

ANDREIA DE JESUS



**UMA ABORDAGEM MORFOLÓGICA PARA A
SEGMENTAÇÃO DE LOGOTIPOS EM
CHEQUES BANCÁRIOS BRASILEIROS
SEM CONHECIMENTO A PRIORI**

Dissertação apresentada ao Programa de Pós-Graduação em Informática Aplicada da Pontifícia Universidade Católica do Paraná como parte dos requisitos para obtenção do título de Mestre em Ciências.

Área de concentração:
Análise e Reconhecimento de Documentos

Orientador:
Jacques Facon

Curitiba
1999



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

Exame de dissertação nº 015

Aos 10 dias do mês de dezembro de 1999, realizou-se a sessão pública de defesa de dissertação "UMA ABORDAGEM MORFOLÓGICA PARA SEGMENTAÇÃO DE LOGOTIPOS EM CHEQUES BANCÁRIOS BRASILEIROS SEM CONHECIMENTO A PRIORI", apresentada por Andreia de Jesus, ano de ingresso 1997, para obtenção do título de Mestre em Ciências. A Banca Examinadora foi composta pelos seguintes professores:

MEMBROS DA BANCA	ASSINATURA
Presidente: Prof. Dr. Jacques Facon (PUCPR)	
Prof. Dr. Abdenaim El Yacoubi (PUCPR)	
Prof. Dr. Neucimar Jeronimo Leite (UNICAMP-SP)	
Prof. Dr. Hae Yong Kim (USP)	

De acordo com as normas regimentais a Banca Examinadora deliberou sobre os conceitos a serem atribuídos e que foram os seguintes:

MEMBROS DA BANCA	CONCEITOS
Presidente: Prof. Dr. Jacques Facon (PUCPR)	APROVADO
Prof. Dr. Abdenaim El Yacoubi (PUCPR)	APROVADO
Prof. Dr. Neucimar Jeronimo Leite (UNICAMP-SP)	APROVADO
Prof. Dr. Hae Yong Kim (USP)	APROVADO
Conceito Final	APROVADO

Observações da Banca Examinadora

Profº Júlio Cesar Nievola
Coordenador do Programa de Pós-Graduação em Informática Aplicada-PUC-PR

Dedico o meu trabalho a minha primeira sobrinha, Ana Flávia, que veio ao mundo no período em que eu desenvolvia a minha pesquisa. O seu nascimento trouxe muita alegria, luz e, principalmente, a certeza de que a vida se renova a cada dia.

Agradecimentos

A Deus, que sempre esteve presente em minha vida.

Aos meus pais, Valdivina e Osmar, que não me deram somente a vida, mas amor, oportunidades e apoio em tudo que eu realizei até hoje.

Aos meus irmãos, Angelita e Marcelo, que no decorrer deste trabalho sempre me incentivaram nos momentos em que eu me sentia cansada e desanimada.

Ao Deon, que acreditou em mim, oferecendo-me a primeira oportunidade de desenvolver um trabalho científico. Com ele aprendi a ser profissional e sobretudo amar a pesquisa.

Ao professor Celso A. Kaestner, que colaborou na orientação inicial deste trabalho; pela atenção, dedicação e amizade.

Ao professor Jacques Facon, pela orientação, dedicação, atenção, apoio e amizade.

A todos os professores que transmitiram os seus conhecimentos e contribuíram para o meu crescimento.

Ao Miguel e a Alessandra, meus companheiros de trabalho e meus amigos, que colaboraram para a realização deste trabalho.

Aos meus amigos e a todas as pessoas que, de alguma forma, apoiaram-me nessa fase da minha vida.

Sumário

Agradecimentos	ii
Sumário	iii
Lista de Figuras	vi
Resumo	x
Abstract	xi

Capítulo 1

Introdução	1
1.1 Desafio	2
1.2 Motivação	3
1.3 Proposta	4
1.4 Contribuição	4
1.5 Organização	5

Capítulo 2

Estado da Arte	6
2.1 Introdução	6
2.2 Técnicas de Segmentação	8
2.2.1 Segmentação por Textura	8
2.2.2 Segmentação por Cor	10
2.2.3 Detecção de Bordas	11
2.2.4 Segmentação por Binarização	13
2.2.5 Redes Neurais	18
2.2.6 Algoritmos Genéticos (AG)	20
2.2.7 Contornos Ativos – <i>Snakes</i>	22
2.2.8 Métodos Baseados na Teoria <i>Fuzzy</i>	26
2.2.9 Morfologia Matemática	29

2.3 Segmentação de Imagens de Documentos.....	29
2.4 Segmentação de Logotipos.....	33
2.5 Análise de Técnicas para Segmentação de Logotipos.....	34

Capítulo 3

Estrutura de um Cheque Bancário e Base de Dados	36
3.1 Introdução.....	36
3.2 Dimensões do Cheque Bancário Brasileiro.....	37
3.3 Elementos que Compõem o Cheque Bancário Brasileiro.....	38
3.4 Preenchimento de Cheques Bancários.....	39
3.5 Normas do Banco Central.....	40
3.6 Base de Dados.....	41

Capítulo 4

Morfologia Matemática	44
4.1 Introdução.....	44
4.2 Operadores Morfológicos Binários.....	47
4.2.1 Erosão Binária.....	47
4.2.2 Dilatação Binária.....	48
4.2.3 Abertura.....	50
4.2.4 Fechamento.....	52
4.2.5 Reconstrução Binária.....	55
4.2.6 Granulometria Binária.....	56
4.3 Operadores Morfológicos em Níveis de Cinza.....	59
4.3.1 Erosão em Níveis de Cinza.....	59
4.3.2 Dilatação em Níveis de Cinza.....	59
4.3.3 Erosão Geodésica.....	61
4.3.4 Reconstrução Dual em Níveis de Cinza.....	61
4.3.5 <i>Fillhole</i>	62

Capítulo 5

Metodologia Aplicada	66
5.1 Introdução.....	66
5.2 Fase de Pré-Processamento.....	70
5.2.1 Etapa de Remoção do Fundo Artístico.....	70

5.2.2 Etapa de Binarização.....	76
5.2.3 Etapa de Remoção de Ruídos	77
5.3 Fase de Localização/Extração do Logotipo.....	79
5.3.1 Etapa de Extração da I-Ésima Maior Porção do Logotipo	79
5.3.2 Etapa de Conexão dos Símbolos e Letras que Compõem o Logotipo	87
5.3.3 Etapa de Extração do Logotipo Dilatado.....	88
5.3.4 Etapa Final de Extração do Logotipo.....	88

Capítulo 6

Análise dos Resultados	98
6.1 Introdução.....	98
6.2 Resultados Quantitativos dos Testes	99
6.3 Análise dos Problemas Encontrados	105
6.3.1 Binarização Insatisfatória	105
6.3.2 Influência de Traços Manuscritos	107
6.3.3 Componentes do Logotipo Classificados com os outros Elementos em Classes Iguais.....	108
6.4 Aplicação da Metodologia Proposta em Outros Tipos de Documento	101

Capítulo 7

Conclusões e Perspectivas	112
7.1 Contribuição do Trabalho.....	113
7.2 Conclusões sobre o Trabalho.....	113
7.3 Perspectivas Futuras.....	113
Referências Bibliográficas	115

Lista de Figuras

FIGURA 1.1 TIPOS DE LOGOTIPOS [KIM, Y. S. & KIM, W.Y., 1998].	3
FIGURA 2.1 IMAGENS DE HISTOGRAMAS	14
FIGURA 2.2 FUNÇÃO DE RELACIONAMENTO TIPO S	28
FIGURA 2.3 UMA HIERARQUIA DAS SUBCATEGORIAS DO PROCESSAMENTO DE DOCUMENTOS, LISTANDO OS TIPOS DE COMPONENTES DO DOCUMENTO [GORMAN & KASTURI, 1995].	30
FIGURA 3.1 LIMITES DE TAMANHO PARA OS CHEQUES BANCÁRIOS BRASILEIROS EM PIXELS	37
FIGURA 3.2 MODELO PARA A DISTRIBUIÇÃO DE INFORMAÇÕES NOS CHEQUES BANCÁRIOS BRASILEIROS [KOERICH, 1997].	38
FIGURA 3.3 EXEMPLO DE UM CHEQUE MOSTRANDO OS CAMPOS E SEU PREENCHIMENTO	39
FIGURA 4.1 TIPOS DE ELEMENTOS ESTRUTURANTES BÁSICOS	44
FIGURA 4.2 EXEMPLOS DE INTERAÇÃO DO ELEMENTO ESTRUTURANTE EM UMA IMAGEM BINÁRIA	45
FIGURA 4.3 EXEMPLO DE INTERAÇÃO ENTRE O ELEMENTO ESTRUTURANTE E A IMAGEM EM NÍVEIS DE CINZA	46
FIGURA 4.4 RESULTADO DA EROÇÃO E DA DILATAÇÃO COM ELEMENTO ESTRUTURANTE QUADRADO E 2 ITERAÇÕES. (A) IMAGEM ORIGINAL; (B) IMAGEM ERODIDA; (C) IMAGEM DILATADA.	50
FIGURA 4.5 RESULTADO DA ABERTURA E DO FECHAMENTO COM ELEMENTO ESTRUTURANTE VERTICAL E 4 ITERAÇÕES. (A) IMAGEM ORIGINAL; (B) ABERTURA; (C) FECHAMENTO.	54
FIGURA 4.6 EXEMPLO DE RECONSTRUÇÃO BINÁRIA COM ELEMENTO ESTRUTURANTE QUADRADO.	55
FIGURA 4.7 GRANULOMETRIA POR ABERTURA COM RECONSTRUÇÃO.	58
FIGURA 4.8 RESULTADO DA EROÇÃO E DA DILATAÇÃO EM NÍVEIS DE CINZA COM ELEMENTO ESTRUTURANTE QUADRADO PLANAR E 2 ITERAÇÕES. (A) IMAGEM ORIGINAL; (B) IMAGEM ERODIDA; (C) IMAGEM DILATADA.	60

FIGURA 4.9 RECONSTRUÇÃO DUAL EM NÍVEIS DE CINZA DA MÁSCARA S COM O MARCADOR Z E ELEMENTO ESTRUTURANTE QUADRADO PLANAR.	62
FIGURA 4.10 REMOÇÃO DE TODOS OS MÍNIMOS INTERNOS DA IMAGEM S PELA RECONSTRUÇÃO MORFOLÓGICA DA FUNÇÃO MARCADOR Z E ELEMENTO ESTRUTURANTE QUADRADO.	63
FIGURA 4.11 (A) IMAGEM EM NÍVEIS DE CINZA; (B) RESULTADO DO PROCESSO FILLHOLE; (C) RESULTADO DA SUBTRAÇÃO ENTRE A IMAGEM (A) E A IMAGEM (B).	65
FIGURA 5.1 ESTRUTURA DO SISTEMA DE SEGMENTAÇÃO AUTOMÁTICA DE LOGOTIPOS EM IMAGENS DE CHEQUES BANCÁRIOS BRASILEIROS.	67
FIGURA 5.2 ÁREA DE TRABALHO DO KHOROS.	68
FIGURA 5.3 ÁREA DE TRABALHO DO BROPI.	69
FIGURA 5.4 IMAGENS BINÁRIAS COM FUNDO ARTÍSTICO.	71
FIGURA 5.5 (A) IMAGEM EM NÍVEIS DE CINZA DO CHEQUE BANCÁRIO; (B) RESULTADO DO PROCESSO FILLHOLE; (C) IMAGEM EM NÍVEIS DE CINZA DO CHEQUE BANCÁRIO COM O FUNDO ARTÍSTICO REMOVIDO.	73
FIGURA 5.6 IMAGEM DE CHEQUE COM TRAÇOS MANUSCRITOS QUE TOCAM NA BORDA E SOBREPÕEM O LOGOTIPO.	74
FIGURA 5.7 IMAGEM RESULTANTE DA SUBTRAÇÃO ENTRE O RESULTADO DO FILLHOLE E A IMAGEM 5.4.	74
FIGURA 5.8 IMAGEM DE CHEQUE COM A ADIÇÃO DE UMA BORDA BRANCA.	75
FIGURA 5.9 IMAGEM DE CHEQUE COM O FUNDO ARTÍSTICO REMOVIDO APÓS A ADIÇÃO DA BORDA BRANCA.	75
FIGURA 5.10 (A) IMAGEM EM NÍVEIS DE CINZA; (B) IMAGEM BINARIZADA PELO ALGORITMO DE OTSU.	77
FIGURA 5.11 (A) IMAGEM ERODIDA; (B) IMAGEM RECONSTRUÍDA E COM OS RUÍDOS ELIMINADOS.	79
FIGURA 5.12 EXEMPLO DE OBTENÇÃO DE UMA PORÇÃO DO LOGOTIPO COM UMA HIPÓTESE.	82
FIGURA 5.13 EXEMPLO DE OBTENÇÃO DE UMA PORÇÃO DO LOGOTIPO COM DUAS HIPÓTESES.	85
FIGURA 5.14 (A) PORÇÃO DE UM LOGOTIPO OBTIDA NA ÚLTIMA CLASSE DO PROCESSO GRANULOMETRICO APLICADO NA FIGURA 5.11B.	86
FIGURA 5.15 – IMAGEM DE CHEQUE COM TODOS OS ELEMENTOS CONECTADOS.	87
FIGURA 5.16 – LOGOTIPO NA VERSÃO DILATADA.	88
FIGURA 5.17 LOGOTIPO SEGMENTADO.	89

FIGURA 5.18 IMAGEM ORIGINAL EM NÍVEIS DE CINZA DE UM CHEQUE PREENCHIDO A MÁQUINA	90
FIGURA 5.19 IMAGEM DE UM CHEQUE PREENCHIDO A MÁQUINA COM O FUNDO ARTÍSTICO REMOVIDO	90
FIGURA 5.20 IMAGEM BINÁRIA DE UM CHEQUE PREENCHIDO A MÁQUINA	91
FIGURA 5.21 IMAGEM DE UM CHEQUE PREENCHIDO A MÁQUINA COM OS RUÍDOS REMOVIDOS	91
FIGURA 5.22 PORÇÃO ISOLADA DO LOGOTIPO OBTIDA NA I-ÉSIMA CLASSE GERADA PELO PROCESSO GRANULOMÉTRICO, APLICADO EM UMA IMAGEM DE CHEQUE PREENCHIDO A MÁQUINA.	92
FIGURA 5.23 IMAGEM DE UM CHEQUE PREENCHIDO A MÁQUINA COM TODOS OS ELEMENTOS CONECTADOS	92
FIGURA 5.24 VERSÃO DILATADA DE UM LOGOTIPO DE UM CHEQUE PREENCHIDO A MÁQUINA	93
FIGURA 5.25 LOGOTIPO SEGMENTADO DE UM CHEQUE PREENCHIDO A MÁQUINA.....	93
FIGURA 5.26 IMAGEM ORIGINAL EM NÍVEIS DE CINZA DE UM CHEQUE EM BRANCO	94
FIGURA 5.27 IMAGEM DE UM CHEQUE EM BRANCO COM O FUNDO ARTÍSTICO REMOVIDO .	94
FIGURA 5.28 IMAGEM BINÁRIA DE UM CHEQUE EM BRANCO	95
FIGURA 5.29 IMAGEM DE UM CHEQUE EM BRANCO COM OS RUÍDOS REMOVIDOS	95
FIGURA 5.30 PORÇÃO ISOLADA DO LOGOTIPO OBTIDA NA I-ÉSIMA CLASSE GERADA PELO PROCESSO GRANULOMÉTRICO, APLICADO EM UMA IMAGEM DE CHEQUE EM BRANCO.	96
FIGURA 5.31 IMAGEM DE UM CHEQUE EM BRANCO COM TODOS OS ELEMENTOS CONECTADOS	96
FIGURA 5.32 VERSÃO DILATADA DE UM LOGOTIPO DE UM CHEQUE EM BRANCO	97
FIGURA 5.33 LOGOTIPO SEGMENTADO DE UM CHEQUE EM BRANCO	97
FIGURA 6.1 EXEMPLO DE SEGMENTAÇÃO DE BOA QUALIDADE. (A) IMAGEM DE CHEQUE BANCÁRIO; (B) LOGOTIPO SEGMENTADO.....	100
FIGURA 6.2 EXEMPLO DE SEGMENTAÇÃO PARCIAL. (A) IMAGEM DE CHEQUE BANCÁRIO; (B) UMA PORÇÃO DO LOGOTIPO.....	101
FIGURA 6.3 EXEMPLO DE SEGMENTAÇÃO RUIM. (A) IMAGEM DE CHEQUE BANCÁRIO; (B) LOGOTIPO SEGMENTADO JUNTO COM OUTROS COMPONENTES.....	102
FIGURA 6.4 (A) IMAGEM EM NÍVEIS DE CINZA; (B) IMAGEM BINARIZADA DE FORMA INSATISFATÓRIA.....	106
FIGURA 6.5 SEGMENTAÇÃO DE UM COMPONENTE DO LOGOTIPO POUCO SIGNIFICATIVO.	106

FIGURA 6.6 IMAGEM BINÁRIA DE CHEQUE BANCÁRIO COM LINHAS SOBREPOSTAS	107
FIGURA 6.7 SEGMENTAÇÃO INDEVIDA DO LÓGOTIPO (DA FIGURA 6.6) POR CAUSA DAS LINHAS SOBREPOSTAS NO CHEQUE.....	108
FIGURA 6.8 (A) IMAGEM PARCIAL DE UM CHEQUE BANCÁRIO; (B), (C) E (D) APRESENTAM COMPONENTES DO LOGOTIPO CLASSIFICADOS, PELO PROCESSO GRANULOMÉTRICO, JUNTO COM OUTROS ELEMENTOS.	110
FIGURA 6.9 IMAGEM DE FAX	111

Resumo

Uma das preocupações, no desenvolvimento de uma aplicação que envolve a análise automática de imagens, é encontrar uma segmentação significativa. Usualmente, são usadas muitas heurísticas para o tamanho e posição dos componentes a serem extraídos. O uso de muitas heurísticas acaba limitando a segmentação para um determinado tipo de aplicação. Mas, segmentar sem conhecimento prévio, empregando poucas heurísticas e poucas hipóteses, é um grande desafio.

Na área de reconhecimento de logotipos, as maiores dos estudos estão baseadas na segmentação manuais dos logotipos, o que impossibilita a aplicação dessas técnicas num contexto geral e/ou com outros tipos de documentos. Portanto, a proposta deste trabalho é tentar localizar e extrair automaticamente logotipos em imagens de cheques bancários brasileiros, utilizando o mínimo de parâmetros heurísticos, visando a obter como resultados uma segmentação exata ou a mais exata possível para todas as situações, permitindo assim o seu reuso em outras aplicações. A técnica utilizada para esse propósito foi a morfologia matemática.

A metodologia proposta foi validada em uma base de dados de 418 imagens de cheques bancários brasileiros. Os testes mostraram a viabilidade dessa abordagem na redução do número de parâmetros heurísticos para a localização e a extração automática de logotipos. Além disso, verificou-se também que essa abordagem viabiliza a extração de logotipos de modelos diferentes.

Palavras-chaves: segmentação, cheques bancários brasileiros, logotipo, granulometria, Morfologia Matemática.

Resumo

Uma das preocupações, no desenvolvimento de uma aplicação que envolve a análise automática de imagens, é encontrar uma segmentação significativa. Usualmente, são usadas muitas heurísticas para o tamanho e posição dos componentes a serem extraídos. O uso de muitas heurísticas acaba limitando a segmentação para um determinado tipo de aplicação. Mas, segmentar sem conhecimento prévio, empregando poucas heurísticas e poucas hipóteses, é um grande desafio.

Na área de reconhecimento de logotipos, as maiorias dos estudos estão baseadas na segmentação manuais dos logotipos, o que impossibilita a aplicação dessas técnicas num contexto geral e/ou com outros tipos de documentos. Portanto, a proposta deste trabalho é tentar localizar e extrair automaticamente logotipos em imagens de cheques bancários brasileiros, utilizando o mínimo de parâmetros heurísticos, visando a obter como resultados uma segmentação exata ou a mais exata possível para todas as situações, permitindo assim o seu reuso em outras aplicações. A técnica utilizada para esse propósito foi a morfologia matemática.

A metodologia proposta foi validada em uma base de dados de 418 imagens de cheques bancários brasileiros. Os testes mostraram a viabilidade dessa abordagem na redução do número de parâmetros heurísticos para a localização e a extração automática de logotipos. Além disso, verificou-se também que essa abordagem viabiliza a extração de logotipos de modelos diferentes.

Palavras-chaves: segmentação, cheques bancários brasileiros, logotipo, granulometria, Morfologia Matemática.

Abstract

One of the major challenges in image processing might be segmentation of significant data. In segmentation many heuristics, which are related to the size and position of the components to be extracted, are generally used. The use of many heuristics ends up by limiting segmentation for a given type of application. Therefore, the design of a segmentation technique with few a priori knowledge and heuristic parameters is a great challenge.

In the field of logo segmentation, most of the works are based on manual logo segmentation. This fact reduces a generalization of the methods in the case of other kinds of documents. The aim of this work is to employ mathematical morphology to automatically locate and extract logos from the image of Brazilian bank checks. The objective is to minimize the number of heuristic parameters in order to obtain the most precise possible segmentation in every situation, thereby allowing this approach to be reused in other applications.

The proposed approach was tested on a database of 418 Brazilian bank checks. The results have shown the feasibility of this approach to reduce the number of heuristic parameters in automatic logo location and extraction. This approach was also shown to allow the extraction of different types of logos, in our case, word-in-mark and composite-mark logos.

Keywords: segmentation, Brazilian bank checks, logo, granulometries, mathematical morphology.

Abstract

One of the major challenges in image processing might be segmentation of significant data. In segmentation many heuristics, which are related to the size and position of the components to be extracted, are generally used. The use of many heuristics ends up by limiting segmentation for a given type of application. Therefore, the design of a segmentation technique with few a priori knowledge and heuristic parameters is a great challenge.

In the field of logo segmentation, most of the works are based on manual logo segmentation. This fact reduces a generalization of the methods in the case of other kinds of documents. The aim of this work is to employ mathematical morphology to automatically locate and extract logos from the image of Brazilian bank checks. The objective is to minimize the number of heuristic parameters in order to obtain the most precise possible segmentation in every situation, thereby allowing this approach to be reused in other applications.

The proposed approach was tested on a database of 418 Brazilian bank checks. The results have shown the feasibility of this approach to reduce the number of heuristic parameters in automatic logo location and extraction. This approach was also shown to allow the extraction of different types of logos, in our case, word-in-mark and composite-mark logos.

Keywords: segmentation, Brazilian bank checks, logo, granulometries, mathematical morphology.

Capítulo 1

Introdução

Um documento pode ser visto como qualquer base de conhecimento fixada materialmente e disposta de maneira que se possa utilizar para consulta, estudo, prova, etc. Ou seja, os documentos são um meio para a transmissão de conhecimento. De fato, muito conhecimento é adquirido de documentos como: relatórios técnicos, arquivos governamentais, jornais, periódicos, revistas, cartas, cheques bancários e outros.

A necessidade de digitalizar os documentos é crescente devido à grande quantidade de papéis manipulados todos os dias no mundo inteiro. Com os avanços tecnológicos, os computadores estão ficando cada vez mais rápidos e baratos, isso está facilitando o armazenamento, a busca, a recuperação e a disseminação de documentos.

A área de análise de imagens de documentos tem como objetivo reconhecer texto, gráficos e figuras em imagens e extrair a informação pretendida como um humano faria. As áreas onde os sistemas de análise de documentos podem ser empregados com sucesso são inúmeras. Por exemplo, eles podem ser usados por um escritório que manipula uma porção muito grande de papéis impressos; em bibliotecas, os sistemas podem resolver problemas e falhas, tais como: números limitados de cópias, perda de documentos, degradação do material impresso.

A aquisição automática de conhecimento de documentos tornou-se um importante objeto de estudo, e muitos pesquisadores estão tentando encontrar novas técnicas de processamento de documentos. Na realidade, é muito difícil desenvolver um sistema geral capaz de processar todos os tipos de documentos (cartas, jornais, revistas, ofícios, etc.), porque cada documento possui uma particularidade própria.

Este trabalho de pesquisa está focado na segmentação automática de logotipos em cheques bancários brasileiros. O desafio, a motivação, a proposta e a contribuição deste trabalho serão descritos nas seções abaixo.

1.1 Desafio

Os principais desafios encontrados neste trabalho são:

- Poucos estudos sobre segmentação (extração) automática de logotipos são encontrados na literatura. Muitos autores concentraram seus estudos na etapa de reconhecimento dos logotipos. Contudo, o reconhecimento não pode ser realizado sem existir antes uma fase de conhecimento e, conhecer significa extrair características dos padrões a serem processados.
- Existem quatro tipos de logotipos [KIM, Y. S. & KIM, W.Y., 1998]:
 - *Word-in-mark* é um logotipo que contém somente caracteres ou palavras na marca;
 - *Device-mark* contém somente elementos gráficos ou figurativos;
 - *Composite-mark* consiste de caracteres ou palavras e elementos gráficos;
 - *Complex-mark* contém uma imagem complexa (Figura 1.1).

Mas os trabalhos encontrados na literatura geralmente processam somente um tipo de logotipo [KIM, Y. S. & KIM, W.Y., 1998].

- Na segmentação, usualmente, são usadas muitas heurísticas para o tamanho e posição dos componentes a serem extraídos. O uso de muitas heurísticas acaba limitando a segmentação para um determinado tipo de aplicação. Mas, segmentar sem conhecimento prévio, empregando poucas heurísticas e poucas hipóteses, é um grande desafio.



Figura 1.1 Tipos de Logotipos [KIM, Y. S. & KIM, W.Y., 1998].

1.2 Motivação

As duas principais motivações para a realização deste trabalho são:

- Na área de processamento de imagens de documentos, o reconhecimento de logotipos tem chamado muita a atenção, pois reconhecendo estes objetos obtêm-se informações semânticas sobre os documentos, o que facilita a análise do conteúdo e o reconhecimento das classes dos mesmos, além de ajudar na decisão de analisar ou não os componentes textuais. Várias são as aplicações que utilizam essas informações para tomar decisões de processamento, as quais podem ser encontradas na literatura. Spitz [SPITZ, 1996] mostra que detecção de logotipos em imagens compactadas pode ser utilizada para sinalizar sistemas que processam o recebimento de fax, enviando sinais de classificação de segurança do documento, informações de roteamento, linguagem do documento ou a necessidade de acabamento (*rendering*) colorido. Suda [SUDA, et al., 1997] utilizou o reconhecimento de logotipos e/ou palavras em cartas comerciais para fazer a discriminação entre classes de layouts similares. Estes trabalhos mostram que os logotipos podem auxiliar em diferentes aplicações.

- O processamento de cheques bancários é um campo importante da análise de imagens de documentos devido à grande quantidade de transações de cheques que ocorrem no mundo inteiro. Do ponto de vista de pesquisa, as imagens de cheques permitem analisar padrões diferentes como: fundo artístico, informações pré-impressas, informações manuscritas, logotipo.

1.3 Proposta

Uma das preocupações, no desenvolvimento de uma aplicação que envolve a análise automática de imagens, é encontrar uma segmentação significativa. As maiorias dos estudos na área de reconhecimento de logotipos baseiam-se na segmentação manual dos logotipos, o que impossibilita a aplicação dessas técnicas num contexto geral e/ou com outros tipos de documentos.

A proposta deste trabalho é tentar localizar e extrair automaticamente logotipos em imagens de cheques bancários brasileiros, utilizando o mínimo de parâmetros heurísticos, visando a obter como resultado uma segmentação exata ou a mais exata possível para todas as situações. A técnica utilizada para esse propósito foi a morfologia matemática.

Como logotipo é uma marca saliente perceptível, maior do que componentes textuais individuais [DOERMANN et al., 1993], um único conhecimento prévio: **o logotipo é maior do que todas as informações presentes no cheque bancário.**

1.4 Contribuição

As principais contribuições deste trabalho são:

- Apresentação de uma metodologia de segmentação que não utiliza conhecimento a priori sobre o tipo de documento e aplica o mínimo possível de heurísticas, permitindo, assim, o seu reuso em outras aplicações.
- Processamento de dois modelos de logotipos: *word-in-mark* e *composite-mark*.

- Contribuição para o desenvolvimento de trabalhos futuros na área de segmentação automática de logotipos.

1.5 Organização

Esta dissertação foi dividida em sete capítulos dispostos da seguinte forma:.

Este capítulo apresentou a importância do processamento de documentos e os desafios, a proposta, objetivos e contribuição deste trabalho.

O Capítulo II indica o papel da segmentação dentro de um sistema de reconhecimento e descreve, de forma sucinta, algumas técnicas de segmentação e trabalhos correlatos.

O Capítulo III expõe a estrutura do cheque bancário brasileiro, as normas estabelecidas pelo Banco Central e a base de dados utilizada para validar a metodologia proposta.

O Capítulo IV apresenta os princípios básicos da morfologia matemática e descreve as operações morfológicas que foram utilizadas na segmentação automática de logotipos em imagens de cheques bancários brasileiros.

O Capítulo V descreve as duas fases do processo automático de segmentação de logotipos e suas respectivas etapas.

O Capítulo VI mostra os resultados obtidos com a metodologia proposta. Os problemas encontrados e resultados quantitativos são apresentados.

O capítulo VII apresenta as conclusões do trabalho, as contribuições, a importância da segmentação de logotipos e trabalhos futuros.

Capítulo 2

Estado da Arte

2.1 Introdução

Um dos maiores desafios do processamento de imagens talvez seja a segmentação da imagem para dados significativos. A grande dificuldade reside no fato de não se conhecer de antemão o número e tipo de estruturas que se encontram na imagem. Essas estruturas são identificadas a partir de sua geometria, forma, topologia, textura, cor ou brilho sendo escolhidas aquelas que possibilitam a melhor identificação.

A segmentação de imagens é um processo que tipicamente particiona o domínio espacial de uma imagem em subconjuntos mutuamente exclusivos, chamados regiões, onde cada região 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.

Haralick propôs [HARALICK & SHAPIRO, 1985] as seguintes propriedades para um bom processo de segmentação:

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

As aplicações na área de segmentação de imagens são variadas: segmentação de imagens de satélites [ERTHAL, et.al, 1991], segmentação de componentes de cartas náuticas [FERNANDES, 1997], extração do contorno da válvula mitral [GONÇALVES, 1998], segmentação de jornais [WILLIAMS & ALDER, 1996], localização de *strings* numéricas em mapas [CHEN & WANG, 1997], identificação de um logotipo em um documento [SPITZ, 1996], entre outras. Em todas essas aplicações, o desempenho final da análise da imagem está diretamente relacionado com a eficiência do processo de segmentação.

Entre as diferentes áreas de aplicação em que a segmentação atua, encontram-se variados tipos de imagem:

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

Em geral, uma imagem pode ser descrita por uma função bidimensional $F(x, y)$, onde (x, y) é a coordenada espacial e $F(x, y)$ o valor da característica em (x, y) . Dependendo do tipo da imagem, o valor da característica pode ser intensidade de luz, profundidade, intensidade de onda de rádio ou temperatura. A imagem digital, por outro lado, é uma função discreta bidimensional $f(x, y)$ digitalizada pela coordenada espacial e pela magnitude do valor da característica. Uma imagem digital pode ser vista como uma matriz bidimensional, cujos índices linha e coluna identificam um ponto (pixel) na imagem, e o valor do elemento da matriz correspondente identifica o nível de intensidade da característica.

Na literatura, encontram-se várias técnicas de segmentação, mas não existe um método simples que seja bom para todos os tipos de imagens, e nem todos os métodos são igualmente bons para um tipo particular de imagem. As metodologias baseiam-se na textura [RUZON, 1997], cor [BRITTO, 1996], detecção de bordas [PELI & MALAH,

1982], técnicas de binarização [MATTANA, 1999], redes neurais [BLANZ & GISH, 1990], algoritmos genéticos [VISA, 1998], contornos ativos [KASS et al., 1987], fuzzy [KELLER et al., 1986], morfologia matemática [OKADA & MALAYAPPAN, 1997], entre outros.

Na seção 2.2, serão apresentadas, de forma sucinta, algumas técnicas de segmentação e trabalhos correlatos.

2.2 Técnicas de Segmentação

2.2.1 Segmentação por Textura

Para Mark [RUZON, 1997], a textura é: “um ou mais padrões locais básicos que são repetidos periodicamente”. Mas, não está claro exatamente qual é o padrão e como ele é repetido. E, além disso, não está claro se a textura é uma propriedade inerente de todas as coisas, ou se alguns objetos ou regiões são desprovidos totalmente de texturas.

Segmentar texturas [RUZON, 1997] é particionar uma imagem em regiões, onde cada região contém uma textura simples e distinta de seus vizinhos.

A textura pode ser representada por modelos que fornecem um meio para transformar uma janela de uma imagem em um conjunto de números. Esses modelos de textura podem ser classificados em três classes principais: estruturas de imagens piramidais, campos randômicos e métodos estatísticos.

O objetivo das estruturas de imagens piramidais é gerar um número de parâmetros homogêneos que representem a resposta de um banco de filtros em escalas variadas e possíveis orientações diferentes. Existem muitos tipos de filtros que podem ser usados para este propósito, como os filtros Laplaciano e Gabor.

Os modelos de campos randômicos assumem que os valores dos pixels são escolhidos por um processo estocástico bidimensional. As duas classes principais são: Campos Randômicos de Markov Gaussiano (Gaussian Markov Random Fields - GMRF's): utilizam a distribuição de probabilidade Gaussiana;

- Campos Randômicos de Gibbs (Gibbs Random Fields - GRF's): utilizam a distribuição de probabilidade exponencial.

Ambas obedecem às propriedades de Markov, que declara o seguinte: “a probabilidade que o pixel tem de ter um certo nível de cinza dado todos os outros pixels da imagem é igual à probabilidade que o pixel tem de ter o mesmo nível de cinza dado somente os seus pixels vizinhos” [RUZON, 1997].

A classe dos métodos estatísticos agrupa várias técnicas:

- Matriz de co-ocorrência: representa a distância entre pixels e a dependência espacial dos níveis de cinza para diferentes orientações.
- Transformada Discreta de Fourier e Transformada Discreta de Coseno: são dois exemplos de uma classe geral de técnicas matemáticas chamadas transformadas ortogonais. Geralmente, um conjunto de funções ortogonais fundamentais é gerado, e, então, cada função é convoluída com a imagem para produzir um conjunto de números para cada ponto.
- Fractais: são aplicados, às vezes, para a análise e segmentação de textura, porque eles possuem uma escala própria. Isto significa que uma análise fractal pode produzir a dimensão fractal da textura e indicar a escala. Desta forma, evita-se ter de adivinhar em que escala a análise deve proceder.

Na literatura, encontram-se vários trabalhos relacionados com a segmentação de texturas. Kruizinga e Petkov [KRUIZINGA & PETKOV, 1998] usaram a saída dos operadores de textura modificados biologicamente (*grating cell operators*) como características de textura e fizeram a comparação com as matrizes de co-ocorrência e a energia de Gabor, que são as características de texturas mais usadas. A comparação foi feita com um método quantitativo, baseado no critério de Fisher. Junto com os resultados de classificação dos experimentos, a comparação mostrou a superioridade do novo operador para problemas de textura orientada.

Cruz [CRUZ, 1998] apresentou uma nova metodologia para a segmentação de texturas. A metodologia é baseada nas características estatísticas de alta ordem (matriz de co-ocorrência e momentos), técnicas de fusão de dados e o método de verossimilhança para encontrar clusters. As características isoladas, matriz de co-ocorrência e momentos, tiveram sempre os mesmos resultados (87.88% e 87.87%

respectivamente). Já as técnicas de fusão de dados aumentaram em 12% a performance em relação às técnicas isoladas.

2.2.2 Segmentação por Cor

A cor é uma característica com grande potencial discriminatório. Em diversas situações, o homem toma por base essa informação para diferenciar objetos ou estruturas. Há casos, onde a cor é adicionada ao processo artificialmente, através do uso de corantes. Isso é comum em estudos microscópicos, quando se torna necessário o destaque de estruturas de interesse específico.

Na análise automática de imagens, o processamento da cor é de grande importância na identificação e extração de características e objetos. Este fato deve-se principalmente ao aumento da capacidade de discernimento dos objetos em imagens coloridas. Entretanto, para aplicar a cor em processamento de imagens, é fundamental que se tenha conhecimento dos principais sistemas de representação desta informação. Estes, de uma forma geral, fornecem uma padronização para a cor, permitindo a sua quantificação, independente das condições psicofísicas dos observadores. Vários são os sistemas para a representação da cor encontrados na literatura; as duas principais abordagens são:

- RGB: consiste na utilização de um espaço tridimensional onde cada cor é obtida a partir das cores primárias vermelha, verde e azul, respectivamente baixa, média e alta frequências do espectro visual pelo olho humano.
- HSI: H é a matiz (descreve a cor pura : vermelha, azul, amarela); S a saturação da cor (grau de pureza); I é a componente que representa a intensidade de luz (brilho) da cor.

Ohta [OHTA & SAKAI, 1980] realizou um estudo comparativo entre diversos sistemas de representação de cor, buscando as componentes mais significativas para a segmentação de imagens coloridas. No processo de comparação, foi considerada a qualidade da segmentação e o cálculo envolvido na transformação para se obter as novas componentes da cor, a partir do sistema RGB. Além de diversas componentes de

sistemas de representação, tais como R, G, B, H, S e I, foram propostas I1, I2 e I3 definidas pelas fórmulas $I1 = (R + B + G)/3$, $I2 = (R - B)/2$, $I3 = (2G - R - B)/4$. Essas foram consideradas significativas nessa ordem e mostraram ser as mais eficientes durante os experimentos realizados. Um dos motivos foi a simplificada transformação linear necessária para se obter as novas componentes a partir do sistema RGB.

BRITTO propôs [BRITTO, 1996] um algoritmo para segmentação de cor. O algoritmo caracteriza-se pelo uso do sistema de representação da cor HSI e toma por base propriedades estatísticas da imagem, através da transformada de *Karhunen-Loève*, que melhor distribui a informação para a devida segmentação. Este algoritmo é aplicado em um processo de auxílio a diagnóstico médico, baseado na quantificação de mucina a partir de imagens de biópsia retal realizadas em ratos da linhagem *Sprague-Dawley*. Basicamente, o seu objetivo é a extração de informações de uma cena, tomando por base o atributo cor.

A análise espectral é outra técnica de segmentação de imagem colorida, em que uma informação a priori sobre os objetos coloridos é usada para classificar os pixels. Porém, em muitas aplicações da vida real, o conhecimento a priori sobre as cores dos objetos é muito difícil de coletar.

Do ponto de vista computacional, o processamento da informação cor implica em considerável aumento dos custos, além de precisar de um espaço de armazenamento três vezes maior. Em algumas aplicações, observa-se a transformação de imagens originalmente coloridas para níveis de cinza. Mas, nem sempre o contraste da imagem monocromática é suficiente para se obter bons resultados no processamento de imagens, principalmente quando o objetivo é a segmentação [BRITTO, 1996].

2.2.3 Detecção de Bordas

A segmentação também pode ser obtida através da detecção de bordas de várias regiões. Este método normalmente tenta localizar pontos com mudanças bruscas nos valores de intensidade de nível de cinza.

Há uma grande variedade de métodos disponíveis na literatura para detectar bordas. Davis [DAVIS, 1975] classificou as técnicas de detecção de bordas em duas categorias: sequencial e paralela. Na técnica sequencial, a decisão se um pixel é um

pixel de borda ou não depende do resultado do detector em alguns pixels examinados anteriormente. No caso da técnica paralela, a decisão, se um ponto é uma borda ou não, é feita com base no ponto em consideração e alguns de seus pontos vizinhos. A performance do método de detecção de bordas sequencial depende da escolha de um ponto inicial apropriado, pois os resultados de pontos anteriores influenciam a seleção e o resultado do próximo ponto.

Há diferentes tipos de operadores diferenciais paralelos como:

- Gradientes de *Roberts*, de *Sobel* e de *Prewitt*: eles não respondem somente a bordas, mas também a pontos isolados. Estes operadores são chamados de derivada da primeira.
- O operador Laplaciano: responde intensamente a ângulos, linhas e pontos isolados. Assim, para imagens ruidosas, a menos que ela tenha um contraste baixo, o ruído produzirá valores de Laplaciano mais altos do que as bordas. Este operador é chamado de derivada segunda.

De acordo com Canny [CANNY, 1986], um bom detector de borda tem de ter as seguintes propriedades:

- Probabilidade baixa para marcação de pontos que não pertençam à borda e a não marcação de pontos que pertençam à borda (boa detecção);
- Pontos marcados como bordas podem estar tão próximos quanto possível do centro de bordas verdadeiras (boa localização);
- Uma e somente uma resposta para um ponto de borda único (resposta única).

Uma boa detecção pode ser obtida pela Maximização da Razão Sinal/Ruído (Maximizing Signal to Noise Ratio – SNR), enquanto para uma boa localização, Canny mediu como o inverso da variância entre o máximo da resposta e a posição real da transição. Para maximizar simultaneamente ambos, boa detecção e critério de localização, Canny [CANNY, 1986] maximizou o produto do SNR e a recíproca do desvio padrão (aproximado) do deslocamento de pontos de borda. A maximização do produto é feita sujeita à coação, que elimina respostas múltiplas para pontos de borda únicos.

As referências [GOKMEN & LI, 1990], [REICHENBACH et al., 1990], [COCQUEREZ & PHILIPP, 1995] apresentam estudos de detecção de borda.

2.2.4 Segmentação por Binarização

Binarização é uma técnica antiga, simples e popular para segmentação de imagens. Ela converte uma imagem originalmente em tons de cinza para uma imagem binária. Pode ser feita com base na informação global (histograma de nível de cinza da imagem de entrada) ou com base na informação local (matriz de co-ocorrência) da imagem. A binarização global ocorre quando um único limiar é aplicado na imagem de entrada; e a binarização local, quando a imagem é particionada em sub-regiões e um limiar é determinado para cada uma destas sub-regiões [PAL N. R. & PAL S. K., 1993].

A binarização também pode ser classificada como:

- Binarização em dois níveis: a imagem é particionada em duas regiões. Objeto (preto) e fundo (branco).
- Multibinarização: quando uma imagem é composta de vários objetos com diferentes características de superfície onde são necessários vários limiares para a segmentação.

Se uma imagem é composta de regiões com diferentes intervalos de níveis de cinza, isto é, as regiões são distintas, o histograma da imagem usualmente mostra diferentes picos (Figura 2.1a). Cada pico refere-se a uma região e picos adjacentes são provavelmente separados por um vale. Mas, normalmente, a situação não é esta, e a seleção do limiar não é uma tarefa trivial (Figura 2.1b).

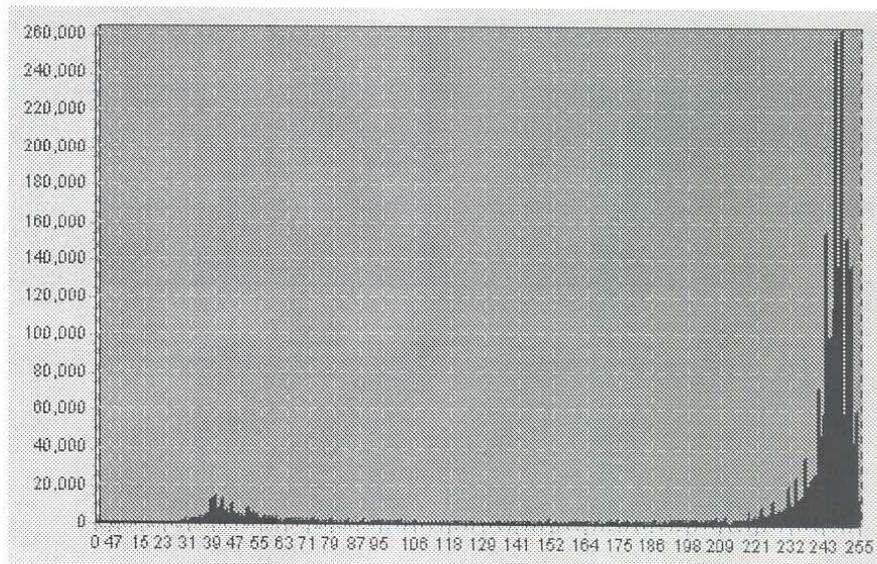


Figura 2.1 - a

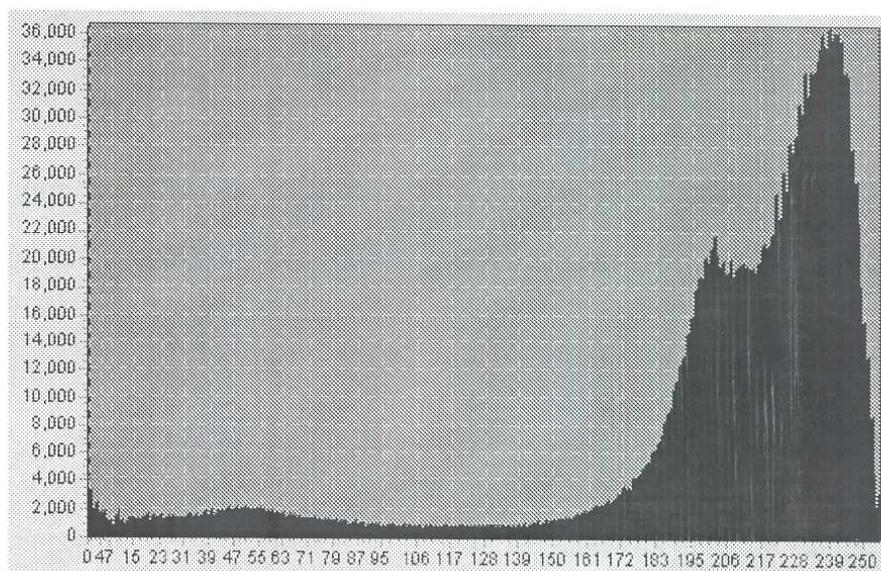


Figura 2.1 - b

Figura 2.1 Imagens de histogramas

Para fazer a seleção do limiar existem vários métodos disponíveis:

- Otsu [OTSU, 1979]: este método baseia-se na análise discriminante. A operação de binarização é obtida pelo particionamento dos pixels de uma imagem em duas classes $C_0 = \{0, 1, 2, \dots, t\}$ e $C_1 = \{t+1, t+2, \dots, l-1\}$, isto é, objeto e fundo.

Otsu demonstrou que o valor ótimo do limiar t^* pode ser determinado pela maximização do critério de separabilidade da função $\eta = \frac{\sigma_B^2}{\sigma_T^2}$, onde η é a variância entre classes dividida pela variância total: $t^* = \text{ArgMax } \sigma_B^2$,

$$\begin{aligned} \sigma_T^2 &= \sum_{i=0}^{l-1} (i - \mu_T)^2 P_i & \mu_T &= \sum_{i=0}^{l-1} i p_i, \\ \sigma_B^2 &= \omega_0 \omega_1 (\mu_1 - \mu_0)^2 & \omega_1 &= 1 - \omega_0 & \omega_0 &= \sum_{i=0}^t P_i, \\ \mu_1 &= \frac{\mu_T - \mu_0}{1 - \omega_0} & \mu_0 &= \frac{\mu_T}{\omega_0} & \mu_t &= \sum_{i=0}^t i p_i, \end{aligned}$$

sendo que ω_0 e ω_1 correspondem às variâncias nas classes C_0 e C_1 ; μ_0 e μ_1 correspondem às médias das classes, respectivamente. σ_B^2 , σ_T^2 as variâncias inter-classes e total, respectivamente.

- Kittler e Illingworth [SAHOO et al., 1988]: prevêm o histograma como uma estimativa da função densidade de probabilidade $p(g)$ da população, que compreende os níveis de cinza dos objetos e do fundo da seguinte maneira:

$$p(g) = \frac{1}{\sigma_1 \sqrt{2\pi}} * e^{-\left(\frac{(g-\mu_1)^2}{2\sigma_1^2}\right)} + \frac{1}{\sigma_2 \sqrt{2\pi}} * e^{-\left(\frac{(g-\mu_2)^2}{2\sigma_2^2}\right)}$$

Os parâmetros μ_i , σ_i^2 e as probabilidades a priori p_i com $i = 1, 2, \dots$ da densidade de probabilidade $p(g)$, associados a uma imagem a ser binarizada, não são usualmente conhecidos. Kittler e Illingworth introduziram uma função critério $J(t)$:

$$J(t) = 1 + 2(p_1(t) \log(\sigma_1(t)) + p_2(t) \log(\sigma_2(t))) - 2(p_1(t) \log(p_1(t)) + p_2(t) \log(p_2(t)))$$

onde os parâmetros $\mu_i(t)$, $\sigma_i^2(t)$ e as probabilidades a priori $p_i(t)$ com $i = 1, 2, \dots$ são definidos como:

$$p_1(t) = \sum_{g=0}^t h(g) \quad \text{e} \quad p_2(t) = \sum_{g=t+1}^{255} h(g)$$

$$\mu_1(t) = \frac{\sum_{g=0}^t gh(g)}{p_1(t)} \quad \text{e} \quad \mu_2(t) = \frac{\sum_{g=t+1}^{255} gh(g)}{p_2(t)}$$

$$\sigma_1^2 = \frac{\sum_{g=0}^t h(g)(g - \mu_1(t))^2}{p_1(t)} \quad \text{e} \quad \sigma_2^2 = \frac{\sum_{g=t+1}^{255} h(g)(g - \mu_2(t))^2}{p_2(t)}$$

onde g é o nível de cinza da imagem e $h(g)$ o valor do histograma para cada nível de cinza. Assim, o limiar T é obtido a partir da minimização da função $J(t)$.

- Pun [PUN, 1981]: sugeriu um algoritmo que tenta achar o limiar t que maximiza a função $H = H_b + H_w$ onde: $H_b(t) = -\sum_{i=0}^t p_i \log(p_i)$ é a entropia associada, para uma imagem com 256 níveis de cinza, com os pixels pretos binarizados com um limiar t . De forma análoga, a entropia dos pixels brancos é $H_w(t) = -\sum_{i=t+1}^{255} p_i \log(p_i)$. Pun demonstrou que basta maximizar a função $f(t)$:

$$f(t) = \frac{H_t}{H_T} \frac{\log(P_t)}{\log(\text{Max}\{p_0, p_1, \dots, p_t\})} + \left[1 - \frac{H_t}{H_T} \right] \frac{\log(1 - P_t)}{\log(\text{Max}\{p_{t+1}, p_{t+2}, \dots, p_{255}\})}$$

com

$$H_t = -\sum_{i=0}^t p_i \log(p_i), \quad H_T = -\sum_{i=0}^{255} p_i \log(p_i) \quad \text{e} \quad P_t = \sum_{i=0}^t p_i.$$

onde H_b , H_T e P_t representam, respectivamente, a entropia dos pixels pretos, a entropia total e a probabilidade cumulativa até o nível de cinza t .

- Kapur [KAPUR et al., 1985]: baseia-se no método de Pun. Kapur procura definir a probabilidade de distribuição A de um objeto contra a probabilidade de distribuição B do fundo, da seguinte forma:

$$A: \frac{p_0}{p_t}, \frac{p_1}{p_t}, \dots, \frac{p_t}{p_t}$$

$$B: \frac{p_{t+1}}{(1-p_t)}, \frac{p_{t+2}}{(1-p_t)}, \dots, \frac{p_{255}}{(1-p_t)}$$

O valor de limiar ótimo é o valor de t que maximiza a expressão $H = H_b(T) + H_w$

(T), onde as entropias dos pixels pretos é $H_b(t) = -\sum_{i=0}^t \frac{p_i}{P_t} \log\left(\frac{p_i}{P_t}\right)$ e dos pixels

brancos é $H_w(t) = -\sum_{i=t+1}^{255} \frac{p_i}{1-P_t} \log\left(\frac{p_i}{1-P_t}\right)$.

- Wong e Sahoo [WONG & SAHOO, 1989]: maximizam a entropia posterior de uma imagem particionada para constranger a medida de uniformidade de Levine e Nazif [LEVINE & NAZIF, 1985] e uma medida de formato. A entropia posterior é maximizada pelo $\min(s1, s2)$ e $\max(s1, s2)$ a fim de obter o limiar para a segmentação, onde $s1$ e $s2$ são os níveis de limiares onde as medidas de uniformidade e formato atingem os valores máximos, respectivamente.
- Bernsen [BERSEN, 1986]: o método de Bernsen é local ou adaptativo. Neste método, para cada pixel (x, y) , o limiar $T(x, y) = \frac{Z_{\min} + Z_{\max}}{2}$ é utilizado, onde Z_{\min} e Z_{\max} são os valores dos níveis máximos e mínimos da escala de cinza em uma vizinhança quadrada de dimensões $r \times r$ centralizada em (x, y) .

Todos esses métodos possuem uma desvantagem em comum: eles utilizam no cálculo somente a informação do histograma (ignorando os detalhes espaciais). Existem muitos esquemas de binarização que usam informação espacial. Por exemplo:

- Medida “busyness” de Weszka e Rosenfeld [WESZKA & ROSENFELD, 1978]: esta medida depende da ocorrência de pixels adjacentes na imagem. A medida “busyness” é minimizada para atingir o limiar para a segmentação
- Deravi e Pal [DERAVI & PAL, 1983]: minimizam a probabilidade condicional de transição através do limite entre duas regiões. Este método também usa a informação local contida na matriz de co-ocorrência da imagem.

Os métodos que binarizam o histograma utilizando detalhes espaciais geram resultados de segmentação mais significativos do que os métodos que usam somente a informação do histograma.

Mattana [MATTANA, 1999] propôs um estudo sobre os métodos de segmentação por binarização para cheques bancários brasileiros, visando à análise da qualidade final das imagens obtidas. Vários algoritmos de binarização globais e locais adaptativos foram aplicados, e a qualidade dos resultados foi avaliada por uma metodologia baseada no reconhecimento por OCR (Optical Character Recognition) das imagens binarizadas. Nos métodos globais, os que destacaram os melhores resultados foi Otsu com uma taxa de 86.12% para cheques não preenchidos e 66.1% para cheques preenchidos a máquina; e o método fuzzy da abordagem de Yager com uma taxa de 86.60% e 59.55%, respectivamente. Já nos métodos locais adaptativos, o método de Bernsen forneceu uma taxa de 92.79% e 84.03%, respectivamente.

2.2.5 Redes Neurais

Uma rede neural é uma coleção de unidades de processamento, chamadas de neurônios, dispostos em uma estrutura de uma ou de multi-camadas, conectados uns aos outros de uma maneira paralela; reagem coletivamente a um dado estímulo e produzem uma única resposta global. Uma rede neural tem a habilidade de aprender um dado tipo de conhecimento, armazenar este conhecimento e usá-lo mais tarde para executar a sua tarefa. Algumas das suas vantagens são: paralelismo, memória distribuída, capacidade de adaptação, tolerância a falhas, robustez diante de ruídos e saída em tempo real.

Vários autores estudaram a segmentação de imagens usando redes neurais. Blanz e Gish [BLANZ & GISH, 1990] usaram uma rede *feedforward* de três camadas para segmentação de imagens, onde o número de neurônios na camada de entrada depende do número de características de entrada para cada pixel, e o número de neurônios na camada de saída é igual ao número de classes. Babaguchi [BABAGUCHI et al., 1990] usou uma rede multi-camada treinada com o *backpropagation* para binarizar imagens. A entrada da rede é um histograma, enquanto a saída é o limiar desejado. Para a fase de treinamento, é preciso um conjunto grande de amostras de imagens com limiares conhecidos que produzam saídas adequadas. Mas, na prática, é muito difícil obter um número grande de amostras de imagens com limiares conhecidos.

Outra situação onde as redes neurais são bastante aplicadas é na segmentação de imagens em ambientes ruidosos. Um algoritmo robusto, para a extração de objetos em cenas altamente corrompidas por ruídos usando a rede neural tipo *Hopfield*, foi desenvolvido nas referências [GHOSH et al., 1991][GHOSH et al., 1992]. A função de energia da rede foi construída de tal maneira que, no estado estável da rede, ela extrai regiões compactas da cena ruidosa. A rede multi-camada [GHOSH et al., 1993], também foi aplicada para segmentar imagens ruidosas. Nesta rede, cada neurônio na camada i ($i > 1$) está conectado a um neurônio correspondente na camada $(i - 1)$ e alguns de seus neurônios vizinhos (na camada $i - 1$). A saída dos neurônios, na camada de saída, foi vista como um conjunto *fuzzy* (seção 2.2.8). As regras de atualizações dos pesos foram derivadas para minimizar o *fuzziness* do sistema. Para esse algoritmo, a arquitetura da rede impõe o sistema a considerar a informação de contexto. Além disso, esse algoritmo integra as vantagens das redes neurais (robustez) e dos conjuntos fuzzy (decisão de conhecimento impreciso/incompleto).

A abordagem neural também é aplicada na segmentação de texturas e na segmentação de imagens coloridas. Goltsev [GOLTSEV, 1996] usou a arquitetura de uma rede neural com organização *assembly* para o problema de segmentação de texturas em cenas naturais, enquanto Moreira [MOREIRA & COSTA, 1996] aplicou as redes neurais na segmentação de imagens coloridas.

Andrade [ANDRADE et al., 1994] modelou a segmentação como um problema de satisfação de restrições (PSR) resolvível através de uma rede neuronal sujeita a certas restrições espaciais (*constraint satisfaction neural network* ou, abreviadamente,

CSNN) [CHEN et al., 1991] apud [ANDRADE et al., 1994]. O problema de satisfação de restrições é, neste caso, interpretado como o processo de atribuição de rótulos aos pixels da imagem. Como resultado, os pixels pertencentes a cada região homogênea da imagem, segundo algum critério de similaridade previamente definido, recebem rótulos iguais por região. O algoritmo encontra uma solução ótima em tempo limitado, é ideal para máquinas paralelas e viabiliza a tarefa de segmentação em tempo real.

2.2.6 Algoritmos Genéticos (AG)

Os algoritmos genéticos são inspirados na teoria da evolução de Darwin. Eles pertencem à classe dos métodos probabilísticos de busca e otimização, embora não sejam aleatórios. Utiliza-se o conceito de probabilidade, mas os AG's não são simples buscas. Pelo contrário, eles tentam dirigir as buscas para regiões do espaço onde é provável que os pontos ótimos estejam.

De modo geral, os AG's têm as seguintes características [SRINIVAS & PATNAIK, 1994]:

- Operam numa população (conjunto) de pontos, e não a partir de um ponto isolado;
- Operam num espaço de soluções codificadas, e não num espaço de busca diretamente;
- Necessitam somente de informação sobre o valor de uma função objetivo para cada membro da população, e não requerem derivadas ou qualquer outro tipo de conhecimento;
- Usam transições probabilísticas, e não regras determinísticas.

Um dos problemas dos AG's é a determinação de parâmetros. No algoritmo básico, o usuário deve definir o tamanho da população, além das probabilidades de recombinação (um processo aleatório que ocorre com probabilidade fixa) e mutação (equivalente à busca aleatória). Em AG's mais sofisticados, há ainda mais parâmetros, comprometendo parte da robustez dos algoritmos. Infelizmente, não há regras claras para a escolha desses parâmetros.

Um algoritmo genético básico é apresentado abaixo:

1. [Iniciar] Gerar um conjunto de soluções adequadas para o problema. Essas soluções são representadas por cromossomos e o conjunto delas forma uma população.
2. [Adequabilidade] Avaliar a adequabilidade (*fitness*) $f(x)$ de cada cromossomo x na população.
3. [Nova população] Criar uma nova população repetindo os seguintes passos até que fique completa.
 - [Seleção] Selecionar dois cromossomos pais da população, de acordo com a sua adequabilidade (quanto maior a adequabilidade, maior a chance de ser selecionado);
 - [Recombinação] Troca-se fragmentos entre pares de cromossomos. Isso emula o fenômeno chamado de “*crossover*”;
 - [Mutação] Seleciona-se uma posição no cromossomo e muda-se o valor do gene correspondente aleatoriamente para outro alelo possível. O processo é geralmente controlado por um parâmetro fixo que indica a probabilidade de um gene sofrer mutação;
 - [Reconhecer] Identificar a nova prole na nova população.
4. [Substituir] Usar a nova população gerada para rodar o algoritmo;
5. [Teste] Se a condição final é satisfatória, pare e retorne à melhor solução na população corrente;
6. [Loop] Ir para o passo dois.

Os algoritmos genéticos também estão sendo aplicados na segmentação de imagens. Visa [VISA, 1998] desenvolveu um novo método baseado em algoritmos genéticos para fazer o pós-processamento em imagens de texturas segmentadas. O algoritmo genético é usado para extrair regras das imagens segmentadas. Essas regras são checadas e usadas em um pós-processamento para melhorar a segmentação.

2.2.7 Contornos Ativos – Snakes

Kass [KASS et al., 1987], em 1987, divulgou a comunidade científica o modelo de contornos ativos. Uma representação de contornos da imagem que permitisse a interação com processos de alto nível era o principal objetivo. Então, foi introduzido um modelo baseado em funções energéticas, cuja minimização traria um conjunto de soluções possíveis devido aos vários mínimos locais encontrados. Os processos de alto nível são utilizados para escolher uma solução de um conjunto encontrado. Esses processos são introduzidos através da adição de novos termos energéticos ao modelo, fazendo com que este se adapte às características das imagens desejadas. Durante o processo de minimização, o contorno ativo tem um comportamento parecido com uma cobra a rastejar, por isso a denominação de *Snake*.

O modelo de contornos ativos define-se como um mapeamento do contorno ativo, uma curva aberta ou fechada $[0, 1]$, no espaço \mathfrak{R}^2 :

$$\Omega = [0,1] \rightarrow \mathfrak{R}^2$$

$$s \rightarrow v(s) = (x(s), y(s))$$

O modelo tem associado um espaço de deformações admissíveis Ad e um determinado funcional $J(v(s))$:

$$J : Ad \in \mathfrak{R}^2 \rightarrow \mathfrak{R}$$

$$v(s) \rightarrow J(v(s)) = \frac{1}{2} \int_0^1 (E_{interna}(s) + E_{imagem}(s) + E_{constrangimento}(s)) ds$$

$E_{interna}$ é a energia interna total aplicada no contorno ativo. E_{imagem} corresponde às forças externas do contorno ativo. $E_{constrangimento}$ é a energia de constrangimento e é criada por campos de energia artificiais impostos pelo usuário ou por um agente de controle de alto nível.

A energia interna é definida como:

$$E_{\text{interna}} = \alpha \left| \frac{\partial v}{\partial s} \right|^2 + \beta \left| \frac{\partial^2 v}{\partial s^2} \right|^2,$$

onde α controla a quantia de alongamento que a snake está disposta a suportar e β controla a quantia de curvatura que ela permite.

A energia da imagem da snake é definida como:

$$E_{\text{imagem}} = w_1 I(x, y) + w_2 |\nabla I(x, y)|^2 + \dots,$$

onde w_1 é chamado coeficiente de linha e w_2 é chamado coeficiente de borda.

A energia de constrangimento externa é definida como:

$$E_{\text{constrangimento}} = -\sum_i k_i (x_{1i} - x_{2i}) + \sum_j \max \left(\text{pico}_j, \frac{1}{r_j} \right),$$

onde o termo k_i representa fatores de elasticidade externa e o termo r_j é chamado fator vulcão.

Vejam as principais vantagens e desvantagens dos contornos ativos [GONÇALVES, 1998].

Vantagens:

- Possibilidade de integração, numa única extração de contornos, com: dados da imagem, uma estimativa inicial do contorno, propriedades do contorno desejado, constrangimentos;

- O contorno ativo é modelado diretamente como uma curva, mantendo-se como tal durante todo o processo iterativo. Assim sendo, as características da curva (comprimento, curvatura e “adequação”) podem ser modificadas e calculadas no algoritmo;
- Existência de um critério ótimo, que envolve as características do contorno ativo e da imagem, que persegue uma solução ótima na minimização da energia do contorno ativo;
- Pode ser utilizado em aplicações 2D e estendido para 3D.

Desvantagens:

- O método converge para uma solução exata? O processo iterativo, ao encontrar um mínimo local, toma essa solução não ótima como ótima, não determinando o mínimo pretendido;
- O modelo dos contornos ativos carece de uma inicialização próxima do contorno a extrair, portanto não resolve independentemente o problema, porque poderá ficar obstruído em mínimos locais durante a minimização da sua energia;
- O contorno ativo deriva na forma de curvatura mínima quando não estão presentes forças exteriores; num ponto, no caso de um contorno ativo fechado, e numa linha, no caso de um contorno ativo aberto. É necessário englobar no modelo forças externas que influenciem o contorno ativo para que haja atração às características da imagem que se pretende extrair.

Desde que Kass, Witkin e Terzopoulos introduziram o modelo dos contornos ativos, várias foram as aplicações do método e também vários os tipos de imagens utilizados. O maior domínio de aplicação é o campo da bio-medicina, sendo a cardiovascular a preferida. Na literatura, encontram-se vários trabalhos que aplicam o modelo dos contornos ativos para extrair/segmentar as características desejadas:

- Eviatar [EVIATAR & SOMORJAI, 1996] utilizou o princípio dos contornos ativos para introduzir um modelo simplificado de detecção dos contornos do córtex, a partir de imagens de ressonância magnética do cérebro.

- Davatzikos [DAVATZIKOS & PRINCE, 1995] desenvolveu uma nova teoria sobre contornos ativos para detectar as reentrâncias do córtex.
- Chalana [CHALANA et al., 1996] utilizou contornos ativos para detectar duas fronteiras no coração, epicárdicas e endocárdicas, a partir da mesma imagem. Verifica-se a introdução de contornos ativos múltiplos em uma imagem para extrair características diferentes.
- Gonçalves [GONÇALVES, 1998] apresentou um ambiente de trabalho para que, durante uma cirurgia no coração, seja possível extrair e parametrizar os contornos da válvula mitral. O sistema é baseado na aplicação dos modelos de contornos ativos, servindo como apoio ao cirurgião, na sua decisão sobre o implante a aplicar. Gonçalves efetuou uma comparação com os vários algoritmos de implementação dos contornos ativos. Os algoritmos propostos por Kass [KASS et al., 1987], Amini [AMINI et al., 1990], Cohen [COHEN, 1991], Williams [WILLIAMS & SHAH, 1992] e Eviatar [EVIATAR & SOMORJAI, 1996] foram implementados. Os critérios utilizados foram o esforço e tempo computacional, a simplicidade de afinação de parâmetros, a inicialização do contorno ativo e os resultados obtidos nas imagens de testes. O algoritmo que reuniu as melhores características, necessárias à extração dos contornos da válvula mitral e à implementação no ambiente de trabalho desenvolvido, foi o método de Kass, modificado por algumas características do método de Cohen. O algoritmo é bastante rápido, não incluindo um esforço computacional elevado. A inicialização do contorno ativo é efetuada exteriormente ao objeto do qual se pretende extrair os contornos. Apesar de existirem somente quatro parâmetros para afinar, este processo requer alguma técnica, tal como os outros algoritmos, de forma a se obter resultados satisfatórios.

Bamford [BAMFORD & LOVELL, 1998] fez uma análise *bayesiana* da segmentação de núcleos de células, usando a pesquisa Viterbi, baseada em contornos ativos.

2.2.8 Métodos Baseados na Teoria Fuzzy

Zadeh [ZADEH] introduziu o conceito de conjuntos *fuzzy*, onde o conhecimento impreciso pode ser usado para definir um evento. Um conjunto *fuzzy* A é representado como:

$$A = \{\mu_A(x_i)/x_i, i = 1, 2, \dots, n\}$$

onde $\mu_A(x_i)$ indica o grau do elemento x_i ser membro do conjunto A.

A relevância da teoria dos conjuntos *fuzzy*, em problemas de reconhecimento de padrões, tem sido adequadamente tratada na literatura [BEZDEK, 1981][BEZDEK & PAL, 1992]. Em vários estágios de um sistema de reconhecimento de padrões, a teoria dos conjuntos *fuzzy* pode ser incorporada para manipular incertezas. Estas incertezas originam-se de deficiências de informação, a qual pode ser incompleta, imprecisa, vaga, contraditória, mal definida.

A incerteza em um padrão de imagem pode ser interpretada em termos de ambigüidade do conceito de nível de cinza (*grayness*) ou ambigüidade espacial (geométrica) ou ambas. A ambigüidade de *grayness* significa a incerteza na decisão se um pixel é branco ou preto. A ambigüidade geométrica refere-se à incerteza do formato e geometria de uma região dentro da imagem.

As técnicas convencionais de análise e reconhecimento de imagens consistem na segmentação da imagem em regiões significativas, extraindo suas bordas e esqueletos, processando várias características/propriedades (por exemplo: área, perímetro, centróide, etc.) e primitivas (por exemplo: linha, curva, cantos), e o relacionamento entre as regiões; finalmente, desenvolvem regras/gramáticas de decisão para descrever, interpretar e/ou classificar a imagem e suas sub-regiões. No sistema convencional, cada uma dessas operações envolve decisões claras (isto é, sim ou não, branco ou preto, 0 ou 1) sobre regiões, características, primitivas, propriedades, relacionamentos e interpretações.

Como as regiões de uma imagem nem sempre são definidas claramente, incertezas podem surgir em todas as tarefas citadas acima. Uma decisão em um nível

particular terá um impacto em todas as atividades dos níveis mais altos. Por exemplo, o problema de extrair objetos em uma cena. A questão é: “Como definir exatamente o objeto ou a região do objeto na cena quando sua borda está mal definida?” Uma binarização para a extração do objeto propagará a incerteza associada para estágios subsequentes (por exemplo, afinamento, extração do esqueleto, seleção de primitivas) e isto pode, sucessivamente, afetar a análise e reconhecimento de características.

Portanto, é conveniente permitir que segmentos ou contornos ou esqueletos sejam subconjuntos *fuzzy* da imagem; os subconjuntos são caracterizados pela possibilidade (grau) de cada pixel pertencer a eles. Similarmente, para descrever e interpretar informações estruturais mal definidas em um padrão é natural definir linha, ângulo, curva, etc. e o relacionamento entre elas, usando rótulos de conjuntos *fuzzy*. A produção de regras de uma gramática pode similarmente ser nebulosa (*fuzzified*) para calcular a incerteza no relacionamento entre as primitivas; desse modo, aumenta o poder produtivo de uma gramática para reconhecimento sintático de um padrão.

Na literatura, encontram-se muitos métodos de segmentação *fuzzy*, alguns deles são: binarização *fuzzy*, agrupamento *fuzzy* e detecção de bordas *fuzzy*.

Diferentes técnicas de binarização do histograma, em que ambas as versões *fuzzy* e não *fuzzy* minimizam a ambigüidade *grayness* (entropia global, índice de nebulosidade, índice de clareza) e ambigüidade geométrica (compacidade *fuzzy*) de uma imagem foram descritas nas referências [PAL & KING, 1980] [PAL & ROSENFELD, 1988]. Estes algoritmos utilizam diferentes funções de associação tipo S (Figura 2.2) para definir “regiões de objetos” *fuzzy* e então selecionar uma que esteja associada com o valor mínimo (ótimo) das medidas citadas acima. O problema é definir uma função de associação apropriada. Xie e Bedrosian [XIE & BEDROSIAN, 1988] tentaram determinar funções de associação para imagens em níveis de cinza.

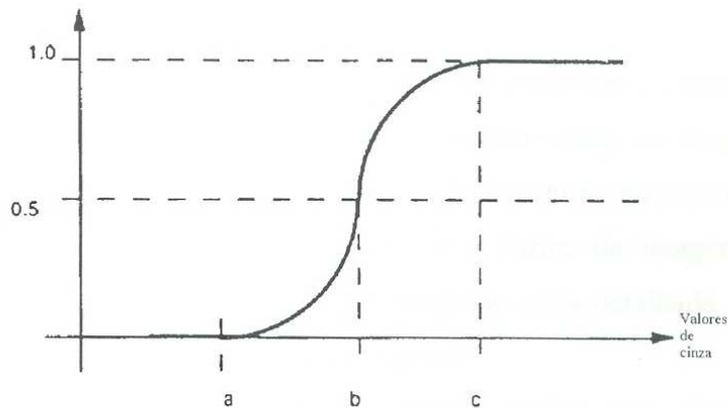


Figura 2.2 Função de relacionamento tipo S

O *Algoritmo de Agrupamento Fuzzy C-Means* (AFC) [BEZDEK, 1981] também foi bastante usado na segmentação de imagens [HUNTSBERGER et al., 1985][TRIVEDI & BEZDEK, 1986][HALL et al., 1992]. O algoritmo AFC usa uma otimização iterativa de uma função objetivo baseada na medida de similaridade ponderada entre os pixels na imagem e cada um dos centros *c-cluster*. Um extremo local desta função objetivo indica um agrupamento ótimo dos dados de entrada. A função objetivo que é minimizada é a seguinte:

$$W_m(U, V) = \sum_{k=1}^n \sum_{i=1}^C (\mu_{ik})^m (d_{ik})^2,$$

onde μ_{ik} é o valor de associação *fuzzy* do *kth* pixel no *ith* agrupamento, d_{ik} é a distância específica, m controla a natureza do clustering, V é o conjunto de centros *c-cluster* e U é a partição *fuzzy* da imagem.

Nas referências [PAL N. R. & PAL S. K., 1989] [PAL, 1990] a entropia de um conjunto *fuzzy*, definida por uma função de associação adaptativa, sobre a vizinhança de um pixel (x, y) , é usada como uma medida de borda em (x, y) . O uso de uma função de associação adaptativa torna o algoritmo de detecção de bordas robusto.

2.2.9 Morfologia Matemática

A idéia central da morfologia matemática é examinar a estrutura geométrica de uma imagem, casando-a com pequenos padrões em várias localizações. Variando o tamanho e as formas dos padrões, denominados elementos estruturantes, pode-se obter informações úteis sobre a forma de diferentes partes da imagem, bem como do relacionamento entre essas partes. Uma introdução mais detalhada sobre morfologia matemática será apresentada no próximo capítulo.

Autores como Fernandes [FERNANDES, 1997], Neto [NETO, 1998], Ohta [OHTA & SAKAI, 1980] e outros desenvolveram metodologias de segmentação de imagens, utilizando morfologia matemática.

Fernandes [FERNANDES, 1997], a partir da teoria da morfologia matemática, segmentou componente gráfico e textual de cartas náuticas, tais com rios, lagoas, estradas de ferro, canais, caracteres, etc. As operações desenvolvidas por Fernandes podem ser associadas, por exemplo, a um sistema eletrônico de manipulação e reconhecimento de cartas náuticas, visando, entre outros, ao armazenamento da informação em bancos de dados geográficos ou hidrográficos. As funções definidas são específicas às formas e estruturas dos componentes das imagens de cartas náuticas.

Neto [NETO, 1998] utilizou operações morfológicas para determinar as regiões de interesse em imagens de impressões digitais. A segmentação apresentou bons resultados, apesar da diversidade das imagens. Problemas de translação nas regiões de interesse das imagens foram contornados. Mas, o método de segmentação proposto por Neto mostrou-se bastante sensível a efeitos de rotação.

A próxima seção apresentará a aplicação da segmentação na área de análise e processamento de documentos.

2.3 Segmentação de Imagens de Documentos

Como descrito no Capítulo 1, um documento pode ser visto como qualquer base de conhecimento, fixada materialmente e disposta de maneira que se possa utilizar para consulta, estudo, prova, etc. Essa definição vem ao encontro do tipo de estudo

desenvolvido no campo de Processamento de Documentos. Nele, o documento “fixado materialmente” serve de base para que informações pertinentes ao seu conteúdo sejam extraídas. O objetivo é reconhecer o texto e os elementos gráficos contidos na imagem desse documento, chegando a formular uma análise análoga à que um observador humano seria capaz de fazer. Existem vários motivos para que se processe a extração de componentes de documentos; os principais deles baseiam-se na reedição da informação, armazenamento e redistribuição, e na recombinação com outras informações.

Para que tal análise seja feita, o Processamento de Documentos é categorizado de duas formas (Figura 2.3):

- A primeira delas está particularmente interessada em zonas de texto, seja por Reconhecimento Automático de Caracteres (OCR), seja pela análise da composição da página de um documento (palavras, linhas de texto, parágrafos, etc);
- A outra está interessada em elementos gráficos de documentos como símbolos em diagramas, fotografias, partituras musicais, desenhos de engenharia, logotipos de companhias, etc.

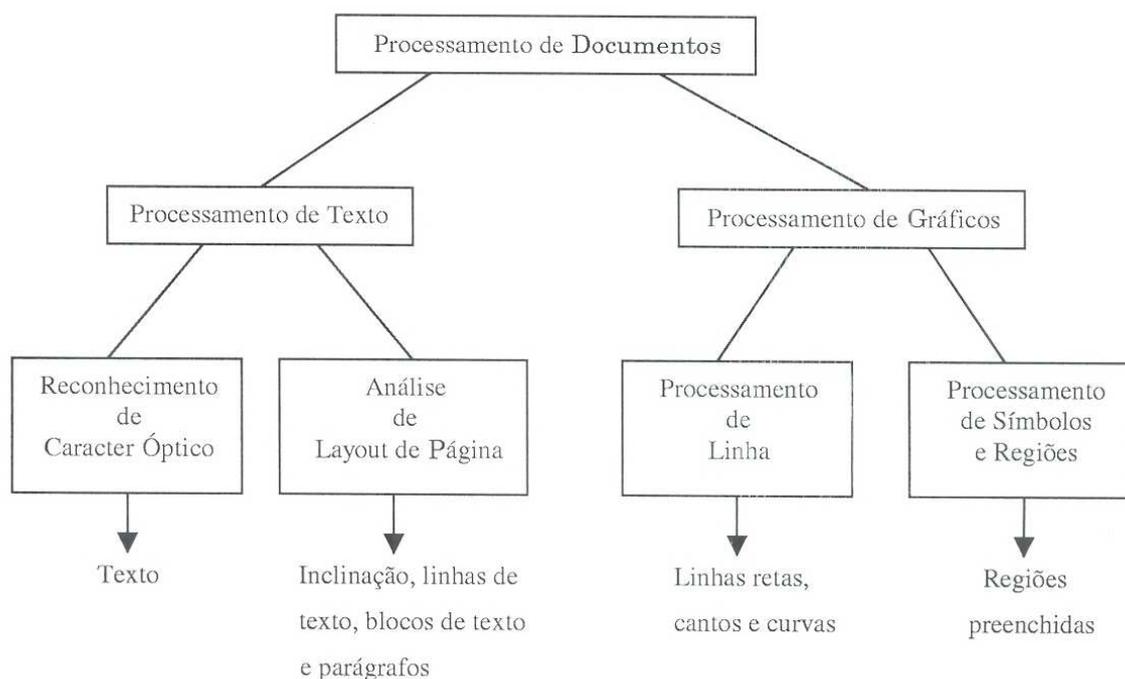


Figura 2.3 Uma hierarquia das subcategorias do Processamento de Documentos, listando os tipos de componentes do documento [GORMAN & KASTURI, 1995]

Os componentes (inclinação, curvas, cantos, regiões preenchidas e linhas) do documento são usados como características intermediárias que auxiliam no reconhecimento final dos caracteres por OCR e, linhas e símbolos para a análise gráfica. Autores como Hinds [HINDS et al., 1990], Tang [TANG et al., 1997], Medioni [MEDIONI & YASUMOTO, 1987], Gorman [GORMAN, 1988] apresentaram métodos para a detecção da inclinação do documento, extração de linhas de referência, detecção de cantos e detecção de características curvilíneas, respectivamente.

A maioria dos documentos consiste de uma mistura de texto, gráfico e figuras. Portanto, as imagens de documentos precisam ser segmentadas para identificar cada uma destas regiões a fim de que técnicas de reconhecimento, específicas de cada categoria (texto, gráfico), sejam aplicadas.

Autores como Jain [JAIN & BHATTACHARJEE, 1992] e Patel [PATEL, 1996] propuseram métodos para a segmentação de páginas de documento. Ambos consideram o documento como uma região texturizada, onde as regiões texto possuem uma textura diferente das regiões não texto. Jain [JAIN & BHATTACHARJEE, 1992] aplicou os filtros bidimensionais de Gabor para extrair as características de cada uma dessas regiões e o método não assumiu nenhum conhecimento a priori do conteúdo, ou estilos de fontes do documento. Patel [PATEL, 1996] usou um conhecimento a priori sobre o formato do documento e aplicou o método de convolução para gerar as imagens de características de textura. Os coeficientes das máscaras de convolução foram obtidos usando uma rede neural artificial de camada única. Os coeficientes destas máscaras foram “treinados” com exemplos de imagens de documentos e tem um potencial mais poderoso do que as máscaras com coeficientes pré-selecionados. Um esquema de binarização baseado na medida de entropia foi usado para segmentar as imagens de características em regiões homogêneas.

Devido às particularidades de cada documento (carta, ofício, mapas, cheques bancários, etc.), as metodologias de segmentação tendem a ser específicas para cada tipo de documento. O cheque bancário é um documento que já gerou muitos estudos relacionados à segmentação (por exemplo: segmentação da data, da assinatura) mas, em geral, as metodologias que foram desenvolvidas não são aplicáveis para outros tipos de documentos.

Okada e Malayappan [OKADA & MALAYAPPAN, 1997] propuseram um método para extrair os componentes de entrada de um cheque bancário. O método utilizou um par de imagens: uma imagem de cheque em branco (imagem de referência) e uma imagem de cheque preenchido (imagem amostra). As duas imagens diferem somente em regiões específicas do cheque reservadas para inserir componentes manuscritos e pré-impessos. Os autores aplicaram a subtração morfológica para extrair esses componentes. O esquema de subtração foi dividido em dois passos: fusão em níveis de cinza (*gray-level space - GLS*) e subtração lógica. Este método pode ser aplicado em outros formatos, mas a principal desvantagem é que as imagens de referência e amostra não podem ser muito diferentes.

Santos [SANTOS, 1997] fez um estudo sobre métodos e técnicas para a segmentação de imagens de cheques bancários. A tabela 2.1 mostra os resultados de algumas análises feitas pelo autor.

Método	Aplicação	Vantagens	Desvantagens
Transformada De Hough	Detecção e eliminação de linhas; detecção do ângulo de inclinação; segmentação de blocos.	Versatilidade	Limitada pela complexidade do problema
Wavelets	Detecção e eliminação de linhas.	Elimina o fundo dos cheques durante o processo de eliminação das linhas.	Alto grau de complexidade na implementação; precisa de outra ferramenta para eliminar as linhas
Filtro de Kalman	Detecção e eliminação de linhas.	Pequeno grau de degradação da informação.	Necessita de um grande número de parâmetros
Morfologia Matemática	Eliminação de linhas e reconstrução da informação perdida	Versatilidade. Dispensa qualquer conhecimento sobre a imagem.	Altamente intuitivo

Tabela 2.1 Análise de técnicas para a segmentação de cheques bancários [SANTOS, 1997].

“A segmentação de imagens de documentos parece ser uma das áreas mais promissoras no âmbito de Processamento de Imagens, no que se refere ao seu desenvolvimento científico. A necessidade emergente de processos ágeis e eficazes vem

fazendo com que as técnicas existentes na área de Processamento de Documentos correspondam cada vez mais rapidamente a essas necessidades” [SANTOS, 1997].

2.4 Segmentação de Logotipos

Logotipos são caracterizados como textos e símbolos gráficos misturados e, quando são reconhecidos, disparam uma associação do objeto em que eles estão anexados com um dado grupo ou organização.

Como descrito no Capítulo 1, os logotipos podem ser divididos em quatro tipos: *word-in-mark*, *device-mark*, *composite-mark* e *complex-mark*. Geralmente, os trabalhos encontrados na literatura, como o trabalho de Y.S.Kim & W.Y.Kim [KIM, Y.S. & KIM, W.Y., 1998] processam somente um tipo de logotipo.

No domínio de documentos, é grande o interesse no reconhecimento de logotipos, pois esses elementos são usados para distinguir documentos similares, porém diferentes, ou para indicar características específicas do documento em questão.

Nesse domínio, duas tarefas análogas de reconhecimento de logotipos são de interesse e requerem indexação em um espaço grande de possibilidades:

- Dado um documento que contém um logotipo, classificá-lo como pertencente a um conjunto finito de logotipos, ou concluir que ele não está presente na base de dados;
- Dado um logotipo representativo (conhecido ou desconhecido), indexá-lo à base de dados de documentos e extrair todos os documentos que contenham este logotipo.

Vários autores concentraram seus estudos na etapa de reconhecimento dos logotipos. Mas, o reconhecimento não pode ser realizado sem existir antes uma fase de conhecimento e, conhecer significa extrair características dos padrões a serem processados. Porém na literatura, pouco se fala da segmentação (extração) automática de logotipos.

Cesarini [CESARINI et al., 1997], Suda [SUDA et al., 1997] e Soffer [SOFFER & SAMET, 1998] despenderam esforços na etapa de reconhecimento, mas a localização e segmentação dos logotipos foram realizadas de forma manual. Doermann [DOERMANN et al., 1993], em seu trabalho, aplicou, na etapa de localização do

logotipo, o algoritmo de segmentação de página, mas a segmentação completa e exata não foi possível para todas as situações, porque houve casos em que os logotipos foram gerados com partes ausentes, regiões corrompidas ou com a presença de componentes que não faziam parte do logotipo.

2.5 Análise de Técnicas para Segmentação de Logotipos

O objetivo desse trabalho é propor uma metodologia de segmentação de logotipos, em cheques bancários brasileiros, que aplica um número reduzido de parâmetros heurísticos. Desta forma, a metodologia poderá ser aplicada não somente em cheques bancários, mas em outros tipos de documentos (fax, cartas, ofícios, entre outros). Para isso, a técnica a ser aplicada não pode, principalmente, depender de:

- Etapa de aprendizagem: limita a aplicação ao tipo de documento presente na base de treinamento;
- Informações a priori sobre a localização do logotipo: limita a aplicação para um único tipo de documento. Isto acontece porque a posição do logotipo varia entre os diferentes tipos de documentos.
- Informações a priori sobre a estrutura do logotipo: limita a aplicação para um tipo particular de logotipo.
- Escolha de um modelo matemático: limita a aplicação para o tipo de informação presente na base de dados. Isto acontece porque muitas vezes a escolha do modelo matemático é feita com base nas informações presentes na base de dados.

Com base nos itens citados acima, a seguinte análise é feita sobre as técnicas de segmentação:

- Segmentação por textura: os cheques bancários possuem texturas variadas, isto dificulta a escolha de um modelo de textura. Além disso, esse modelo de textura seria baseado nas características do cheque bancário, impedindo a aplicação para outros tipos de documentos.
- Segmentação por cor: os cheques bancários possuem diferentes padrões de cores, isto dificulta a escolha de um sistema de representação de cor. E no ponto

de vista computacional, o processamento da informação cor implica em aumento dos custos e precisa de um espaço de armazenamento grande.

- Detecção de bordas: essa técnica não é viável para a aplicação em questão, porque as bordas não especificam uma característica do logotipo, outros elementos contidos no cheque também apresentam bordas.
- Redes neurais: esta técnica necessita da fase de aprendizagem, apresenta um custo computacional alto e precisa de informações a priori sobre a estrutura do logotipo.
- Algoritmos genéticos: o principal problema dessa técnica é a determinação de parâmetros. Um desses parâmetros é a função de *fitness*. Esta função está diretamente relacionada com os dados em estudo.
- Contornos ativos (*snakes*): necessita de informação a priori sobre a localização dos elementos para fazer uma segmentação prévia. Além disso, apresenta cálculos matemáticos complexos e pesados.
- **Morfologia matemática:** é a técnica mais adequada para o objetivo em questão. Dispensa informações a priori sobre o tipo de documento e localização do logotipo; não possui cálculos complexos; não necessita de aprendizagem.

A única informação prévia utilizada nesse trabalho é: o logotipo é maior do que todas as informações presentes no cheque bancário. Esta informação para a morfologia matemática é muito útil, porque esta técnica extrai informações relacionadas à geometria e à topologia da imagem.

Capítulo 3

Estrutura de um Cheque Bancário e Base de Dados

3.1 Introdução

A figura do cheque bancário aparece, na prática bancária, como um instrumento que os bancos colocam à disposição dos clientes para movimentar os fundos disponíveis em seu poder e efetuar pagamentos sem a necessidade do manuseio da moeda. Um cheque representa uma ordem incondicional em direito que é emitida por uma pessoa em favor de outra pessoa, é assinado pela pessoa que emite, e exige que o banco pague na apresentação uma determinada quantia em dinheiro [KOERICH, 1997].

Neste capítulo serão descritos os seguintes itens:

- As dimensões do cheque bancário;
- Elementos que compõem o cheque bancário brasileiro;
- Preenchimento de cheques bancários;
- Normas do Banco Central;
- Base de dados que será utilizada para validar a metodologia proposta.

3.2 Dimensões do Cheque Bancário Brasileiro

As dimensões dos cheques bancários brasileiros são estabelecidas pelo Banco Central. Segundo o manual de Normas de Instruções [BANCO, 1983], as dimensões dos cheques bancários devem obedecer às medidas em milímetros descritas na tabela abaixo.

Medida	Comprimento	Largura
Valor Médio	175	80
Valor Mínimo	171	76
Valor Máximo	176	81

Tabela 3.1 Normas de dimensões de um cheque bancário brasileiro

A Figura 3.1 mostra os limites de tamanho máximo e mínimo em pixels para os cheques bancários brasileiros digitalizados em 200 dpi.

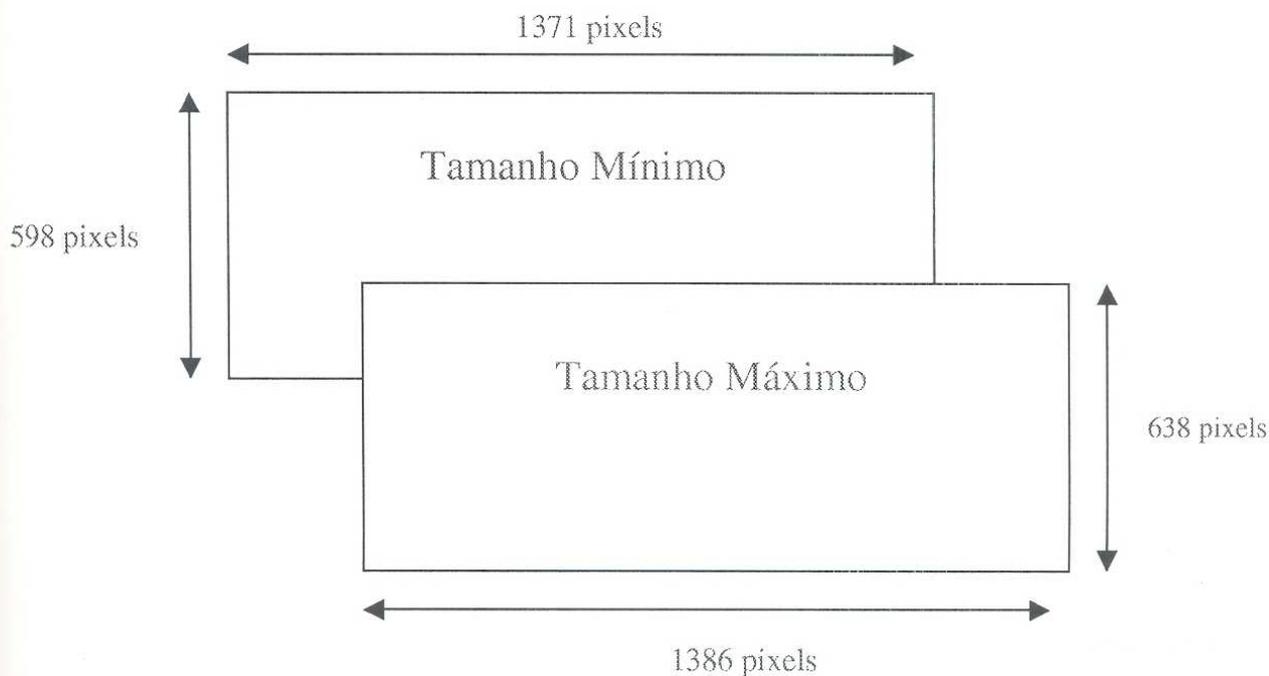


Figura 3.1 Limites de tamanho para os cheques bancários brasileiros em pixels

3.3 Elementos que Compõem o Cheque Bancário Brasileiro

O Conselho Monetário Nacional instituiu um modelo único de cheque visando a uma maior segurança e eficiência nos serviços que os bancos prestam à coletividade. A localização dos elementos dentro da área do cheque e seus campos de preenchimento também estão definidos no Manual de Normas e Instruções.

As informações relevantes do cheque são constituídas de campos pré-impessos no corpo do cheque, contendo informações do banco como: número de compensação, número do banco, número da agência, número do cheque, dígitos verificadores, logotipo do banco, nome do banco, endereço da agência. Há também informações pré-impresas relativas ao proprietário da conta como: número da conta, valor numérico em moeda corrente da folha do cheque, valor por extenso, nome do favorecido, local e data da emissão, assinatura do cliente, nome do cliente, número do CPF/CGC, além de outras informações como telefone ou RG. O cheque ainda possui, em seu canto inferior, uma linha MICR CMC -7 (código de barras) destinado à leitura digital dos seguintes dados: número de compensação, número do banco, número da agência, número do cheque, dígitos verificadores. Os elementos citados acima estão exemplificados na Figura 3.2

Número de Compensação, Número do Banco, Número da Agência, Número da Conta, Número do cheque e Dígitos Verificadores		Valor Numérico
Valor por Extenso		
Nome do Favorecido		
Logotipo do Banco	Local e Data de Emissão	
Identificação da Agência (Nome, Número, Endereço e Telefone) Informações sobre o cliente - limite de crédito	Assinatura, Nome do Cliente, Número de CPF/CGC, outras informações (telefone - RG - Conta Conjunta)	
Linha MICR CMC - 7		

Figura 3.2 Modelo para a distribuição de informações nos cheques bancários brasileiros [KOERICH, 1997].

3.4 Preenchimento de Cheques Bancários

Conforme as normas do Banco Central, o preenchimento dos campos de um cheque pode ser feito a mão ou por meios mecânicos, como impressoras ou máquinas de escrever, pelo próprio emitente ou por terceiro [COVELLO, 1996].

Os campos de preenchimento de um cheque são: valor numérico, valor por extenso, beneficiário, local, data e assinatura do cliente. O layout do cheque possui linhas pré-impressas para delimitação dos campos e orientação do preenchimento. A Figura 3.3 mostra os campos e o preenchimento a mão de um cheque bancário.

Beneficiário		Valor por Extenso			Valor Numérico			
Comp	Banco	Agência	CI	Conta	C2	Cheque No	C3	R\$
009	409	0282	6		9	EX 600412	1	100,00
Pague por este cheque a quantia de		Cem Reais						
a		Banco Central						centavos acima
UNIBANCO		Coba, 28			de março			de 1998
Cheque Especial		Cecilia Pelin						
Unibanco - União de Bancos Brasileiros S.A.		tel: 2246989						
0439 CINELANDIA		RG: 3843044-2 (92)						
RUA LUIS XAVIER 79								
CURITIBA PR								
		Local e Data			Assinatura			

Figura 3.3 Exemplo de um cheque mostrando os campos e seu preenchimento

O elemento valor numérico (R\$) é destinado à especificação do valor do cheque em algarismos. A expressão “Pague por este cheque a quantia de” seguida de duas linhas é reservada para a especificação do valor por extenso, sendo os centavos opcionais, conforme as normas do Banco Central. Uma terceira linha, iniciada com a preposição “a” e terminada com a expressão “ou a sua ordem” ou se for o caso, “não a ordem” é reservada para a indicação do nome do beneficiário do cheque. À direita, abaixo do nome do beneficiário, devem ser impressas as linhas reservadas à indicação

do local e data de emissão do cheque e à assinatura do emitente [BANCO, 1983]. Na seção 3.5, serão descritas as normas estabelecidas pelo Banco Central relacionadas aos elementos citados acima, com exceção do nome do beneficiário.

Outras características do cheque como: fundo artístico, faixa de segurança, impressão com tinta ultra-violeta estão descritas em [MATTANA, 1999].

3.5 Normas do Banco Central

Algumas normas estabelecidas pelo Banco Central [BANCO, 1983] relacionadas ao valor numérico, valor por extenso, data e local da emissão e assinatura são descritas abaixo:

- Em casos de divergências quando os valores numéricos e por extenso são indicados, prevalece o valor por extenso. Quando esses valores são indicados mais de uma vez, prevalece o menor valor, conforme a lei nº 7.357, de setembro de 1985, art.12º.
- Na emissão de cheques, é facultativa a grafia por extenso dos valores correspondentes aos centavos, conforme o decreto nº 99.179, de outubro de 1990.
- A assinatura do cheque é permitida que seja impressa por um processo mecânico, ou seja, por chancela mecânica que reproduz uma assinatura exata do próprio punho, conforme resolução 885.
- Em casos de não ser indicado o nome do lugar da emissão, considera-se emitido o cheque no lugar indicado junto ao nome do emitente, conforme a lei nº 7.357, de setembro de 1985, art. 2º.
- Cheque pode ser devolvido quando apresentar algumas das seguintes irregularidades: sem data de emissão, com o mês grafado numericamente, porque hoje se escreve somente o mês por extenso (decreto nº 22.393 de janeiro de 1993), ausência de assinatura, não registro do valor por extenso.

3.6 Base de Dados

A criação de uma base de dados de cheques é uma tarefa difícil. Porque é preciso ter uma instituição financeira que esteja disposta a ceder cheques reais, mas por motivos de segurança e proteção, sobretudo por contrato com seus clientes, não o faz; a não ser que exista um acordo entre um centro de pesquisas e a própria instituição. Como no Brasil existem cerca de 220 bancos credenciados, um contrato desta natureza com todas as instituições torna-se inviável. Por isso, as publicações voltadas ao processamento de cheques bancários apresentam resultados obtidos a partir de pequenas bases de dados, muitas vezes restritas regionalmente [KOERICH, 1997]. Exemplos de bases de dados existentes e disponíveis são descritas na tabela 3.2.

Bases de Dados Existentes		
Base de Dado	Tipo	Local
CEDAR	Base de dados real; dígitos e palavras manuscritas; contexto de correio.	Center of Excellence in Document Analysis Recognition, Buffalo, USA
CENPARMI	Base de laboratório e base real; dígitos e palavras manuscritas; contexto de cheques bancários.	Center of Pattern Analysis and Machine Intelligence, Montreal, Canadá
NIST	Dígitos e palavras manuscritas; contexto geral.	National Institute Standards and Technology, USA
LRPRC	Base de dados real de cheques bancários brasileiros; contempla 13 instituições bancárias	Laboratório de Reconhecimento de padrões e Redes Neurais de Computadores da Universidade Estadual de Campinas – Brasil [KOERICH, 1997]
LARDOC I	Base de dados de laboratório; palavras e dígitos manuscritos; contexto de datas.	Laboratório de Análise e Reconhecimento de Documentos Pontifícia Universidade Católica do Paraná, Brasil
LARDOC II	Base de dados real de cheques bancários brasileiros; contempla 18 instituições bancárias.	Laboratório de Análise e Reconhecimento de Documentos Pontifícia Universidade Católica do Paraná, Brasil

Tabela 3.2 Algumas bases de dados existentes

A metodologia proposta foi validada com imagens da base de dados LARDOC II. Esta base possui imagens coloridas de cheques bancários brasileiros. Todos os cheques foram adquiridos em 300 dpi's e 16 milhões de cores. Outras informações sobre a aquisição dos cheques encontram-se em [MATTANA, 1999].

Como a tarefa de extrair informações em imagens coloridas é complexa e exige um esforço computacional grande, as imagens da base de dados LARDOC II foram transformadas em imagens em tons de cinza.

Nos testes, foram utilizadas 58 imagens de cheques em branco, 190 imagens de cheques preenchidos a máquina e 230 imagens de cheques preenchidos a mão. Um total de 478 imagens de cheques bancários brasileiros. As tabelas 3.3, 3.4 e 3.5 descrevem a base de dados.

Banco	Codificação	Quantidade
Banco Meridional	MER	01
Banco de Crédito Nacional SA	BCN	01
Banco América do Sul	BAS	02
Banco Bradesco	BRA	02
Banco HSBC Bamerindus SA	HSBC	02
Unibanco	UNI	04
Banco Itaú	ITA	04
Banco do Brasil	BB	04
Caixa Econômica Federal	CEF	06
Banco Real	REA	09
Banco Banestado	BAN	09
Banco Bamerindus SA	BAM	14
Total de cheques sem preenchimento		58

Tabela 3.3 Descrição da base de dados de cheques bancários sem preenchimento

Banco	Codificação	Quantidade
Banco Santander SA	SAN	01
Banco Bandeirantes	BBS	01
Banco Bamerindus SA	BAM	02
Unibanco	UNI	03
Banco Real	REA	03
Banco Bradesco	BRA	16
Banco Itaú	ITA	21
Banco HSBC Bamerindus SA	HSBC	22
Caixa Econômica Federal	CEF	32
Banco Banestado	BAN	41
Banco do Brasil	BB	48
Total de cheques preenchidos a máquina		190

Tabela 3.4 Descrição da base de dados de cheques bancários preenchidos a máquina

Banco	Codificação	Quantidade
Banco do Crédito Nacional SA	BCN	01
Banco Banestes	BES	01
Banco Noroeste SA	BNO	02
Banco Bamerindus SA	BAM	04
Banco América do Sul	BAS	04
Banco Excel	EXC	06
Unibanco	UNI	08
Banco Bandeirantes	BBS	09
Citybank SA	CIT	10
Banco Bradesco	BRA	11
Banco Itaú	ITA	16
Banco Real	REA	19
Caixa Econômica Federal	CEF	22
Banco do Brasil	BB	28
Banco HSBC Bamerindus SA	HSBC	31
Banco Banestado	BAN	58
Total de cheques preenchidos a mão		230

Tabela 3.5 Descrição da base de dados de cheques bancários preenchidos a mão

Capítulo 4

Morfologia Matemática

4.1 Introdução

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, qualquer que seja a área em que se aplique (biologia, sociologia, gramática, etc).

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.

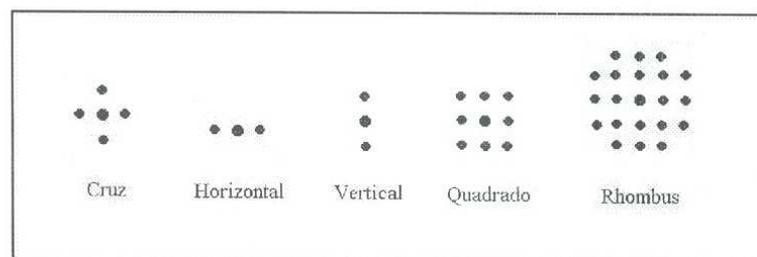


Figura 4.1 Tipos de elementos estruturantes básicos

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.

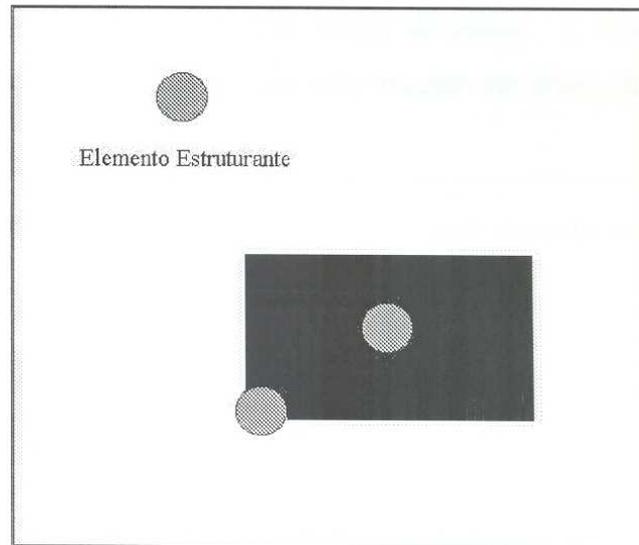


Figura 4.2 Exemplos de interação do elemento estruturante em uma imagem binária

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. A Figura 4.3 mostra a interação do elemento estruturante com uma imagem em níveis de cinza.

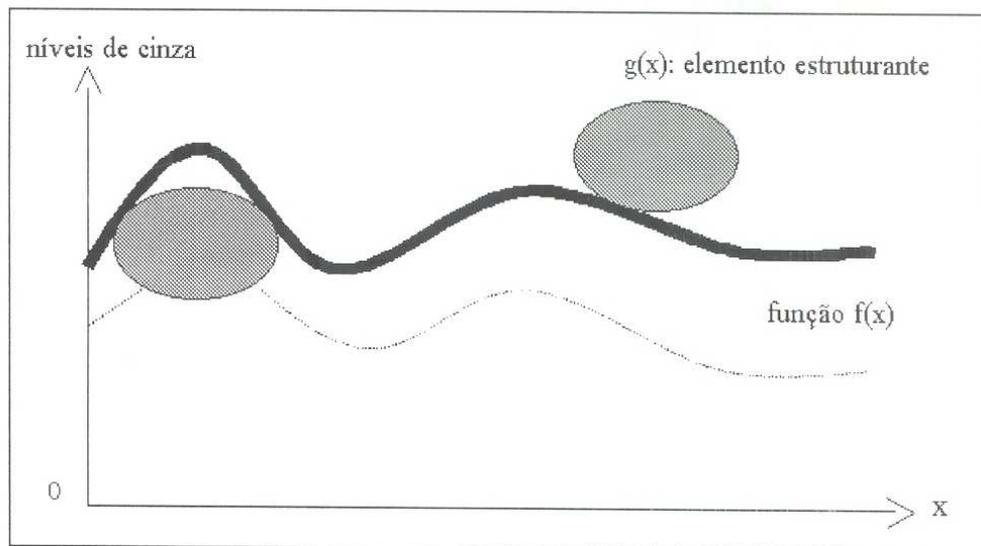


Figura 4.3 Exemplo de interação entre o elemento estruturante e a imagem em níveis de cinza

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 é 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, pelos autores, de maneiras diferentes [BANON & BARRERA, 1994]. Este trabalho adotou o formalismo de Minkowski [SERRA, 1982] e as definições foram extraídas em [FACON, 1996][SERRA, 1982].

4.2 Operadores Morfológicos Binários

4.2.1 Erosão Binária

A erosão de um conjunto X pelo elemento estruturante B é a seguinte:

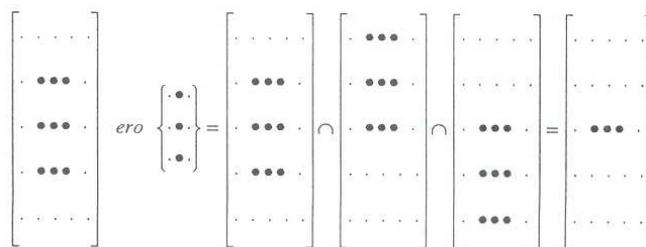
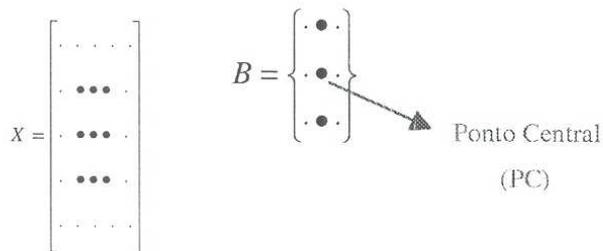
$$\varepsilon^B(X) = X \ominus \tilde{B} = \bigcap_{b \in \tilde{B}} X_b$$

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.

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} . Vejamos o exemplo 4.1.

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

Exemplo 4.1:



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 (Figura 4.4a):

- Diminuição das partículas (pixels pretos);
- Eliminação de grãos (pixels pretos) de tamanho inferior ao tamanho do elemento estruturante;
- Aumento dos buracos (pixels brancos);
- Separação de grãos próximos.

4.2.2 Dilatação Binária

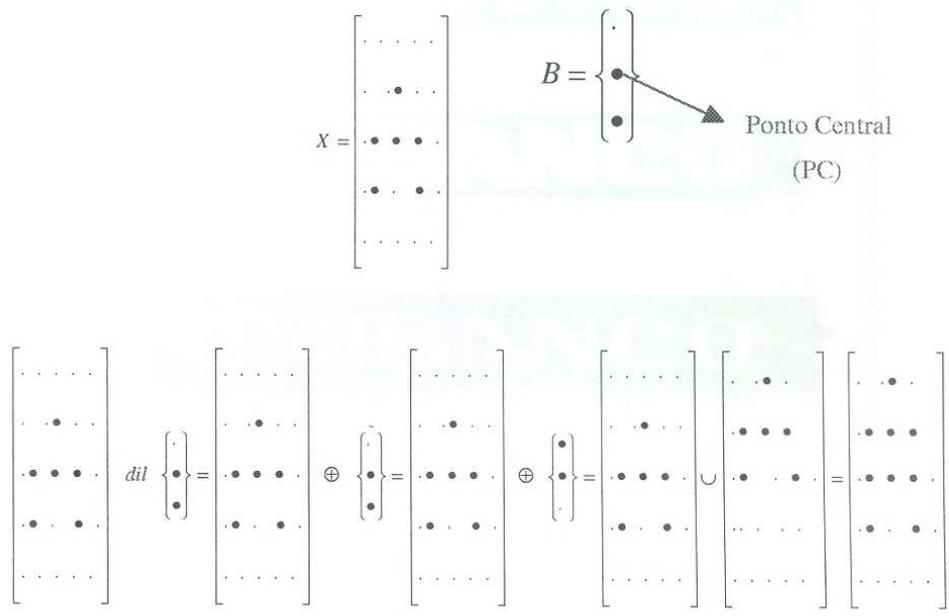
A dilatação de um conjunto X pelo elemento estruturante B é a seguinte:

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

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

Conforme a definição acima, o conjunto a ser dilatado, X , é deslocado em função das posições permitidas pelo elemento estruturante B . Os deslocamentos são realizados em relação ao ponto central de \tilde{B} . Vejamos o exemplo 4.2.

Exemplo 4.2:



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 (Figura 4.4c):

- Expandir partículas;
- Preencher pequenos buracos;
- Conectar grãos próximos.



Figura 4.4 Resultado da erosão e da dilatação com elemento estruturante quadrado e 2 iterações. (a) imagem original; (b) imagem erodida; (c) imagem dilatada.

4.2.3 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, podemos prever que esta operação consiste em erodir e, depois, dilatar o resultado da erosão.

A abertura de um conjunto X pelo elemento estruturante B é:

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

O exemplo 4.3 apresenta uma imagem X com dois conjuntos conectados. A operação de abertura é aplicada para separar estes conjuntos.

O segundo passo consiste em processar o resultado anterior por adição de Minhowski com B. Essa adição resulta da união de X com X, deslocado segundo as posições permitidas por B.

$$\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \oplus \begin{Bmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{Bmatrix} =$$

$$\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}$$

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

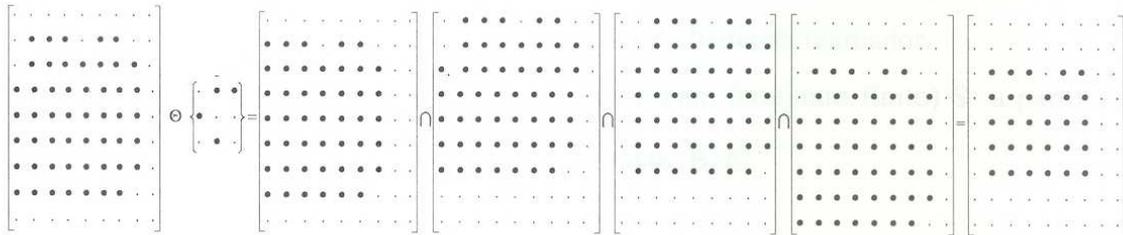
Os efeitos obtidos pela abertura binária são (Figura 4.5b):

- 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;
- O conjunto aberto é mais regular que o conjunto inicial;
- O conjunto aberto é menos rico em detalhes que o conjunto inicial.

4.2.4 Fechamento

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

O segundo passo consiste em processar, por adição de Minowski com B, o resultado anterior. Essa subtração funda-se na interseção de X com X, deslocado segundo as três posições permitidas por B.



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 (Figura 4.5c):

- 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;
- O conjunto fechado é mais regular que o conjunto inicial;
- Conjunto fechado é menos rico em detalhes que o conjunto inicial.



Figura 4.5 Resultado da abertura e do fechamento com elemento estruturante vertical e 4 iterações. (a) imagem original; (b) abertura; (c) fechamento.

4.2.5 Reconstrução Binária

O processo de reconstrução consiste em recuperar um conjunto S a partir de um conjunto Z , onde S 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 Z chama-se marcador.

A reconstrução binária $\rho_S(z)$ de uma máscara (conjunto finito) S a partir do marcador Z ($Z \subseteq S$), usando o elemento estruturante B , é:

$$\rho_S(z) = \lim_{n \rightarrow \infty} \underbrace{\delta_{cS}^B(\dots \delta_{cS}^B(Z))}_n$$

onde $\delta_{cS}^B(Z) = \delta^B(Z) \cap S$ representa a dilatação condicional de Z relativamente a S .

A reconstrução permite eliminar os padrões indesejáveis sem deformar os desejáveis. A Figura 4.6 ilustra o processo de reconstrução binária.



Figura 4.6 Exemplo de reconstrução binária com elemento estruturante quadrado

Nota-se, na Figura 4.6, que somente os elementos que faziam parte do marcador foram recuperados.

4.2.6 Granulometria Binária

A granulometria pode ser comparada com um processo de peneiramento. Peneirar consiste em separar substâncias reduzidas ao estado de fragmento, que apresentam vários tamanhos. Efetuar um bom peneiramento depende do tamanho da malha da peneira usada. Após peneirar, são obtidos dois conteúdos: o primeiro com os objetos menores que a malha da peneira, o segundo com o que sobrou do peneiramento, de tamanho maior que a malha.

Matheron [MATHERON, 1975] propôs um conjunto de regras que, se forem verificadas, permitem obter uma boa granulometria. Para isto, existem três axiomas:

Sejam X a imagem a ser analisada e $T^{(\lambda)}(X)$ a transformação que permite realizar uma análise granulométrica, $T^{(\lambda)}(X)$ representa exatamente o refugo da peneira de tamanho λ .

1. a transformação morfológica deve ser anti-extensiva; isso significa que o conjunto transformado deve ser menor que o de origem ou seja:

$$\forall \lambda > 0, \quad T^{(\lambda)}(X) \subset X \quad \forall X$$

2. a transformação morfológica deve ser crescente, ou seja:

$$\forall \lambda > 0, \quad Y \subset X \Rightarrow T^{(\lambda)}(Y) \subset T^{(\lambda)}(X) \quad \forall X$$

3. consideremos a transformação de uma imagem X a partir de duas transformações morfológicas sucessivas de parâmetros respectivos λ_1 e λ_2 . O resultado final deve ser idêntico, qualquer que seja a seqüência de transformações empregadas. Além disso, o resultado deve ser idêntico ao obtido pela transformação de maior parâmetro λ :

$$\begin{aligned} \forall \lambda_1, \lambda_2 > 0, \quad T^{(\lambda_1)}(T^{(\lambda_2)}(X)) &= T^{(\lambda_2)}(T^{(\lambda_1)}(X)) \\ &= T^{\sup(\lambda_1, \lambda_2)}(X) \quad \forall X \end{aligned}$$

Suponha agora que as transformações consideradas são representadas por imagens binárias ou conjuntos discretos. Neste contexto, a granulometria é uma seqüência de aberturas ϕ_n , indexadas sobre um inteiro $n \geq 0$. Cada abertura (correspondendo a um tamanho de malha) remove mais do que a anterior, até alcançar o conjunto vazio:

$$\forall X, \forall n \geq m \geq 0, \phi^{nB}(X) \subseteq \phi^{mB}(X)$$

A granulometria por abertura pode gerar deformações nos conjuntos, o que resulta em função e densidade de distribuições incorretas dificultando, portanto, a interpretação. Isto ocorre porque os conjuntos podem ter diversas formas, ou seja, a imagem contém vários padrões e/ou ruídos, dificultando o processo de peneiramento. Peneirar, por exemplo, estruturas redondas e alongadas a partir de uma única seqüência de elementos estruturantes pode resultar em resultados duvidosos. Conjuntos sobrepostos podem gerar o mesmo efeito. Várias soluções são possíveis. Uma delas é modificar o processo de peneiramento por abertura, introduzindo o processo de reconstrução binária a partir de marcadores binários. Sabemos que, após a abertura, o conjunto que “sobrevive” existe realmente na peneira. Somente a sua forma e as suas dimensões foram alteradas. A idéia consiste, portanto, em extrair o marcador de cada conjunto restante (por exemplo, pelo processo de última erosão) para depois reconstruir o verdadeiro conjunto que normalmente deveria aparecer.

A Figura 4.7 apresenta o resultado de processo de granulometria por abertura com reconstrução binária e elemento estruturante vertical. Os elementos da imagem distribuíram-se em cinco classes.

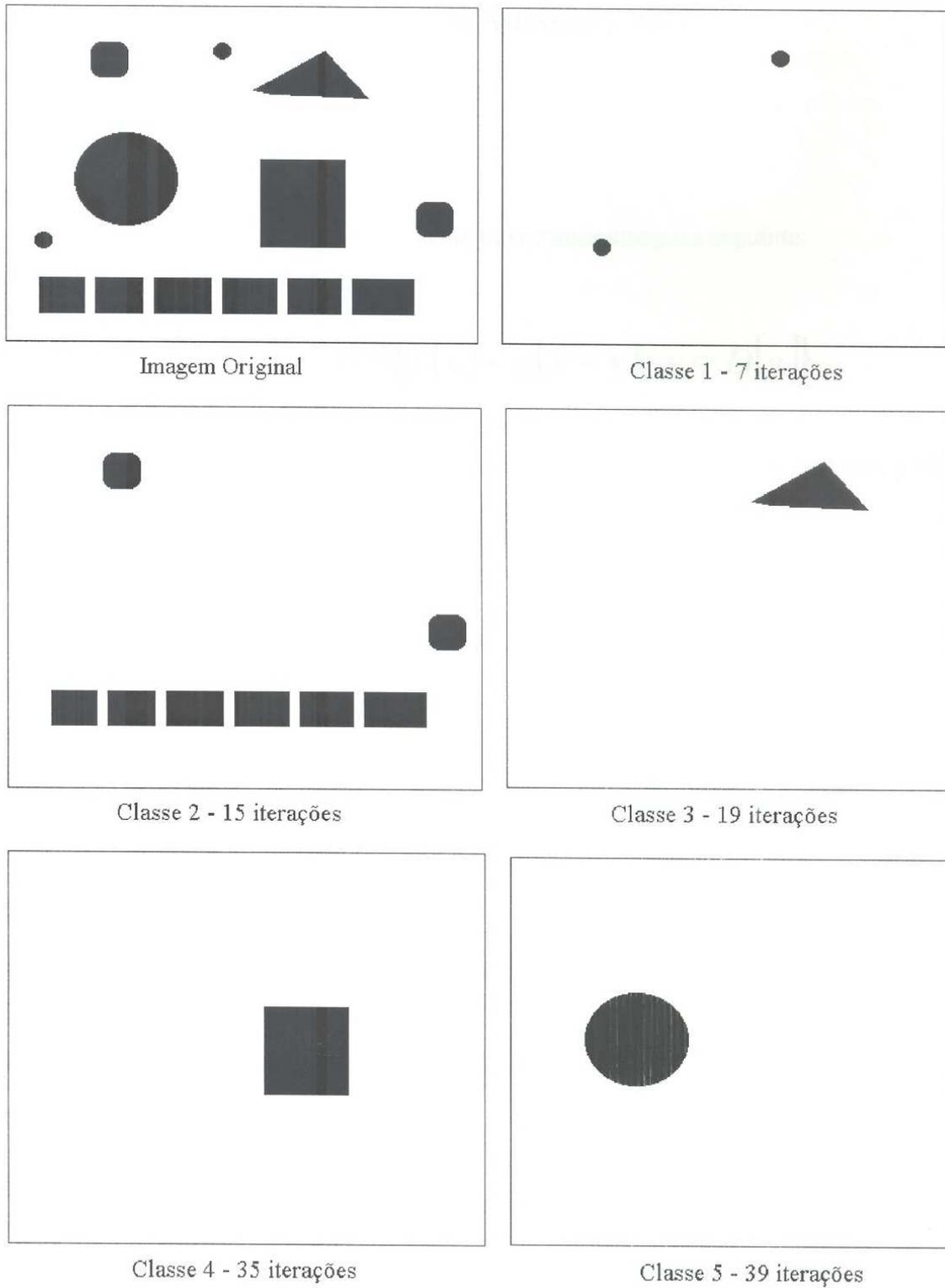


Figura 4.7 Granulometria por abertura com reconstrução.

4.3 Operadores Morfológicos em Níveis de Cinza

4.3.1 Erosão em Níveis de Cinza

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

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

Min equivale a mínimo, x é o ponto a ser processado na imagem original, y são os pontos envolvidos pelo elemento estruturante e $D[g]$ é o domínio do elemento estruturante.

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

Os efeitos da erosão em níveis de cinza são (Figura 4.8b):

- Escurecer a imagem;
- Alargar e expandir os vales (padrões escuros);
- Conectar vales próximos;
- Reduzir e às vezes reduzir picos (padrões claros);
- Separar picos próximos.

4.3.2 Dilatação em Níveis de Cinza

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

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

Max equivale a máximo, x é o ponto a ser processado na imagem original, y são os pontos envolvidos pelo elemento estruturante e $D[g]$ é o domínio do elemento estruturante.

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

Os efeitos da dilatação em níveis de cinza são (Figura 4.8c):

- Clarear a imagem;
- Alargar e expandir os picos (padrões claros);
- Conectar picos próximos;
- Reduzir e às vezes eliminar vales (padrões escuros);
- Separar vales.

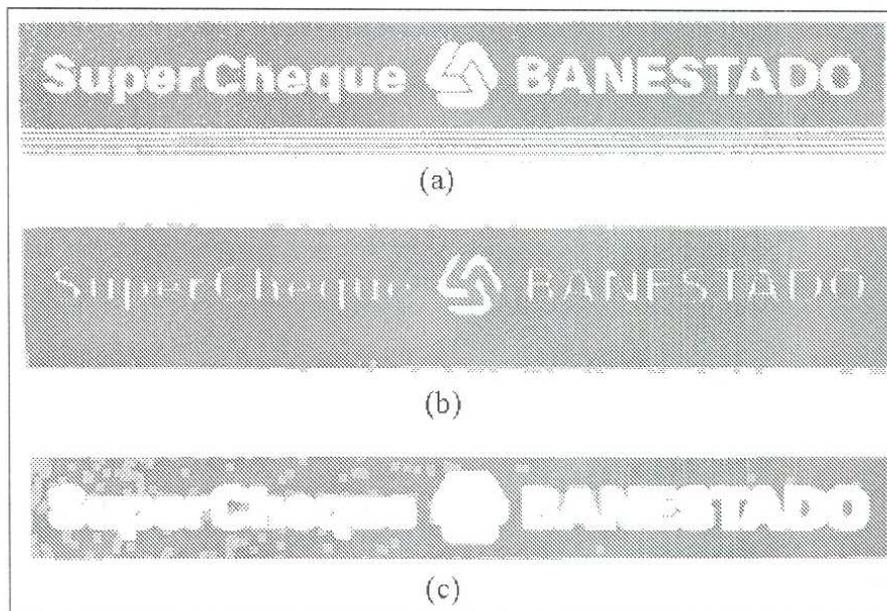


Figura 4.8 Resultado da erosão e da dilatação em níveis de cinza com elemento estruturante quadrado planar e 2 iterações. (a) imagem original; (b) imagem erodida; (c) imagem dilatada.

4.3.3 Erosão Geodésica

A erosão geodésica de um sinal marcador f em relação à máscara g é a seguinte:

$$\varepsilon_g^{(1B)}(f) = \text{Max}(\varepsilon^{(1B)}(f), g)$$

onde Max é o valor máximo entre $\varepsilon^{(1B)}(f)$ e g pixel a pixel.

4.3.4 Reconstrução em Níveis de Cinza Dual ou por Erosão

A reconstrução em níveis de cinza é a seqüência lógica da reconstrução binária. A meta é similar, tentar reconstituir uma imagem que sofreu modificações no decorrer do tratamento. Porém, a grande diferença da reconstrução em níveis de cinza em relação à binária é que a imagem não será “perfeitamente” reconstruída.

Há vários tipos de reconstrução em níveis de cinza. Este trabalho apresenta a definição da reconstrução dual em níveis de cinza, porque o operador *Fillhole*, aplicado na metodologia proposta, utiliza esse tipo de reconstrução.

Sejam duas imagens S e Z em níveis de cinza definidas no mesmo domínio, tomando os valores discretos $k \in [0, n]$ e respeitando a relação de ordem $Z \geq S$. Por dualidade pode-se definir a reconstrução dual em níveis de cinza (ou reconstrução por erosão).

$$\rho_S(Z) = \text{Min}_{n \geq 1} [\varepsilon_S^{nB}(Z)] = \text{Min}_{n \geq 1} \underbrace{\varepsilon_S^B(\dots \varepsilon_S^B(Z))}_n$$

onde B é o elemento estruturante e Min é mínimo.

A Figura 4.9 ilustra o processo de reconstrução dual em níveis de cinza.

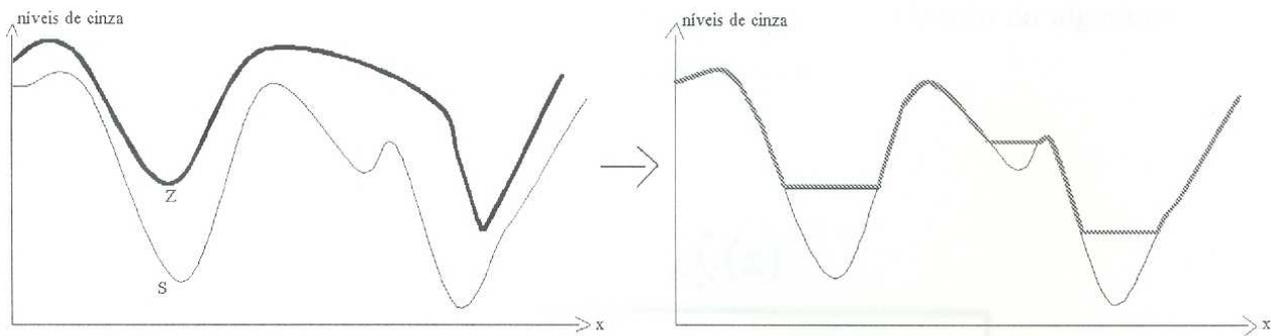


Figura 4.9 Reconstrução Dual em níveis de cinza da máscara S com o marcador Z e elemento estruturante quadrado planar.

4.3.5 *Fillhole*

Os furos de uma imagem podem ser interpretados como vales ou mínimos locais não conectados com a borda da imagem.

Estes furos podem ser removidos ou preenchidos usando uma técnica de imposição de mínimo. A imagem marcadora Z usada na reconstrução morfológica dual ou por erosão é construída da seguinte maneira: a imagem Z tem as mesmas dimensões que a imagem original S . Z é preenchido pelo maior valor em níveis de cinza encontrado na imagem original, exceto ao longo de sua borda onde os valores de níveis de cinza da imagem original são mantidos:

$$FILL(S) = \rho_s^*(Z)$$

onde

$$Z = \begin{cases} S(x) & \text{se } x \text{ encontra-se na borda de } S, \\ t_{\max} & \text{caso contrário.} \end{cases}$$

t_{\max} é o valor máximo de nível de cinza da imagem.

Esta reconstrução morfológica corresponde a uma generalização do algoritmo usado para preencher furos de objetos binários (Figura 4.10).

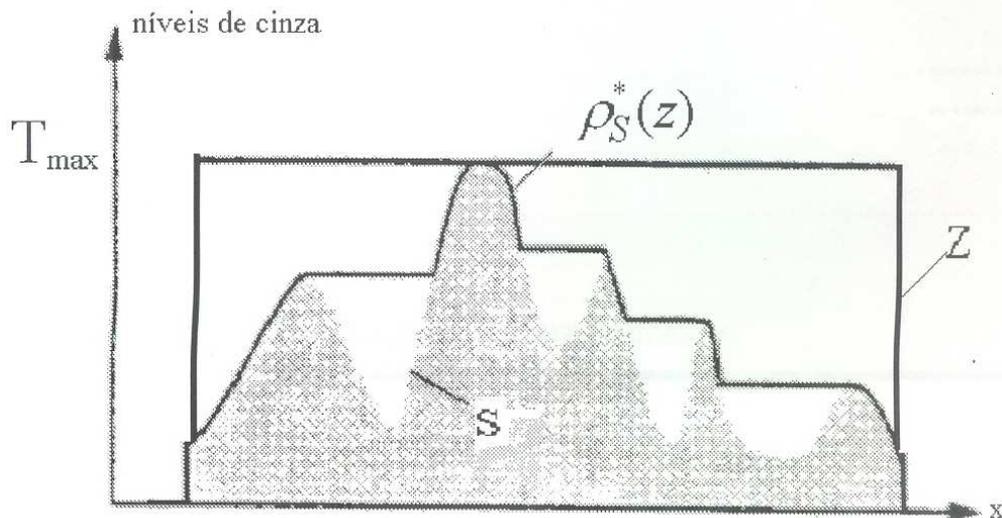


Figura 4.10 Remoção de todos os mínimos internos da imagem S pela reconstrução morfológica da função marcador Z e elemento estruturante quadrado.

Os objetos extraídos pela binarização de uma imagem frequentemente apresentam alguns furos devido ao ruído. O procedimento *fillhole* é adequado para filtrar esses objetos. Em alguns casos, é recomendado preencher os furos diretamente em imagens em níveis de cinza.

No caso das imagens em níveis de cinza de cheques bancários, o processo *Fillhole* pode ser aplicado para separar o fundo artístico das informações relevantes. Este processo gera como resultado uma imagem contento o fundo artístico do cheque (Figura 4.11b). Aplicando a subtração, entre a imagem original (Figura 4.11a) e a imagem resultante do processo *Fillhole* (Figura 4.11b), obtém-se as informações relevantes do cheque (Figura 4.11c). Esse procedimento é uma nova proposta para a remoção do fundo artístico de cheques bancários.

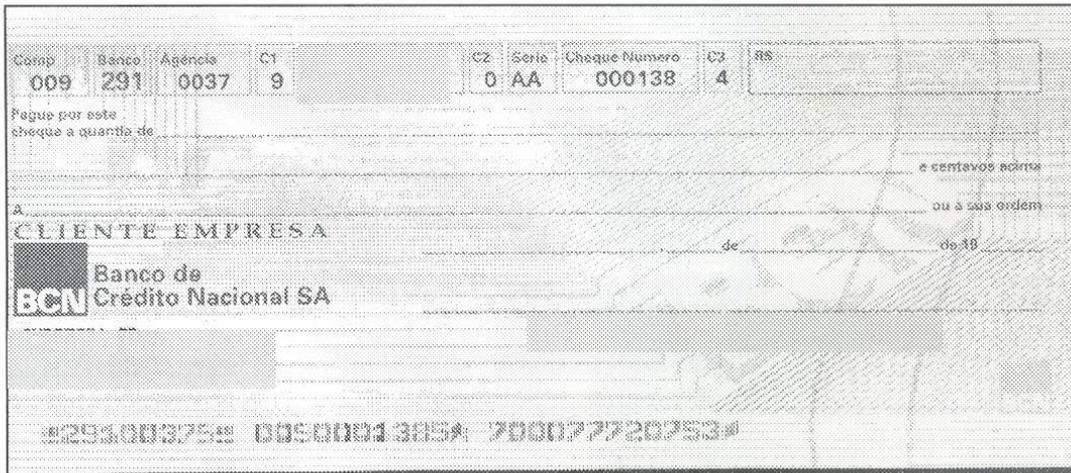


Figura 4.11 - a

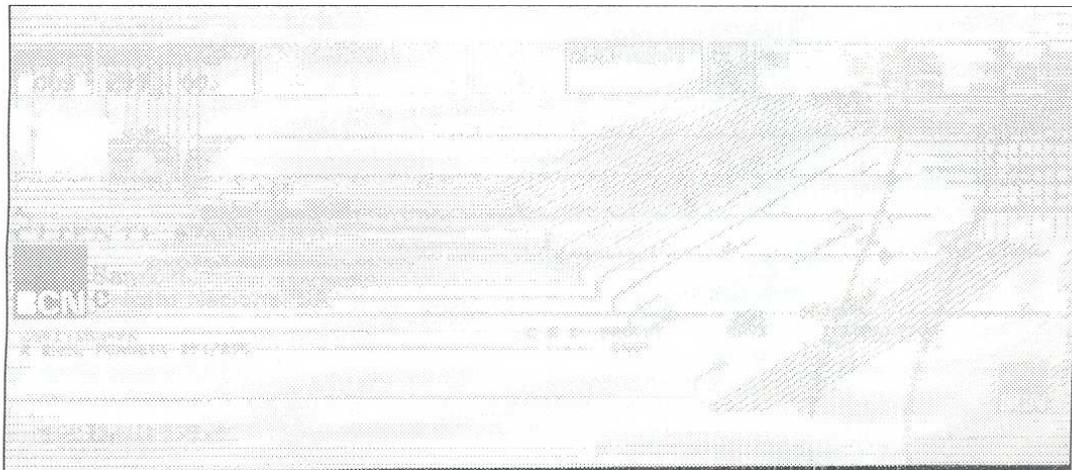


Figura 4.11 - b



Figura 4.11 - c

Figura 4.11 (a) Imagem em níveis de cinza; (b) Resultado do processo Fillhole; (c) Resultado da subtração entre a imagem (a) e a imagem (b).

O próximo capítulo descreverá a metodologia proposta, que aplica operadores morfológicos, para segmentar automaticamente logotipos em imagens de cheques bancários brasileiros.

Capítulo 5

Metodologia Aplicada

5.1 Introdução

Na segmentação, geralmente são usadas muitas heurísticas para o tamanho e posição dos componentes a serem extraídos. Os logotipos (que estão tipicamente confinados em uma região compacta e isolada do documento, e não estão ligados na estrutura do texto como outros gráficos), freqüentemente, são separados de outros componentes gráficos por sua posição na página e usando o conhecimento prévio do tipo do documento (memorando, carta, etc..) [DOERMANN et al., 1993]. O uso de muitas heurísticas acaba limitando a segmentação para um determinado tipo de aplicação.

A metodologia aplicada neste estudo emprega a morfologia matemática para localizar logotipos em imagens de cheques bancários brasileiros e extraí-los automaticamente. O objetivo é minimizar o número de parâmetros heurísticos, visando a obter uma segmentação mais exata possível para todas as situações, permitindo, assim, o reuso dessa abordagem para outras aplicações. É importante lembrar que segmentar sem conhecimento prévio, empregando poucas heurísticas e poucas hipóteses, é um grande desafio.

A proposta do logotipo é fornecer uma ligação simbólica com uma organização. Ele é uma marca saliente perceptível, maior do que componentes textuais individuais [DOERMANN et al., 1993] e encontra-se separado do resto do texto. Na base de dados de cheques bancários brasileiros, verificou-se nas 478 imagens, que o logotipo realmente é um componente saliente. Portanto, apesar da complexidade das imagens dos

cheques, aplicou-se um único conhecimento prévio: **o logotipo é maior do que todas as informações presentes no cheque bancário**. O conhecimento prévio “**o logotipo encontra-se separado do resto do texto**” não foi aplicado porque isto também se verifica para outros elementos (por exemplo, tabelas e figuras).

As duas fases principais do processo automático de segmentação de logotipos em cheques bancários brasileiros são: pré-processamento e extração/localização do logotipo. A Figura 5.1 apresenta a estrutura do sistema empregada neste trabalho.

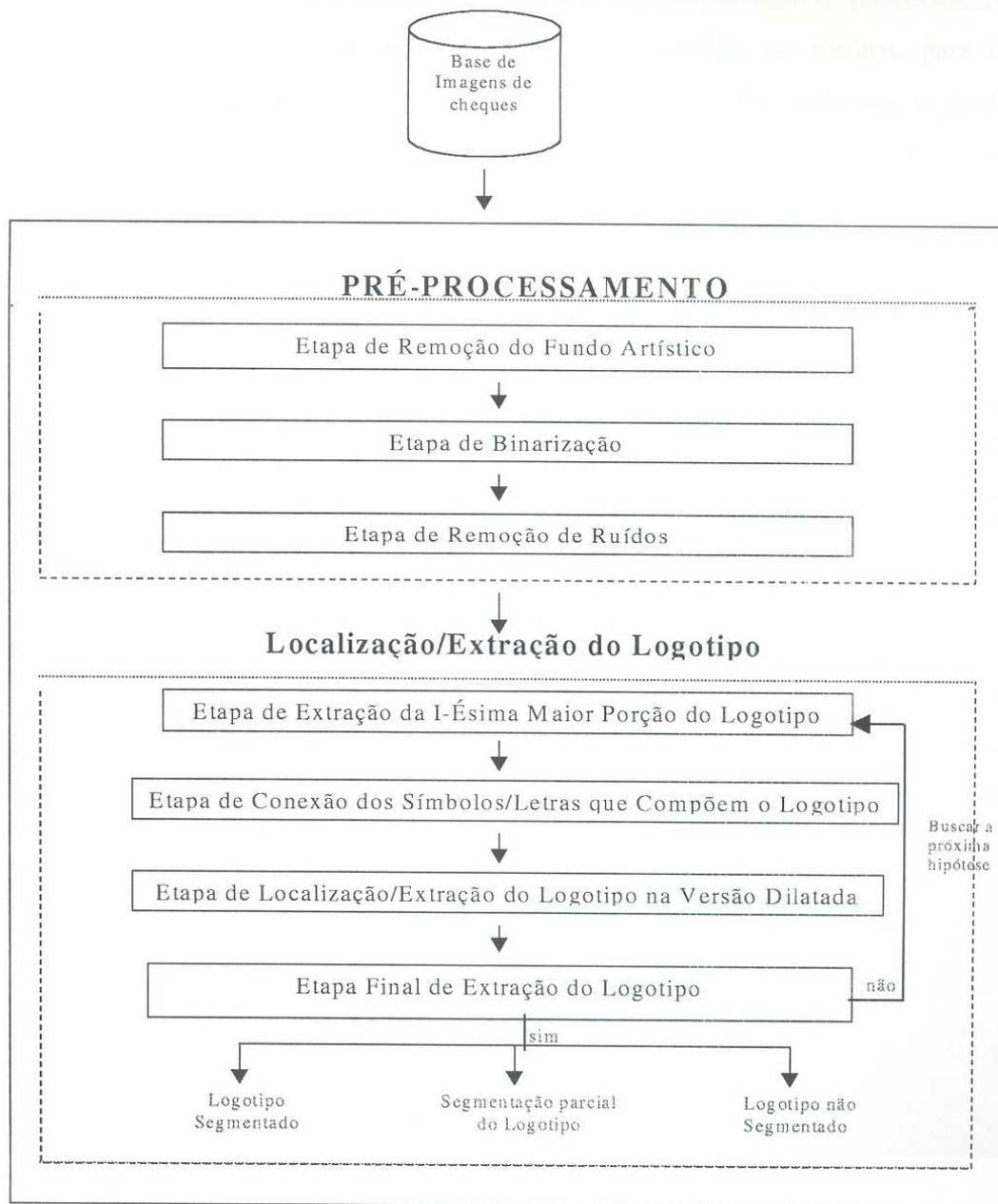


Figura 5.1 Estrutura do sistema de segmentação automática de logotipos em imagens de cheques bancários brasileiros.

Para o processamento das etapas do sistema foram utilizadas duas ferramentas o Khoros e o Bropi. O Khoros é um sistema para processamento digital de imagens, desenvolvido por um grupo de universidades, que consiste em um conjunto de programas, cada um realizando uma determinada operação, ou processamento. Além desses programas, que formam uma biblioteca de processamento digital de imagens, o sistema conta também com um ambiente para a manipulação dessas rotinas, através da construção de diagramas que indicam a seqüência e interação entre os processos. Novas rotinas de processamento de imagens podem ser incluídas no Khoros, para isto é necessário criar um novo programa, que implementa o novo processo, o qual será incluído no sistema. O Khoros contém aproximadamente 900.000 linhas de código, e possui um pouco mais que 300 programas e bibliotecas diferentes que estão organizados em *toolboxes*. Cada *toolbox* contém capacidades diferentes; as *toolboxes* podem ser misturadas e combinadas de acordo com as necessidades do usuário (Figura 5.2). O Khoros emprega o sistema operacional UNIX [KHOROS].

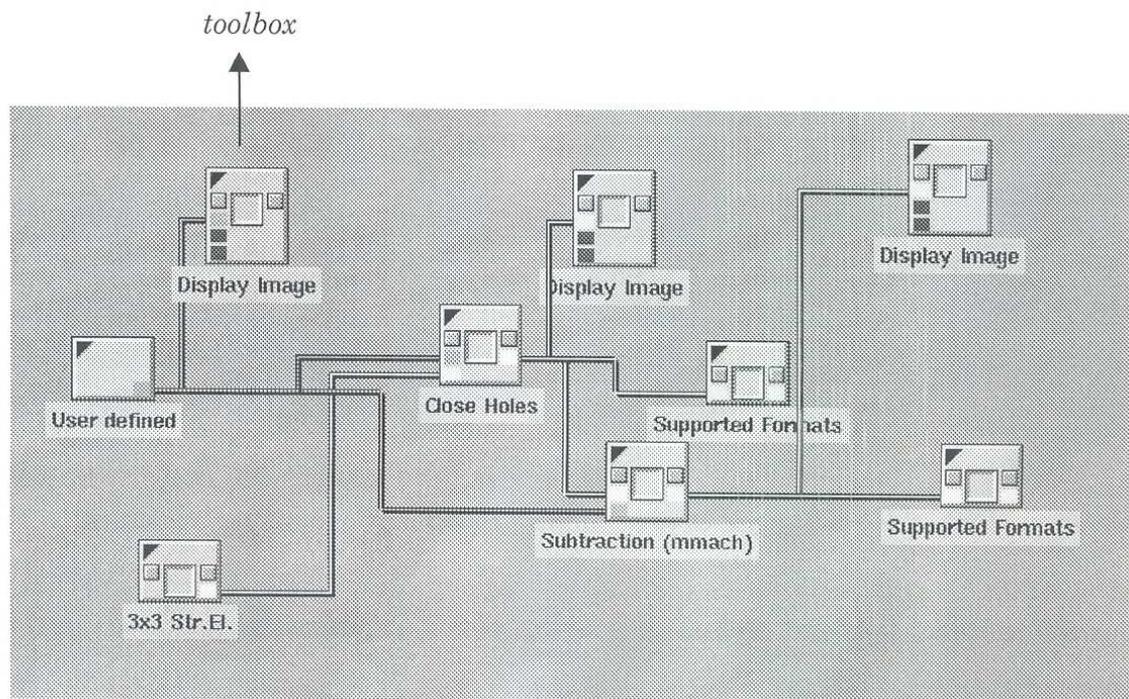


Figura 5.2 Área de trabalho do Khoros

O Bropi é uma biblioteca de rotinas de processamento de imagens. Esta biblioteca foi desenvolvida em ambiente Windows 95/NT, utilizando-se da biblioteca de classes MFC 4.2 (*Microsoft Foundation Classes*). O objetivo principal da estrutura da biblioteca é disponibilizar uma hierarquia de classes que defina os principais elementos envolvidos no processamento digital de imagens e, com isto, facilitar o desenvolvimento de novos processos a serem inseridos nessa biblioteca [MATRAKAS & BORTOLOZZI, 1998]. Esta ferramenta foi desenvolvida no laboratório LARDOC (Laboratório de Análise e Reconhecimento de Documentos) da Pontifícia Universidade Católica do Paraná. A Figura 5.3 apresenta a área de trabalho do Bropi.

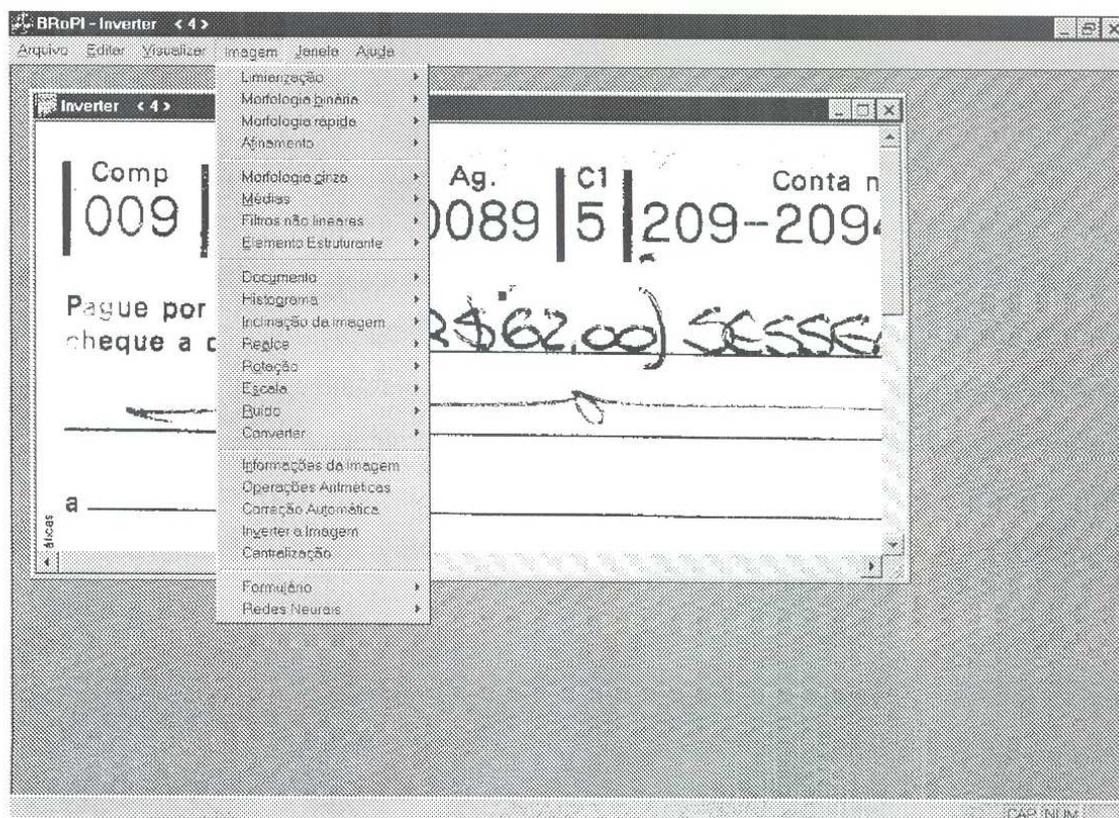


Figura 5.3 Área de trabalho do Bropi

A etapa de Remoção do Fundo Artístico, da fase de Pré-Processamento, utilizou as ferramentas existentes no Khoros. A seqüência de *toolboxes* aplicada está definida na Figura 5.2. O tempo de processamento, nessa etapa, em média é de 15 minutos. Ela foi processada em um Pentium II – 350 com 128 Mbytes de memória RAM e imagens 2053 X 898 X 16M. As demais etapas utilizaram as ferramentas existentes no Bropi, exceto a

etapa de Extração da I-Ésima Maior Porção do Logotipo, da fase de Localização/Extração do Logotipo, que foi implementada e inserida no Bropi. Estas etapas foram processadas em um computador Pentium 233 com 64 Mbytes de memória RAM. O tempo de processamento, em média, dessas etapas é de 45 segundos ao total.

Nas próximas seções, serão descritas e justificadas as etapas da metodologia de segmentação de logotipos em cheques bancários brasileiros.

5.2 Fase de Pré-Processamento

5.2.1 Etapa de Remoção do Fundo Artístico

As imagens dos cheques bancários brasileiros são muito complexas por causa dos layouts de fundo, os quais garantem a autenticidade dos mesmos. A tarefa de extrair informações do cheque é complicada, porque as informações relevantes se misturam com o layout de fundo. Portanto, faz-se necessária a remoção prévia do fundo artístico.

Por apresentar somente duas informações (preto e branco) as imagens binárias não são adequadas para realizar a remoção do fundo artístico. Neste caso, os elementos do cheque e o fundo artístico são identificados pela mesma informação (pixel preto). A Figura 5.4 exemplifica este fato.



Figura 5.4 - a

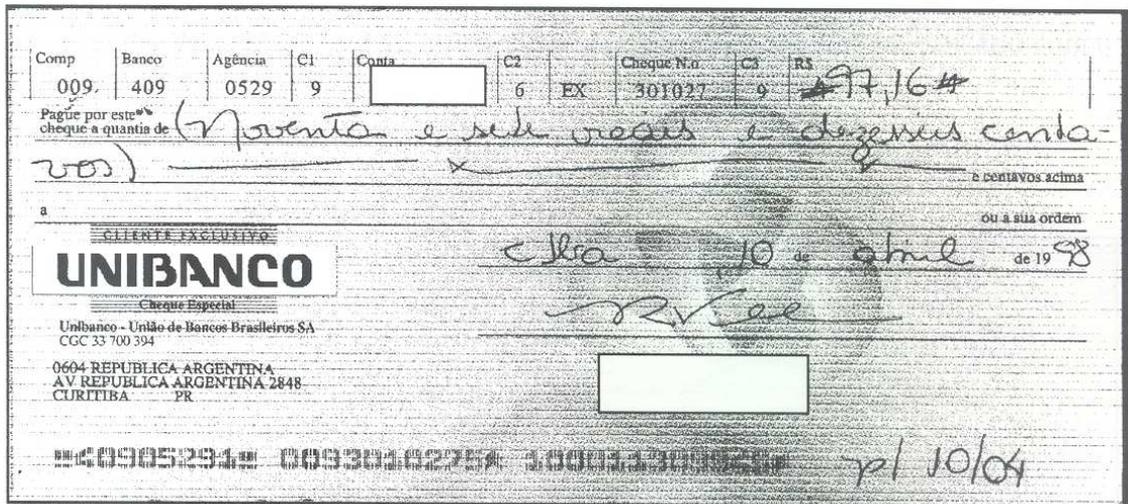


Figura 5.4 - b

Figura 5.4 Imagens Binárias com Fundo Artístico

Verifica-se, nas imagens de cheques da Figura 5.4, que as informações relevantes estão misturadas com o layout de fundo. Isto torna inviável a remoção do fundo artístico em imagens binárias. Nas imagens em níveis de cinza, as informações são representadas por 256 níveis de cinza. Os valores de cinza do fundo artístico diferem dos valores de cinza das informações relevantes do cheque. Isto torna viável a remoção do fundo artístico nessas imagens.

As ferramentas de remoção direta do fundo (remove o fundo artístico e mantém as informações relevantes) foram rejeitadas porque apresentaram resultados insatisfatórios, tais como:

- Remoção insuficiente do fundo artístico;
- Após o processo de binarização verificou-se aglomeração de ruídos;
- Degradação do logotipo.

Optou-se, portanto, pela abordagem inversa de conservar o fundo antes de removê-lo. Concluiu-se que, por ser automática e não heurística, a melhor ferramenta para a abordagem escolhida é o operador morfológico *Fillhole*. Esse operador permite conservar a maioria dos componentes do fundo. A remoção é efetuada por subtração

entre a imagem original em níveis de cinza do cheque bancário e a imagem resultante do operador morfológico *Fillhole*. A Figura 5.5c apresenta um cheque bancário com o fundo artístico removido.



Figura 5.5 - a



Figura 5.5 - b

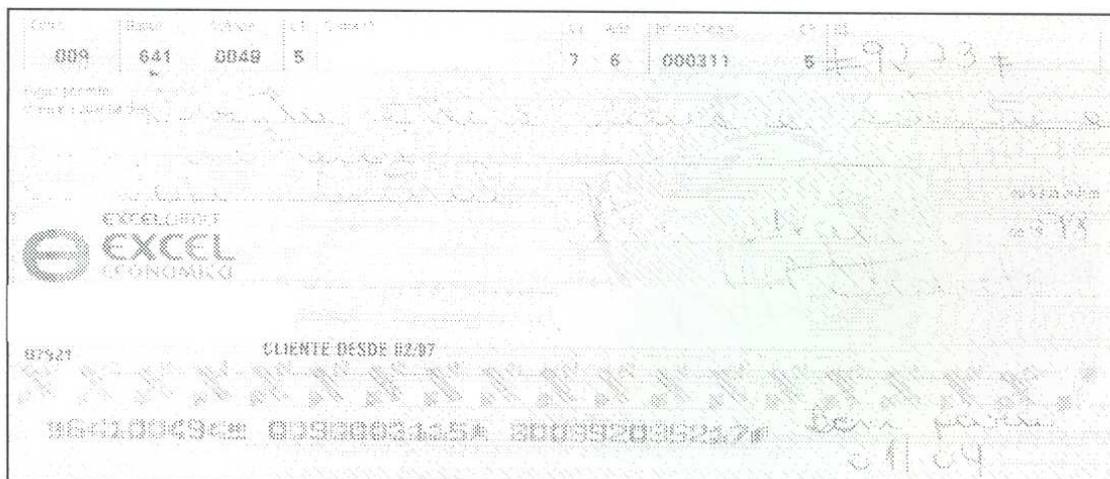


Imagem 5.5 – c

Figura 5.5 (a) Imagem em níveis de cinza do cheque bancário; (b) Resultado do processo Fillhole; (c) Imagem em níveis de cinza do cheque bancário com o fundo artístico removido

No entanto, um problema foi encontrado nesta etapa. Alguns cheques preenchidos possuem traços manuscritos que tocam a borda da imagem e sobrepõem o logotipo (Figura 5.6). Como o processo *Fillhole* preenche todos os elementos que tocam a borda, por subtração o logotipo desaparece (Figura 5.7).



Figura 5.6 Imagem de cheque com traços manuscritos que tocam na borda e sobrepõem o logotipo

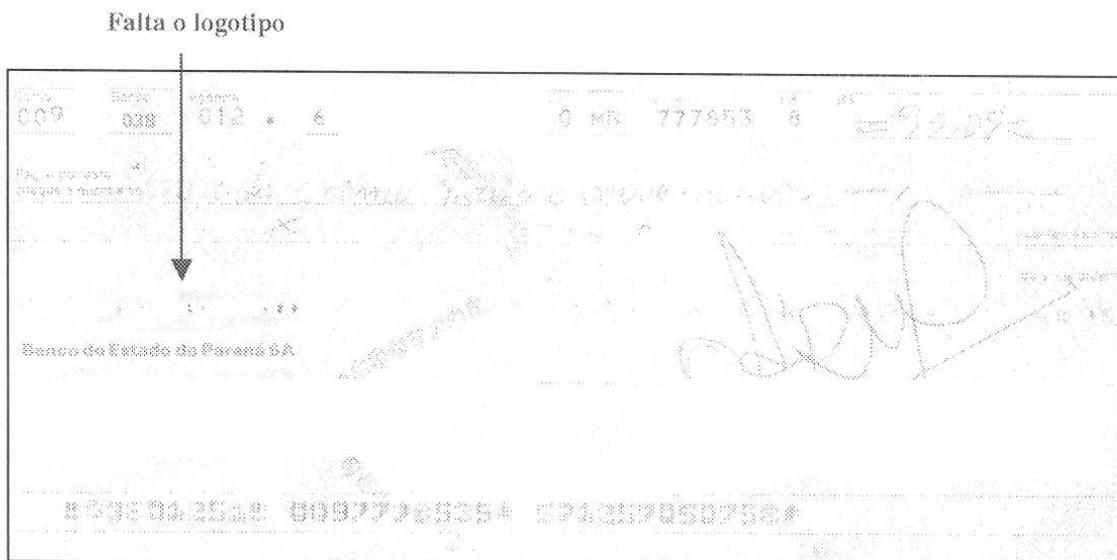


Figura 5.7 Imagem resultante da subtração entre o resultado do Fillhole e a imagem 5.4

Para evitar que o processo *Fillhole* construa logotipos com traços sobrepostos tocando a borda da imagem, esses foram separados dessa borda pela adição de uma

5.2.2 Etapa de Binarização

Como já foi dito no Capítulo 2, binarização é o método que transforma imagens em tons de cinza em imagens binárias (preto/branco). Mas, nem sempre imagens binarizadas são de boa qualidade. A avaliação desta qualidade é um desafio. Quando é feita visualmente, a imagem binária de um cheque bancário é considerada “boa” se for possível ler corretamente os componentes segmentados, sem dificuldades [MATTANA, 1999].

Mattana [MATTANA, 1999] estudou a eficiência de vários algoritmos de binarização sobre imagens de cheques bancários brasileiros, e o algoritmo global que apresentou os melhores resultados em tempo reduzido para este tipo de aplicação foi o Otsu. Este algoritmo foi aplicado na binarização das imagens em níveis de cinza dos cheques bancários brasileiros (Figura 5.10).

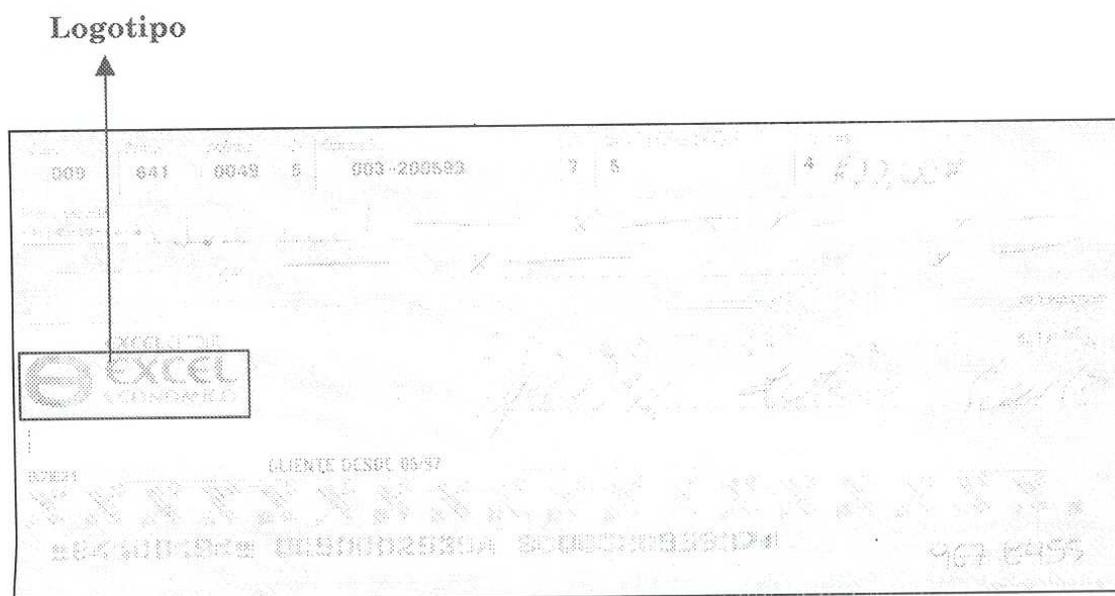


Figura 5.10 - a

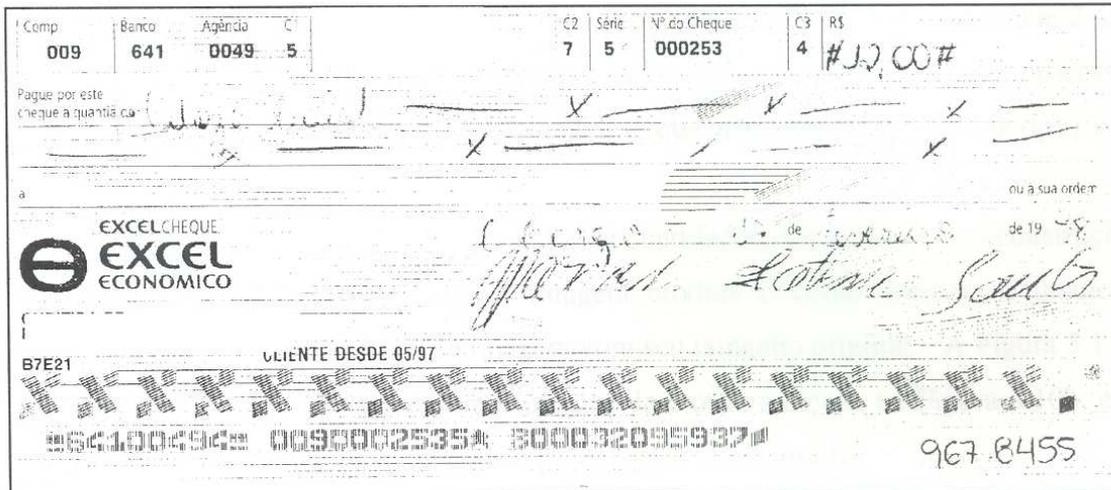


Figura 5.10 - b

Figura 5.10 (a) Imagem em níveis de cinza; (b) Imagem binarizada pelo algoritmo de Otsu

5.2.3 Etapa de Remoção de Ruídos

Para eliminar os ruídos provenientes dos resíduos do fundo do cheque e do algoritmo de Otsu, e eliminar as informações irrelevantes, é aplicada a operação morfológica de erosão. Só que, nesta técnica, é preciso escolher o elemento estruturante a ser aplicado (quadrado, horizontal, vertical, rhombus, cruz) e definir o número de iterações a serem processadas. Para efetuar essa operação de forma objetiva e não heurística, empregou-se a abordagem granulométrica binária.

As imagens dos cheques foram erodidas variando o tipo de elemento estruturante e o número de iterações. A granulometria foi aplicada nas imagens erodidas com objetivo de analisar como as informações relevantes, nesse caso o logotipo, eram distribuídas nas classes. Os resultados foram analisados, de forma visual, e concluiu-se que, por eliminar as informações indesejáveis sem prejudicar as informações relevantes, 2 iterações de erosão com o elemento estruturante cruz é a solução mais adequada. Verificou-se que elementos estruturantes menores (vertical ou horizontal) que o cruz não eliminam as informações indesejáveis; e elementos estruturantes maiores (quadrado ou rhombus) que o cruz eliminam as informações indesejáveis, mas

prejudicam as informações relevantes. Em relação ao número de iterações, quando menor que 2, as informações indesejáveis não são eliminadas; quando maior que 2, os componentes dos logotipos ficam desgastados e são classificados, pelo processo granulométrico, junto com outros componentes (que não pertencem ao logotipo) em uma mesma classe.

Como a erosão modifica o tamanho das entidades, foi aplicada a reconstrução binária (utilizando como marcador a imagem erodida e como máscara a imagem original) para obter as entidades da imagem com seu tamanho original. A Figura 5.11a apresenta o resultado da operação morfológica de erosão aplicada na Figura 5.10b, e a Figura 5.11b apresenta a imagem após o processo de reconstrução.

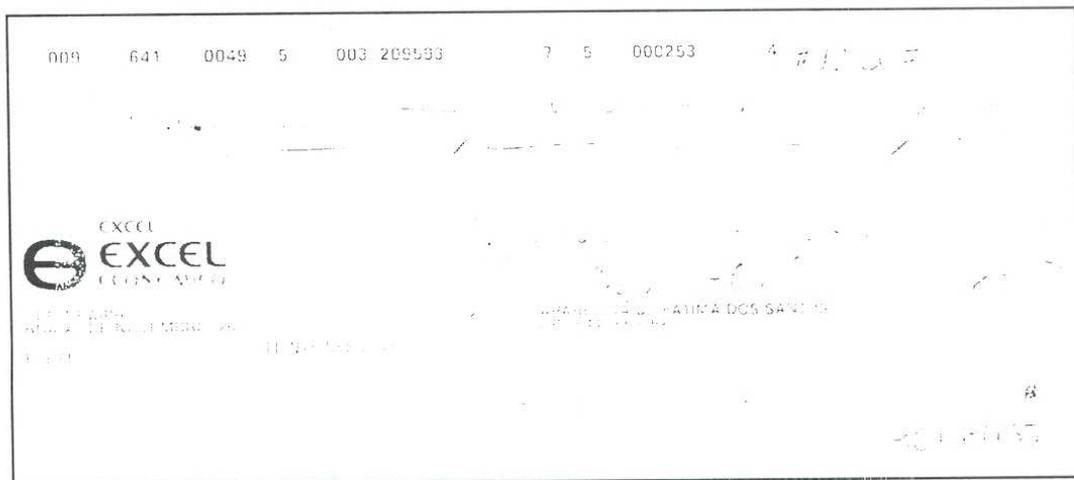


Figura 5.11 - a

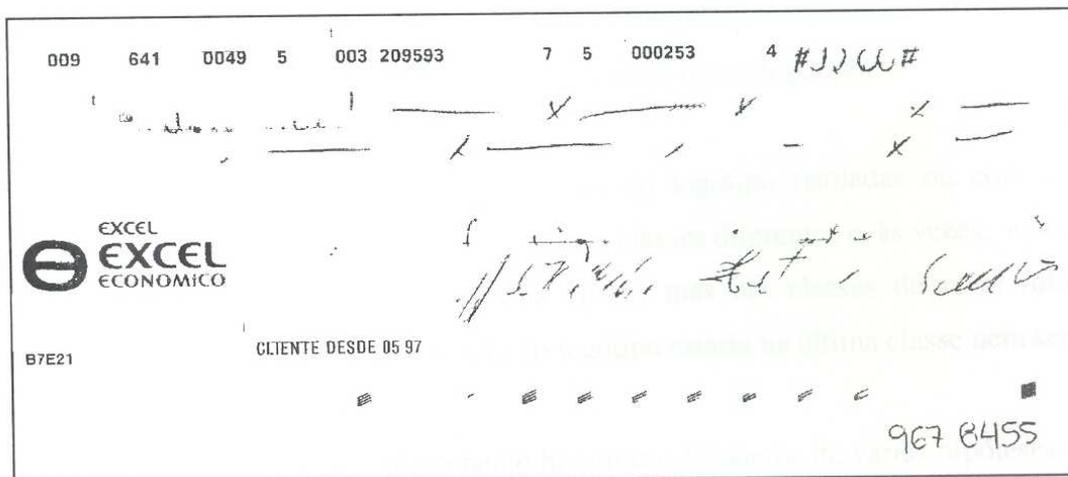


Figura 5.11 - b

Figura 5.11 (a) Imagem erodida; (b) Imagem reconstruída e com os ruídos eliminados

5.3 Fase de Localização/Extração do Logotipo

Devido à grande variedade de logotipos em cheques bancários, localizá-los e extraí-los, eliminando as informações irrelevantes e reconstruindo o logotipo, é extremamente heurístico.

Neste trabalho, a extração automática de logotipos apóia-se numa única evidência: **o logotipo é o maior componente encontrado no cheque bancário.** Partindo dessa hipótese, pode-se, portanto, pensar em peneirar a imagem até descobrir esse maior componente. Por utilizar o critério de grandeza na ausência de parâmetros heurísticos, empregou-se novamente a granulometria binária.

As etapas da fase de localização/extração de logotipos são descritas abaixo.

5.3.1 Etapa de Extração da I-Ésima Maior Porção do Logotipo

A última classe gerada no processo granulométrico, teoricamente, deve conter o logotipo por inteiro. Porém, na prática, isso não ocorre devido a:

- Desgastes sofridos durante o processo de binarização;

- O tamanho dos componentes no logotipo ser irregular;
- Alguns componentes manuscritos serem maiores que o logotipo.

O resultado é a obtenção de porções do logotipo (isoladas ou com outros elementos que não fazem parte do logotipo) em classes diferentes e, às vezes, nenhuma dessas porções é classificada na última classe, mas em classes de nível inferior. Portanto, a hipótese de que uma porção do logotipo estaria na última classe nem sempre seria verdadeira.

Como na área de reconhecimento há um uso freqüente de várias hipóteses e há um paradoxo que diz: “**segmentar para reconhecer ou reconhecer para segmentar?**”, optou-se aplicar várias hipóteses para segmentar. Assim, a qualidade do resultado da segmentação pode ser verificada através da etapa de reconhecimento.

No caso deste trabalho, cada classe gerada pelo processo granulométrico representará uma hipótese. A última classe é a primeira hipótese (o logotipo é o maior componente do cheque), a penúltima classe a segunda hipótese (o logotipo é o segundo maior componente do cheque), e assim por diante (Figura 5.12 e 5.13). A etapa de reconhecimento decidirá qual hipótese gera o melhor resultado.

O objetivo é encontrar uma classe que contenha uma porção isolada do logotipo para que o processo de reconstrução binária, nas próximas etapas (seção 5.2.3 a seção 5.2.4), possa reconstruí-lo no seu formato original.

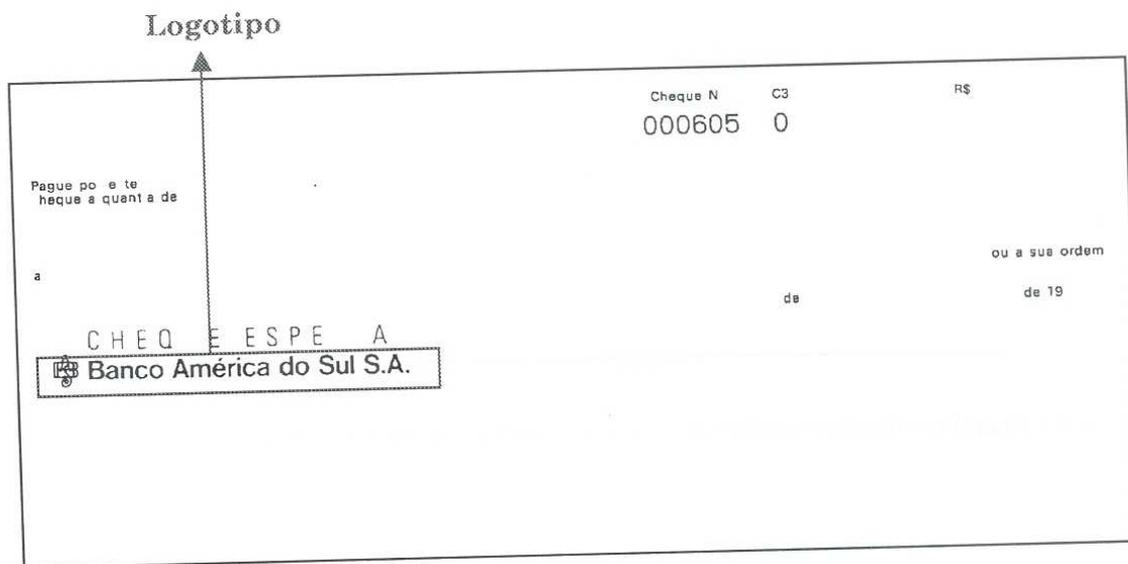


Figura 5.12 – (a) Porção de uma imagem binária de cheque bancário.

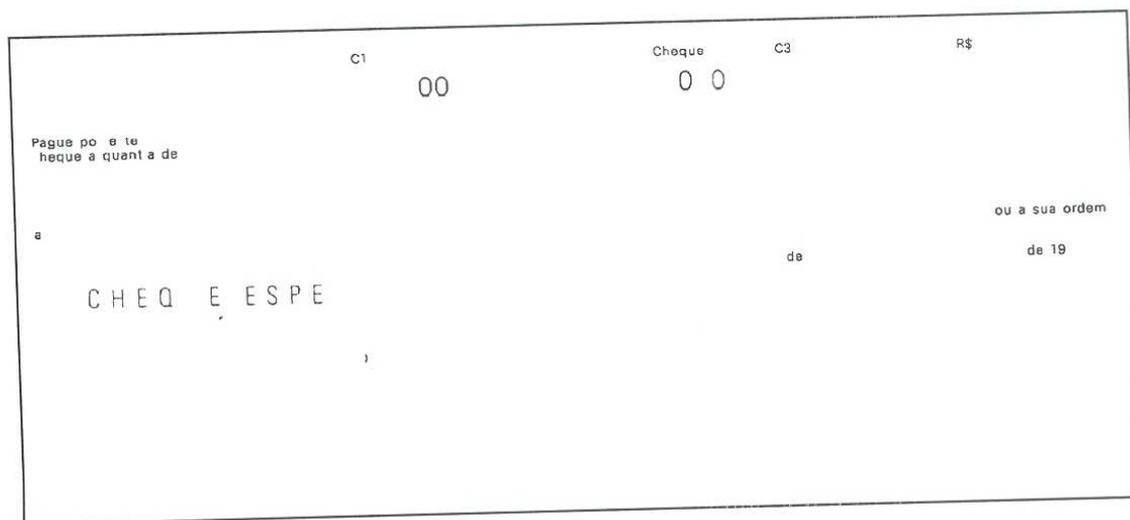


Figura 5.12 – (b) Primeira classe gerada pelo processo granulométrico aplicado na Figura 5.11a.

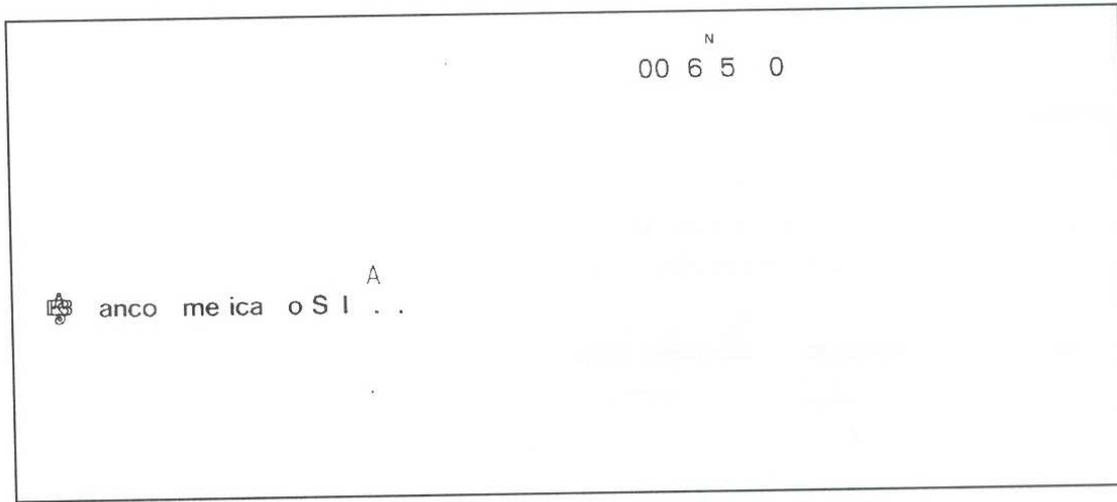


Figura 5.12 – (c) Segunda classe gerada pelo processo granulométrico aplicado na Figura 5.11a.

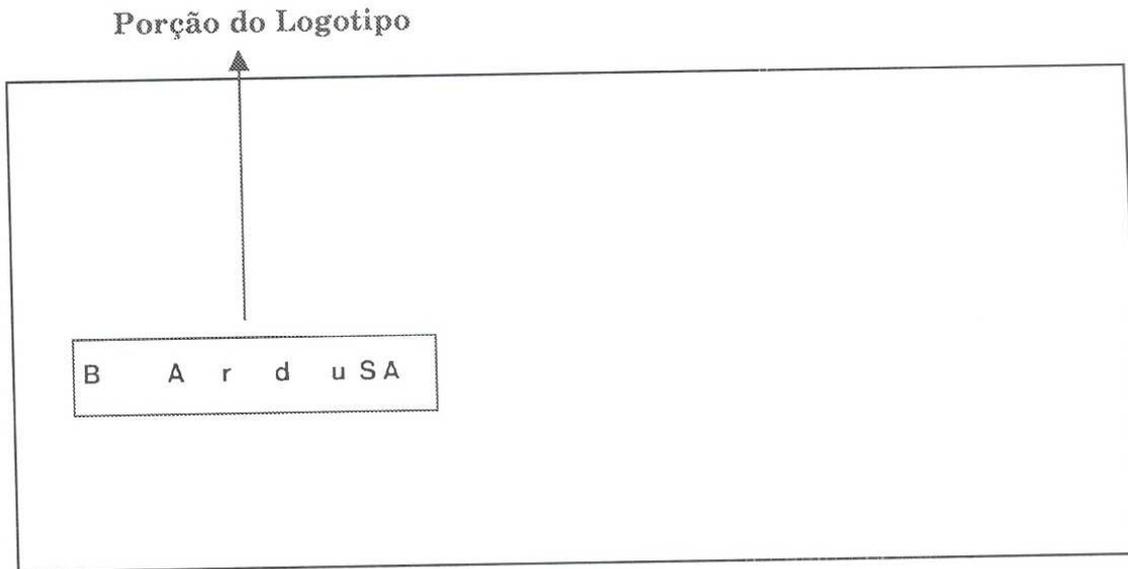


Figura 5.12 – (d) Última classe gerada pelo processo granulométrico aplicado na Figura 5.11a.

Figura 5.12 Exemplo de obtenção de uma porção do logotipo com uma hipótese.



Figura 5.13 – (a) Porção de uma imagem binária de um cheque bancário.

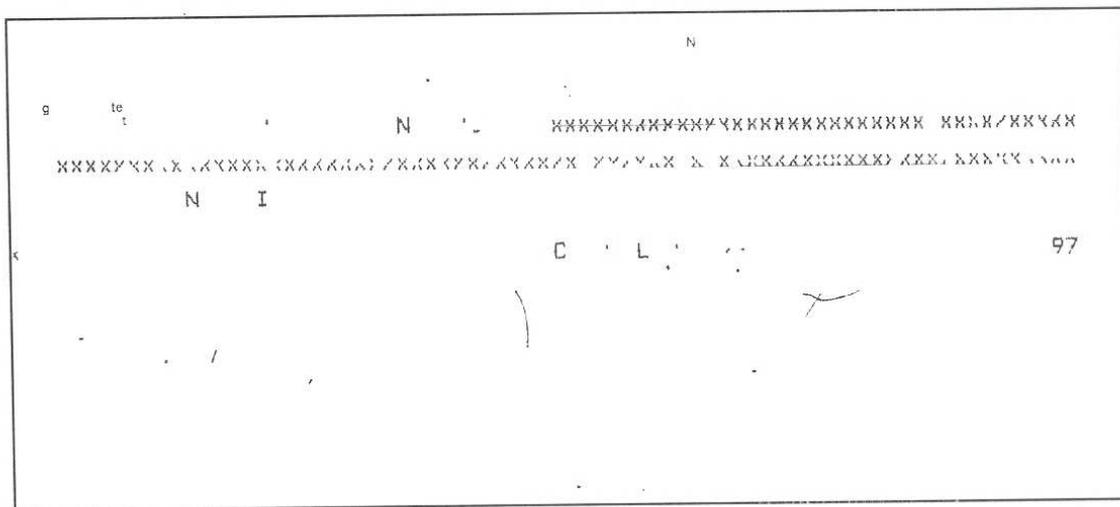


Figura 5.13 – (b) Primeira classe gerada pelo processo granulométrico aplicado na Figura 5.13a.

251701 9 (45,00)
QUARENTA E CINCO REAIS
X X X
A P CHEMI E C A LTDA
AMF ARGO 22 DEZEMBRO

Figura 5.13 – (c) Segunda classe gerada pelo processo granulométrico aplicado na Figura 5.13a.

A C



Figura 5.13 – (d) Terceira classe gerada pelo processo granulométrico aplicado na Figura 5.13a.

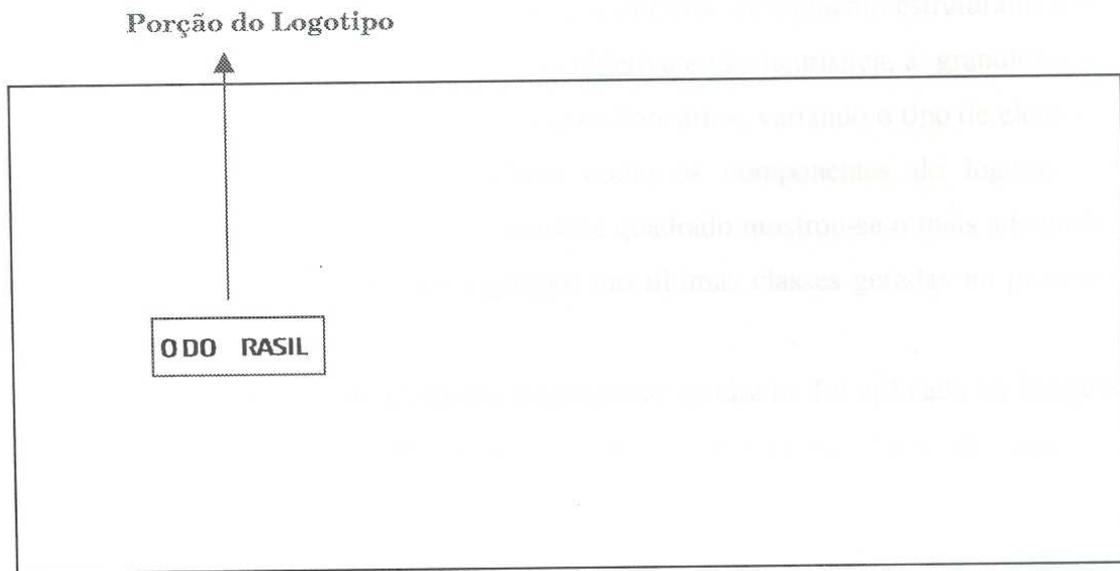


Figura 5.13 – (e) Penúltima classe gerada pelo processo granulométrico aplicado na Figura 5.13a.

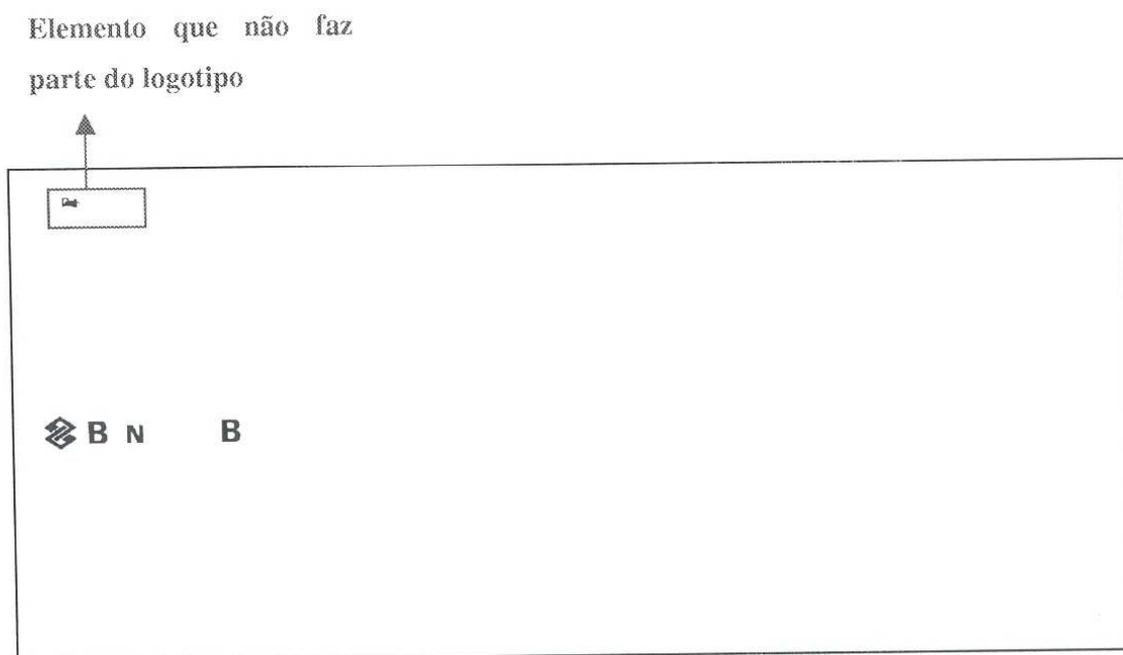


Figura 5.13 – (f) Última classe gerada pelo processo granulométrico aplicado na Figura 5.13a.

Figura 5.13 Exemplo de obtenção de uma porção do logotipo com duas hipóteses.

Na técnica granulométrica, é preciso definir o tipo de elemento estruturante a ser aplicado. Para efetuar essa escolha de forma objetiva e não heurística, a granulometria foi aplicada novamente nas imagens de cheques bancários, variando o tipo de elemento estruturante. O objetivo agora é analisar como os componentes do logotipo se distribuem nas classes. O elemento estruturante quadrado mostrou-se o mais adequado, por extrair as maiores porções dos logotipos nas últimas classes geradas no processo granulométrico.

A granulometria, com elemento estruturante quadrado, foi aplicada na imagem sem ruídos (Figura 5.11b). O elemento extraído na i -ésima classe do processo granulométrico é apresentado na Figura 5.14.

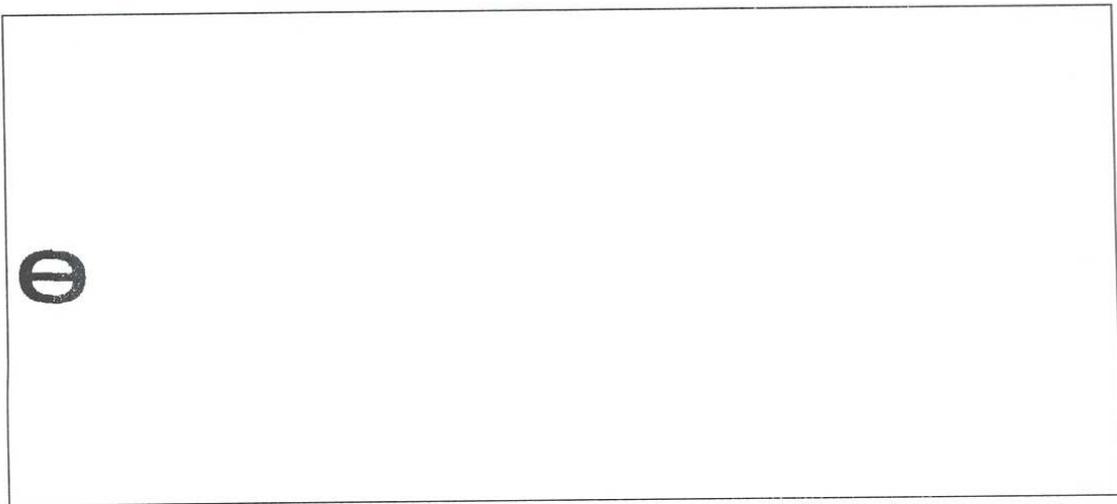


Figura 5.14 (a) Porção de um logotipo obtida na última classe do processo granulométrico aplicado na Figura 5.11b.

5.3.2 Etapa de Conexão dos Símbolos e Letras que Compõem o Logotipo

Para extrair o logotipo em sua integridade através da reconstrução binária, é fundamental que todos os símbolos e letras que compõem o logotipo estejam conectados na imagem de máscara. Nesse caso, a máscara é a imagem do cheque sem ruídos (Figura 5.11b). A conexão dos elementos foi feita de forma heurística, dilatando horizontalmente a imagem (Figura 5.11b) com 20 iterações (Figura 5.15).

Em alguns logotipos, a disposição dos seus elementos encontra-se na vertical. Portanto, para fazer a conexão desses elementos é preciso utilizar a dilatação vertical. O problema reside no fato que o critério de seleção do número adequado de dilatações verticais é altamente dependente do cheque tratado. Testes foram realizados com a dilatação vertical em uma amostra de cheques (18 imagens e cada imagem representa uma instituição bancária) variando o número de iterações e verificou-se que: ou os elementos do logotipo foram conectados, ou junto com os elementos foram agregados outros componentes, ou simplesmente os elementos não foram conectados. Por não encontrar um número exato de iterações que satisfizesse todos os casos a dilatação vertical não foi aplicada. No caso de logotipos com elementos dispostos na posição vertical ocorrerá uma segmentação parcial desse logotipo (veja Capítulo 6).

Logotipo com os elementos conectados

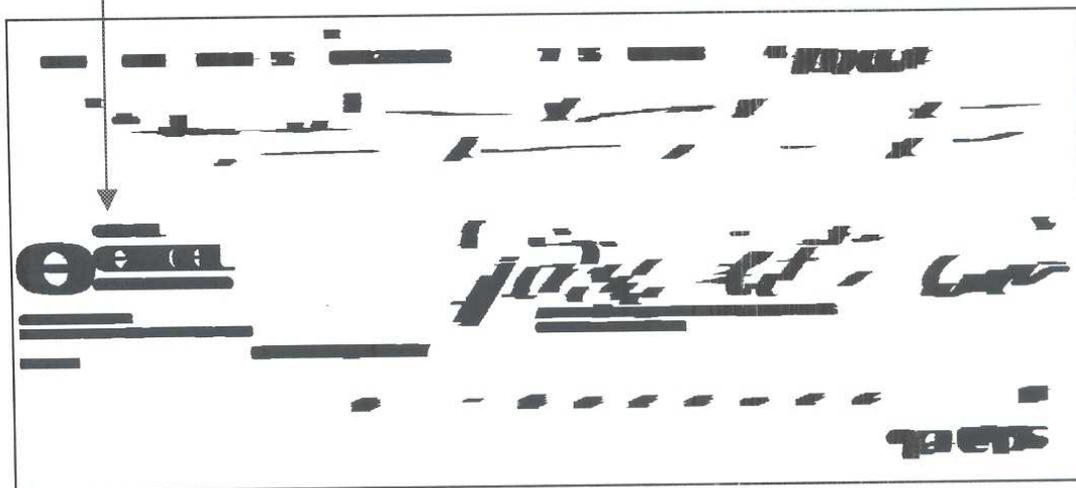


Figura 5.15 – Imagem de cheque com todos os elementos conectados

5.3.3 Etapa de Extração do Logotipo Dilatado

A extração da versão dilatada do logotipo é feita por reconstrução binária. O marcador é a imagem que contém a maior porção do logotipo (Figura 5.14b) e a máscara é a imagem que contém todos os símbolos e letras do logotipo conectados (Figura 5.15). A imagem que contém o logotipo na versão dilatada é apresentada na Figura 5.16.

5.3.4 Etapa Final de Extração do Logotipo

Finalmente, a extração do logotipo em seu formato original é feita por intersecção entre a imagem binária filtrada (Figura 5.11b) e a imagem contendo a versão dilatada do logotipo (Figura 5.16). A segmentação do logotipo em sua integridade é apresentada na Figura 5.17.

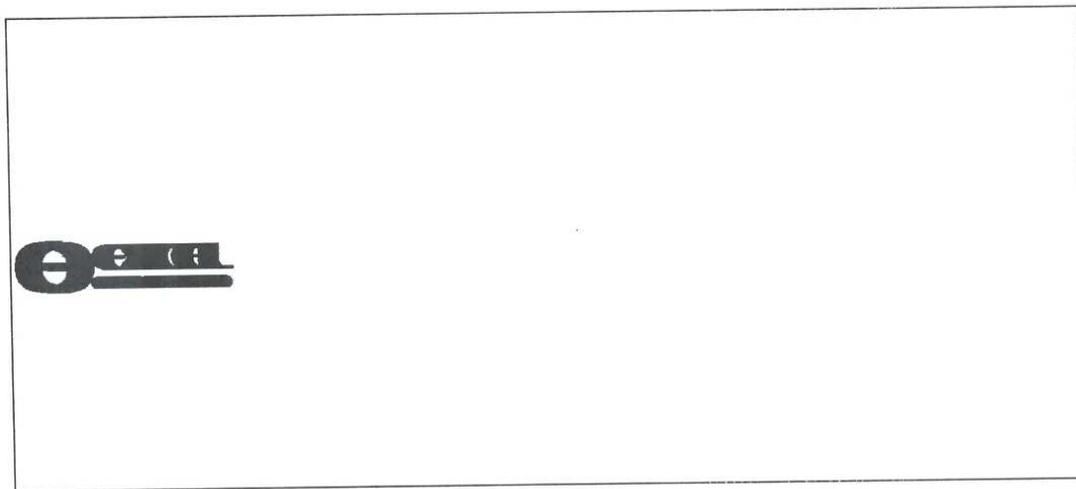


Figura 5.16 – Logotipo na versão dilatada

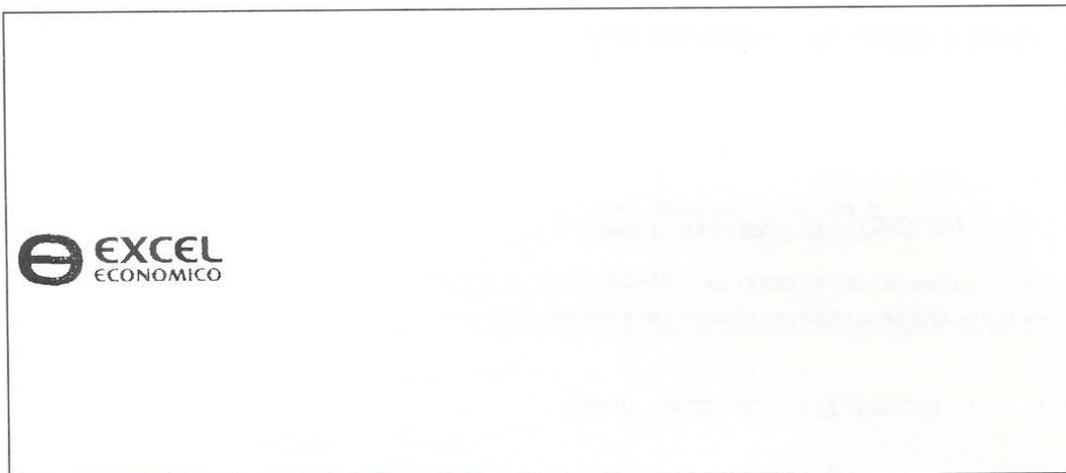


Figura 5.17 Logotipo segmentado

Este capítulo apresentou a metodologia proposta e suas etapas para segmentar logotipos de cheques bancários brasileiros. O exemplo apresentado no texto foi de um cheque preenchido a mão. Exemplos de segmentação de logotipos em cheques preenchidos a máquina e cheques em branco são ilustrados no final deste capítulo. Os problemas encontrados e os resultados estatísticos dos testes alcançados com essa metodologia serão discutidos no próximo capítulo.

- Exemplo de segmentação de logotipos em cheque preenchido a máquina:



Figura 5.18 Imagem original em níveis de cinza de um cheque preenchido a máquina



Figura 5.19 Imagem de um cheque preenchido a máquina com o fundo artístico removido



Figura 5.20 Imagem binária de um cheque preenchido a máquina

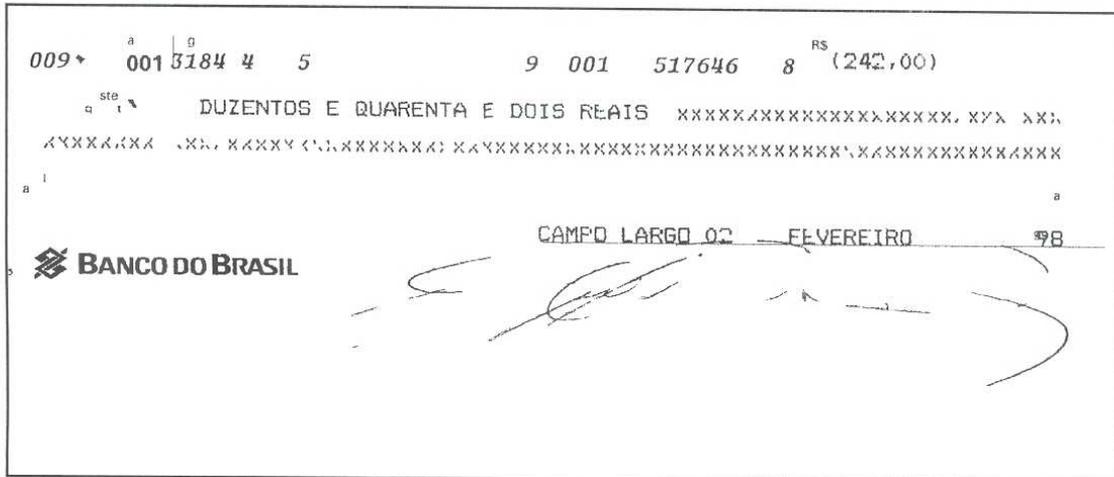


Figura 5.21 Imagem de um cheque preenchido a máquina com os ruídos removidos

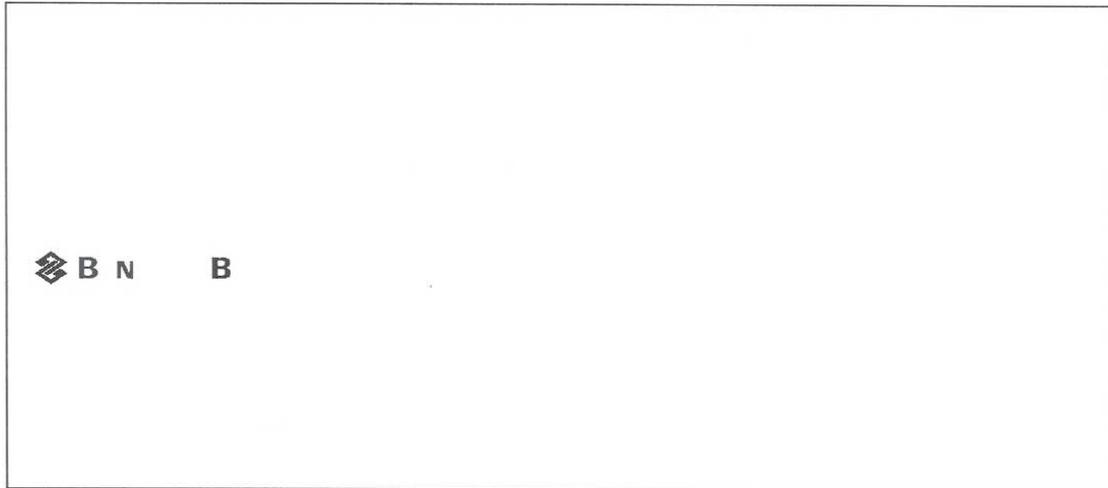


Figura 5.22 Porção isolada do logotipo obtida na i-ésima classe gerada pelo processo granulométrico, aplicado em uma imagem de cheque preenchido a máquina.

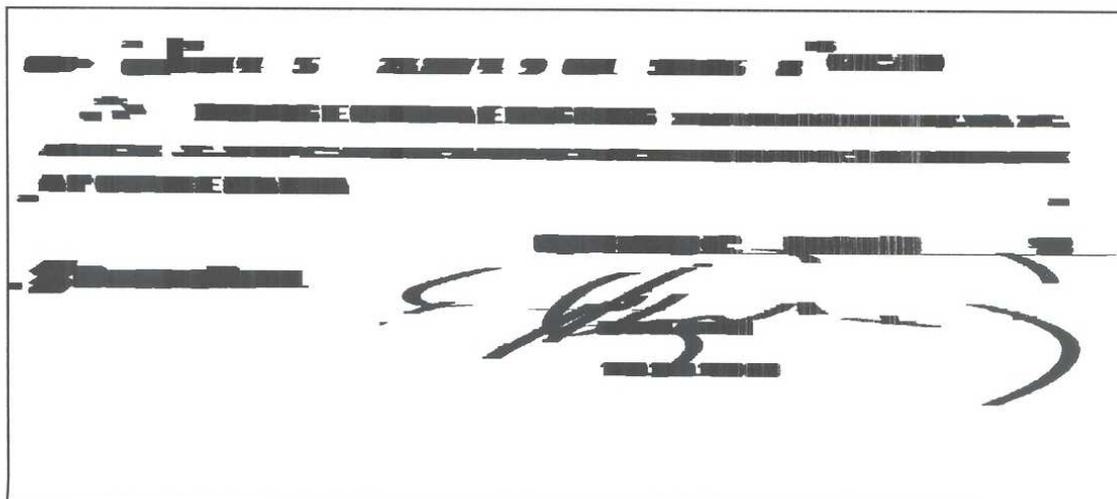


Figura 5.23 Imagem de um cheque preenchido a máquina com todos os elementos conectados

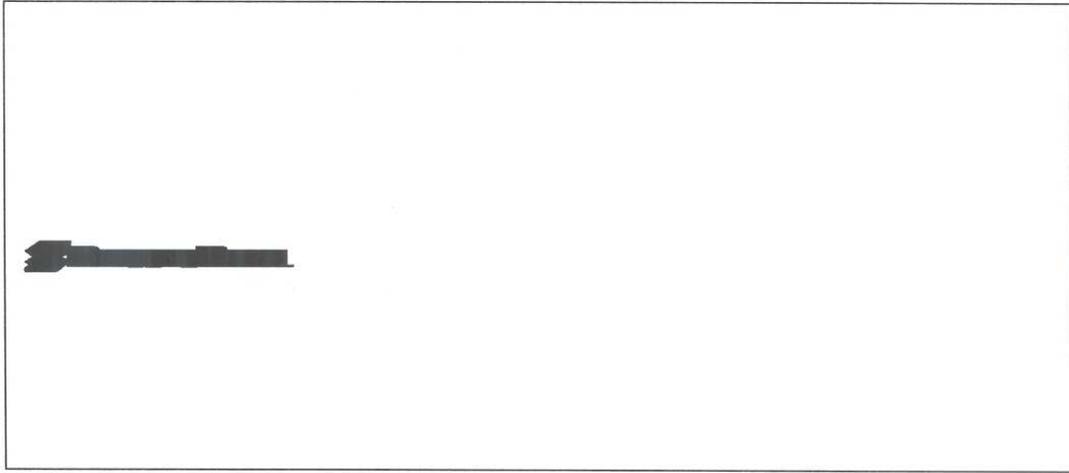


Figura 5.24 Versão dilatada de um logotipo de um cheque preenchido a máquina

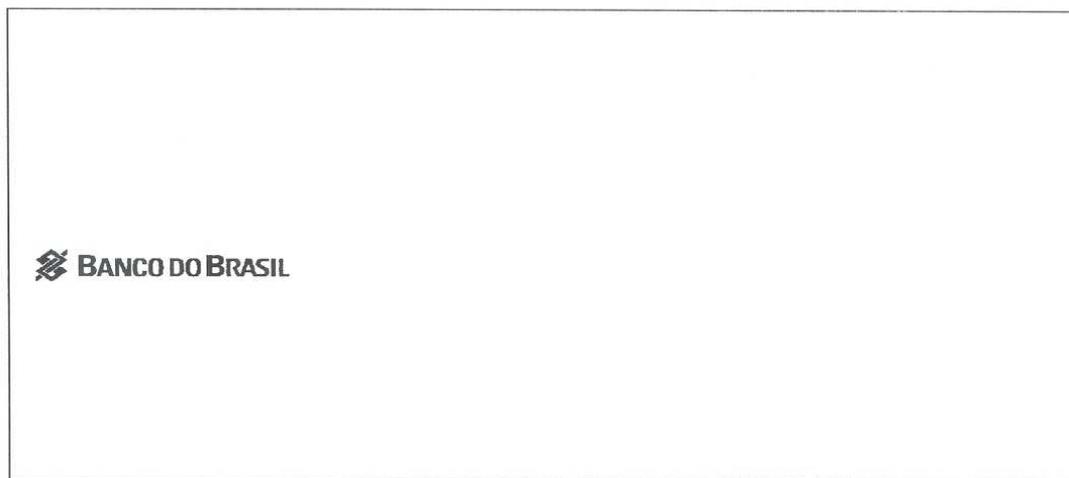


Figura 5.25 Logotipo segmentado de um cheque preenchido a máquina

- Exemplo de segmentação de logotipos em cheque em branco:

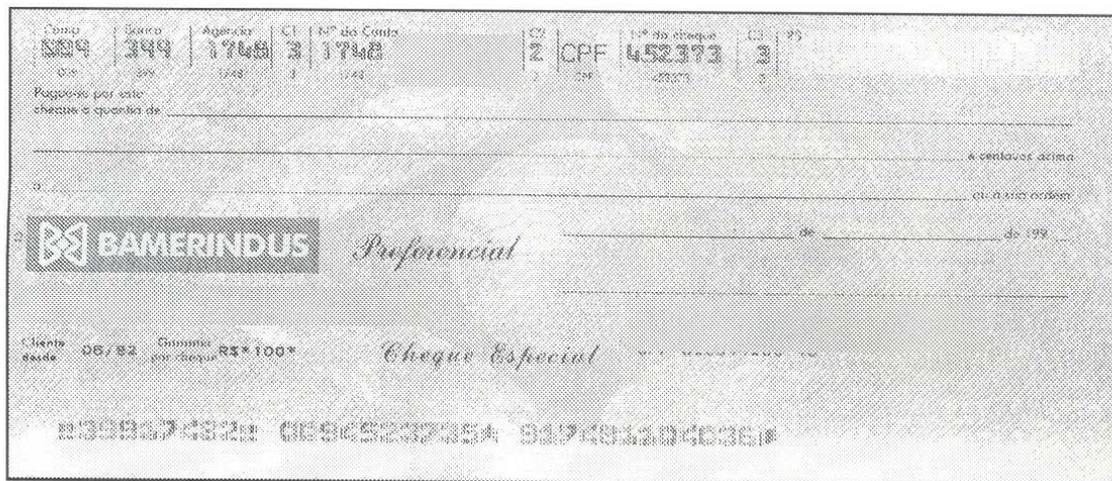


Figura 5.26 Imagem original em níveis de cinza de um cheque em branco

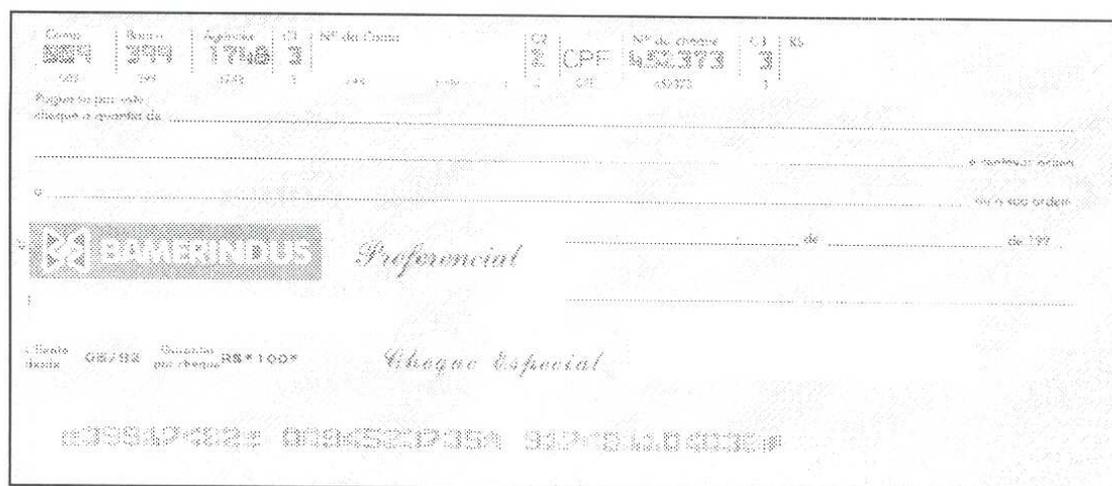


Figura 5.27 Imagem de um cheque em branco com o fundo artístico removido

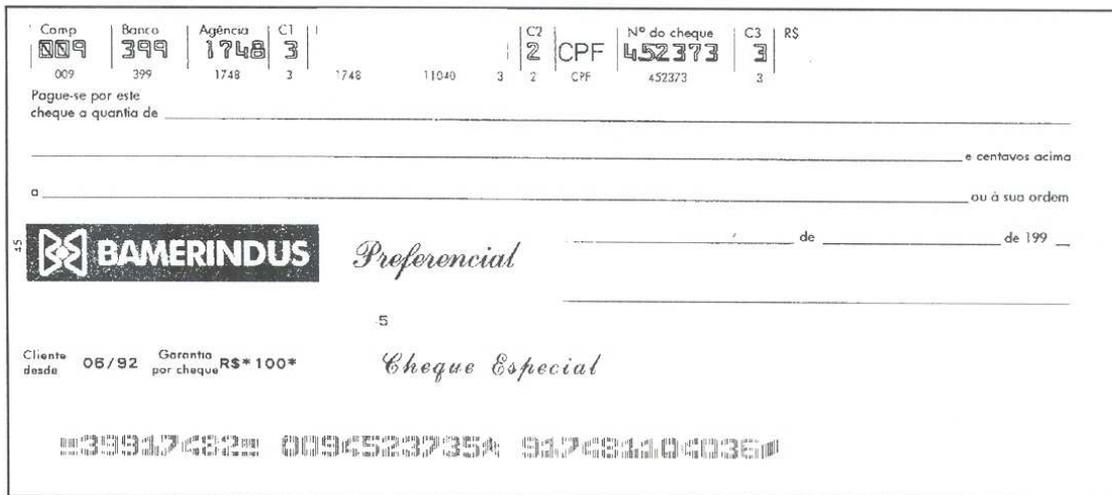


Figura 5.28 Imagem binária de um cheque em branco

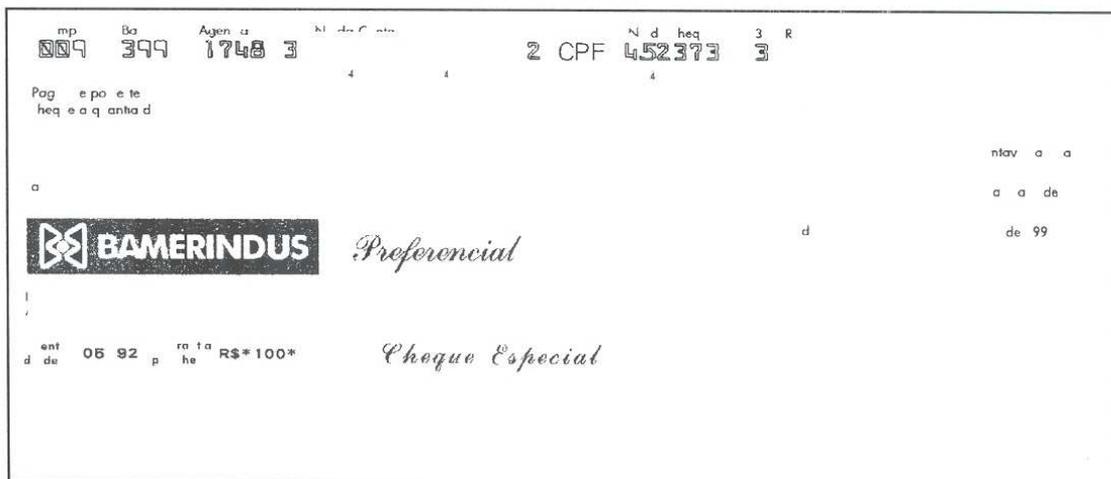


Figura 5.29 Imagem de um cheque em branco com os ruídos removidos

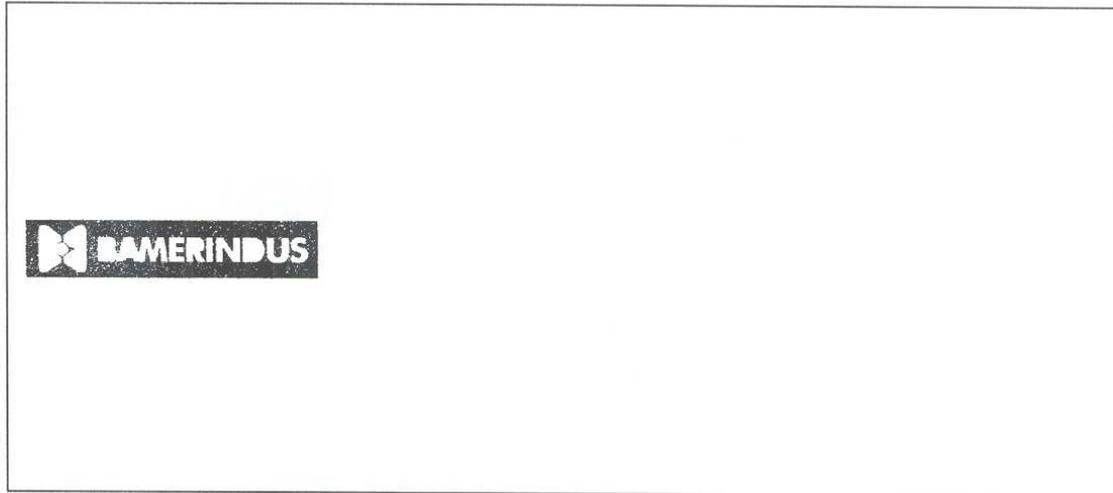


Figura 5.30 Porção isolada do logotipo obtida na i -ésima classe gerada pelo processo granulométrico, aplicado em uma imagem de cheque em branco.

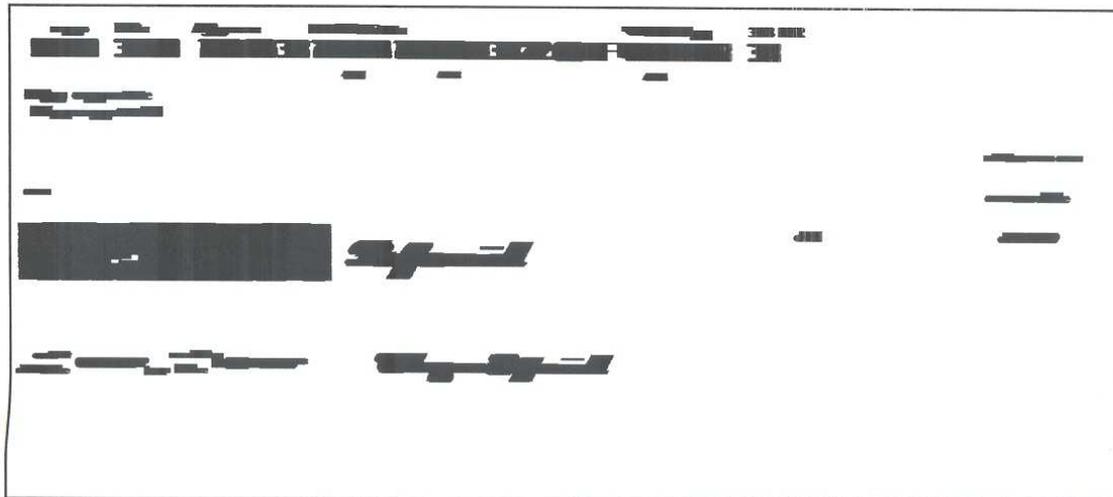


Figura 5.31 Imagem de um cheque em branco com todos os elementos conectados

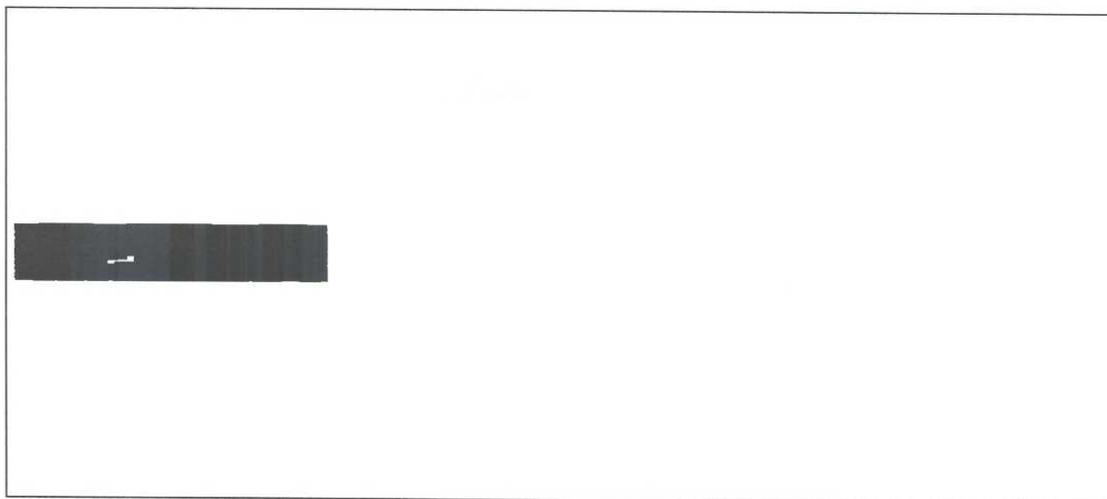


Figura 5.32 Versão dilatada de um logotipo de um cheque em branco

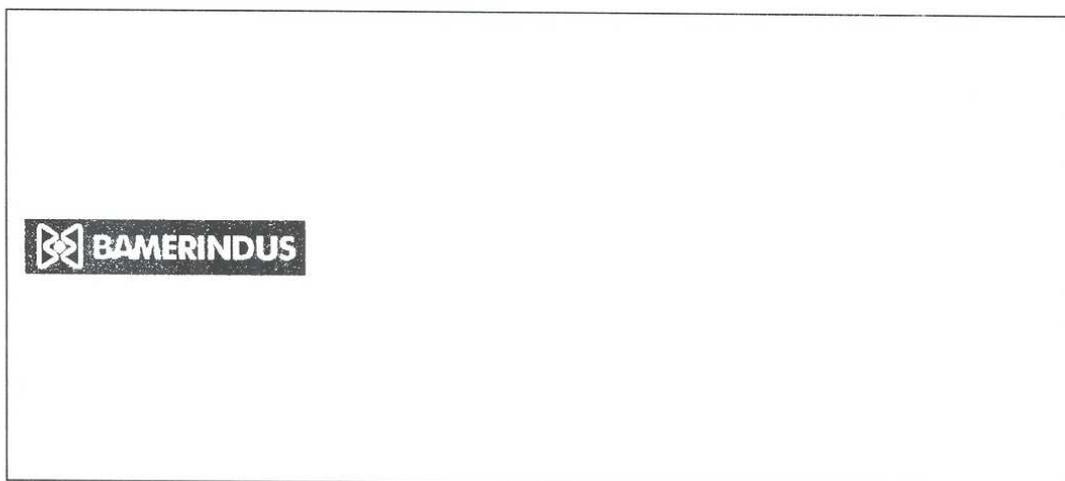


Figura 5.33 Logotipo segmentado de um cheque em branco

Capítulo 6

Análise dos Resultados

6.1 Introdução

A metodologia proposta foi testada em uma base de imagens em níveis de cinza de cheques bancários brasileiros, misturando logotipos do tipo *word-in-mark* e *composite-mark* [KIM Y. S. & KIM W. Y., 1998]. A base de dados, como citado no Capítulo 3, contém 478 cheques bancários preenchidos a mão, preenchidos a máquina e cheques não preenchidos (em branco), contemplando 18 instituições bancárias brasileiras.

A diferença entre um cheque preenchido e um cheque em branco pode ser definida da seguinte forma: um cheque em branco ou vazio pode ser visto como um documento comum (por exemplo: envelope de correio ou carta) com algumas particularidades: fundo artístico complexo e dados pré-impressos com tamanhos diversos; um cheque preenchido contém as mesmas particularidades de um cheque em branco mais a adição de dados manuscritos e/ou pré-impressos.

Para uma instituição financeira, as informações relevantes de um cheque bancário são os dados entrados pelo usuário (data, valor numérico, assinatura e outros). Portanto, o processamento de cheques em branco não é importante para estas instituições. Mas, como o cheque em branco pode ser visto como um documento comum, o seu processamento pode ser interessante para a área de documentos porque ele permite analisar padrões diferentes (informações manuscritas e pré-impressas, fundo artístico, logotipo) em um mesmo documento. Desta forma, é possível avaliar como as

técnicas de processamento de imagens se comportam diante de padrões diferentes misturados, podendo ser estendidas para outros tipos de documentos.

Neste capítulo serão apresentados os resultados quantitativos dos testes que foram efetuados, além de analisar os principais problemas encontrados.

6.2 Resultados Quantitativos dos Testes

A análise dos resultados dos testes foi feita de forma visual. A qualidade do resultado da segmentação foi considerada da seguinte forma:

- Completa: quando o logotipo é segmentado em sua integridade (Figura 6.1);
- Parcial: quando o logotipo é segmentado com a ausência de um algum componente (Figura 6.2);
- Ruim: quando o logotipo é segmentado junto com outros componentes que não fazem parte dele ou quando outro componente qualquer é segmentado no lugar do logotipo (Figura 6.3).

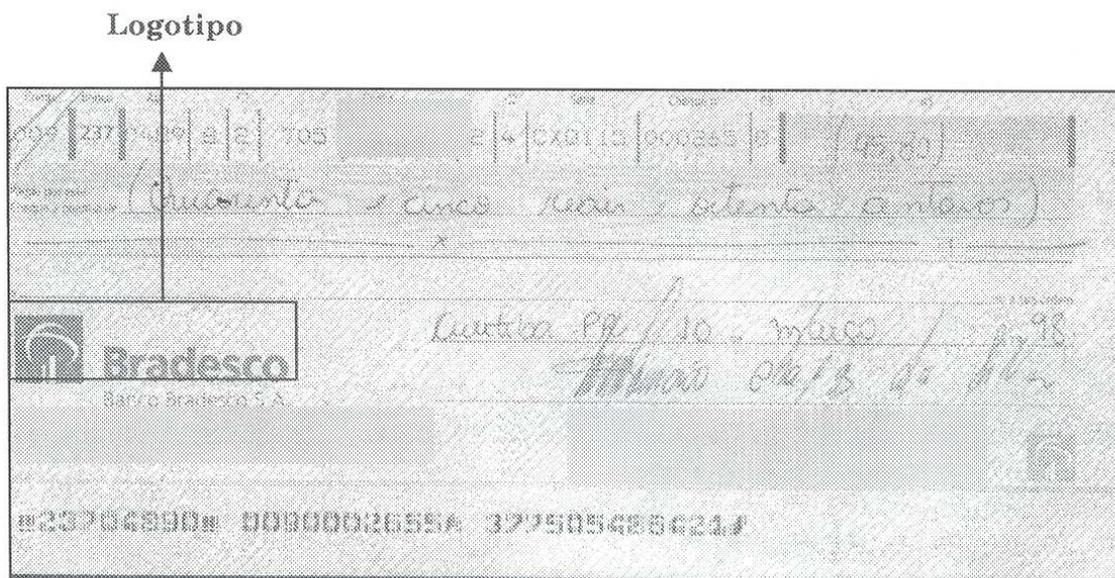


Figura 6.1 - a

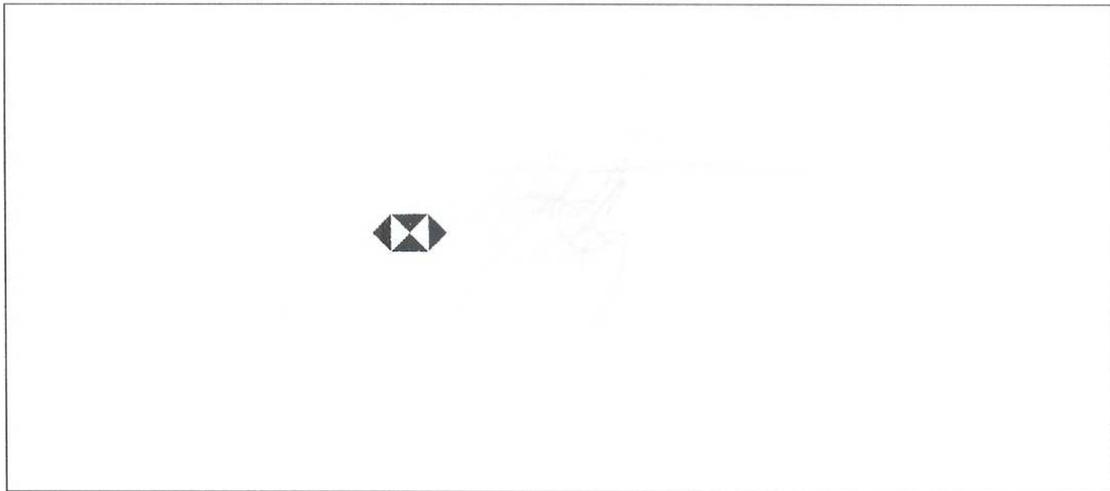


Figura 6.2 – b

Figura 6.2 Exemplo de segmentação parcial. (a) imagem de cheque bancário; (b) uma porção do logotipo.

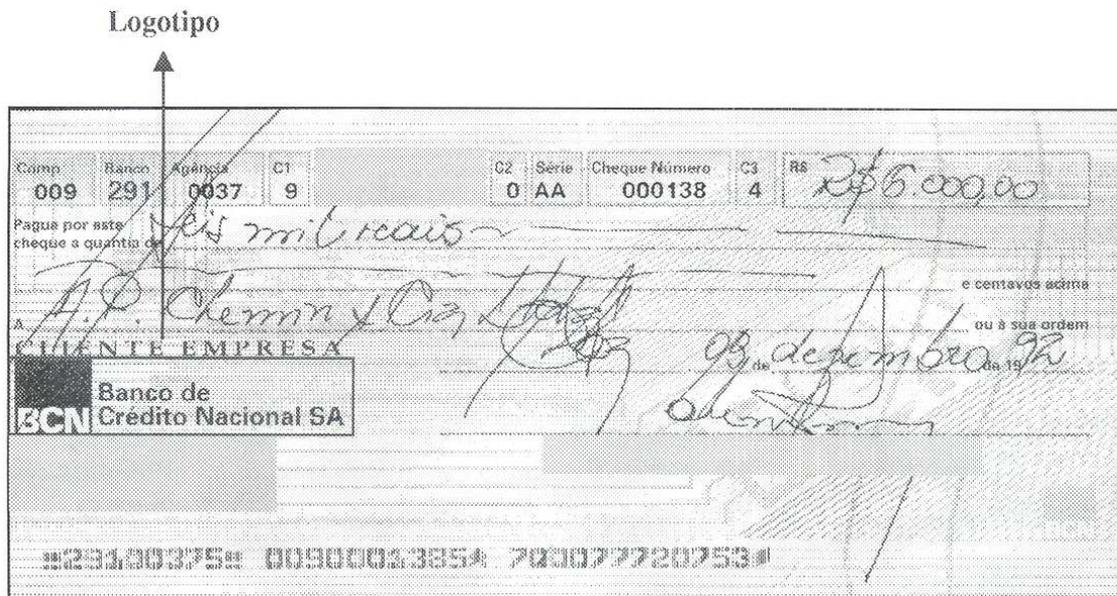


Figura 6.3 – a

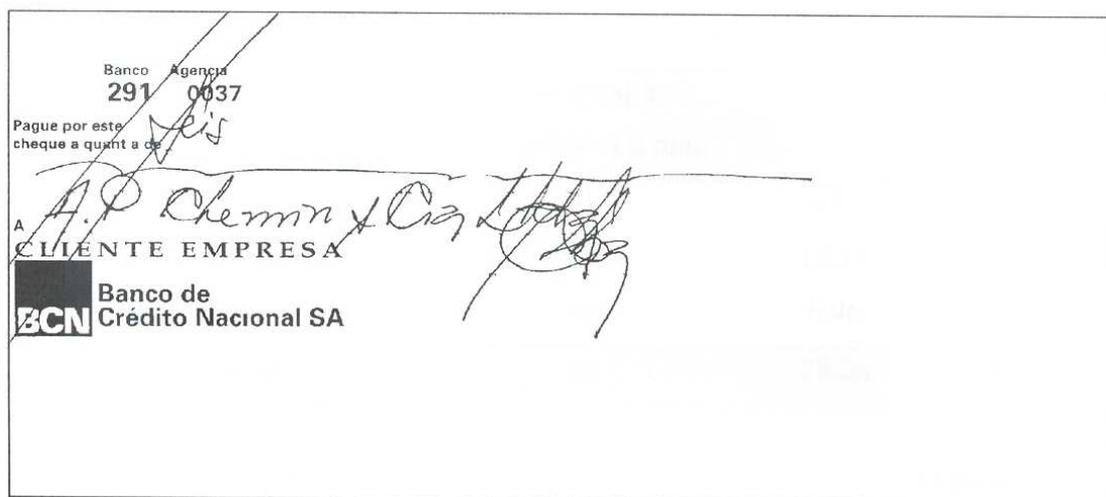


Figura 6.3 – b

Figura 6.3 Exemplo de segmentação ruim. (a) imagem de cheque bancário; (b) logotipo segmentado junto com outros componentes.

Os problemas encontrados com a metodologia proposta foram:

- (i) Binarização insatisfatória;
- (ii) Influência de traços manuscritos;
- (iii) Componentes do logotipo classificados com outros elementos em classes iguais.

Das 230 imagens de cheques preenchidos a mão, 28.26% apresentaram um dos problemas citados (Tabela 6.1). Das 190 imagens de cheques preenchidos a máquina, 15.79% apresentaram problemas (Tabela 6.2). E das 58 imagens de cheques em branco, 5% apresentaram um dos problemas citados (Tabela 6.3).

Problemas	Nº de cheques Preenchidos a mão	%
(i)	21	9.13
(ii)	28	12.17
(iii)	16	6.96
Total	65	28.26

Tabela 6.1 Resultados quantitativos da segmentação com qualidade ruim - cheques preenchidos a mão.

Problemas	Nº de cheques Preenchidos a máquina	%
(i)	21	11.05
(ii)	01	0.53
(iii)	08	4.21
Total	30	15.79

Tabela 6.2 Resultados quantitativos da segmentação com qualidade ruim - cheques preenchidos a máquina.

Problemas	Nº de cheques em branco	%
(i)	01	1.72
(ii)	00	0.0
(iii)	04	6.90
Total	5	8.62

Tabela 6.3 Resultados quantitativos da segmentação com qualidade ruim - cheques em branco.

A metodologia proposta para segmentação automática de logotipos, aplicada nas 478 imagens da base de dados, apresentou **59.62%** de segmentação completa e **17.16%** de segmentação parcial utilizando uma hipótese e, **2.09%** de segmentação completa e **0.21%** de segmentação parcial utilizando duas hipótese. Isto equivale a **79.08%** de toda a base de dados. A tabela 6.4 apresenta os resultados quantitativos parciais para cada tipo de cheque.

Tipo de Cheque	Quant.	Nº de Hipóteses	Segmentação Completa (nº de cheques)	Segmentação Parcial (nº de cheques)	% Completa	% Parcial
Preenchido a mão	230	Última classe	136	25	59.13	10.87
		Duas Últimas classes	03	01	1.31	0.43
Preenchido a máquina	190	Última classe	109	48	57.37	25.26
		Duas Últimas classes	03	00	1.58	0.0
em branco	58	Última classe	40	09	68.96	15.52
		Duas Últimas classes	04	00	6.90	0.0
478			295	83	61.72	17.36

Tabela 6.4 Resultados quantitativos da segmentação parcial e completa.

Pode-se observar, na tabela 6.4, que os cheques em branco atingiram os melhores resultados. Isto ocorreu porque a metodologia mostrou-se sensível diante de alguns padrões manuscritos. Portanto, ela precisa ser melhorada com técnicas mais apuradas para reduzir a influência desses componentes. Mas, a metodologia mostrou-se robusta diante de padrões diferentes misturados (fundo artístico, componentes pré-impresos de tamanhos diferentes, logotipo), podendo ser aplicada em outros tipos de documentos. Um exemplo da aplicação da metodologia em outro documento será mostrado na seção 6.4.

Na base de dados em questão foi necessário no máximo duas hipóteses (última e penúltima classe gerada pelo processo granulométrico) para encontrar uma porção isolada do logotipo. Como a granulometria utiliza a grandeza para classificar, isto prova que o logotipo é o maior ou o segundo maior componente de um cheque bancário.

6.3 Análise dos Problemas Encontrados

6.3.1 Binarização Insatisfatória

Como nem sempre imagens binarizadas são de boa qualidade, algumas imagens de cheques bancários foram binarizadas de forma insatisfatória, gerando degradações e redução do tamanho em alguns logotipos. A Figura 6.4 exemplifica este problema.

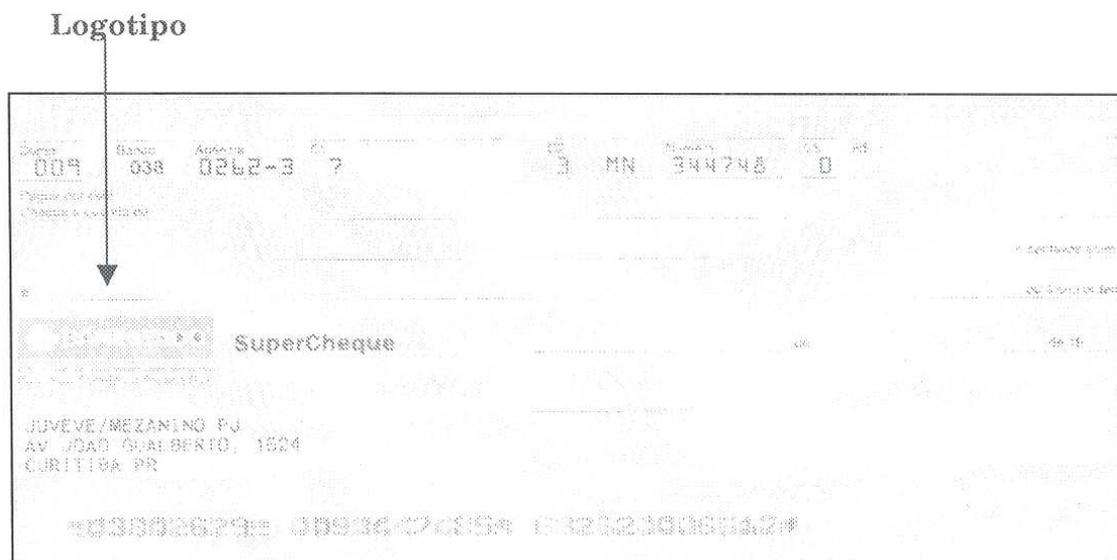


Figura 6.4 - a



Figura 6.4 - b

Figura 6.4 (a) Imagem em níveis de cinza; (b) Imagem binarizada de forma insatisfatória.

Na Figura 6.4b, verifica-se que a técnica de binarização degradou consideravelmente o logotipo e, devido a isto, ele deixou de ser o maior componente do cheque bancário. Portanto, em alguns casos, outros componentes que não fazem parte do logotipo são segmentados; em outros casos, uma parte pouco significativa do logotipo é segmentada (Figura 6.5).

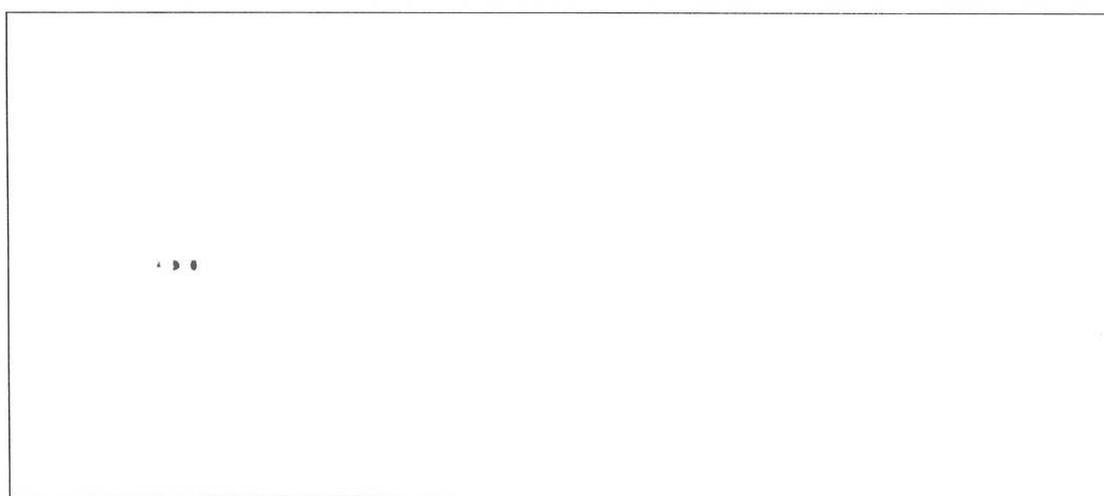


Figura 6.5 Segmentação de um componente do logotipo pouco significativo.

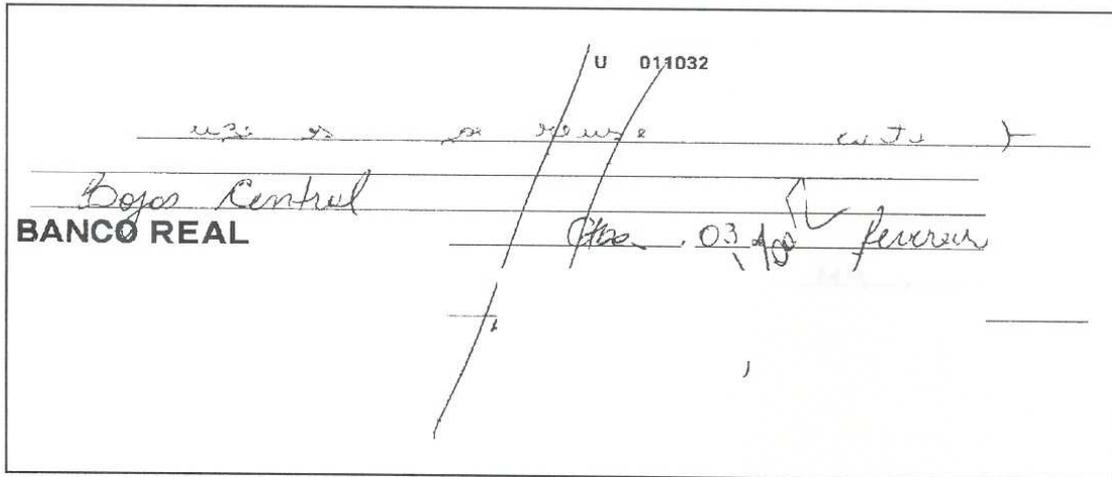


Figura 6.7 Segmentação indevida do logotipo (da Figura 6.6) por causa das linhas sobrepostas no cheque.

6.3.3 Componentes do Logotipo Classificados com os outros Elementos em Classes Iguais

Em alguns casos, o processo granulométrico classificou os componentes do logotipo junto com outros elementos. Isto afetou diretamente a performance das demais etapas da metodologia, e outros componentes além do logotipo foram segmentados. A Figura 6.8 exemplifica este tipo de problema.

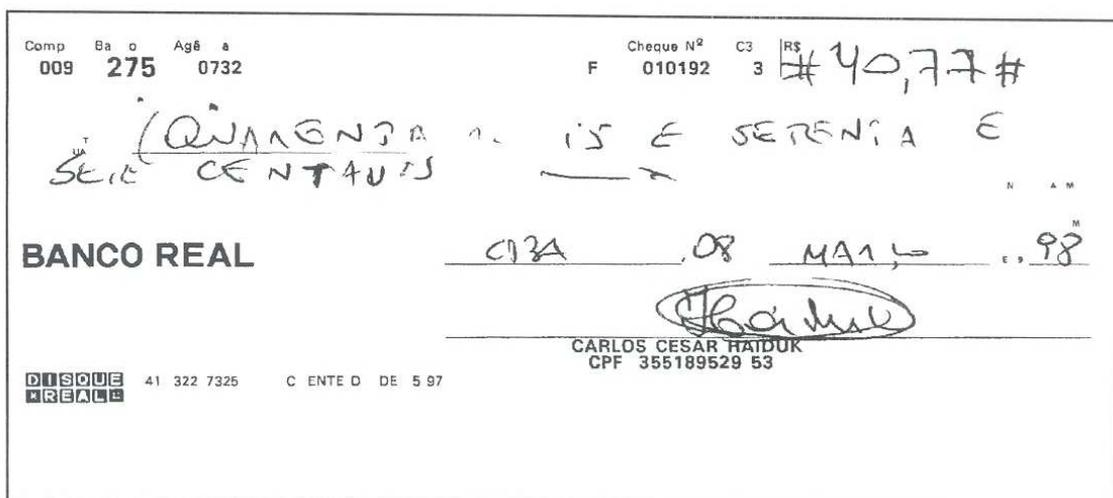


Figura 6.8 - a

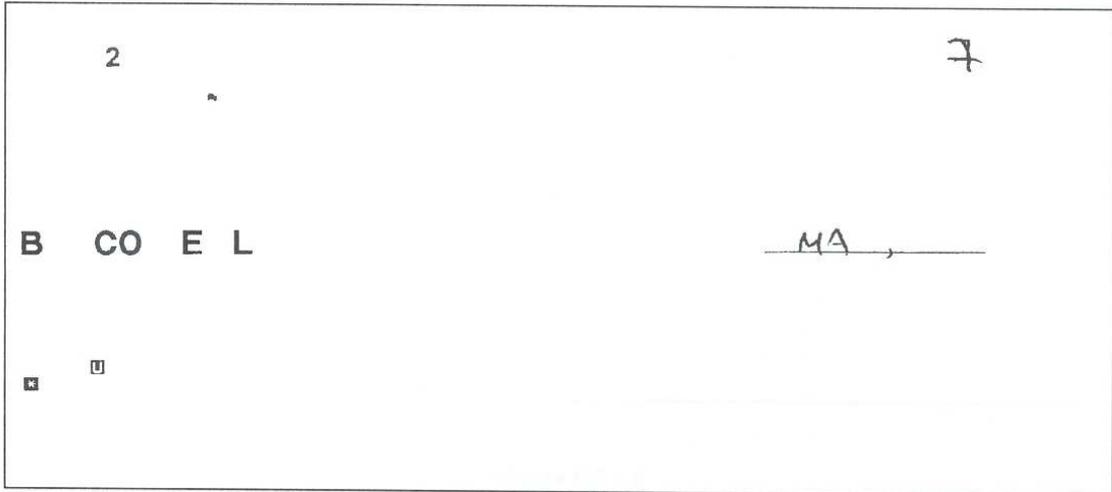


Figura 6.8 – b

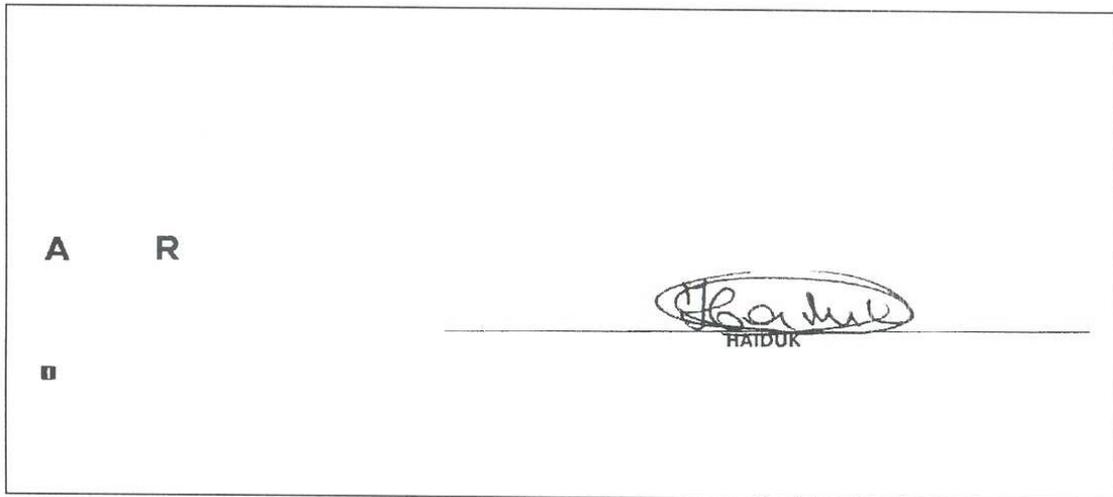


Figura 6.8 – c

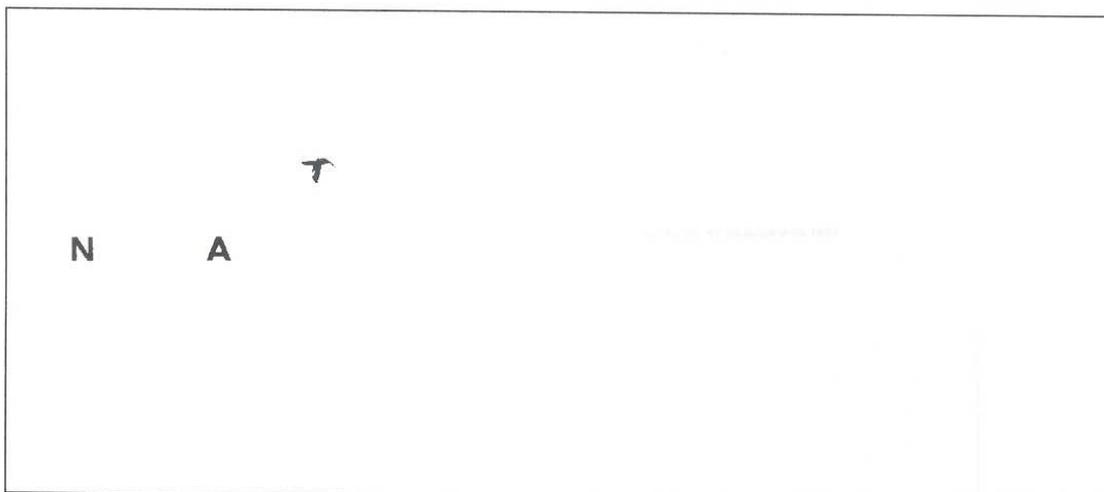


Figura 6.8 – d

Figura 6.8 (a) imagem parcial de um cheque bancário; (b), (c) e (d) apresentam componentes do logotipo classificados, pelo processo granulométrico, junto com outros elementos.

6.4 Aplicação da Metodologia Proposta em Outros Tipos de Documento

Para a área de processamento de documentos, os resultados alcançados com os cheques em branco foram muito interessantes, porque a metodologia proposta mostrou-se robusta para casos onde há padrões diferentes misturados (fundo artístico complexo, informações pré-impresas de diferentes tamanhos, logotipo), podendo ser estendida para outros tipos de documentos (carta, memorando, ofício, fax, envelope de correio entre outros) mais simples ou tão complexos quanto um cheque bancário.

Um exemplo é as imagens de fax. Elas não possuem um fundo artístico, portanto são mais simples do que as imagens de cheques bancários (Figura 6.9). A metodologia proposta foi aplicada em algumas dessas imagens e obteve sucesso na segmentação do logotipo.

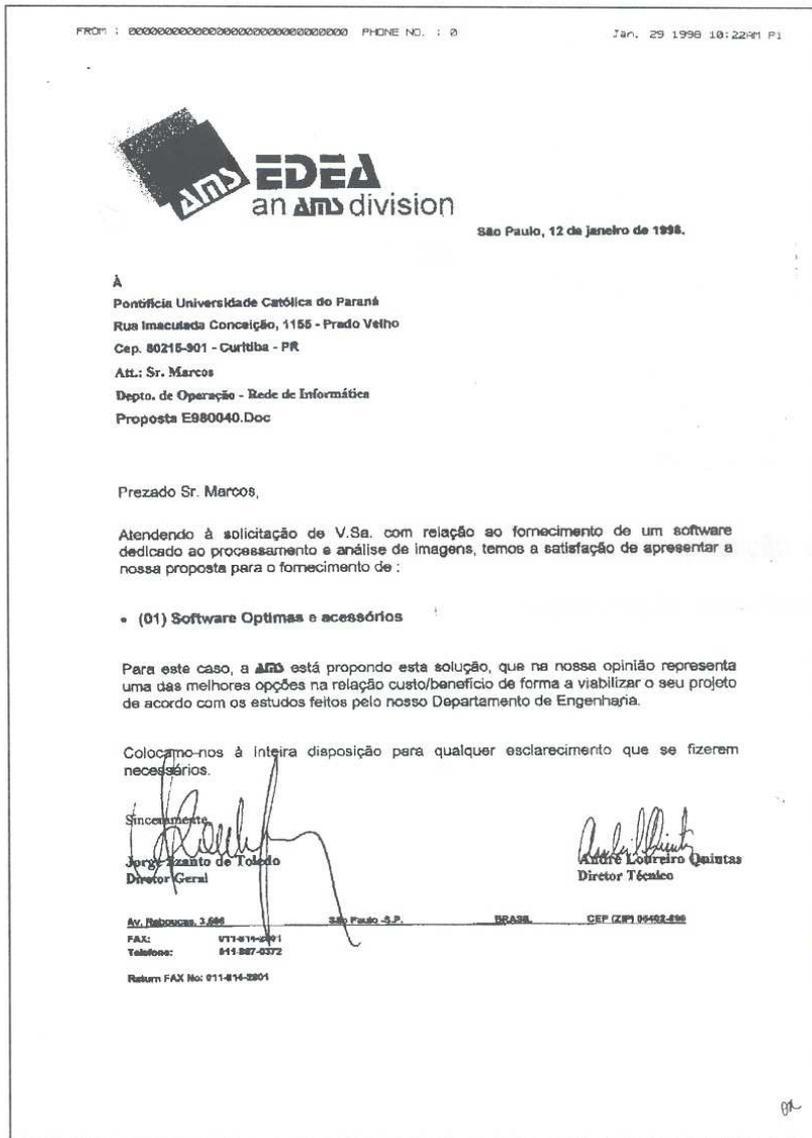


Figura 6.9 Imagem de fax

O objetivo principal desse trabalho foi minimizar o uso de heurísticas na segmentação de logotipos em cheques bancários brasileiros, visando a obter uma segmentação mais exata possível para todas as situações, permitindo, assim, o reuso dessa abordagem para outras aplicações.

Capítulo 7

Conclusões e Perspectivas

Uma das preocupações no desenvolvimento de uma aplicação que envolve a análise automática de imagens é encontrar uma segmentação significativa. A grande dificuldade reside no fato de não se conhecer de antemão o número e tipo de estruturas que se encontram na imagem. Geralmente, são usadas muitas heurísticas para o tamanho e posição dos componentes a serem extraídos, isso limita a segmentação para um determinado tipo de aplicação.

A abordagem apresentada neste estudo empregou a morfologia matemática para localizar e extrair automaticamente logotipos de imagens de cheques bancários brasileiros. O objetivo principal do trabalho foi reduzir o número de parâmetros heurísticos, visando a obter como resultado uma segmentação exata ou a mais exata possível para todas as situações, permitindo assim o seu reuso em outras aplicações.

No entanto, segmentar sem conhecimento prévio, empregando poucas heurísticas e poucas hipóteses, é um desafio. No caso de cheques bancários brasileiros, apesar da complexidade das imagens, aplicou-se um único conhecimento prévio: **o logotipo é maior do que todas as informações presentes no cheque bancário.**

As contribuições e conclusões do trabalho, e as perspectivas futuras serão apresentadas nas próximas seções.

7.1 Contribuição do Trabalho

Como já foi descrito no Capítulo 1 as principais contribuições do trabalho são:

- Uma metodologia de segmentação que não utiliza conhecimento a priori sobre o tipo de documento e aplica o mínimo possível de heurísticas, permitindo, assim, o seu reuso em outras aplicações.
- Processamento de dois modelos de logotipos: *word-in-mark* e *composite-mark*.
- Contribuição para o desenvolvimento de trabalhos futuros na área de segmentação automática de logotipos.

7.2 Conclusões sobre o Trabalho

Os testes realizados, na base de dados de 478 imagens de cheques bancários brasileiros, mostraram a viabilidade da abordagem proposta na redução do número de parâmetros heurísticos para a localização e a extração automática de logotipos. Além disso, verificou-se também que essa abordagem viabiliza a extração de logotipos de modelos diferentes, em nosso caso, logotipos *word-in-mark* e *composite-mark*.

A metodologia mostrou-se robusta nos casos onde há padrões diferentes misturados (fundo artístico complexo, informações pré-impressas de diferentes tamanhos, logotipo). Entretanto, quando padrões manuscritos são inseridos a metodologia torna-se sensível.

7.3 Perspectivas Futuras

No domínio de documentos, é grande o interesse no reconhecimento de logotipos, pois esses elementos são usados para distinguir documentos similares, porém diferentes, ou para indicar características específicas do documento em questão.

Mas, o reconhecimento não pode ser realizado sem existir antes uma fase de conhecimento e, conhecer significa extrair características dos padrões a serem processados. Portanto, o objetivo da segmentação/extração do logotipo é distingui-lo dos outros elementos existentes na imagem, para submetê-lo posteriormente a uma fase de reconhecimento.

As perspectivas futuras são as seguintes:

- A aplicação ou o desenvolvimento de um método de reconhecimento de logotipos.
- Processamento das duas tarefas de interesse no domínio reconhecimento de logotipo:
 1. Dado um documento que contém um logotipo, classificá-lo como pertencente a um conjunto finito de logotipos, ou concluir que ele não está presente na base de dados;
 2. Dado um logotipo representativo (conhecido ou desconhecido), indexá-lo à base de dados de documentos e extrair todos os documentos que contenham este logotipo.
- Melhorias na metodologia de segmentação proposta para tentar reduzir a influência que os componentes manuscritos exercem sobre a performance do método.

Referências Bibliográficas

- [AMINI et al., 1990] Amini A., Weymouth T., Jain R., **Using Dynamic Programming for Solving Variational Problems in Vision**, IEEE Trans. Pattern Anal. Machine Intell, vol.12, nº9, pp. 855-867, Sep, 1990.
- [ANDRADE et al., 1994] Andrade M.C., Peixoto F. C. e Araújo A. A., **Segmentação de Imagens Através de Rede Neuronal por Satisfação de Restrição em Ambiente Paralelo**, 7th Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens, pp. 47-52, 1994.
- [BABAGUCHI et al., 1990] Babaguchi K., Yamada K., Kise K. and Tezuka T., **Connectionist Model Binarization**, Proc. 10th ICPR, pp. 51-56, 1990.
- [BAMFORD & LOVELL, 1998] Bamford P., Lovell B., **Bayesian Analysis of Cell Nucleus Segmentation by a Viterbi Search Based Active Contour**, Proceeding of the 14th International Conference on Pattern Recognition, ICPR'98, Brisbane, Australia, August 16-20, pp. 133-135, 1998.
- [BANCO, 1983] Banco Central de Brasil, **Manual de Normas e Instrução**, 1993.
- [BANON & BARRERA, 1994] Banon Gerald J. and Barrera J., **Bases da Morfologia Matemática para a Análise de Imagens Binárias**, 9th Escola de Computação, 230p, Recife, 1994.

- [BERSEN, 1986] Bersen J., **Dynamic Thresholding of Gray-Level Images**, Proc. 8th Int'l Conf. on Pattern Recognition, pp. 1251-1255, Paris, France, 1986.
- [BEZDEK & PAL, 1992] Bezdek J. C. and Pal S. K., **Fuzzy Models for Pattern Recognition: Methods that Search for Structures in Data**, IEEE Press, New York, 1992.
- [BEZDEK, 1981] Bezdek J. C., **Pattern Recognition with Fuzzy Objective Function Algorithms**, Plenum Press, New York, 1981.
- [BLANZ & GISH, 1990] Blanz W. E. and Gish S. L., **A Connectionist Classifier Architecture Applied to Image Segmentation**. Proc. 10th ICPR, pp. 279-277, 1990.
- [BRITTO, 1996] Britto Alceu S. J., **Segmentação por Cor Utilizando a Transformada de Karhunen-Loève**. Dissertação de Mestrado, Centro Federal Tecnológico do Paraná – Curso de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Curitiba, 1996.
- [CANNY, 1986] Canny J. F., **A Computational Approach to Edge Detection**, IEEE Trans. Pattern Analysis Mach. Intell. 8, pp. 679-698, 1986.
- [CESARINI et al., 1997] Cesarini F., Francesconi E., Gori M., Marimai S., Sheng J. Q. and Soda G., **A Neural-Based Architecture for Spot-Noisy Logo Recognition**. Proceeding of IV International Conference on Document Analysis and Recognition, pp. 175-179, Ulm, 1997.
- [CHALANA et al., 1996] Chalana V., Linker D. T., Haynor D. R., Kim Y., **A Multiple Active Contour Model for Cardiac Boundary Detection on Echocardiographic Sequences**, IEEE Trans. Medical Imaging, vol.15, n°3, pp. 290-298, June 1996.

- [CHEN & WANG, 1997] Chen L.H. and Wnag J. Y., **A System for Extracting and Recognizing Numeral Strings on Maps**, Proceeding of 4th International Conference on Document Analysis and Recognition, vol.1, pp.337-341, 1997.
- [CHEN et al., 1991] Chen C. T. et al., **Medical Image Segmentation by a Constraint Satisfaction Neural Network**, IEEE Transactions on Nuclear Science, vol.28, n°2, pp. 678-686, 1991.
- [COCQUEREZ & PHILIPP, 1995] Cocquerez J. P. et S. Philipp, **Analyse d'images: filtrage et segmentation**, 457p., editora Masson, 1995.
- [COHEN, 1991] Cohen L., **Note on Active Contour Models and Ballons**, CVGIP: Image Understanding, vol. 53, n°2, pp. 211-218, 1991.
- [COVELLO, 1996] Covello S. C., **Prática de Cheque**, Livraria e Editora Universitária de Direito, São Paulo, 1996.
- [CRUZ, 1998] Cruz C. A., **Unsupervised Texture Segmentation Using Stochastic Version of EM Algorithm and Data Fusion**, pp 1005-1009, 1998.
- [DAVATZIKOS & PRINCE, 1995] Davatzikos C. A., Prince, J. L., **An Active Contour Model for Mapping the Cortex**, IEEE Trans. Medical Imaging, vol.14, n°1, pp. 65-80, March 1995.
- [DAVIS, 1975] Davis L. S., **A survey of Edge Detection Techniques**, Comput. Graphics Image Process. 4, pp. 248-270, 1975.
- [DERAVI & PAL, 1983] Deravi F. and Pal S.K., **Gray Level Thresholding Using Second-Order Statistics**, Pattern Recognition Lett.1, pp.417-422, 1983.

- [DOERMANN et al., 1993] Doermann David S., Rivlin E. and Weiss I. **Logo Recognition Using Geometric Invariants**, Technical Report CS-TR-3145, University of Maryland College Park, MD, pp. 894-897, 1993.
- [ERTHAL et.al, 1991] Erthal Guaraci J. et.al, **Um Sistema de Segmentação e Classificação de Imagens de Satélite**, 4th Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens (SIBGRAP '91), pp.237-240, 1991.
- [EVIATAR & SOMORJAI, 1996] Eviatar H. and Somorjai R.L., **A Fast Simple Contour Algorithm for Biomedical Images**, in Pattern Recognition Letters, 17, pp. 969-974, 1996.
- [FACON, 1996] Facon J., **Mathematical Morphology: Theory and Examples**, in Portuguese, Curitiba, Brazil, October, 1996.
- [FERNANDES, 1997] Fernandes Carlos R. C., **Segmentação Morfológica de Componentes de Cartas Náuticas**, Dissertação de Mestrado, Universidade Estadual de Campinas, Instituto de Computação, 1997.
- [GHOSH et al., 1991] Ghosh A., Pal N. R. and Pal S. K., **Object Extration Using a Self Organization Neural Network**, Intelligent Robotics, M. Vidyasagar and M. Trivedi, eds, pp. 686-697, Tata McGraw-Hill, New Delhi, 1991.
- [GHOSH et al., 1992] Ghosh A., Pal N. R. and Pal S. K., **Object Background Classification Using Hopfield Type Neural Network**, Int. J. Pattern Recognition Artif. Intell. 6(5), pp. 989-1008, 1992.
- [GHOSH et al., 1993] Ghosh A., Pal N. R. and Pal S. K., **Self-Organization for Object Extration Using Multilayer Neural Networks and Fuzziness Measure**, IEEE Trans. Fuzzy Sust. 1(1), pp. 54-68, 1993.

- [GOKMEN & LI, 1990] Gokmen M. and Li C. C., **Edge Detection with Iteratively Refined Regularization**, Proc. 10th ICPR, pp. 690-693, 1990.
- [GOLTSEV, 1996] Goltsev A., **Na Assembly Neural Network for Texture Segmentation**, Neural Networks, vol.9, n^o4, pp. 643-653, 1996.
- [GONÇALVES, 1998] Gonçalves Paulo Jorge S., **Extração e Medição dos Contornos da Válvula Mitral Utilizando Modelos Activos de Contornos**, Dissertação de Mestrado em Engenharia Mecânica – Perfil de Sistemas, 1998.
- [GORMAN & KASTURI, 1995] Gorman Lawrence O' and Kasturi R., **Document Image Analysis**, IEEE Computer Society Press, The Institute of electrical and Etelectronics Engineers, INC, Califórnia, pp.2, 1995.
- [GORMAN, 1988] Gorman L. O', **Curvilinear Features Detection from Curvature Estimation**, Proc. 9th Int'l Conf. Pattern Recognition, pp. 1116-1119, 1988.
- [HALL et al., 1992] Hall L. O., Bensaid A. M., Clarke L. P., Velthuizen R. P., Silbiger M. and Bezdek J.C., **A Comparison of Neural Network and Fuzzy Clustering Techniques in Segmentation Magnetic Resonance Images of the Brain**, IEEE Trans. Neural Network 3(5), pp. 672-681, 1992.
- [HARALICK & SHAPIRO, 1985] Haralick R. M. and Shapiro L. G., **Image Segmentation Techniques**, vol.29, pp. 100-132, 1985.
- [HINDS et al., 1990] Hinds S. C., Fisher J. L. and Amato D. P. D', **A Document Skew Detection Method Using Run – Length Encoding and the Hough Transform**, Proc. 10th Int'l Conf. Pattern Recognition, pp. 464-468, 1990.

- [HUNTSBERGER et al., 1985] Huntsberger C. L., Jacobs C. L. and Cannon R. L., **Iterative Fuzzy Image Segmentation**, Pattern Recognition 18, pp. 131-138, 1985.
- [JAIN & BHATTACHARJEE, 1992] Jain Anil K. and Bhattacharjee S., **Text Segmentation Using Gabor Filters for Automatic Document Processing**, Machine Vision and Applications, vol.5, pp 169-184, 1992.
- [KAPUR et al., 1985] Kapur N. J., Sahoo P.K. and Wong A. K. C., **A New Method for Gray-Level Picture Thresholding**, Computer Vision, Graphics and Image Processing, vol.29, pp. 273-285, 1985.
- [KASS et al., 1987] Kass M., Witkin A., Terzopoulos D., **Snakes: Active Contour Models**, First Int. Conf. on Comp. Vision, London, pp. 259-268, June, 1987.
- [KELLER et al., 1986] Keller J., Qiu H. and Tahani H., **The Fuzzy Integral in Image Segmentation**, Proc. NAFIPS – 86. New Orleans, Louisiana, pp. 324-338, June, 1986.
- [KHOROS] <http://www.icaen.uiowa.edu/~dip/KHOROS/khoros.html>
- [KIM Y. S. & KIM W. Y., 1998] Kim Y. S. and Kim W. Y., **Content-Based Trademark Retrieval System Using a Visually Salient Features**, Image and Vision Computing, pp. 931-939, 1998.
- [KOERICH, 1997] Koerich A. L., **Uma proposta para o Processamento de Imagens de Cheques Bancários: Extração e Armazenamento de Informações**, Dissertação de Msc. na Universidade Estadual de Campinas, SP, Fevereiro, 1997.

- [KRUIZINGA & PETKOV, 1998] Kruizinga P. and Petkov N., **Grating Cell Operator Features for Oriented Texture Segmentation**, Proceeding of the 14th International Conference on Pattern Recognition, ICPR'98, Brisbane, Australia, August 16-20, pp 1010-1014, 1998.
- [LEVINE & NAZIF, 1985] Levine M. D. and Nazif A. M., **Dynamic Measurement of Computer Generated Image Segmentation**, IEEE Trans. Pattern Analysis Mach. Intell. 7, pp.155-164, 1985.
- [MATHERON, 1975] Matheron G., **Random Sets and Integral Geometry**, J. Wiley, New York, 1975.
- [MATRAKAS & BORTOLOZZI, 1998] Matrakas M.D. and Bortolozzi F., **Projeto de uma Biblioteca para Processamento de Imagens de Documentos Orientada a Objetos**, Seminário Internacional de Gestão de Documentos (ISDM'98), pp.236-242, 1998.
- [MATTANA, 1999] Mattana M. F., **Avaliação da Qualidade da Segmentação por Binarização de Cheques Bancários a partir do Processo de Reconhecimento**. Dissertação de Mestrado, Pontifícia Universidade Católica do Paraná, Curitiba, 1999.
- [MEDIONI & YASUMOTO, 1987] Medioni G., Yasumoto Y., **Corner Detection and Curve Representation Using Cubic B-Splines**, Computer Vision, Graphics and Image Processing, vol.39, pp. 267-278, 1987.
- [MOREIRA & COSTA, 1996] Moreira J. and Costa Fabiano F., **Neural-Based Color Image Segmentation and Classification Using Self-Organizing Maps**, 4th SIBGRAPI, pp 47-54, 1996.

- [NETO, 1998] Neto Hugo V., **Reconhecimento Automático de Impressões Digitais Utilizando Wavelets e Redes Neurais Artificiais**, Dissertação de Mestrado, Centro Federal de Educação Tecnológica do Paraná, Curso de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Curitiba, 1998.
- [OHTA & SAKAI, 1980] Ohta Y. and Sakai T., **Color Information for Region Segmentation**, Comput. Graphcs Image Process. 13, pp. 224-241, 1980.
- [OKADA & MALAYAPPAN, 1997] Okada M. and Malayappan S., **Extraction of User Entered Components from a Personal Bankcheck Using Morphological Subtraction**, International Journal of Pattern Recognition and Artificial Intelligence, vol.11 n°5, pp. 699-715, 1997.
- [OTSU, 1979] Otsu N., **A Threshold Selection Method from Gray-Level Histograms**, IEEE Trans. Sys. Man and Cybernnetics, v. SMC9, n°1, pp. 62-66, 1979.
- [PAL & KING, 1980] Pal S. K. and King R. A., **Image Enhancement Using Fuzzy Sets**, Electron. Lett. 16, pp. 376-378, 1980.
- [PAL & ROSENFELD, 1988] Pal S. K. and Rosenfeld A., **Image Enhancement and Thresholding by Optimization of Fuzzy Compactness**, Pattern Recognition Lett. 7, pp. 77-86, 1988.
- [PAL N. R. & PAL S. K., 1989] Pal N. R. and Pal S. K., **Some Information Measures on Fuzzy Sets and their Applications to Image Processing**, Proc. Naconecs-89, pp. 94-96, Tata McGraw-Hill, New Delhi, 1989.

- [PAL N. R. & PAL S. K., 1993] Pal Nikhil R. and Pal Sankar k., **A Review on Image Segmentation Techniques Pattern Recognition**, vol.26 n°9, pp. 1277-1294, 1993.
- [PAL, 1990] Pal N. P., **On Image Information Measures and Object Extraction**, Ph. D. Dissertation, Indian Statistical Institute, Calcutta, 1990.
- [PATEL, 1996] Patel D., **Page Segmentation for Document Image Analysis using a Neural Networks**, Optical Engeneering, vol.35 n°7, pp. 1854-1861, 1996.
- [PELI & MALAH, 1982] Peli T. and Malah D., **A Study of Edge Detection Algorithms**, Comput. Graphics Image Process, pp. 1-21, 1982.
- [PUN, 1981] Pun T., **Entropic Thresholding – The New Approach**, Computer Graphics and Image Processing, vol.16, pp.210-239, 1981.
- [REICHENBACH et al., 1990] Reichenbach S. E., Park S. K. and Gartenberg R. A., **Optimal Small Kernels for Edge Detection**, Proc. 10th ICPR, pp. 57-63, 1990.
- [RUZON, 1997] Ruzon Mark A., **Texture Segmentation. An Introduction Primer**, http://robotics.stanford.edu/~ruzon/text_seg/
- [SAHOO et al., 1988] Sahoo P. K., Soltani S. and Wong A. K. C., **A Survey of Thresholding Techniques**, Computer Vision, Graphics and Image Processing, vol.41, pp 233-260, 1988.
- [SANTOS, 1997] Santos José Eduardo B., **Estudo sobre Métodos e Técnicas para a Segmentação de Imagens de Cheques Bancários**, Dissertação de Mestrado em Ciências, Centro Federal de Educação Tecnológica do Paraná, 1997.

- [SERRA, 1982] Serra J., **Image Analysis and Mathematical Morphology**. Academic Press, London, 1982.
- [SOFFER & SAMET, 1998] Soffer A. and Samet H., **Using Negative Shape Features for Logo Similarity Matching**. Proceeding of the 14th International Conference on Pattern Recognition, volume I, pp. 571-573, Brisbane, Australia, August, 1998.
- [SPITZ, 1996] Spitz A. L., **Logotype Detection in Compressed Images Using Alignment Signatures**, Fifth Annual Symposium on Document Analysis and Information Retrieval, April 15-17, University of Nevada – Las Vegas, 1996.
- [SRINIVAS & PATNAIK, 1994] Srinivas M. and Patnaik L. M., **Genetic Algorithms: A Survey**, Computer, vol.27, n°6, pp. 17-26, 1994.
- [SUDA et al., 1997] Suda P.; Bridoux C. and Manderlechner G., **Logo and Word Matching Using a General Approach to Signal Registration.**, 4th International Conference Documents Analysis and Recognition (ICDAR'97), pp. 61-65, Paris, 1997.
- [TANG et al., 1997] Tang Yuan Y., Ma H., Liu J., Li Bing F. and Xi D., **Multiresolution Analysis in Extraction of Reference Lines from Documents with Gray Level Background**, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.19, n°8, August, 1997.
- [TRIVEDI & BEZDEK, 1986] Trivedi M. and Bezdek J. C., **Low Level Segmentation of Aerial Images with Fuzzy Clustering**, IEEE Trans. Syst. Man Cybern. 16(4), pp. 589-598, 1986.

- [VISA, 1998] Visa A., **A Genetic Algorithm Based Method to Improve Image Segmentation**, Proceeding of the 14th International Conference on Pattern Recognition, ICPR'98, Brisbane, Australia, August 16-20, pp. 1015-1017, 1998.
- [WESZKA & ROSENFELD, 1978] Weszka J. S. and Rosenfeld A., **Threshold Evaluation Techniques**, IEEE Trans. Syst. Man Cybern. 8, pp. 622-629, 1978.
- [WILLIAMS & ALDER, 1996] Williams P.S. and Alder M.D., **Generic Texture Analysis Applied to Newspaper Segmentation**, <http://odin.cc.uwa.edu.au/~williams/research/papers/icnn96/icnn96.html>
- [WILLIAMS & SHAH, 1992] Williams D., Shah M., **A Fast Algorithm for Active Contours and Curvature Estimation**, CVGIP: Image Understanding, vol.55, n°1, pp. 14-26, 1992.
- [WONG & SAHOO, 1989] Wong A.K.C. and Sahoo P.K., **A Gray Level Threshold Selection Method Based on Maximum Entropy Principle**, IEEE Trans. Syst. Man Cybern. 19, pp. 866-871, 1989.
- [XIE & BEDROSIAN, 1988] Xie W. X. and Bedrosian S. D., **Experimentally Derived Fuzzy Membership Function for Gray Level Images**, J. Franklin Inst. 325, pp. 154-164, 1988.
- [ZADEH] <http://theory.lcs.mit.edu/~iandc/Authors/zadehla.html>