadores, resultado do processo descrito anteriormente, e como máscara a imagem resultante do processo de eliminação da complexidade da imagem, é aplicado o processo de reconstrução dual [Facon, 1996].

Teoricamente, o resultado desta reconstrução, mostrado na figura 5.16, deve apresentar as linhas com os conteúdos que compõem o endereço dos envelopes postais, fato que demonstra a manutenção dos blocos de interesse da imagem. Porém, poderá acontecer na reconstrução ruídos que por ventura não foram totalmente eliminados, ficando alguns marcadores.

As rotinas que compõem o método proposto, estão otimizadas através da sua implementação, considerando-se a utilização de memória e performance dos algoritmos.

Neste capítulo foi apresentada a metodologia aplicada para desenvolver o processamento necessário à obtenção da segmentação de uma imagem em níveis de cinza. Dentro deste processamento, destaca-se o pré-processamento, cuja finalidade é melhorar a

qualidade da imagem escaneada, o processamento da imagem e a segmentação que têm por objetivo encontrar e localizar os blocos de endereço dos envelopes postais.

No capítulo seguinte serão apresentados os resultados obtidos na aplicação do método aqui proposto sobre a base de dados criada para execução do trabalho. Também serão apresentadas as análises estatísticas as quais abrem a discussão a respeito do desempenho do método em relação aos resultados alcançados.

Capítulo 6

Resultados Experimentais

O trabalho com análise de imagens de envelopes postais exige a utilização de uma base de imagens para o seu desenvolvimento. A disponibilidade de uma biblioteca contendo um número significativo de imagens é um fator crítico para qualquer projeto de pesquisa experimental nesta área.

A análise dos resultados obtidos na experimentação realizada durante o desenvolvimento da presente dissertação, ficaria melhor se compreendida antes, porém destacadas e descritas as características principais da base de imagens que lhe serviu como objeto de trabalho.

6.1. Base de Imagens Postais da PUCPR

A necessidade de criação de uma base de dados, contendo imagens que reflitam a realidade brasileira e a necessidade da utilização de um sistema de análise de imagens de envelopes no Brasil, fomentaram a criação da base de imagens, ou biblioteca de imagens da PUCPR. Eis aí uma tarefa difícil, pois tais imagens são de caráter sigiloso, podendo ser utilizadas apenas para a finalidade de pesquisa. Posto isso, o seu uso fica restrito ao LUCI – Laboratório Unificado de Ciências da Imagem da PUCPR.

6.1.1. Descrição da Base de Imagens da PUCPR

A base de imagens, cedida pelos Correios à PUCPR, é composta por 50.000 imagens postais dos Correios brasileiros, de formatos variados e divididos em quatro partes, conforme descrito a seguir.

Tabela 6.2: Características da base SE 256 níveis de cinza.

Tipo de objeto	Quantidade de imagens	Quantidade de CD's	Características
Formato semi-embaraçoso	20.000	45	TIFF 5.0 Compressed 256 tonalidades de cinza 100% escala 200 dpi

A figura 6.2 é uma imagem de envelope postal manuscrito, no formato semi-embaraçoso de fundo complexo, na qual é objeto de estudo do trabalho em questão, disponível no LUCI para esta e outras pesquisas.



Figura 6.2: Imagem tipo SE 256 NC 200dpi.

Tabela 6.3: Características da base FN 16 tonalidades de cinza.

Tipo de objeto	Quantidade de imagens	Quantidade de CD's	Características
Formato normal	5.000	01	TIFF 5.0 Compressed 16 tonalidades de cinza 100% escala 150 dpi

Na figura 6.3, temos uma imagem de envelope postal classificado como formato normal manuscrito, com 16 níveis de cinza, também disponível no LUCI para futuros trabalhos de pesquisa.



Figura 6.3: Imagem tipo FN 16 NC 150dpi.

Tabela 6.4: Características da base SE 16 tonalidades de cinza.

Tipo de objeto	Quantidade de imagens	Quantidade de CD's	Características
Formato semi-embarçoso	5.000	04	TIFF 5.0 Compressed 16 tonalidades de cinza 100% escala 150 dpi

A figura 6.4 é uma imagem postal de formato semi-embarçoso, pré-impresso disponível para pesquisa no LUCI.



Figura 6.4: Imagem tipo SE 16 NC 150dpi.

6.1.2. Definição da Base de Imagens

No momento em que as imagens foram digitalizadas e armazenadas em arquivos magnéticos, foram selecionadas 20.000 imagens no formato semi-embarçoso em 256 níveis de cinza com escala de 100% e 200 dpi de envelopes postais manuscritos e pré-impresos.

Desta sub-base foi, então, retirada a amostra para testes destinada a este trabalho, selecionando apenas as imagens manuscritas, através da abertura de cada arquivo e identificando-as como manuscrita ou não. Tal procedimento foi aplicado em 10.000 das 20.000 imagens da referida sub-base, sendo que o percentual de envelopes postais manuscritos, em cada CD, composto de aproximadamente 450 imagens.

6.1.3. Base de Testes

A base de testes em questão compõe-se de 120 imagens digitalizadas, correspondentes à parte externa dos envelopes postais, oriundas dos Correios, nos padrões já especificados, sendo que foram considerados os envelopes postais classificados como 'semi-embarçosos' ou de fundo complexo, manuscritos em níveis de cinza e de diferentes tamanhos. Tal escolha se justifica pelo fato de apresentarem estruturas físicas bastante semelhantes.

A base das imagens para os testes foi dividida em outras duas sub-bases, sendo que, a utilizada para testes durante o desenvolvimento do algoritmo contém 40 imagens, sendo 1/3 da base de testes e a que foi utilizada para a validação final das rotinas e verificação do método proposto contém as 80 imagens restantes, ou seja, 2/3 da base de testes.

6.2. Resultados Gerais

Os resultados a seguir estão de acordo com os critérios de avaliação construídos visualmente pelo pesquisador, o qual se utilizou das imagens obtidas após processamento completo. Estes critérios foram submetidos às análises realizadas por um grupo de especialistas na área, participantes de programas desenvolvidos na PUCPR, os quais referendaram os referidos critérios.

Os critérios assim construídos apresentam a seguinte descrição:

- a) **Muito bom (MB):** foram consideradas as imagens que apresentaram a localização de todas as linhas de endereço. Caso haja a presença de alguns ruídos, os mesmos poderão ser desconsiderados, desde que não comprometam o futuro processo de reconhecimento dos blocos;

- b) **Bom(B)**: foram consideradas as imagens nas quais tenham sido perdidas partes de uma das linhas do endereçamento, mas não prejudica o resultado da informação, podendo ser recuperável. Podem apresentar alguns ruídos desde que os mesmos não interfiram nos processos futuros;
- c) **Regular(REG)**: foram consideradas as imagens nas quais foram perdidas partes de mais de uma linha do bloco do endereçamento, que poderá prejudicar a leitura da informação;
- d) **Ruim(R)**: neste caso, o processo, ao invés de limpar a imagem fez com que desaparecesse grande parte do endereçamento ou, então, não tornou possível a eliminação de ruídos que comprometem o processamento da mesma.

Após a completa realização dos testes, as 80 imagens, que constituem a amostra da base dos testes destinada à validação, puderam ser classificadas como mostra a tabela 6.5, a seguir:

Tabela 6.5 :Resumo estatístico.

Técnica		Muito bom	Bom	Regular	Ruim	Total
Deteção de bordas	Figuras	37	26	7	10	80
	Percentual	46,25%	32,5%	8,75%	12,5%	100%
Tophat	Figuras	28	31	10	11	80
	Percentual	35%	38,75%	12,5%	3,75%	100%

Tabela 6.6: Resumo acumulado dos critérios muito bom e bom

Técnica		Bom
Deteção de bordas	Figuras	63
	Percentual	78,75%
Tophat	Figura	59
	Percentual	73,75%

6.3. Análise dos Resultados

A seguir estão expostos e comentados exemplos do processamento completo de imagens classificadas dentro de cada critério, a partir da aplicação das técnicas de detecção de bordas e Tophat. As fases do processamento podem ser identificadas pela legenda abaixo.

Detecção de bordas:

- a) imagem original,
- b) detecção de borda com elemento estruturante quadrado,
- c) erosão com elemento estruturante quadrado e 3 iterações,
- d) limiarização por Otsu,
- e) fechamento com elemento estruturante horizontal e 10 iterações,
- f) abertura com elemento estruturante vertical e 3 iterações,
- g) localização dos blocos, e
- h) reconstrução da imagem final.

Tophat:

- a) imagem original,
- b) tophat fechamento com elemento estruturante quadrado e 3 iterações,
- c) inversão da imagem,
- d) limiarização por Otsu,
- e) dilatação com elemento estruturante quadrado e 3 iterações,
- f) fechamento elemento estruturante horizontal e 20 iterações,
- g) abertura com elemento estruturante vertical e 3 iterações,
- h) localização dos blocos, e
- i) reconstrução da imagem final.

6.3.1. Detecção de bordas

Na figura 6.5 temos a seqüência do processo da detecção de bordas, no critério muito bom.

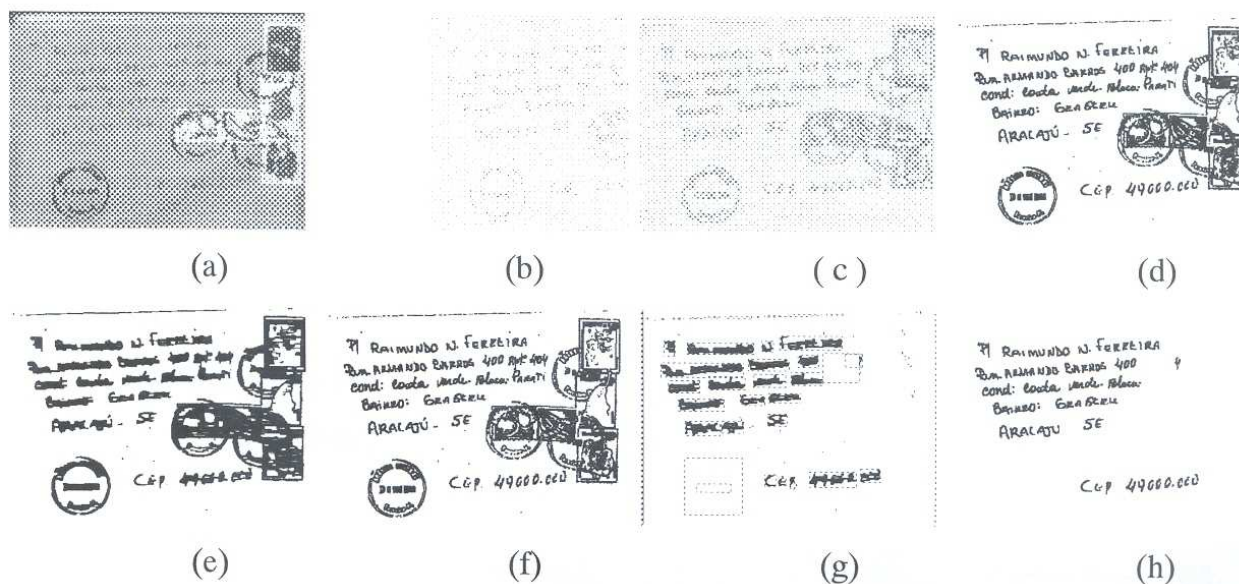


Figura 6.5: Imagens do critério muito bom por detecção de bordas.

Observando, atentamente, as imagens obtidas em cada fase do experimento, conclui-se que a imagem final (h) pode ser incluída neste critério porque não contém os blocos correspondentes às imagens dos selos, carimbos e os ruídos, o qual faziam parte da imagem original. Pode-se observar que é uma imagem limpa e clara, ficando os blocos disponível para uma futura rotulação.

A figura 6.6 é mostrado a seqüência e o resultado do processo para o critério considerado bom (h).

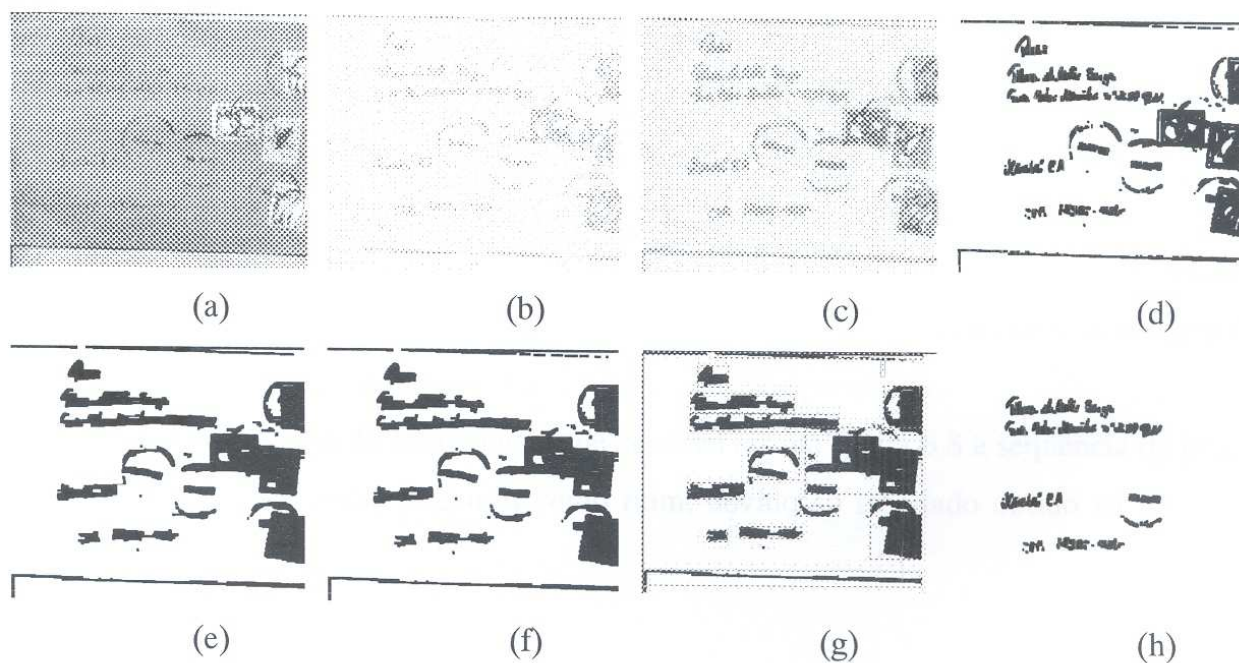


Figura 6.6: Imagens do critério bom por detecção de bordas.

A visualização das imagens, obtidas no processo acima, fornece elementos suficientes para classificá-la neste critério, pois é perceptível a perda de informações na primeira linha do endereçamento. Todavia, os ruídos ainda existentes localizam-se em regiões separadas dos blocos de endereço. Ao final, a imagem apresenta-se limpa, com blocos desconexos, o que facilita a obtenção da marcação precisa de blocos, entretanto, com perda de informações diminui a qualidade nos processos futuros.

Nesta seqüência da figura 6.7, temos como resultado final no item (h) o que é considerado como regular.

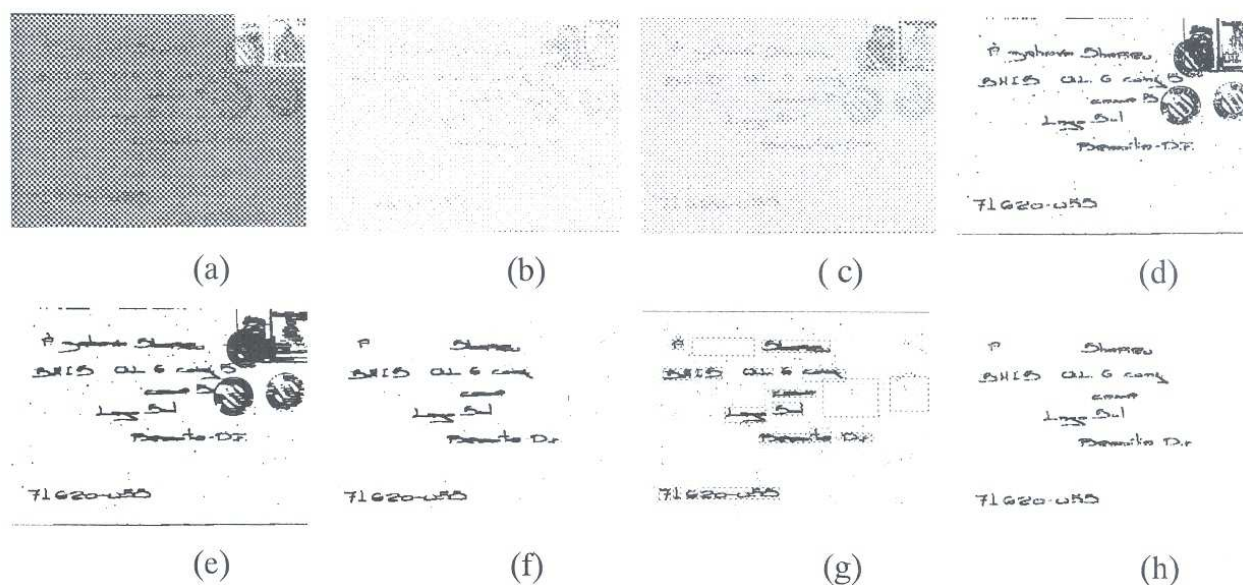


Figura 6.7: Imagens do critério regular por detecção de bordas.

Um olhar atento e crítico, dirigido à seqüência de imagens acima, permite classificar a imagem final como regular, visto que são visíveis as perdas de informações nos blocos da primeira, segunda e terceira linhas do endereçamento. São perdidos os números que completam o endereçamento, pois os mesmos foram incorporados aos blocos das imagens de carimbos e selos, cujos tamanhos excedem ao estabelecido, sendo eliminados. A imagem final pode até ser dita limpa, mas faltam-lhes informações indispensáveis.

Na mesma linha de raciocínio, pode-se observar na figura 6.8 a seqüência de imagens que compõe o critério que julgamos como ruim, devido ao resultado obtido na imagem do item (h).

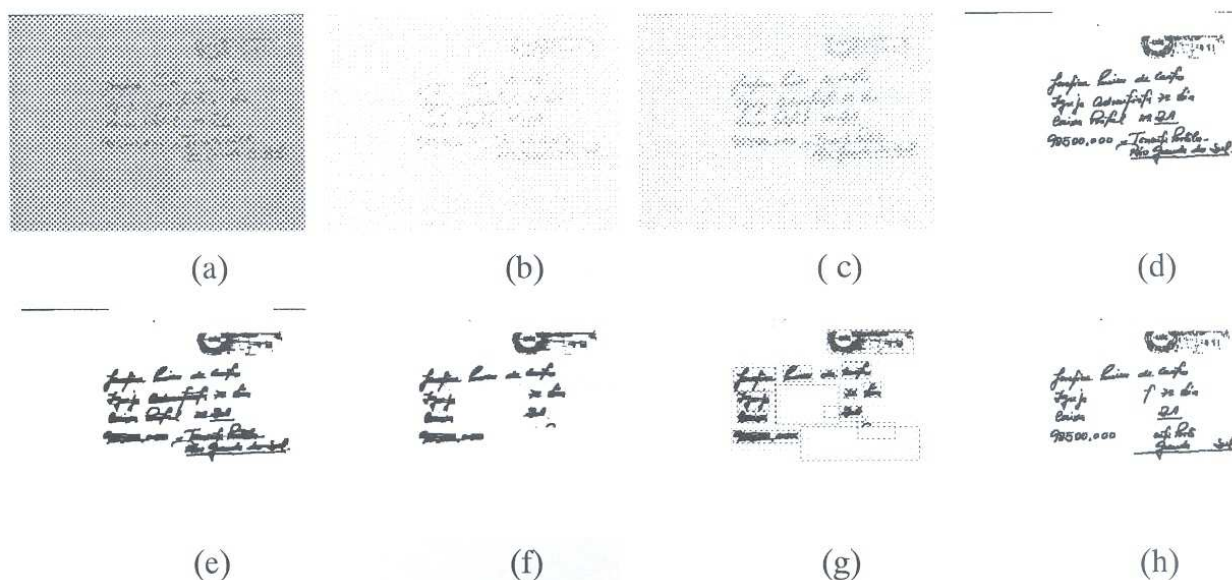


Figura 6.8: Imagens do critério ruim por detecção de bordas.

A referida imagem, após o processamento, pode ser avaliada como ruim, pois são muito evidentes as perdas ocorridas durante o processo, em 3 das 4 linhas do endereçamento. Tais perdas se dão devido ao traçado gráfico, o qual é alongado e ocasiona a aproximação extrema dos blocos formados, sua conexão e conseqüente eliminação. Ocorrem superposições entre blocos de ruídos e de endereçamento. Nota-se ainda a não eliminação do bloco correspondente ao carimbo no canto superior direito. A imagem final apresenta-se com muitas falhas, as quais comprometem severamente o processamento futuro desta imagem.

6.3.2. Tophat

Na figura 6.9 temos a seqüência do processo tophat, tendo como resultado final no item (i) o critério considerado como muito bom.

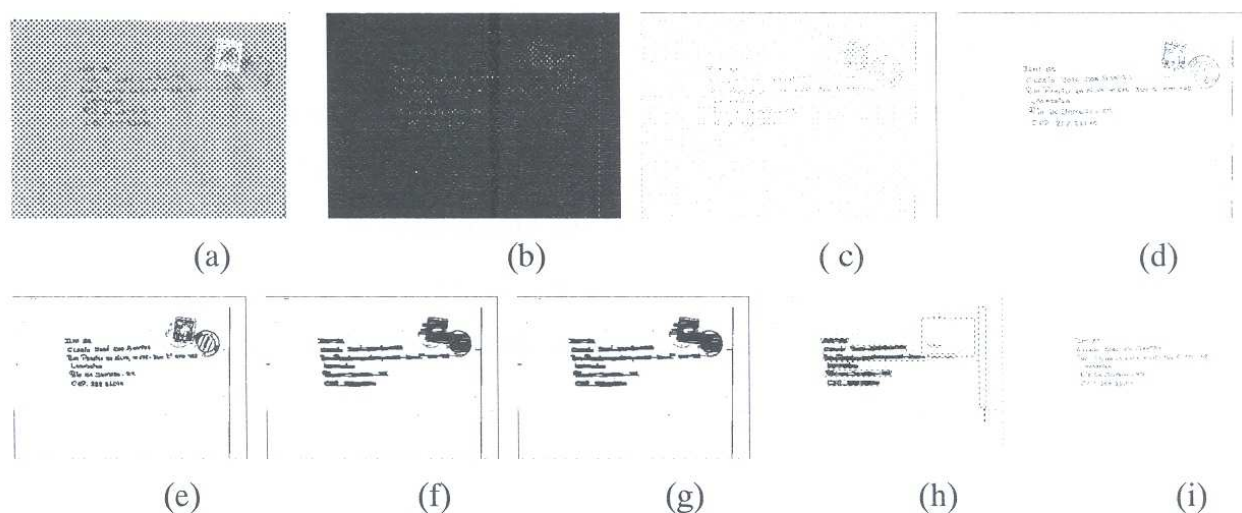


Figura 6.9: Imagens do critério muito bom por Tophat.

O conjunto de imagens acima oferece uma imagem final de muito boa qualidade, pois os ruídos referentes a carimbos e selos foram devidamente eliminados e não ocorreu nenhuma perda de informações nos blocos de endereçamento. O traçado das letras é bastante regular e as palavras estão bem distribuídas no envelope. Carimbos e selos encontram-se em região adequada do envelope, não há portanto superposição dos elementos que formam os blocos, por isso, a imagem se mostra nítida, clara e completa.

Na figura 6.10 tem-se a seqüência e o resultado do processo para o critério considerado bom, ver item (i).

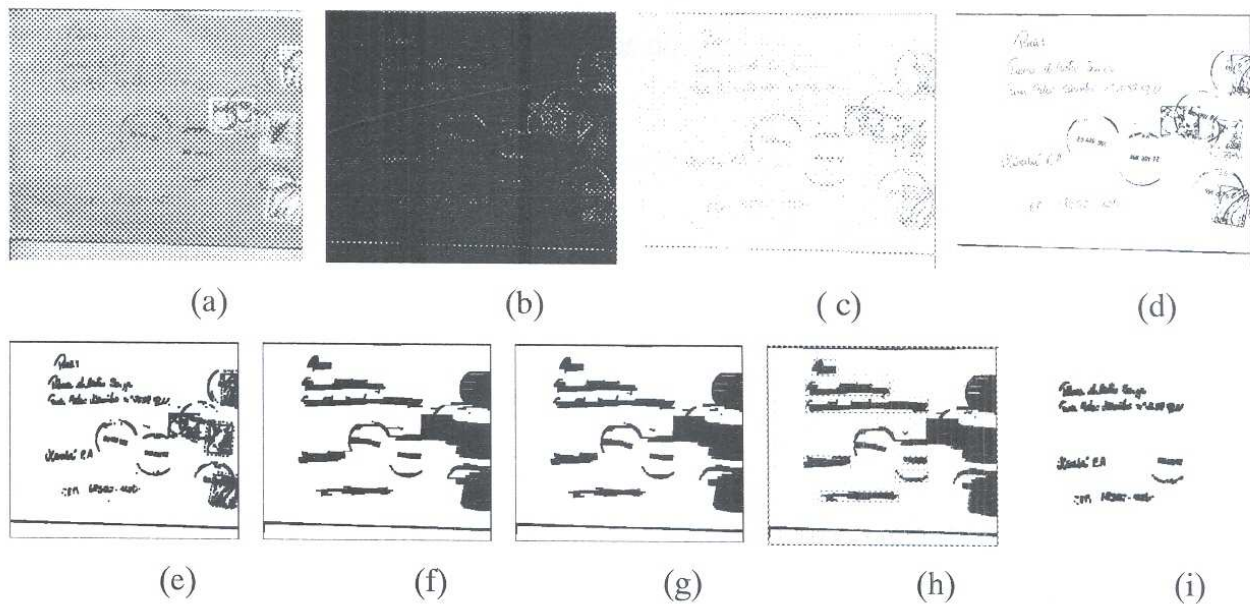


Figura 6.10: Imagens do critério bom por Tophat.

A imagem final, obtida nesta seqüência de imagens, permite classificar tal imagem como boa, pois é possível observar que ocorreram perdas de informação na primeira linha do endereçamento devido à altura da letra, a qual faz com que o tamanho do bloco exceda os limites indicados. Percebe-se, ainda, que não foi possível a eliminação de um bloco de ruído na região centro-direita da imagem, a qual acredita-se o não comprometimento da imagem em um futuro processo de reconhecimento das linhas do bloco-endereço.

Nesta seqüência da figura 6.11, temos como resultado final sub-item (i) o que é considerado como regular dentro deste critério e avaliação.

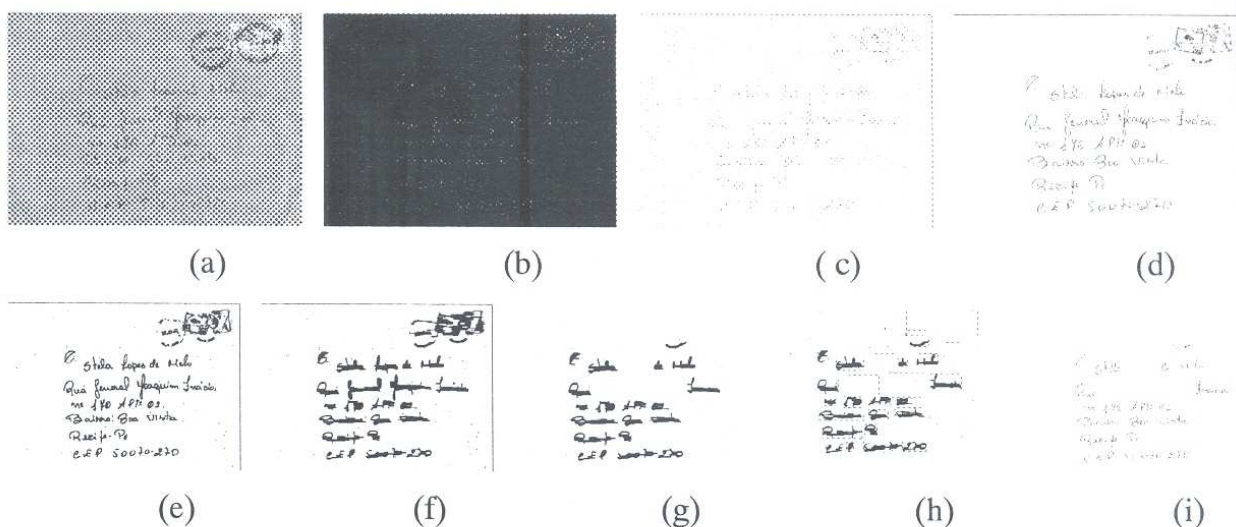


Figura 6.11: Imagens do critério regular por Tophat.

Nesta seqüência a imagem final sofre perdas na primeira e terceira linhas do endereçamento. O provável motivo é o traçado alongado das letras promovendo a conexão dos blocos, os quais, por excesso de tamanho são eliminados. Nota-se, também, que não foi possível eliminar totalmente os blocos de ruídos no canto superior direito da imagem. A imagem apresenta pouca clareza e muitos ruídos, o que traz um comprometimento acentuado da imagem para processamentos futuros.

Considerando a mesma linha de raciocínio, pode-se observar na figura 6.12 a seqüência de imagens que compõe o critério que julgamos como ruim no processo tophat, devido ao resultado final (i) que foi obtido nesta imagem.

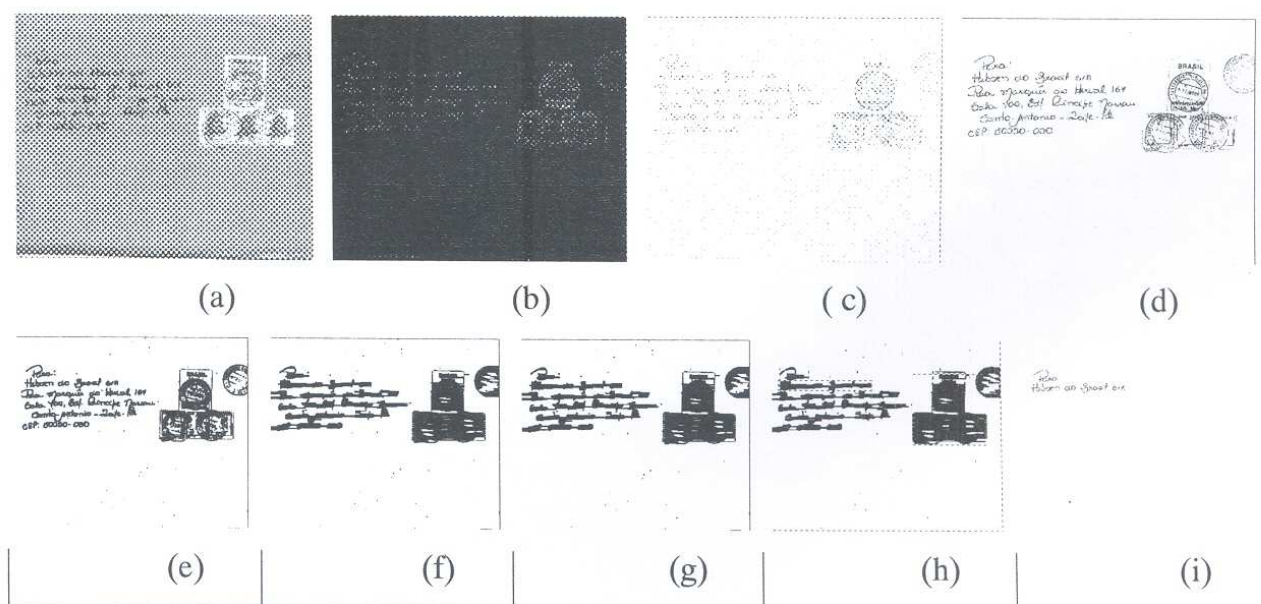


Figura 6.12: Imagens do critério ruim por Tophat.

A seqüência de imagens acima mostra que esta imagem contém duas linhas cujos blocos tocam a borda da imagem, o que determina a sua eliminação e, ainda, há letras que se sobrepõem nas linhas superior ou inferior. O resultado final é uma imagem que contém apenas duas das seis linhas do endereçamento, a qual não se faz interessante para processamentos futuros.

Há que se considerar, ainda, dois fatores observados durante os experimentos, os quais podem servir de apoio para as possíveis conclusões:

a) O tempo médio utilizado no processamento da imagem para a detecção de bordas foi de 32 segundos, ao passo que no tophat foi de 48 segundos, testes estes de validação, foram realizados em um equipamento Pentium II 700 MHz com 256 Mb de memória RAM. Tais resultados foram obtidos (*cronometrado*) numa experimentação aplicada a uma subamostra de 10% das imagens de validação destinada para os testes deste processo.

b) Vale a pena, ainda, comentar que o número de processos a serem realizados na detecção de bordas é inferior em uma unidade, em relação ao tophat.

Neste capítulo foram apresentados os resultados obtidos pela aplicação da metodologia proposta no presente trabalho. A seguir serão apresentadas as conclusões a partir dos dados obtidos e sua comparação.

Conclusão

7.1. Considerações Finais

A abordagem apresentada neste estudo empregou como objetivo principal a morfologia matemática para localizar e segmentar automaticamente o bloco-endereço de envelopes postais manuscritos em níveis de cinza, visando obter como resultado uma segmentação do bloco-endereço das imagens postais, e traz como proposta a implementação de um algoritmo que automatize o processo proposto. A aplicação paralela das técnicas de Detecção de bordas e Tophat permitiu não só a implementação do algoritmo, mas também a comparação da aplicabilidade e rendimento de cada uma das técnicas em imagens definidas anteriormente.

7.2. Conclusões

Observando-se a Tabela de dados exposta em 6.2 e considerando os resultados demonstrados em 6.3, pode-se levantar dados que norteiem um estudo comparativo entre as duas técnicas aplicadas.

A detecção de bordas apresenta-se como um método eficiente no que diz respeito ao aproveitamento dos blocos, ou seja, há menor perda de informações, obtendo-se maior clareza na imagem final, devido à melhor desconexão dos blocos pela aplicação das técnicas de fechamento horizontal e abertura vertical, o que se manifestou em 78,75% das imagens destinadas aos teste, o qual podem ser considerados como satisfatório quando são levadas em consideração as diferentes situações das imagens analisadas, conforme vimos no capítulo anterior, e também suas complexidades.

Já na aplicação do tophat, proporcionou imagens menos nítidas e com menor aproveitamento do bloco-endereço, sendo que nesta técnica foi utilizado um processo a mais que a detecção de borda durante o processamento pós binarizado, considerando que o trabalho desenvolvido foi utilizado a mesma base de testes, tendo o resultado considerado bom com 73,75% das imagens testadas.

7.3. Recomendações e Sugestões

Os resultados obtidos neste trabalho podem ser melhorados na medida em que, na etapa do pré-processamento, sejam implementados processos que aumentem o contraste da região manuscrita da imagem dos envelopes postais.

Os estudos necessários para a determinação dos melhores métodos para o pré-processamento da imagem são apontados como continuação deste trabalho, bem como a realização de um estudo algorítmico que promova um melhor desempenho das rotinas dos processos de localização do blocos de endereço em imagens postais. Como sugestão também, é importante rever o processo de reconstrução binária que em alguns casos reconstruiu ruídos, que, para um futuro reconhecimento dos componentes do bloco, poderá prejudicar em parte este processo. Tais estudos aliados ao presente trabalho podem e devem abrir espaço para o desenvolvimento de processos que automatizem o reconhecimento dos blocos-endereço das imagens postais, que é o objetivo do projeto PAEP.

Referências Bibliográficas

- [BAI90] BAIRD, H.; JONES, S.; FORTUNE, S. *Image segmentation by shape-directed covers*, ICDAR'90 pp820-825, 1990.
- [BOU96] LEBOURGEOIS, FRANK. *Localisation de textes dans une image à niveaux de gris*, 1996.
- [BOU00] LEBOURGEOIS; EMPTOZ, HUBERT. *Segmentation des document imprimés composites em niveaux de gris*, 2000.
- [CHE97] CHENG, HUI; BOUMAN, CHARLES; ALLEBACH JAN P. *Multiscale Document Segmentation*, 1997.
- [DIA96] DIAS, ANTONIO P. *Minimum spanning trees for text segmentation. Fifth Annual Symposium on Document Analysis and Information Retrieval*, pp 51-65, April, 1996.
- [FAC93] FACON, JACQUES. *Processamento e Análise de Imagens*, CEFET-PR, julho 1993.
- [FAC96] FACON, JACQUES. *Morfologia Matemática; Teoria e Exemplos*. Curitiba, Brasil, 1996.
- [O'GORMAN, 1995] O'GORMAN, LAWRENCE; KASTURI, RANGACHAR. *Document image Analysis*, IEEE, 1995.
- [GU98] GU, L., KANEKO, T., TANAKA, N., HARALICK, R. M. *Morphological Segmentation Applied to Character Extraction from Color Cover Images, Mathematical Morphology and Its Applications to Image and Signal Processing*, Kluwer Academic Publishers, pp 367-375, 1998.

- [HEU94] HEUTE, L. *Reconnaissance de Caractères Manuscrit: Application à La Lecture Automatic des Chèques et des Envelopes Postales*. Docteur Thèse de L'Université de Rouen. Rouen, France, déc, 1994, 239p.
- [JAI92] JAIN, ANIL K.; BHATTACHARJEE; CHEN, YAO. *On Texture In Document Images*, 1992.
- [JAI92] JAIN, ANIL K.; BHATTACHARJEE, SUSHIL K. *Addres Block Location on Envelopes Using Gabor Filters: Supervised Method*, 1992.
- [JAI98] JAIN, ANIL K.; YU, BIN. *Document Representation and It's Application to Page Decomposition*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 20, March, 1998.
- [JIN01] JIN, NING; TAN, YUAN Y. *Text Area Localization under Complex-Background Using Wavelet Decomposition*, 2001.
- [LE95] LE X. DANIEL, et al. *Classification of Binary Document Images into Textual or Nontextual Data Blocks Using Neural Network Models*. *Machine Vision and Applications*, pp 289-304, 1995.
- [MAT98] Matrakas, M. D.; Bortolozzi, F.; Projeto de uma Biblioteca para Processamento de Imagens de Documentos Orientada a Objetos, International Seminar on Document Management, pp 236-242, 1998.
- [MOR96] Moraes, Ronei Marcos; Image classification using mathematical morphology, Anais do IX Sibgrapi, pp357-358, 1996.
- [NAG91] Nagy, G.; Seth, S.; Hierarchical representation of optically scanned documents, Proc of 7 ICPR IEEE CS Pres Los Alamitos, California. pp347-349, 1991.
- [NOR95] Normand, N.; Viard-Gaudin, C.; A background based adaptive page segmentation algorithm. pp 138-141, IEEE, 1995.
- [OLI98] Oliveira L. E. S. "Estudo sobre a extração das estruturas lógica e física a partir de imagens de cheques bancários brasileiros", Dissertação de Mestrado, Centro Federal de Educação Tecnológica do Paraná, 1998.

- [PIE01] Pietikäinen, Matti; Okum, Oleg; Edge_based Method for Text Detection from Complex Document Images, 2001.
- [SAI00] Saito, Leandro Kazuo; Segmentação e rotulação de uma Imagem
- [SAN00] Santos, Jair Ferreira dos. O que é pós-moderno. Coleção primeiros passos. Editora Brasiliense; São Paulo, 2000.
- [SRI89] Srihari, S. N.; Govindaraju, V.; Textual image analysis using the Hough Transform, International Journal of Machine Vision and applications. pp 141-153, 1989.
- [SUE94] Suen, Ching Y., Tang, Yuan Y.; Document Structures: A survey, Document Image Analysis, World Scientific Publishing Co. Pte. Ltda, pp 1080 - 1111, 1994.
- [YAC95] Yacoubi, A . El-Yacoubi, J. M. Bertille, and M. Gilloux, "Conjoined Location and Recognition of Street Names with in a Postal Address Deliveri Line", International Conference on Document Analyses and recognition, pp1024, 1027, Vol.2, August 1995, Montréal, Canadá.
- [YAC99] Yacoubi, A . El-Yacoubi, M. Gilloux, R. Sabourin and C. Y. Suen, "Unconstrained Handwritten Word Recognition using Hidden Markov Models," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 21, N° 8, pp. 752-760, August 1999.
- [YUA01] Yuan, Q., Tan, C.L.; Text Extraction from Gray Scale Document Images Using Edge Information, 2001.
- [YU97] Yu, Bin; Jain, Anil K.; Mohiuddin, Moidin; Adres Block Location on Complex Mail Pieces, 1997.
- [XUE99] Xue, Junliamng; Ding Xiaoqing; Liu Changsong; Destination Adres Block Location on Handwritten Chinese Envelope, 1999.