



MARCO FLÁVIO MATTANA

AVALIAÇÃO POR RECONHECIMENTO DA QUALIDADE DA SEGMENTAÇÃO POR BINARIZAÇÃO DE CHEQUES BANCÁRIOS

Dissertação apresentada ao Curso de Mestrado em Informática Aplicada do Departamento de informática da Pontifícia Universidade Católica do Paraná.

Área de Atuação: Processamento e Análise de Imagens
Orientador: Prof. Dr. 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 002

Aos 08 dias do mês de fevereiro de 1999, realizou-se a sessão pública de defesa de dissertação **“Avaliação da Qualidade da Segmentação por Binarização de Cheques Bancários a Partir do Processo de Reconhecimento”**, apresentada por Marco Flávio Mattana, ano de ingresso 1996, 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 (PUC-PR)	
Prof. Dr. Flávio Bortolozzi (PUC-PR)	
Prof. Dr. Nelson D. A. Mascarenhas (UF – São Carlos)	
Prof. Dr. Roberto Lotufo (UNICAMP-SP)	

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 (PUC-PR)	A
Prof. Dr. Flávio Bortolozzi (PUC-PR)	A
Prof. Dr. Nelson D. A. Mascarenhas (UF – São Carlos)	A
Prof. Dr. Roberto Lotufo (UNICAMP-SP)	A
Conceito Final	A

Observações da Banca Examinadora,

O CANDIDATO DEVERÁ APRESENTAR AS CORREÇÕES
APONTADAS NECESSÁRIAS PELA BANCA.
O CANDIDATO TEVE TRABALHO PUBLICADO EM CONGRESSOS

Prof. Robert C. Burnett

Coordenador do Programa de Pós-Graduação em Informática Aplicada-PUC-PR

Ficha Catalográfica

Mattana Marco Flávio, Avaliação por Reconhecimento da Qualidade da Segmentação por Binarização de Cheques Bancários, Curitiba 1999
102p

Dissertação (Mestrado) Pontifícia Universidade Católica do Paraná.

1. Processamento e Análise de imagens 2. Processamento de Documentos - I. Pontifícia Universidade Católica do Paraná II. Centro de Ciências Exatas e de Tecnologia III. Curso de Mestrado em Informática Aplicada

Dedico aos meus pais

AGRADECIMENTOS

Ao orientador deste trabalho e grande amigo, Prof. Dr. Jacques Facon, pela condução do projeto, pelo incentivo, apoio e sobretudo, pela compreensão em alguns momentos de dificuldade;

Ao Prof. Dr. Flávio Bortolozzi pela oportunidade e apoio ao projeto;

Ao Prof. Dr. Robert C. Burnett, Coordenador do PPGIA, pelas palavras de incentivo e amizade;

A todos os Professores do PPGIA pelas palavras de incentivo e apoio;

A Pontifícia Universidade Católica do Paraná e ao Programa de Pós Graduação em Informática Aplicada, pela oportunidade concedida;

Ao Sr. Ary Paulino Chemin Jr, pela amizade e cessão dos cheques bancários que formaram grande parte a base de dados do projeto;

Ao Prof. Alceu de Souza Britto Jr. por ter gentilmente cedido o uso do Software PDI-MAGEM;

A minha esposa Simone, pela compreensão, amor e carinho e pelos sábados e domingos em que me deixou trabalhar "sossegado";

Aos meus filhos Rafael e Diogo que, com argumentos irrepreensíveis, me fizeram parar um pouco e relaxar para jogar no computador;

A todos que contribuíram de forma direta ou indireta para este trabalho;

E a Deus por tudo.

Sumário

LISTA DE FIGURAS	ix
LISTA DE TABELAS	xiii
ABREVIATURAS	xv
LISTA DE FÓRMULAS	xvi
RESUMO	xviii
ABSTRACT	xix
1 INTRODUÇÃO	1
2 ESTRUTURA DO CHEQUE BANCÁRIO BRASILEIRO	4
2.1 INTRODUÇÃO	4
2.2 DIMENSÕES DE CHEQUES BANCÁRIOS BRASILEIROS	4
2.3 PREENCHIMENTO DE CHEQUES BANCÁRIOS	7
2.4 FUNDO ARTÍSTICO	8
2.5 OUTRAS CARACTERÍSTICAS	16
2.5.1 Faixa de segurança impressa	16
2.5.2 Impressão com tinta ultravioleta	17
2.5.3 Outras características	17
2.6 BASE DE DADOS	18
2.6.1 Introdução	18
2.6.2 Aquisição	19
2.6.3 Meio de aquisição	20
2.6.4 Composição	20

3	A COR	24
3.1	INTRODUÇÃO	24
3.2	SISTEMAS DE CORES	25
3.2.1	Introdução	25
3.2.2	Modelos <i>RGB</i> e <i>CMY</i>	25
3.2.3	Modelo <i>HSI</i>	30
3.3	SEGMENTAÇÃO	31
3.3.1	Introdução	31
3.3.2	Segmentação pelo modelo <i>HSI</i> aplicado a cheque bancários	34
4	METODOLOGIA DE SEGMENTAÇÃO POR BINARIZAÇÃO	36
4.1	INTRODUÇÃO	36
4.2	CONSIDERAÇÕES SOBRE A BINARIZAÇÃO	37
4.3	HISTOGRAMA	38
4.4	TÉCNICAS DE BINARIZAÇÃO GLOBAIS VERSUS LOCAIS ADAPTATIVAS	39
4.5	MÉTODOS GLOBAIS	41
4.5.1	Método por Abordagem Fuzzy	42
4.5.1.1	Método de Huang	43
4.5.1.2	Método de Yager	43
4.5.2	Método por Dois Picos	44
4.5.3	Método por Entropia de Pun	44
4.5.4	Método por Entropia de Kapur	45
4.5.5	Método por Entropia de Johannsen	46
4.5.6	Método por Entropia de Li-Lee	46
4.5.7	Método de Kittler-Illingworth	47
4.5.8	Método de Otsu	48
4.5.9	Método por Pixels de Borda	49
4.5.10	Método por Porcentagem	49
4.5.11	Método por Seleção interativa de Thrussel	50
4.5.12	Método por Valor Médio	50
4.6	MÉTODOS LOCAIS OU ADAPTATIVOS	51
4.6.1	Introdução	51
4.6.2	Método de Bernsen	51
4.6.3	Método por Médias Móveis	52

4.6.4	Método por Relaxação de Peleg	53
5	ESTRATÉGIA DE AVALIAÇÃO DE QUALIDADE ADOTADA E RESULTADOS	56
5.1	INTRODUÇÃO	56
5.2	QUALIDADE DA IMAGEM	56
5.3	METODOLOGIA	58
5.4	RESULTADOS DOS MÉTODOS GLOBAIS	63
5.4.1	Método por Valor Médio	63
5.4.2	Métodos por Porcentagem e por Pixels de Borda	65
5.4.3	Método por Dois Picos	65
5.4.4	Método de Kittler-Illingworth	66
5.4.5	Método de Otsu	68
5.4.6	Método por Fuzzy Set - abordagem de Huang	71
5.4.7	Método por Fuzzy Set - Abordagem de Yager	72
5.4.8	Método por Seleção iterativa de Thrussel	75
5.4.9	Método por Entropia de Pun	78
5.4.10	Método por Entropia de Kapur	79
5.4.11	Método por Entropia de Johannsen	82
5.4.12	Método por Entropia de Li-Lee	84
5.5	RESULTADOS DOS MÉTODOS LOCAIS ADAPTATIVOS	87
5.5.1	Método de Bernsen	87
5.5.2	Método por Médias Móveis	89
5.5.3	Método por Relaxação	92
5.6	TAXAS DE RECONHECIMENTO POR BANCO VERSUS MÉTODOS GLOBAIS E LOCAIS	92
5.7	TAXAS GLOBAIS DE RECONHECIMENTO	92
5.8	CONCLUSÕES SOBRE AS TAXAS FINAIS DE RECONHECIMENTO	95
6	CONCLUSÕES E PERSPECTIVAS	97
	REFERÊNCIAS	100

Lista de Figuras

2.1	Exemplo de limites de tamanho máximo, médio e mínimo para 3 diferentes cheques bancários.	5
2.2	Modelo para a distribuição de informações nos cheques bancários.	6
2.3	Exemplo de um cheque mostrando os campos e seu preenchimento.	7
2.4	Variações de impressão, tipologia e corpo do cabeçalho do campo valor por extenso.	9
2.5	Variações de impressão das expressões "ou à sua ordem" e "e centavos acima".	10
2.6	Tipos de fundos artísticos.	11
2.7	Variações na tonalidade de impressão.	12
2.8	Estratificação de classes de clientes por expressões adicionadas ao cheque.	13
2.9	Estratificação de classes de clientes por mudança da cor do fundo artístico e inserção de expressão.	14
2.10	Estratificação de classes de clientes por mudança da cor do fundo artístico sem inserção de expressão.	15
2.11	Impressão randômica de segurança a fim de evitar fraude por estelionato.	16
2.12	Impressão de segurança em ultravioleta vista sob o canal de matiz	17
2.13	Exemplo de um cheque avulso com preenchimento manual/mecânico	18
2.14	Cheques com variação de tonalidades devido a aquisição em tipos diferentes de scanners	21
3.1	Espectro eletromagnético da luz visível.	25
3.2	Representação da absorção e reflexão da luz sobre objetos cromáticos e acromáticos.	26
3.3	Sistema aditivo de cores ou <i>RGB</i>	27
3.4	Sistema subtrativo de cores ou <i>CMY</i>	28
3.5	a) Espaço <i>RGB</i> e b) Espaço <i>CMY</i> segundo distribuição espacial.	29
3.6	Espaço <i>HSI</i>	31
3.7	Segmentação de uma imagem em canais <i>RGB</i>	32
3.8	Segmentação de uma imagem em canais <i>HSI</i>	33

3.9	Segmentação de um cheque em canais <i>HSI</i>	35
4.1	Histograma de uma imagem em níveis de cinza	39
4.2	Histograma típico com divisões de vale e picos	40
4.3	Histograma complexo	40
5.1	Cheque bancário binarizado por dois métodos diferentes	57
5.2	Cheque bancário do Banco do Brasil	59
5.3	Cheque bancário do Banco Banestado S.A.	60
5.4	Cheque do Banco Itaú S.A.: a) colorido b) canal de Intensidade	61
5.5	Visualização no histograma do valor de limiar obtido pelo método por valor médio.	63
5.6	Resultado da binarização do cheque bancário pelo método por valor médio.	64
5.7	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por valor médio; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado.	64
5.8	Visualização no histograma do valor de limiar obtido pelo método por dois picos.	65
5.9	Resultado da binarização do cheque bancário pelo método por dois picos.	65
5.10	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por dois picos; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado.	66
5.11	Visualização no histograma do valor de limiar obtido pelo método de Kittler-Illingworth.	66
5.12	Resultado da binarização do cheque bancário pelo método de Kittler-Illingworth.	67
5.13	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método de Kittler-Illingworth; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado.	67
5.14	Visualização no histograma do valor de limiar obtido pelo método de Otsu.	68
5.15	Resultado da binarização do cheque bancário pelo método de Otsu.	69
5.16	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método de Otsu; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado.	69
5.17	Visualização no histograma do valor de limiar obtido pelo método por Fuzzy Set - abordagem de Huang.	71
5.18	Resultado da binarização do cheque bancário pelo método por Fuzzy Set - abordagem de Huang.	71
5.19	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por Fuzzy Set - abordagem de Huang; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado.	72

5.20	Visualização no histograma do valor de limiar obtido pelo método por Fuzzy Set - abordagem de Yager.	72
5.21	Resultado da binarização do cheque bancário pelo método por Fuzzy Set - abordagem de Yager.	73
5.22	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por Fuzzy Set - abordagem de Yager; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado.	73
5.23	Visualização no histograma do valor de limiar obtido pelo método por seleção iterativa de Thrussel.	75
5.24	Resultado da binarização do cheque bancário pelo método por seleção iterativa de Thrussel.	76
5.25	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por seleção iterativa de Thrussel; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado.	76
5.26	Visualização no histograma do valor de limiar obtido pelo método por entropia de Pun.	78
5.27	Resultado da binarização do cheque bancário pelo método por entropia de Pun. . .	78
5.28	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por entropia de Pun; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado. . . .	79
5.29	Visualização no histograma do valor de limiar obtido pelo método por entropia de Kapur.	79
5.30	Resultado da binarização do cheque bancário pelo método por entropia de Kapur. .	80
5.31	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por entropia de Kapur; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado. . .	80
5.32	Visualização no histograma do valor de limiar obtido pelo método por entropia de Johannsen.	82
5.33	Resultado da binarização do cheque bancário pelo método por entropia de Johannsen. .	82
5.34	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por entropia de Johannsen; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado. .	83
5.35	Visualização no histograma do valor de limiar obtido pelo método por entropia de Li-Lee.	85
5.36	Resultado da binarização do cheque bancário pelo método por entropia de Li-Lee. .	85

5.37	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por entropia de Li-Lee; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado. . .	85
5.38	Resultado da binarização do cheque bancário pelo método de Bernsen.	87
5.39	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método de Bernsen; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado.	88
5.40	Resultado da binarização do cheque bancário pelo método por médias móveis. . . .	90
5.41	Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por médias móveis; b) do reconhecimento por <i>OCR</i> desse cabeçalho binarizado. . . .	90
5.42	Resultado da binarização do cheque bancário pelo método por relaxação.	92
6.1	Resultados da segmentação por binarização: a) cheque bancário do banco Baneste preenchido com caracteres manuscritos, b) método de Yager, c) método de Otsu, d) método de Bernsen	99

Lista de Tabelas

2.1	Normas de dimensões de um cheque bancário	5
2.2	Bases de dados de documentos existentes	19
2.3	Descrição da base de dados de cheques sem preenchimento	21
2.4	Descrição da base de dados de cheques com preenchimento mecânico	22
2.5	Descrição da base de dados de cheques com preenchimento manuscrito	22
5.1	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método de Otsu para cheques não preenchidos.	70
5.2	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método de Otsu para cheques preenchidos a máquina.	70
5.3	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por Fuzzy Set - abordagem de Yager para cheques não preenchidos.	74
5.4	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por Fuzzy Set - abordagem de Yager para cheques preenchidos a máquina.	74
5.5	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por seleção iterativa de Thrussel para cheques não preenchidos.	77
5.6	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por seleção iterativa de Thrussel para cheques preenchidos a máquina.	77
5.7	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por entropia de Kapur para cheques não preenchidos.	81
5.8	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por entropia de Kapur para cheques preenchidos a máquina.	81
5.9	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por entropia de Johannsen para cheques não preenchidos.	83
5.10	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por entropia de Johannsen para cheques preenchidos a máquina.	84

5.11	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por entropia de Li-Lee para cheques não preenchidos.	86
5.12	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por entropia de Li-Lee para cheques preenchidos a máquina.	86
5.13	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método de Bernsen para cheques não preenchidos.	88
5.14	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método de Bernsen para cheques preenchidos a máquina.	89
5.15	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por médias móveis para cheques não preenchidos.	91
5.16	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus método por médias móveis para cheques preenchidos a máquina.	91
5.17	Taxas de reconhecimento por <i>OCR</i> por banco versus métodos globais e locais para cheques não preenchidos.	93
5.18	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus métodos globais e locais para cheques não preenchidos.	93
5.19	Taxas de reconhecimento por <i>OCR</i> por banco versus métodos globais e locais para cheques preenchidos a máquina.	94
5.20	Taxas de reconhecimento (em %) por <i>OCR</i> por banco versus métodos globais e locais para cheques preenchidos a máquina.	94
5.21	Taxas globais de reconhecimento (em %) por <i>OCR</i> de todos os métodos de segmentação por binarização.	95

Lista de Abreviaturas

CGC	Cadastro Geral de Contribuintes
CMY	do inglês, Cyan, Magenta, Yellow ou Cyan, Magenta e Amarelo
CPF	Cadastro de Pessoas Físicas
DPI	do inglês DOTS PER INCH ou pontos por polegadas
HSI	do inglês Hue, Saturation, Intensity ou Matiz, Saturação e Intensidade
Nm	nanômetros ($10^{-7} cm$)
OCR	do inglês, Optical Character Recognition, ou Reconhecimento Óptico de Caracteres
Pixel	do inglês, Picture Element ou Elemento Gráfico
RGB	do inglês, Red, Green, Blue ou Vermelho, Verde e Azul
TX_{CH}	Taxa de reconhecimento de cada cheque
TX_{GA}	Taxa de reconhecimento global por algoritmo
TX_{MBA}	Taxa média de reconhecimento por banco versus algoritmo

Lista de Fórmulas

A e A^c	plano A e seu plano complementar
$D(Q, P) = - \sum_{k=0}^N q_k \log_2(\frac{q_k}{p_k})$	entropia cruzada entre duas distribuições P e Q (página 46)
$D_p(t) = [\sum_g \mu_g(g) - \overline{\mu_g}(g) ^p]^{1/p}$	valor do fuzziness entre dois planos para o valor de limiar t (página 43)
$E(t) = \frac{1}{nm} \sum_g H_f[\mu_g(g)]h(g)$	valor da Entropia do conjunto Fuzzy (página 43)
g	nível de cinza de um pixel
$h(g)$	histograma para um determinado nível de cinza g
$H_b(t)$	entropia dos pixels pretos para o limiar t (página 44)
$H_f(x) = -x \log(x) - (1-x) \log(1-x)$	função de Shannon para classificação da entropia (página 43)
$H_w(t)$	entropia dos pixels brancos para o limiar t (página 44)
$J(t)$	função critério do Método de Kittler-Illingworth (página 47)
m, n	número de colunas e de linhas de uma imagem
μ_0	média dos níveis de cinza em uma imagem de fundo (páginas 47 e 48)
μ_1	média dos níveis de cinza em uma imagem de primeiro plano (páginas 47 e 48)
μ_i	média de uma classe i
$M_{i+1} = M_i - \frac{M_i}{n} + g_{i+1}$	média móvel no pixel $i + 1$ (página 52)

$\eta(t)$	função critério do Método de Li-Lee (página 46)
$p(x_i)$	probabilidade de ocorrência de um símbolo x_i
$p(g)$	função de probabilidade no nível de cinza g
p_i	probabilidade do pixel no nível de cinza i (página 44)
P_t	probabilidade cumulativa até o nível do valor do limiar t (página 44)
p_i^k, q_i^k	probabilidades do pixel de nível de cinza i pertencer à classe preta ou à classe branca (página 53)
$S_b(t)$	entropia dos pixels pretos para o limiar t (página 46)
$S_w(t)$	entropia dos pixels brancos para o limiar t (página 46)
σ^2	desvio padrão de uma classe
$\sigma_B^2(t)$	variância inter-classes (página 48)
$\sigma_W^2(t)$	variância intra-classes (página 48)

Resumo

Este trabalho descreve uma abordagem de avaliação da qualidade por reconhecimento da segmentação por binarização de cheques bancários. Os cheques são adquiridos num formato colorido e os padrões de cor são decompostos segundo o modelo *HSI*. Vários algoritmos de binarização globais e locais adaptativos são aplicados e a qualidade dos resultados é avaliada por uma metodologia baseada no reconhecimento por *OCR* nas imagens binárias das informações segmentadas. Os melhores resultados foram obtidos com os Métodos de Yager e Bernsen para cheques sem preenchimento e Otsu e Bernsen para cheques com preenchimento mecânico.

Resumo

Este trabalho descreve uma abordagem de avaliação da qualidade por reconhecimento da segmentação por binarização de cheques bancários. Os cheques são adquiridos num formato colorido e os padrões de cor são decompostos segundo o modelo *HSI*. Vários algoritmos de binarização globais e locais adaptativos são aplicados e a qualidade dos resultados é avaliada por uma metodologia baseada no reconhecimento por *OCR* nas imagens binárias das informações segmentadas. Os melhores resultados foram obtidos com os Métodos de Yager e Bernsen para cheques sem preenchimento e Otsu e Bernsen para cheques com preenchimento mecânico.

Abstract

This work describes an approach to quantify the quality of thresholding-based segmentation on Brazilian bank checks. The checks are previously acquired by scanner and soon after, their colored patterns are decomposed according to the *HSI* model. Several thresholding algorithms are then applied and an automatic evaluation of the results is carried out. The evaluation methodology is based on "performance assessment" of the character recognition rate from an *OCR*. The best results were obtained with the Methods of Yager and Bernsen for checks without fill and Otsu and Bernsen for checks with mechanical fill.

Abstract

This work describes an approach to quantify the quality of thresholding-based segmentation on Brazilian bank checks. The checks are previously acquired by scanner and soon after, their colored patterns are decomposed according to the *HSI* model. Several thresholding algorithms are then applied and an automatic evaluation of the results is carried out. The evaluation methodology is based on "performance assessment" of the character recognition rate from an *OCR*. The best results were obtained with the Methods of Yager and Bernsen for checks without fill and Otsu and Bernsen for checks with mechanical fill.

Capítulo 1

INTRODUÇÃO

Durante sua evolução, o homem passou por diversos estágios. O mais importante foi, sem dúvida, a invenção do alfabeto e a posterior utilização da palavra escrita. Decorrente desta descoberta, o registro de observações astronômicas, manifestos de cargas e comércio, códigos de leis ou simplesmente o registro da história passaram a ser transcritos em tábuas de barro, blocos de pedra, papiros, cascas de árvores, pergaminhos e, após a invenção do papel pelos chineses, sobre este tipo de suporte que nos serve até hoje.

Johann Gutemberg descobriu a forma de registrar sobre papel caracteres tipográficos quando inventou sua prensa de tipos. Hoje, nós, homens modernos, dependemos exaustivamente destas descobertas; uma vez que manuseamos diariamente, uma série de documentos sob a forma de tipos impressos sobre papel.

Necessitamos, nesta era de informações, rapidez e agilidade para o manuseio destes tipos de informações que devem possuir um mínimo de qualidade.

Nas áreas comerciais, os principais tipos de informações impressas são documentos como listas, contratos, relatórios, fax, notas fiscais, entre outras. Na área monetária, os três principais documentos são o papel moeda, cartões eletrônicos e os cheques bancários.

Este trabalho restringe-se a este último tipo de documento. Apesar de ser um documento de dimensões pequenas, existem muitas informações importantes impressas, pois um cheque bancário é uma ordem de pagamento de um valor, emitido em favor de uma pessoa ou empresa. É possível verificar que a grande maioria das transações comerciais são efetuadas por meio deste tipo de do-

cumento em detrimento do uso do papel moeda devido à própria segurança do sistema. Assim, o volume de cheques emitidos diariamente no Brasil ultrapassa a quantia de centenas de milhares de unidades que são processados nas Câmaras de Compensação de Cheques. Ainda que seja uma forma segura de pagamento, fraudes nestes tipos de documentos são comuns, se bem que os bancos se esforçam em dificultá-las ao máximo incluindo em seus cheques linhas de segurança, fundos artísticos complexos, sistemas de bloqueio de compensação, etc...

Durante todo o processo de desconto ou pagamento de um cheque bancário, o computador se faz presente com exceção em algumas etapas de procedimentos que ainda são executados sob a forma manual, sem a interferência da informática no processo por não haver possibilidade de o computador ainda interagir. As operações automáticas como leitura ótica dos dados do cliente - agência, número da conta, número do cheque - já se encontram automatizadas. As operações manuais restringem-se à digitação do valor do cheque executando sua entrada no banco de dados gerando a transferência dos fundos entre contas, verificações de preenchimento e comprovação da autenticidade da assinatura. Devido à grande quantidade de cheques, seria desejável que houvesse maior rapidez, agilidade e segurança maior no processamento deste tipo de documento.

As operações manuais que ainda carecem de automatização estão sendo pesquisadas no campo de domínio da computação, notadamente no âmbito do processamento de documentos que envolve técnicas de processamento de imagens e reconhecimento de padrões. Cabe ao processamento digital de imagens, filtrar e extrair as informações relevantes dos documentos; e ao reconhecimento de padrões a comparação destas informações sob a forma de reconhecimento.

O objetivo deste trabalho reside no primeiro campo onde se procura um método a fim de segmentar um cheque com confiabilidade, extraindo suas informações relevantes com qualidade.

A segmentação por binarização de imagens em tons de cinza é o primeiro passo na grande maioria dos sistemas de análise de documentos. A seleção de um método de binarização apropriado que resulte em uma imagem com qualidade é um problema típico enfrentado pelos pesquisadores, pois a avaliação desta qualidade acaba sendo feita por critérios visuais. Como ainda não existe o melhor método de binarização, percebe-se que há bastante espaço para pesquisas neste campo. Um problema que requer mais trabalho é identificar métodos de binarização ou enfoques que melhor trabalham com documentos com características particulares [O'Gorman et al (1995)]. Muitos dos

métodos existentes na bibliografia não foram formulados para cheques bancários em particular, e seus desempenhos não são bem conhecidos. Os cheques bancários possuem características, tais como linhas finas, fundo artístico etc., que favorecem um método sobre outro.

Contornando o critério visual da qualidade, [O'Gorman et al (1995)] explica que a questão de qual método é o melhor para usar, está na realidade relacionada com a questão de como melhor quantificar os resultados da binarização. Portanto, o objetivo procura ensaiar os resultados obtidos pela binarização por meio de um *OCR* (do inglês Optical Character Recognition) a fim de se determinar o melhor método de binarização para cheques bancários ou famílias de cheques. A principal contribuição deste trabalho é indicar um método de binarização aplicado a cheques bancários brasileiros que disponibilize uma imagem final com qualidade, sem perda de características relevantes para o pós processamento desta imagem como extração de dígitos ou caracteres, estudos de identificação de assinaturas ou valores por extenso, por exemplo. A estrutura deste trabalho é descrita a seguir.

O Capítulo 2 analisa os cheques bancários brasileiros, suas características próprias como tamanho, dados impressos, fundos artísticos, dados de preenchimentos além de outras informações relevantes.

No Capítulo 3 encontram-se descritos os sistemas de cores *RGB* e *HSI* bem como a segmentação dos cheques por canais de cores utilizando-se o sistema *HSI* a fim de se obter uma imagem em tons de cinza para a aplicação dos algoritmos.

O Capítulo 4 descreve os algoritmos de binarização globais e locais adaptativos existentes na literatura.

O Capítulo 5 ilustra a aplicação dos algoritmos, seus resultados e a estratégia adotada para avaliar a sua qualidade a partir de imagens de cheques bancários. Para ter uma avaliação rigorosa, as imagens manipuladas não foram submetidas a nenhuma técnica de préprocessamento. Sob a forma de tabelas estão relacionadas as taxas de reconhecimento obtidas pelo ensaio do *OCR* sobre as imagens binarizadas.

Por fim, o capítulo 6 descreve as conclusões da pesquisa e formula perspectivas em termos de segmentação por binarização com os algoritmos globais e locais adaptativos existentes na literatura.

Capítulo 2

ESTRUTURA DO CHEQUE BANCÁRIO BRASILEIRO

2.1 INTRODUÇÃO

A movimentação de fundos bancários pode ser feita por diversas maneiras. A maneira mais comum utilizada são cartões eletrônicos e cheques bancários para pagamento sem o manuseio do papel moeda em si. No âmbito deste trabalho, um cheque representa uma ordem incondicional em direito que é emitida por uma pessoa física ou jurídica em favor de outra pessoa, também física ou jurídica, sendo assinado pela pessoa emitente e exige que o banco pague, à sua apresentação, a quantia em dinheiro referenciada no mesmo [BCB (1983)].

Neste capítulo são descritos os elementos que compõem um cheque bancário brasileiro, dimensões, disposições de seus elementos, fundo artístico e suas variações; além de normas estabelecidas pelo Banco central que são relevantes para este trabalho e para o sistema como um todo no qual se insere este trabalho.

2.2 DIMENSÕES DE CHEQUES BANCÁRIOS BRASILEIROS

O Banco Central é o responsável pela normalização dos elementos que constituem um cheque bancário. Segundo o Manual de Normas de Instruções [BCB (1983)], as dimensões dos cheques bancários devem obedecer as medidas em milímetros descritas na tabela 2.1.

A figura 2.1 mostra os limites de tamanho máximo, médio e mínimo para 3 diferentes

Medida	Comprimento	Largura
Valor Médio	175 mm	80 mm
Valor Mínimo	176 mm	81 mm
Valor Máximo	179 mm	84 mm

Tabela 2.1: Normas de dimensões de um cheque bancário

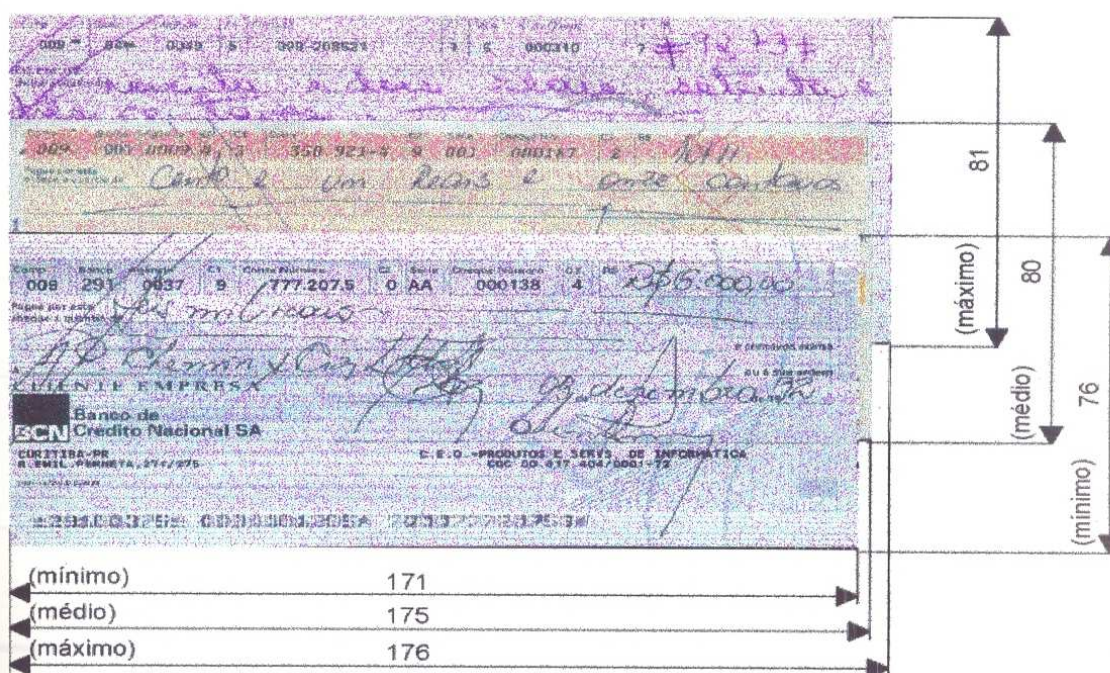


Figura 2.1: Exemplo de limites de tamanho máximo, médio e mínimo para 3 diferentes cheques bancários.

cheques bancários.

O Manual de Normas e Instruções também prevê a localização dos elementos dentro da área do cheque bem como seus campos de preenchimento. As informações relevantes constituem-se de campos pré-impressos 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. Contém ainda informações 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 de emissão, assinatura do cliente, nome, 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 a leitura digital dos dados de número de compensação, número do banco, número da agência, número da conta, número do cheque e dígitos verificadores. Os elementos acima descritos estão exemplificados na figura 2.2.

Número da 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 do CPF/CGC outras informações (Telefone - RG - Conta Conjunta)
Linha MICR CMC - 7		

Figura 2.2: Modelo para a distribuição de informações nos cheques bancários.

2.3 PREENCHIMENTO DE CHEQUES BANCÁRIOS

De acordo com as normas do Banco Central, o preenchimento dos campos pelo cliente pode ser feito à mão, desde que de forma legível, ou por meios mecânicos como impressoras ou máquina de escrever pelo próprio emitente ou por terceiros. O Banco Central indica que o preenchimento deve ser executado nas cores azul ou preta a fim de facilitar o processamento por parte do banco como por exemplo, a microfilmagem dos cheques compensados.

Os campos de preenchimento são os seguintes: valor numérico, valor por extenso, beneficiário, local, data e assinatura do cliente. O lay-out do cheque prevê linhas pré-impressas para delimitação dos campos e orientação do preenchimento. A figura 2.3 mostra os campos e o preenchimento à mão de um cheque.

O cheque é emitido pelo Banco Citibank S.A., Rua Marechal Deodoro, 711, Curitiba, PR. O valor numérico é R\$ 84.23,00 e o valor por extenso é "Oitenta e quatro reais e duas centavos". O beneficiário é "A. P. Almeida e Cia. Ltda." e a data é "08 de abril de 1998". A assinatura é de Marcos Antonio Justino e/ou Wlanete Cassiano de Barros Justino.

Beneficiário			Valor por Extenso			Valor Numérico		
009	745	0010	0	98113011	9	003211	5	R\$ 84.23

Pague por meio de (Oitenta e quatro reais e duas centavos)

A. P. Almeida e Cia. Ltda.

CITIBANK

Banco Citibank S.A.
RUA MARECHAL DEODORO, 711
CURITIBA
CGC-33479023/0017-47 PR

08 de abril de 1998

MARCOS ANTONIO JUSTINO E/OU
WLANETE CASSIANO DE BARROS JUSTINO
CPF-583812039-91

Local e Data

Assinatura

Figura 2.3: Exemplo de um cheque mostrando os campos e seu preenchimento.

O elemento valor numérico, representado pelo símbolo monetário R\$, é destinado à especificação do valor do cheque em algarismos arábicos. Os bancos recomendam o preenchimento deste campo antecedido e precedido por algum caractere que dificulte sua alteração ou rasura como por exemplo, parênteses, asteriscos, cerquilhas, etc.

Impresso, antecedendo o campo de especificação do valor por extenso, encontra-se a frase

"*Pague-se por este cheque a quantia de*" seguida de duas linhas. No preenchimento deste campo é dispensado a especificação do valor do cheque em centavos sendo esta opcional. O Banco Central não admite qualquer outro caracter ou símbolo impresso além das duas linhas.

O campo do beneficiário é iniciado com a preposição "*a*" e finalizada com a expressão ou "*à sua ordem*" ou, se for o caso, "*não à ordem*" e é reservada à indicação do nome do beneficiário do cheque podendo ser uma pessoa física ou jurídica [BCB (1983)]. À direita, abaixo do campo do Beneficiário, são impressas linhas reservadas à indicação do local de emissão do cheque e data. Abaixo deste linha, impressa em igual tamanho da anterior, encontra-se uma linha que delimita o campo de assinatura. Existem variações pequenas no formato de impressão destes campos como mostra a figura 2.4, no tocante ao lay-out de impressão, tipologia e tamanho de corpo.

As variações ainda estão presentes no final da linha de preenchimento do valor extenso onde está presente a expressão "*e centavos acima*" em alguns cheques.

A figura 2.5 mostra as variações do impresso "*ou à sua ordem*" e da expressão "*e centavos acima*".

2.4 FUNDO ARTÍSTICO

O Banco Central prevê a impressão de um fundo artístico para evitar falsificações nos cheques. Este fundo artístico pode conter grafismos de qualquer tipo, desde que não interfiram com a legibilidade dos dados impressos e dados preenchidos pelo cliente. O fundo artístico, normalmente é impresso utilizando-se em média três 3 ou mais tonalidades da mesma cor ou utilizando cores diferentes em degrade suave. Os motivos de impressão variam bastante e podem referir-se a temas regionais ligados à fauna ou flora, grafismos abstratos ou ao logotipo do banco de forma estilizada ou não. A figura 2.6 mostra alguns tipos de fundos artísticos.

A normativa do Banco Central prevê que a impressão do fundo artístico seja composta de 70% de branco e 30% de cor [BCB (1983)], não especificando se o valor de 70% refere-se à tinta branca os 30% restantes ao pigmento colorido para gerar a impressão.

Durante a realização deste trabalho foi possível constatar que este fundo artístico não respeita estritamente a norma, pois existem variações bastante apreciáveis em sua impressão no tocante a densidade do pigmento, pois alguns cheques apresentam seu fundo artístico com tonalidades fortes (como no caso do Banco Itaú SA) e outros com uma tonalidade esmaecida (como no caso do Banco HSBC-Bamerindus). Outro fato observado condiz à partida de impressão dos cheques de um mesmo banco, que pode apresentar variações de tonalidade em sua impressão, talvez por mistura

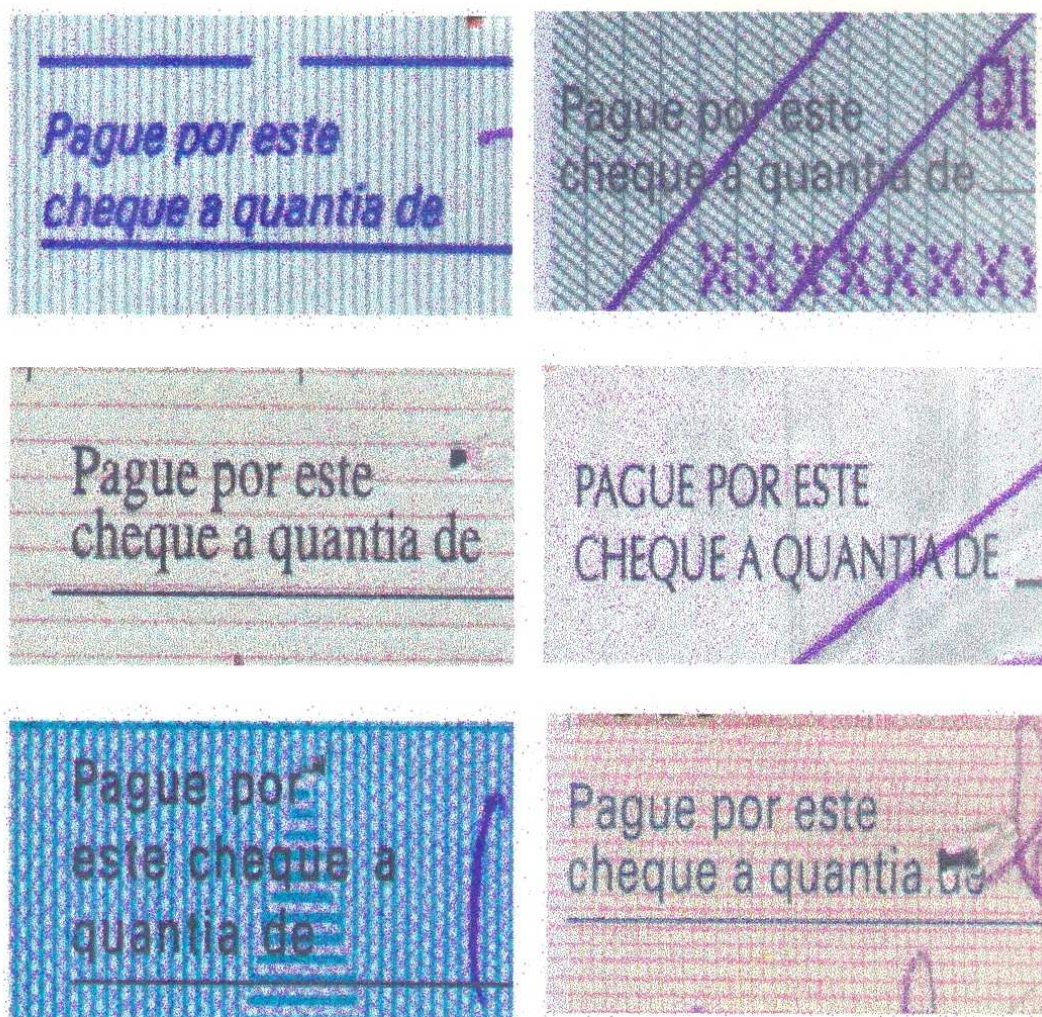


Figura 2.4: Variações de impressão, tipologia e corpo do cabeçalho do campo valor por extenso.

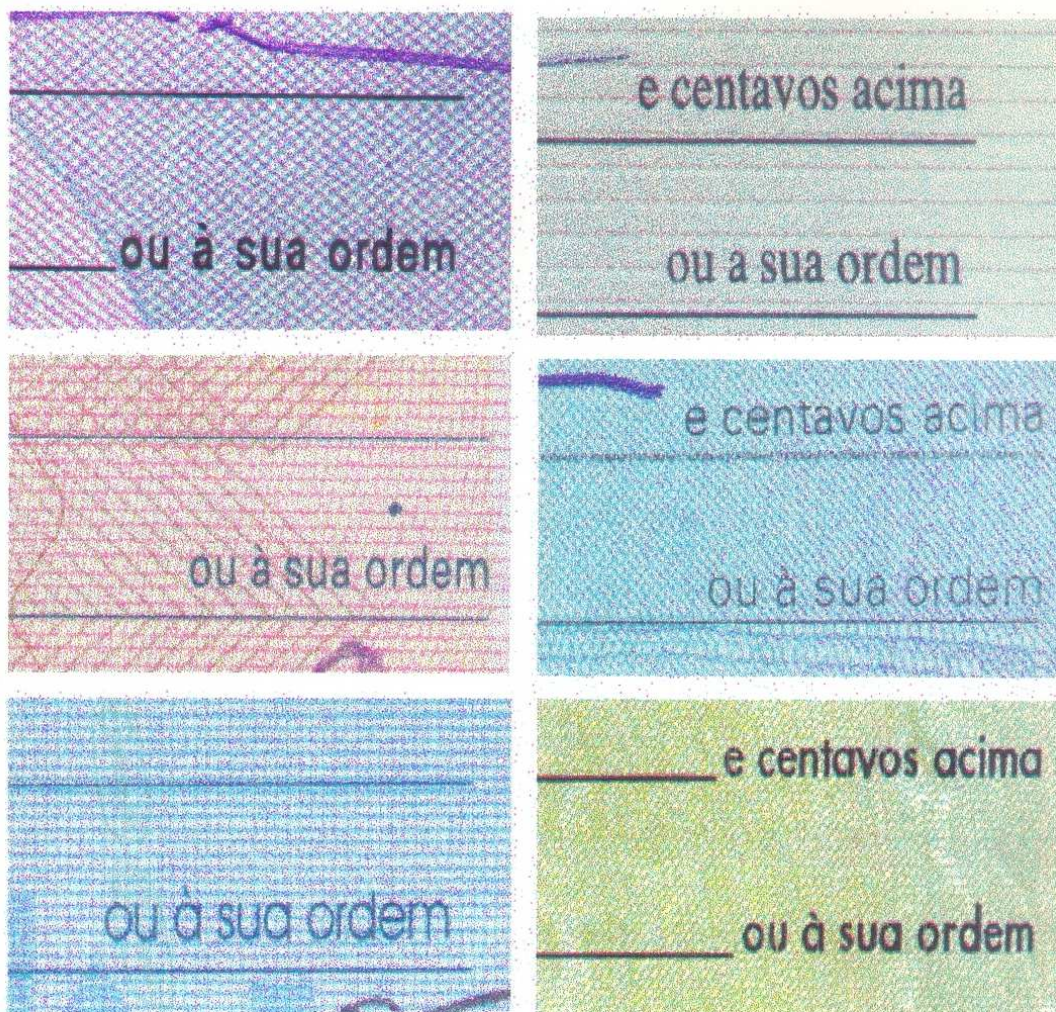


Figura 2.5: Variações de impressão das expressões "ou à sua ordem" e "e centavos acima".



Figura 2.6: Tipos de fondos artísticos.

ou contaminação da tinta utilizada na impressão do fundo ou mesmo mudança de fornecedor da tinta ou papel por parte do banco devido à localização do seu parque gráfico. Esta variação está representada na figura 2.7.



Figura 2.7: Variações na tonalidade de impressão.

Outro fato notado referente ao fundo artístico está associado ao tipo de conta que o cliente possui. Muitos bancos categorizam seus clientes em "*normais*" e "*preferenciais*" criando uma estratificação de classes de sua clientela. Tal estratificação está sujeita às normas do banco e esta baseia-se no saldo médio, movimentação da conta, outros tipos de aplicações que o cliente realiza com o banco, etc. (o que o banco chama de classe ou faixa etária em função do saldo médio do correntista) trazendo certos benefícios ao cliente. Esta individualização ou estratificação de classes está inserida nos cheques, em sua grande maioria, por um adendo no campo do logotipo do banco ou na área reservada à identificação do sacado ou antecedendo o nome do cliente. Esta estratificação de classe de clientes pode ser vista na figura 2.8.

A estratificação também pode ser feita por alteração na cor do fundo artístico. A figura 2.9 mostra essa estratificação por alteração de cor do fundo e expressão adicionada ao lado do campo do logotipo do banco.

Na figura 2.10 pode ser vista a estratificação por mudança de cor do fundo artístico somente.



Figura 2.8: Estratificação de classes de clientes por expressões adicionadas ao cheque.



Figura 2.9: Estratificação de classes de clientes por mudança da cor do fundo artístico e inserção de expressão.

Para evitar falsificações, rasque cuidadosamente nos cheques e utilize sempre a lâmina de corte, utilizando-se sempre do papel com as setas.

Código	Banco	Agência	DV	C1	Conta	C2	Série	Chave no	C3	RS
009	001	2020	3	4	112.686-5	4	001	227763	0	

Pague por este cheque a quantia de

BANCO DO BRASIL

LE - AGUA VERDE - CURITIBA PR
00.000.000/3919-56
01 - BRASILIO ITIBERE 3668

CONDOMINIO EDIFICIO VISCÓ
NDE DE ARAGUAIA
80.563.679/0001-66

110112312312 000022776354 0000000000000000

Código	Banco	Agência	DV	C1	Conta	C2	Série	Chave no	C3	RS
009	001	1876	7	8	208.733-2	8	001	075353	0	

Pague por este cheque a quantia de

BANCO DO BRASIL

XEE - MAL. DEODORO - CURITIBA PR
00.000.000/3679-07
13 - PAB ECT
JOSE LOUREIRO 540

PAULA A B FADE
ABRAO M FADE NETO
111.373.129-04

11011137651 0000753533 0000000000000000

Figura 2.10: Estratificação de classes de clientes por mudança da cor do fundo artístico sem inserção de expressão.

2.5 OUTRAS CARACTERÍSTICAS

2.5.1 Faixa de segurança impressa

Alguns bancos inserem nos seus cheques uma linha de segurança para evitar falsificações. Este tipo de falsificação é feito da seguinte forma: O cliente preenche normalmente um cheque e o assina. O estelionatário, possuindo outro cheque (do mesmo banco) executa um corte, utilizando uma ferramenta afiada, no cheque do cliente descolando as fibras superiores do papel com os campos preenchidos deixando apenas a assinatura e dados de identificação originais no cheque. Procedendo do mesmo modo, recorta o outro cheque e preenche os dados com novos valores e insere sobre o cheque do cliente. O corte fica imperceptível pois o mesmo muitas vezes é executado segundo as linhas do fundo artístico tornando a fraude muitas vezes imperceptível [Pacheco (1998)].

Desta forma, os bancos adotaram as linhas impressas de segurança visando o resguardo dos seus clientes a este tipo de estelionato. Esta linha é disposta no sentido vertical, no canto direito do cheque, cobrindo os campos de Valor numérico, Valor por Extenso, Nome do Favorecido, Local e data de emissão e Assinatura. Esta faixa é impressa segundo ângulos ou sinuosidades aleatórias sendo encontrado o mesmo padrão das linhas a cada mil cheques impressos [Pacheco (1998)]. A figura 2.11 identifica dois tipos de impressão de segurança sobre um mesmo tipo de cheque.



Figura 2.11: Impressão randômica de segurança a fim de evitar fraude por estelionato.

2.5.2 Impressão com tinta ultravioleta

A impressão com tinta ultravioleta evita outro tipo de fraude utilizando-se máquinas reprográficas, onde o cheque é copiado, preenchido (com a assinatura do cliente falsificada) e depois descontado como se fosse um cheque válido. Tal fraude foi utilizada quando apareceram no mercado máquinas de reprografia coloridas. Devido ao comprimento de onda da iluminação utilizada por estas máquinas, na cópia são impressos caracteres como "NULO", "MODELO" ou "SEM VALIDADE", tornando o cheque reproduzido sem valor comercial para transações monetárias. Este tipo de impressão pode ser visto na figura 2.12 e é obtido pela segmentação de cores do cheque segundo o canal *HSI* (detalhado no Capítulo 3). O canal de matiz *H* mostra a impressão em ultravioleta.



Figura 2.12: Impressão de segurança em ultravioleta vista sob o canal de matiz

2.5.3 Outras características

Foi observado em um único cheque de série, o preenchimento manuscrito no campo "Número da Conta" sendo o restante do cheque impresso normalmente. Deve-se notar que a impressão do nome do cliente e *CPF* foram feitos por impressora de fita quase sobre a linha *MIRC CMC-7* e contém ainda o número da conta. Este cheque foi obtido junto a um cliente que também é funcionário do próprio banco possuindo conta na sede da agência bancária. Esta prática, que não excede a 0,1% da totalidade dos cheques compensados em movimento diário por um banco, remonta aos

cheques chamados avulsos em que o preenchimento é feito pelo caixa do próprio banco ou mesmo pelo cliente na falta de talão de cheque por parte do mesmo. O cheque aqui referenciado pode ser vista na figura 2.13.

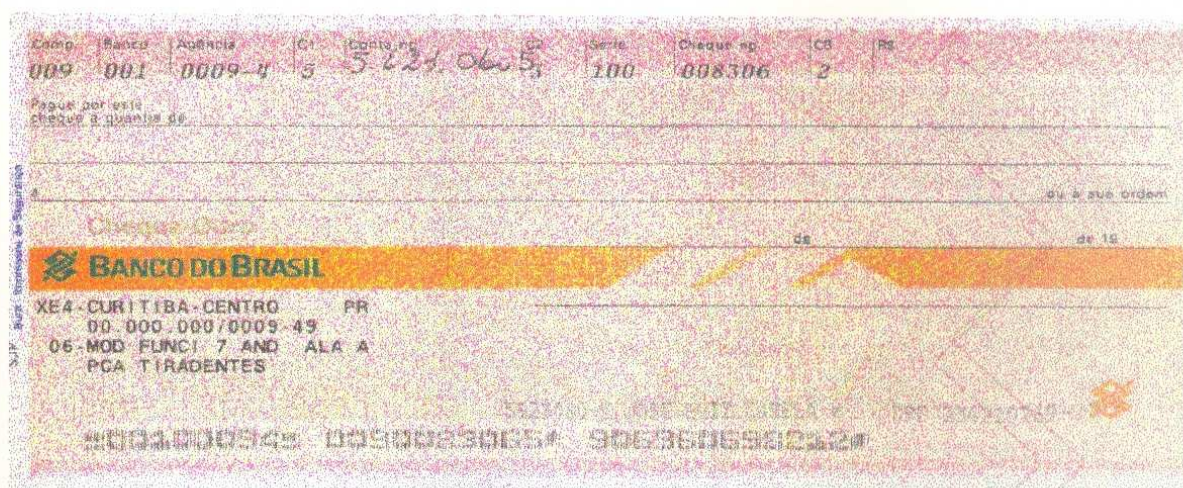


Figura 2.13: Exemplo de um cheque avulso com preenchimento manual/mecânico

2.6 BASE DE DADOS

2.6.1 Introdução

A montagem de uma base de dados apresentava-se como fundamental para o termo deste trabalho de pesquisa. Uma das maiores dificuldades para a criação desta base de dados de cheques bancários, remonta a se encontrar uma instituição financeira que esteja disposta a ceder cheques reais, que 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. Dessa forma, as publicações voltadas ao processamento de cheques bancários apresentam resultados obtidos a partir de pequenas bases de dados, muitas vezes restritas regionalmente [Koerich (1997)].

[Koerich (1997)] afirma que as bases de dados existentes são insuficientes por diversos motivos. Vários conjuntos de dados contêm informações isoladas que não refletem as informações encontradas na prática; outras contêm dados quase que exclusivamente binárias limitando o pesquisador

à taxa de amostragem e ao algoritmo de limiarização fornecido por um digitalizador particular. Isto impede experimentos com pré-processamento ou reconhecimentos no domínio de níveis de cinza da imagem que podem trazer melhorias no processamento das mesmas [Morita (1998)].

Torna-se importante salientar que o objetivo deste trabalho está relacionado com a fase de pré-processamento dos cheques, ou seja a separação dos mesmos. Uma vez que as bases de dados existentes disponíveis, listadas na tabela 2.2, possuam em sua grande maioria elementos em forma binária, necessitou-se proceder a aquisição de uma nova base de dados com cheques coloridos a fim de se alcançar os objetivos deste trabalho.

	BASES DE DADOS EXISTENTES
CEDAR	Center of Excellence in Document Analysis Recognition Buffalo, USA
CENPARMI	Center of Pattern Analysis and Machine Intelligence Montreal, Canadá
NIST	National Institute Standards and Technology USA
LRPRC	Laboratório de Reconhecimento de Padrões e Redes de Computadores Universidade Estadual de Campinas - Brasil
LARDOC	Laboratório de Análise e Reconhecimento de Documentos Pontifícia Católica do Paraná - Brasil

Tabela 2.2: Bases de dados de documentos existentes

2.6.2 Aquisição

Fundamentado pelo descrito acima, procedeu-se à aquisição de uma nova base de dados que contivesse cheques coloridos e não já limiarizados para futuros processamentos. Em um primeiro instante, foram coletados cheques bancários de forma aleatória junto a amigos e alunos, que cederam gentilmente folhas de seus talões de cheques sem preenchimento algum. Com a cessão destas folhas, foi concedido, também, o direito da utilização dos mesmos sob a forma de imagens para a realização deste trabalho.

A montagem da base de dados foi feita por processo de aquisição direta por meio de scanner e arquivada em memória de massa. Em um segundo instante, o Sr. Ary Paulino Chemin Jr., gentilmente cedeu cheques preenchidos por clientes de seu estabelecimento comercial, o que aumentou

sobremaneira a base de dados. De posse das imagens já adquiridas, os cheques foram classificados em três categorias: limpos, manuscritos e preenchidos por meio mecânico perfazendo cerca de 540 cheques. A distribuição dos cheques foi aleatória, não classificada segundo a sua quantidade por banco, mas baseada no movimento comercial do estabelecimento do Sr. Ary Paulino Chemin Jr. que se situa na cidade de Campo Largo, região metropolitana de Curitiba.

Assim, a base de dados é falha no tocante a bancos regionais com por exemplo, bancos estaduais ou bancos setorializados em certos estados que não possuem agência fora de sua região de cobertura.

2.6.3 Meio de aquisição

Os cheques foram adquiridos por meio de um scanner de mesa marca *HewlettPackard*, modelo 3c utilizando-se a folha padrão de calibração de cor que acompanha o scanner. Todos os cheques foram adquiridos em 300 *DPI's* e 16 milhões de cores.

Um fato observado durante a fase de aquisição, foi a existência de variações de tonalidade de cores se a mesma fosse executada em outro scanner. Esta variação de cor, apesar de ser feita a calibração dos dois scanner pela folha padrão dos mesmos, está relacionada com o desgaste mecânico do scanner ou vida útil de sua célula de iluminação. A figura 2.14 mostra o mesmo cheque adquirido por meio de um scanner *HP 3c* e por meio de um scanner *UMAX* modelo *ASTRA 1200 – S*. É possível notar a variação de tonalidade na intensidade da componente vermelha nos dois cheques devido a característica da fonte de iluminação.

2.6.4 Composição

A base de dados é atualmente composta de 66 cheques sem preenchimento, codificados segundo o nome do banco, subdivididos conforme a tabela 2.3. Os cheques com preenchimento manuscrito totalizam 222 cheques conforme a tabela 2.5. Os cheques com preenchimento mecânico totalizam 251 cheques conforme a tabela 2.4.

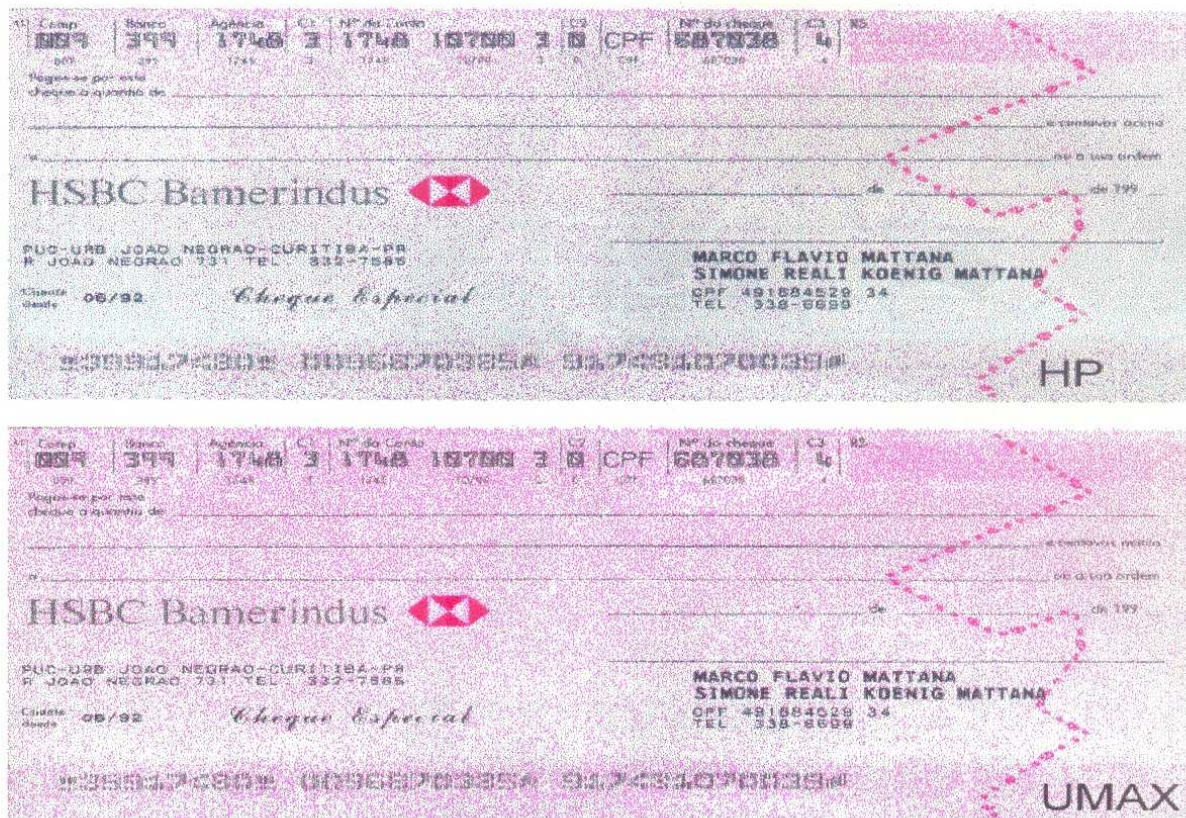


Figura 2.14: Cheques com variação de tonalidades devido a aquisição em tipos diferentes de scanners

Banco	Codificação	Quantidade
Banco Bamerindus SA	BAM	14
Banco Banestado	BAN	11
Banco América do Sul	BAS	02
Banco do Brasil	BB	05
Banco Bradesco	BRA	02
Caixa Econômica federal	CEF	06
Banco HSBC Bamerindus SA	HSBC	02
Banco Itaú	ITA	04
Banco Meridional	MER	01
Banco Real	REA	14
Unibanco	UNI	04

Tabela 2.3: Descrição da base de dados de cheques sem preenchimento

Banco	Codificação	Quantidade
Banco Bamerindus SA	BAM	02
Banco Banestado	BAN	34
Banco América do Sul	BAS	05
Banco do Brasil	BB	37
Banco Bandeirantes	BBS	11
Banco Bradesco	BRA	29
Caixa Econômica federal	CEF	27
Citybank SA	CIT	19
Banco HSBC Bamerindus SA	HSBC	19
Banco Itaú	ITA	29
Banco Real	REA	16
Banco Meridional	MER	11
Unibanco	UNI	12

Tabela 2.4: Descrição da base de dados de cheques com preenchimento mecânico

Banco	Codificação	Quantidade
Banco Cidade SA	BAC	01
Banco Bamerindus SA	BAM	04
Banco Banestado	BAN	57
Banco América do Sul	BAS	03
Banco do Brasil	BB	26
Banco Bandeirantes	BBS	08
Banco de Crédito Nacional SA	BCN	01
Banco Banestes	BES	01
Banco Francês e Brasileiro SA	BFB	01
Banco Noroeste SA	BNO	03
Banco Boa Vista SA	BOA	01
Banco Bradesco	BRA	17
Caixa Econômica federal	CEF	20
Citybank SA	CIT	09
Banco Excel	EXC	06
Banco HSBC Bamerindus SA	HSBC	28
Banco Itaú	ITA	14
Banco Real	REA	13
Banco Santander SA	SAN	01
Unibanco	UNI	08

Tabela 2.5: Descrição da base de dados de cheques com preenchimento manuscrito

Neste Capítulo foram comentadas as principais características dos cheques bancários brasileiros bem como a formação da base de dados para a pesquisa. No próximo Capítulo serão observadas as principais características referentes à cor e sistemas de cores e sua segmentação.

Capítulo 3

A COR

3.1 INTRODUÇÃO

A luz encontra-se no domínio das radiações eletromagnéticas. Estas radiações ou ondas propagam-se em forma de partículas a partir de uma fonte, seguindo trajetórias retilíneas.

A radiação eletromagnética que se situa nos estreitos limites de 400 a 700 nanômetros, onde 1 nanômetro, ou abreviadamente *nm*, corresponde à 1 milionésimo de milímetro ou 10^{-7} centímetros, possui a propriedade de estimular a retina humana, dando a sensação da luz. Assim, esta faixa do espectro de radiação eletromagnética chama-se luz visível e sua cor característica é o branco.

O espectro visível inicia-se com o comprimento de onda de 400 nm onde a luz parece ter uma cor profunda de violeta escuro, tornando-se cada vez mais azul à medida que o comprimento de onda se aproxima de 450 nm. A 500 nm, o azul começa a dar lugar ao azul-verde chegando ao verde puro em 580 nm. A partir deste comprimento de onda o verde começa a se tornar amarelo e próximo a 600 nm o amarelo cede lugar ao alaranjado. Próximo ao comprimento de onda de 650 nm o alaranjado começa a se tingir de vermelho que escurece até se tornar tão obscuro que se torna muito difícil de se detectar. Neste ponto, o comprimento de onda é de 700 nm. [Langford (1984)]

Abaixo de 400 nm, encontra-se a luz ultravioleta e a cima de 700 nm, a infravermelha. A figura 3.1 exemplifica o espectro de luz visível.

A cor nos objetos é formada pela reflexão ou absorção de certos comprimentos de onda, a partir de uma fonte de iluminação da qual emana luz branca. Os objetos que refletem todos os comprimentos de onda de maneira iguais são chamados de acromáticos e são vistos pelo olho

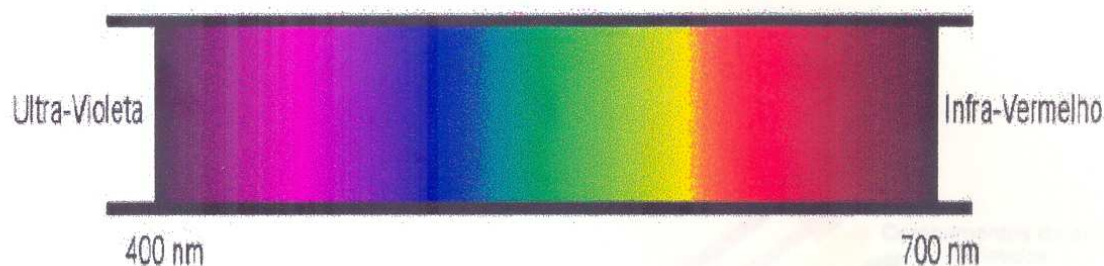


Figura 3.1: Espectro eletromagnético da luz visível.

humano como pretos, brancos ou com tonalidades intermediárias de cinza. Já os objetos que refletem comprimentos de onda de valores diferentes são chamados de cromáticos e apresentam uma cor própria. A figura 3.2 ilustra estas características.

3.2 SISTEMAS DE CORES

3.2.1 Introdução

Na literatura é possível observar a contribuição da cor em algumas aplicações, bem como, constatar a fundamental importância da forma de se representar esta informação na imagem, para se obter resultados significativos no seu processamento. Em [Umbaugh et al (1993)] é descrita a importância da informação cor no diagnóstico automático de tumor de pele. Durante os experimentos, diferentes sistemas de representação de cor são avaliados na identificação de características Deste tipo de tumor. [Ohta et al (1980)] faz também um estudo comparativo de vários sistemas para representação da cor utilizados no processo de segmentação de imagens. Esses diferentes estudos mostram que, basicamente, três sistemas de representação da cor são usados, o modelo *RGB* (do inglês, Red, Green, Blue), o modelo *CMY* (do inglês, Cyan, Magenta, Yellow) e o modelo *HSI* (do inglês, Hue, Saturation, Intensity), matiz, saturação e intensidade.

3.2.2 Modelos *RGB* e *CMY*

A luz branca formada pôr comprimentos de onda de 400 a 700 nanômetros pode ser decomposta em violeta, azul, verde, amarelo, alaranjado e vermelho conforme visto no item anterior. Porém, é possível formar novas cores através da combinação de dois ou mais comprimentos de onda (ou

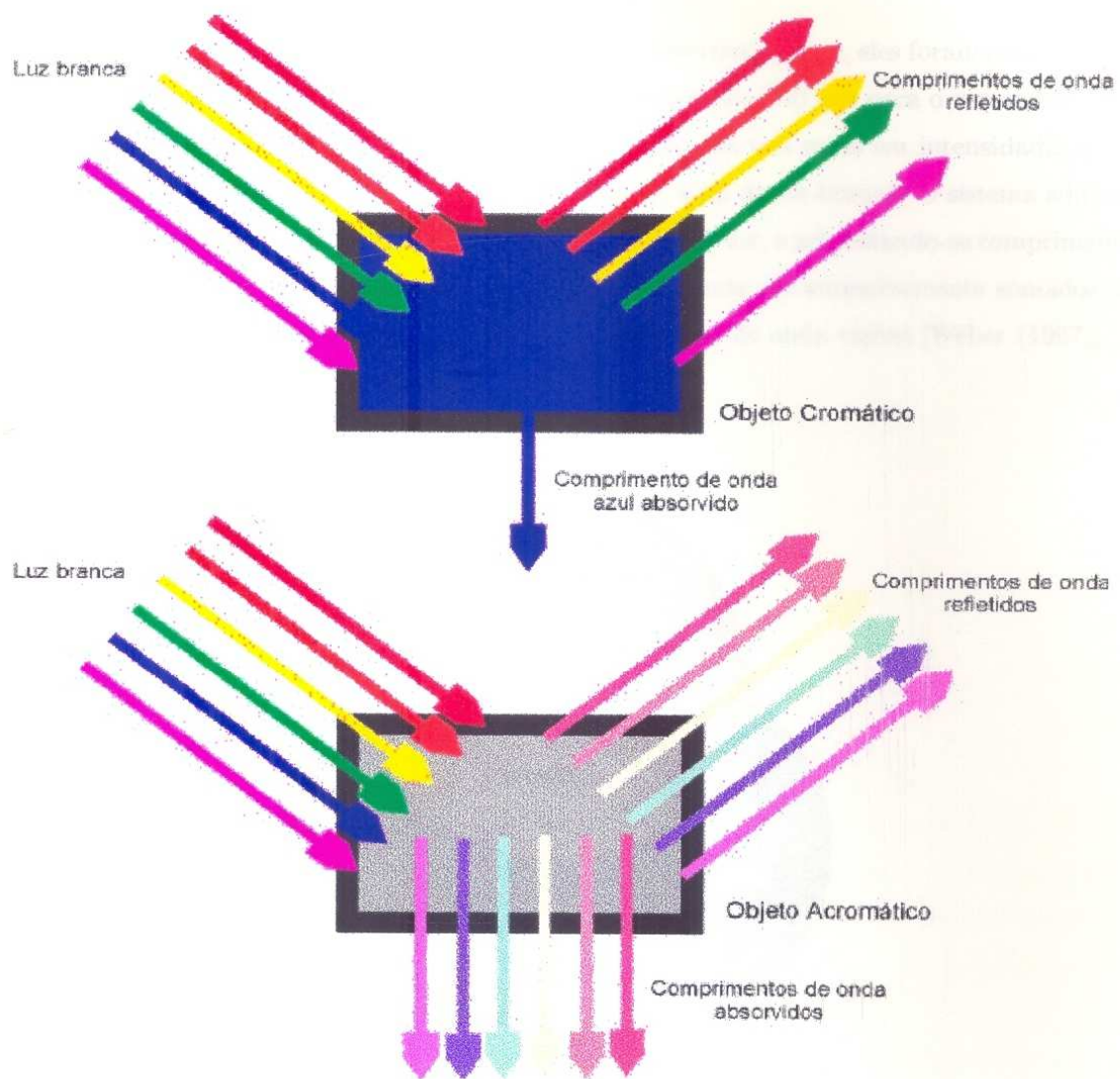


Figura 3.2: Representação da absorção e reflexão da luz sobre objetos cromáticos e acromáticos.

cores) do espectro visível. O conjunto que resulta em um maior número de combinações está composto das cores vermelha, azul e verde. Tal conjunto forma o chamado sistema aditivo de cores [Langford (1984)] e é representado pela sigla *RGB* - do inglês Red, vermelho; Green, verde e Blue, azul.

Uma vez que os comprimentos de onda variem sob determinada cor, eles foram normatizados pelo *C.I.E* (Comissão Internationale de l'Eclairage) como sendo 700 *nm* para o vermelho, 546,1 *nm* para o verde e 435,8 *nm* para o azul. O agrupamento das três cores em intensidades iguais resulta nas chamadas cores secundárias do sistema aditivo, além da cor branca. O sistema aditivo é assim chamado porque, partindo-se do preto - ausência total de luz, e adicionando-se comprimentos de onda obtém-se novas cores onde seus comprimentos de onda são sucessivamente somados até a obtenção do branco que corresponde ao maior comprimento de onda visível [Weber (1997)]. O sistema aditivo está representado na figura 3.3.



Figura 3.3: Sistema aditivo de cores ou *RGB*.

Um outro sistema, complementar ao primeiro no tocante às cores primárias, é chamado de sistema subtrativo, responsável pelas cores dos objetos. Moléculas, aqui chamadas de pigmentos, subtraem comprimentos de onda pôr absorção, refletindo outros comprimentos o que se traduz nas cores características dos objetos. Enquanto que no sistema aditivo, as cores são formadas pôr

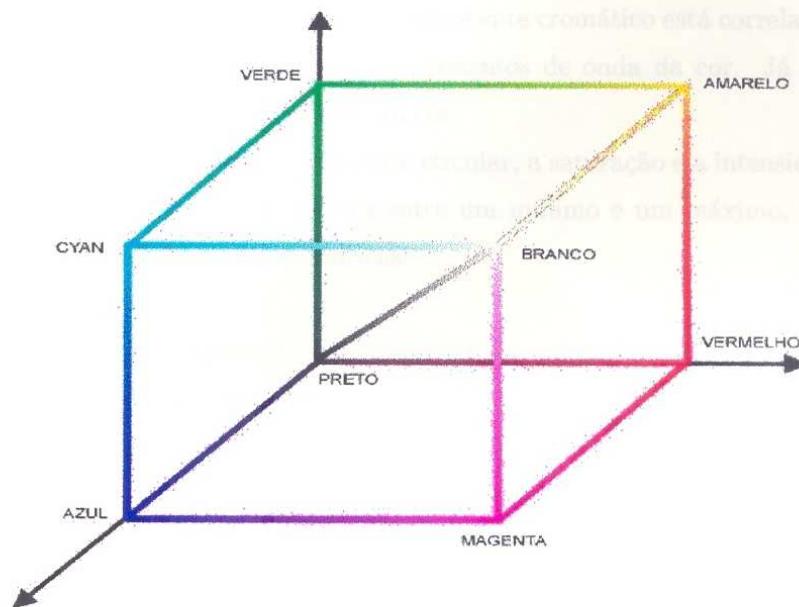
luz, no sistema subtrativo as cores são obtidas pôr pigmentação das superfícies dos objetos tendo como base as cores Cyan, Magenta e Amarelo formando a sigla *CMY* (do inglês, cyan, magenta e yellow). Assim, partindo-se do branco - ausência de cor, obtendo-se as cores primárias (cyan, magenta e amarelo), as cores resultantes - vermelho, verde e azul - chamadas cores secundárias, acabam se degenerando em preto descrevendo os comprimentos de onda refletidos pelo objeto. O sistema subtrativo pode ser visto na figura 3.4.



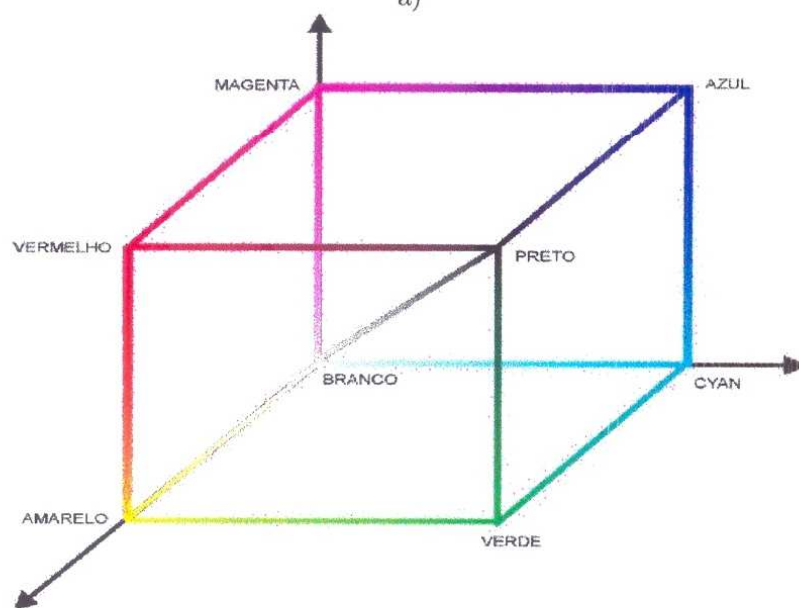
Figura 3.4: Sistema subtrativo de cores ou *CMY*.

Tanto o sistema de cores aditivas (*RGB*) quanto o sistema de cores subtrativas (*CMY*) podem ser, então, descritos de uma maneira mais simples se representados como espaços tridimensionais. A representação dos dois sistemas assume a forma de um cubo onde os vértices são representados pelas cores que compõem os dois sistemas. Estas representações podem ser vistas na figura 3.5.

Além do uso em monitores coloridos, a maioria das câmaras coloridas usadas na aquisição de imagens utilizam o modelo *RGB*, o que o torna um importante modelo de cor no processamento de imagens. Porém, a definição de uma cor a partir das componentes *RGB* em nada se assemelha a forma natural de percepção de cor humana.



a)



b)

Figura 3.5: a) Espaço *RGB* e b) Espaço *CMY* segundo distribuição espacial.

3.2.3 Modelo *HSI*

Outros modelos de cores podem ser descritos pois a percepção humana das cores está relacionado a componentes cromáticos e de intensidade. O componente cromático está correlacionado ao matiz e a saturação; e estes, correlacionados aos comprimentos de onda da cor. Já a amplitude dos comprimentos de onda definem a intensidade da cor.

Tomando-se o matiz como um componente circular, a saturação e a intensidade da cor como valores descritos como uma escala de valores entre um mínimo e um máximo, torna-se possível descrever o sistema com coordenadas cilíndricas.

Fica assim definido o sistema *HSI* (do inglês, Hue; matiz, Saturation, saturação; Intensity, intensidade). Os níveis de cinza situam-se na reta que descreve a altura de um cone duplo invertido e estão compreendidos entre o preto, cujo valor é 0, e o branco, cujo valor é 1 passando pelo cinza formado de 50% de preto e 50% de branco com valor de 0,5 coincidindo com a base comum dos cones. Na base comum dos cones situam-se as cores de máxima saturação, que variam segundo esta base, sua cromaticidade. A matiz da cor é obtida sobre as geratrizes do cones que podem escurecer (se a direção for para baixo) ou clarear (se a direção for para cima). Todas as cores podem ser obtidas segundo este espaço bicônico que definem um volume espacial como representado na figura 3.6.

O Sistema de representação *HSI* busca uma forma mais próxima da manipulação natural de cor utilizada pelo homem. Nesse sistema torna-se mais natural a busca pôr uma determinada cor, dada sua matiz, variando a saturação e intensidade de forma independente. [Ohta et al (1980)] demonstrou que no modelo *HSI*, a informação cor é mais descorrelacionada que no caso do modelo *RGB*. Em várias aplicações o modelo *HSI* se mostrou mais apropriado que outros modelos de representação cromática. Na prática, o modelo *HSI* é deduzido do modelo *RGB* pelas seguintes relações [Gonzales et al (1992)]:

$$\begin{aligned} I &= \frac{R + G + B}{3 * 255} \\ S &= 1 - \frac{3}{(R + G + B)} \text{Min}(R, G, B) \\ H &= \arctan\left(\sqrt{\frac{255((R - G) + (R - B))}{(R - G)^2 + (R - G)(G - B)}}\right) \end{aligned}$$

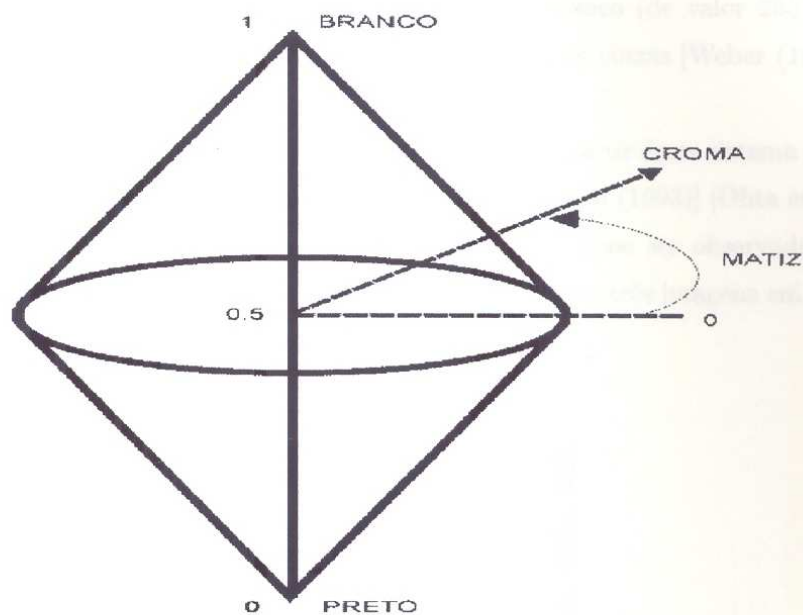


Figura 3.6: Espaço *HSI*.

3.3 SEGMENTAÇÃO

3.3.1 Introdução

No âmbito do Processamento de Imagens, segmentar consiste em identificar e extrair estruturas desejáveis presentes em uma cena, sendo a eficiência de tal processo diretamente relacionada ao desempenho final da análise das imagens.

Não existe uma teoria completa sobre segmentação sendo as várias técnicas "ad hoc" que diferem entre si nas propriedades que enfatizam como regiões fechadas e contornos [Haralick et al(1992)]. A grande dificuldade reside no fato de não se conhecer de antemão o número e tipo de estruturas que se encontram na imagem. Estas estruturas são identificadas a partir de sua geometria, forma, topologia, textura, cor ou brilho sendo escolhidas aquelas que possibilitam a melhor identificação.

Neste caso, o atributo cor possui um grande poder discriminatório sendo, assim, usado com maior frequência na identificação de estruturas de objetos. Assim, o atributo cor, no âmbito dos cheques bancários torna-se o fator importante para a segmentação. Para a obtenção da limiarização de uma imagem colorida, torna-se imprescindível a obtenção de uma outra imagem, esta em escala de cinzas. Assim, em uma imagem, a amplitude de seus elementos discretos, denominados de pixels (do inglês, picture elements) pode ser representada em níveis de cinza, de maneira global,

em 256 níveis partindo-se do preto (cujo valor é 0) até o branco (de valor 255) com os valores intermediários correspondendo as demais variações na escala de cinzas [Weber (1997)].

Em imagens coloridas de padrão *RGB*, a segmentação a partir deste sistema de representação de cores não possibilita um resultado satisfatório [Umbaugh et al (1993)] [Ohta et al (1980)], pois as componentes de cor são fortemente correlacionadas como pode ser observado na figura 3.7. Em imagens com padrão *RGB*, após a segmentação são obtidas três imagens em escala de cinzas representando as componentes azul, verde e vermelha.

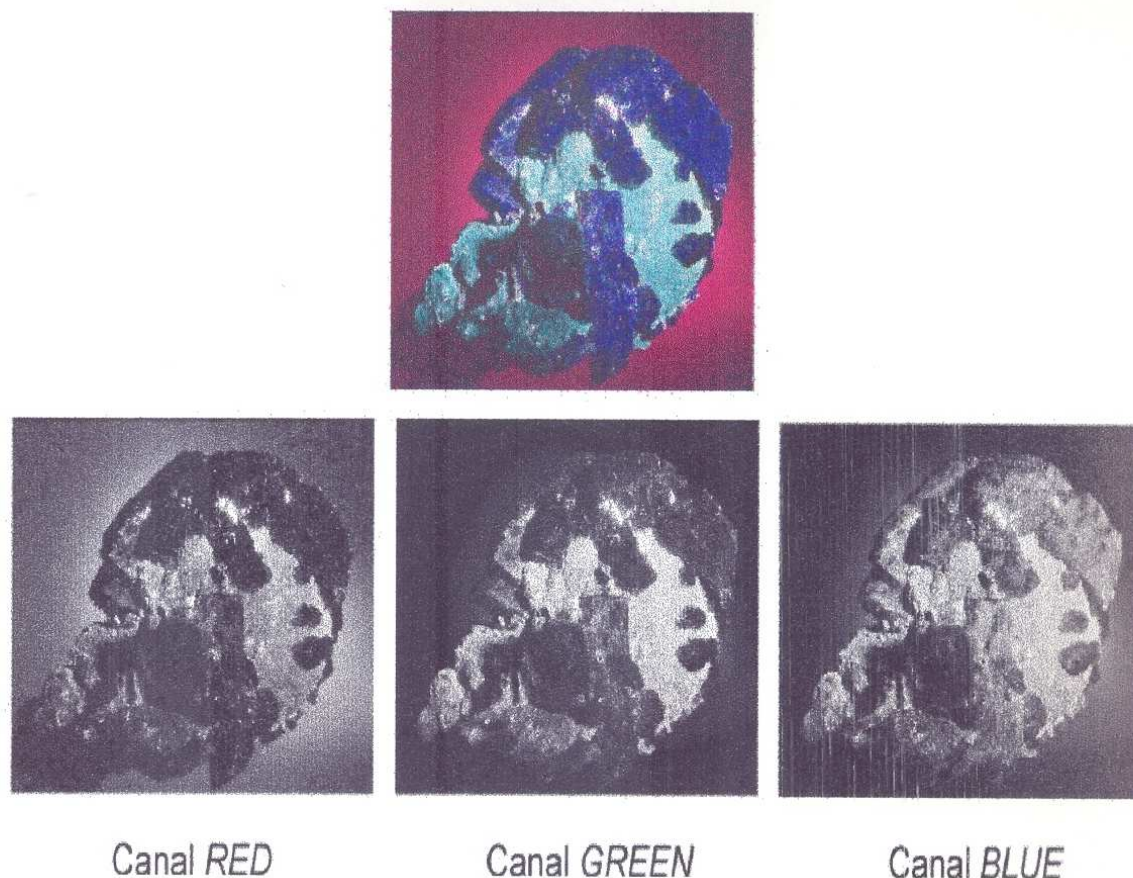


Figura 3.7: Segmentação de uma imagem em canais *RGB*.

O sistema *HSI* possibilita uma segmentação de canais mais descorrelacionada (figura 3.8) [Umbaugh et al (1993)] [Ohta et al (1980)].

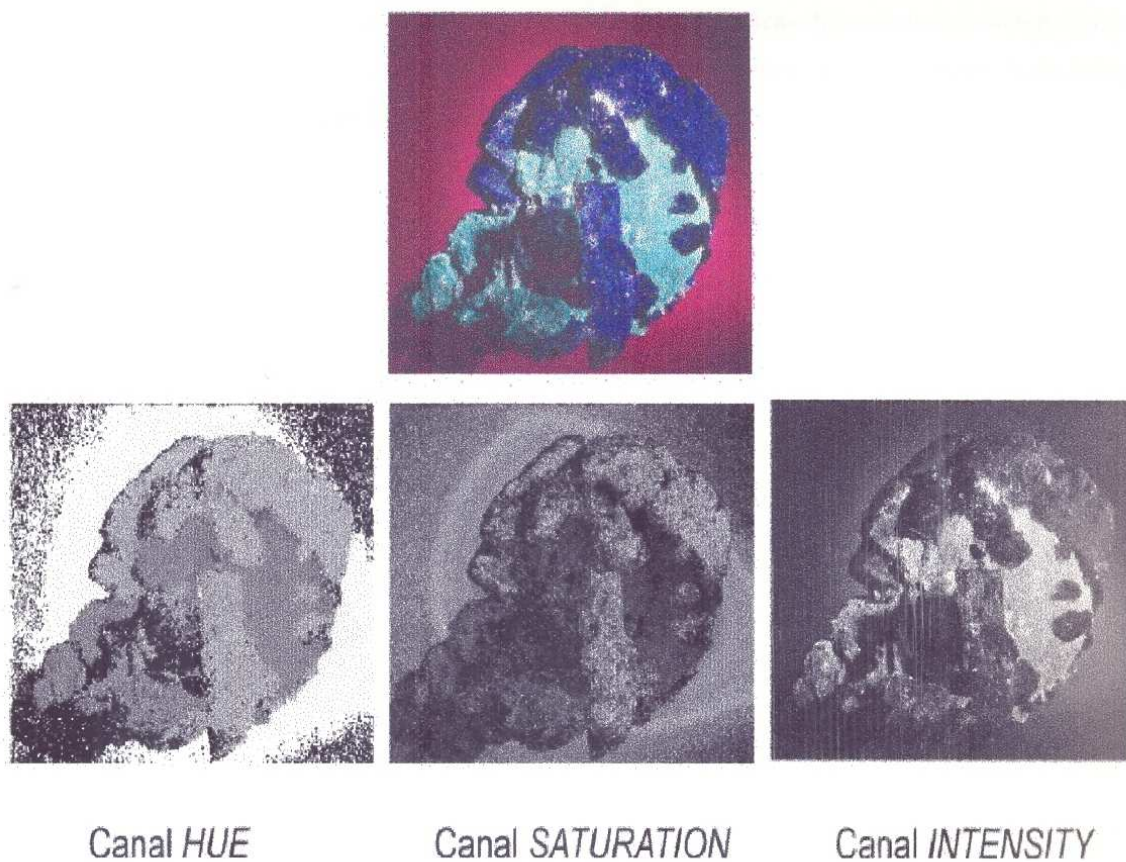


Figura 3.8: Segmentação de uma imagem em canais *HSI*.

3.3.2 Segmentação pelo modelo *HSI* aplicado a cheque bancários

Como o modelo *HSI* possibilita uma segmentação de canais mais descorrelacionada, este sistema foi adotado para o processamento dos cheques bancários. Posteriormente à fase de aquisição, procedeu-se à segmentação das imagens dos cheques pelo sistema *HSI*. O canal considerado como de interesse é o canal *I* (Intensity) que fornece uma imagem na escala de cinzas própria para a aplicação das técnicas de limiarização. Os canais de saturação *S* e de matiz *H* são então, desprezados. A segmentação de um cheque pode ser observada na figura 3.9.

Neste Capítulo foram comentadas as principais características de cores e seus sistemas correspondentes (RGB, CMY e *HSI*). No próximo Capítulo serão observadas as principais metodologias de segmentação de imagens por binarização.



Imagem Original



Canal Matiz
(Hue)



Canal Saturação
(Saturation)



Canal Intensidade
(Intensity)

Figura 3.9: Segmentação de um cheque em canais *HSI*.

Capítulo 4

METODOLOGIA DE SEGMENTAÇÃO POR BINARIZAÇÃO

4.1 INTRODUÇÃO

A digitalização de documentos pôr meio de scanners ou captura digital de vídeo fornecem imagens compostas de pixels como forma de dados de entrada para um sistema de processamento. Estes pixels são elementos de uma matriz cujos valores apontam a intensidade da reflectância versus iluminância da imagem. Se a imagem for composta apenas de duas cores, ou seja, preto e branco, os valores atribuídos a estes pixels serão desligado (valor 0) ou ligado (valor 1). Isto define uma imagem de padrão binário. No caso de imagens em tons de cinza, os valores dos pixels são computados em uma faixa de tons de cinza entre 0 a 255.

A primeira etapa em processamento de documentos é preparar esta imagem para análise futura. Este processamento pode incluir, dependendo da imagem que se deseja pós-processar, binarização (reduz uma imagem em tons de cinza ou mesmo colorida para uma imagem binária), redução de ruído (minimiza formas espúrias da imagem), afinamento e detecção de bordas (permitem a obtenção de forma mais rápida de características pertinentes aos objetos da imagem) [O'Gorman et al (1995)]. Estas operações de processamento são chamadas de processamento de baixo nível ou processamento a nível de pixel.

A binarização, como já visto, é a conversão de uma imagem originalmente em tons de cinza

para uma imagem binária. A principal característica de uma imagem binária é aquela em que ela deve conter todas as informações essenciais referenciando o número, posição e a forma dos objetos, mesmo que contendo um número muito menor de informações se comparando-a a uma imagem em níveis de cinza. Pixels com níveis de cinza similares em uma região próxima normalmente pertencem ao um mesmo objeto.

Devido a fatores como impressão não uniforme do documento, variação de cromaticidade da impressão ou reflectância não uniforme da página do documento ou ainda, como resultado de transições de intensidade entre as bordas das regiões localizadas entre os objetos e o fundo, mesmo se o documento a ser adquirido for binário, a imagem obtida após a aquisição não se encontra apenas em dois níveis, mas em uma grande gama de intensidades. O objetivo principal da binarização é, então, identificar e associar os pixels que pertencem à região dos objetos com uma intensidade (pixel ligado) e os outros pixels, os que pertencem ao fundo com uma intensidade diferente (pixel desligado). A razão essencial para classificar pixels a partir dos níveis de cinza é que se reduz a complexidade dos dados e simplifica-se muitos procedimentos de reconhecimento e classificação posteriores [Parker (1996)].

4.2 CONSIDERAÇÕES SOBRE A BINARIZAÇÃO

No caso de documentos que possuem ótimo contraste entre o seus elementos de objetos impressos sobre um fundo com contraste uniforme, é possível combinar a digitalização e a binarização da imagem com um valor de limiar padrão fornecido pelo próprio scanner. Porém, para documentos que possuam uma gama diversa de variações de contrastes entre objetos e fundo, este valor de limiar fixo acaba pôr não fornecer imagens com uma boa separação entre os elementos constituintes da imagem [O'Gorman et al (1995)].

Segundo [O'Gorman et al (1995)], um mesmo documento se adquirido pôr scanners diferentes, que possuem variações em suas fontes luminosas, apresentarão limiares ótimos diferentes. [O'Gorman et al (1995)] apresenta duas soluções para este caso, ambas empíricas, para a determinação do limiar ótimo. A primeira consiste em determinar o limiar ótimo de maneira manual pôr calibração do scanner a cada imagem adquirida. O segundo método proposto parte de uma aquisição inicial de uma imagem em níveis de cinza com separações específicas de 0 a 255 níveis. Após esta digitalização, define-se o valor ótimo pôr observação visual; ajusta-se este valor ao scan-

ner e procede-se a aquisição do documento. Ambas as formas empíricas demandam uma perda de tempo expressiva e um volume de aquisições bastante razoável.

Os métodos descritos na literatura e avaliados neste trabalho, não dependem do valor de limiar fornecido pelo scanner, mas partem de imagens coloridas adquiridas e segmentadas em 256 níveis de cinza pelo canal *HSI* conforme já exemplificado no Capítulo 3.

4.3 HISTOGRAMA

Seja uma imagem ideal contínua onde a intensidade x pode tomar todos os valores de um intervalo $[a, b]$. Em termos de probabilidade, podemos definir a função de repartição de x , $F(x)$ como sendo a probabilidade P que x tenha um valor inferior ou igual a α , seja $P[x < \alpha] = F(\alpha)$. Usando a noção de limite, podemos dizer que $F'(x) = \lim_{\Delta x \rightarrow 0} \frac{F(x+\Delta x) - F(x)}{\Delta x}$ representa a estimação da função de densidade de probabilidade de x e denota-se $p(x)$. O histograma na primeira ordem de uma imagem contínua é a função de densidade de probabilidade $F'(x)$ quando x é igual a derivada de $F(x)$. No caso de uma imagem digitalizada com 2^m níveis de cinza que contém n pixels, cada um deles toma um valor aleatório de nível de cinza r_k dentro dos k valores existentes. A função de densidade de probabilidade chama-se nesse caso p_k e expressa-se com base na função de repartição $F(r_k)$ como $p_k(r_k) = \frac{(F(r_k) - F(r_{k-1}))}{n} = \frac{n_k}{n}$ onde n_k é o número de pixels para o nível de cinza r_k , e $p_k(r_k)$ representa o histograma discreto.

Pode-se notar que o histograma quando representado por meio de uma tabela fornece, para cada nível de cinza, o número de pixels correspondentes na imagem. Na medida que o histograma fornece, para um nível de cinza dado, somente o número de pixels e não a localização desses, ele permite dar uma descrição global da imagem.

Uma outra definição de histograma pode ser obtida em [Jähne (1995)] que o define da seguinte maneira: de forma geral, a probabilidade da distribuição dos pixels em uma imagem não é conhecida a priori, mas estimada pôr medições. Se a imagem possui uma distribuição tonal homogênea, isto é a posição do pixel não depende de seu nível de cinza, então é possível estimar essa probabilidade de distribuição pelo histograma. Um histograma típico pode ser observado na figura 4.1.

Tomando-se pôr base o histograma de uma imagem, é possível escolher-se o limiar baseado no valor da gama cromática de níveis de cinza. Pôr exemplo, tomando-se o limiar aleatoriamente

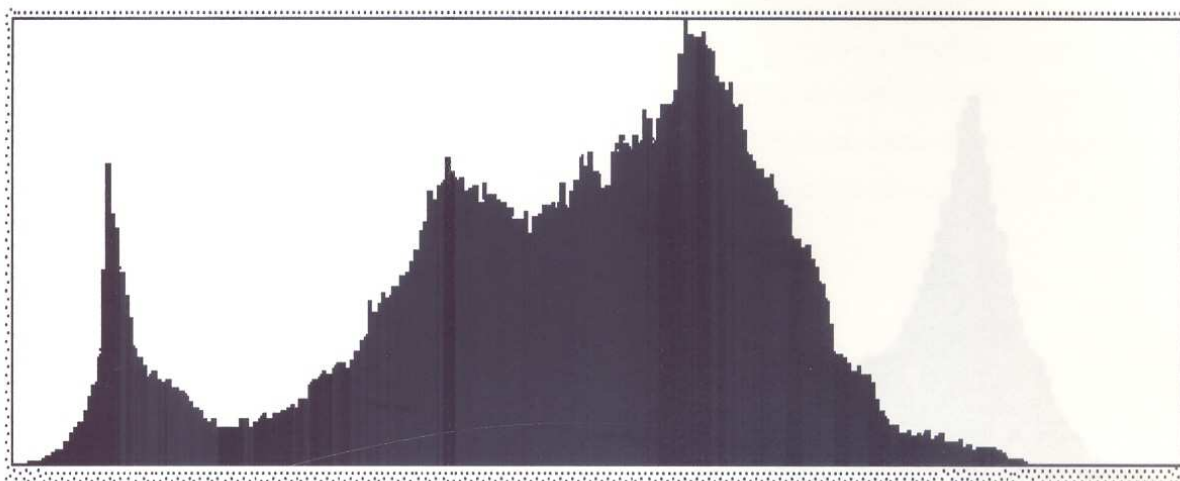


Figura 4.1: Histograma de uma imagem em níveis de cinza

de uma imagem a ser binarizada como sendo 120, isto significa que todos os pixels abaixo deste valor serão assinalados como valor desligado ou 0 - serão convertidos para preto - e os valor acima do limiar , serão assinalados como ligados ou 1 - serão convertidos para branco. A busca desde limiar pode ser facilitada se o histograma apresentar certas características como em um caso ideal, o histograma possuir um agudo vale entre dois picos, representando os objetos e o fundo, respectivamente, assim, o limiar ótimo pode ser escolhido como o valor que representa esse vale como pode ser observado na figura 4.2.

Todavia, para a maior parte das imagens reais, não é somente difícil detectar precisamente o valor do vale, especialmente nos casos em que o vale é plano e largo, ou que o histograma possua ruído, ou ainda, quando os dois picos são extremamente diferentes no tamanho ou quando não se produz um vale visível. Exemplos destes tipos de histogramas podem ser observados na figura 4.3.

4.4 TÉCNICAS DE BINARIZAÇÃO GLOBAIS VERSUS LOCAIS ADAPTATIVAS

Existem muitas técnicas propostas com objetivo de superar todas as dificuldades para encontrar o valor do limiar que podem ser dividida em duas famílias. A primeira delas conhecida como Métodos Globais utiliza um único valor de limiar aplicado sobre a imagem como um todo. Apesar desta técnica fornecer bons resultados para certos tipos de documentos ela pode gerar erros de binarização em certas áreas de delimitação entre fundo e objetos [O'Gorman et al (1995)]. Um

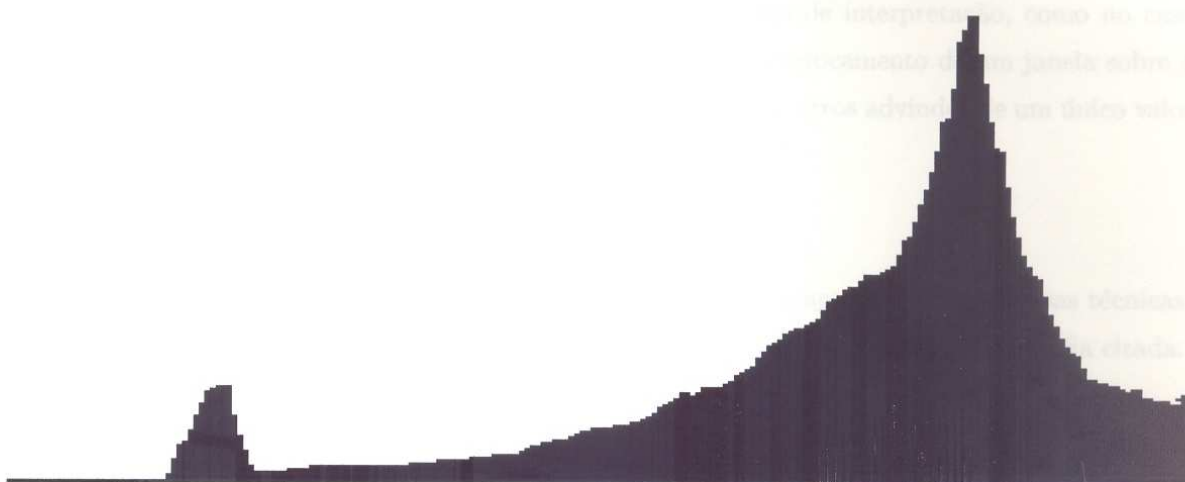


Figura 4.2: Histograma típico com divisões de vale e picos

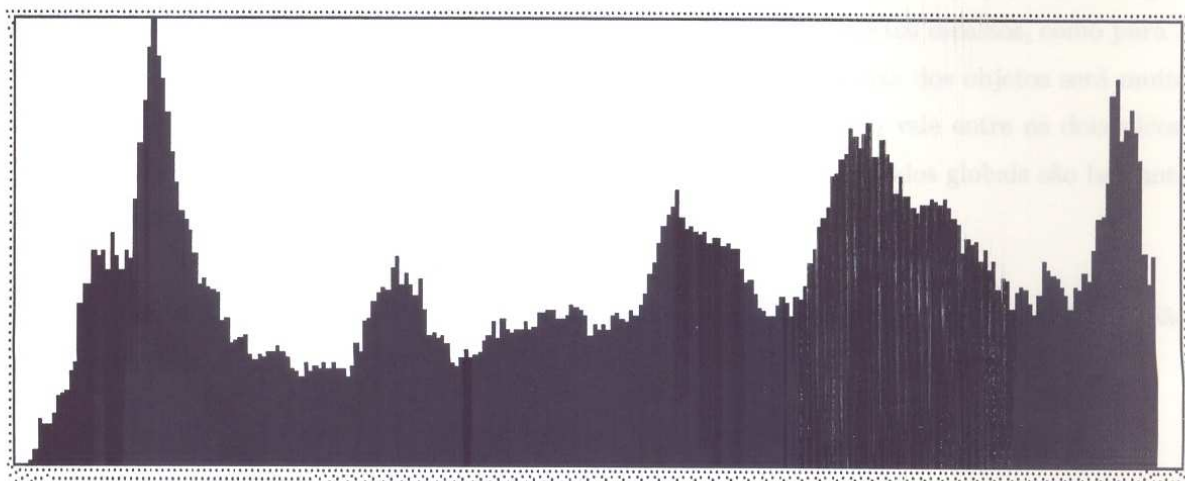


Figura 4.3: Histograma complexo

exemplo seria a tentativa de se obter a binarização de uma imagem que contivesse certas áreas do fundo com mesmos tons de cinza de certas áreas do objeto. A segunda família, conhecida como Métodos Locais Adaptativos foi desenvolvida para evitar erros de interpretação, como no caso descrito acima. Nessa segunda família, as técnicas prevêm o deslocamento de um janela sobre a imagem de onde são computados valores limiares locais, evitando erros advindos de um único valor para toda a imagem.

Observação: Apesar da possível repetição de variáveis empregadas nas diversas técnicas, os métodos apresentados a seguir serão descritos conforme foram descritos na bibliografia citada.

4.5 MÉTODOS GLOBAIS

A maneira mais simples de se obter o valor limiar global para uma imagem é utilizar os valores de intensidade obtidos a partir do histograma dos pixels da imagem.

Para uma imagem com objetos e intensidades de fundo bem diferenciadas, o histograma terá dois cumes distintos. O vale entre esses cumes pode ser achado como o mínimo entre dois máximos, e é escolhido como sendo o limiar que melhor separa os dois cumes.

Existem algumas desvantagens da seleção de um limiar único. A primeira desvantagem é a falta de contraste entre objetos e fundo e a possível presença de ruído. A segunda desvantagem reside no fato de que, especialmente para uma imagem que contém objetos escassos, como para a maioria das imagens de gráficos, o pico que representa os níveis de cinza dos objetos será muito menor que o do fundo. Assim, estas diferenças tornam difícil achar o vale entre os dois picos. Apesar destas desvantagens, devido à rapidez de processamento, os métodos globais são bastante utilizados para binarizar imagens.

A aplicação desses métodos à cheques bancários, serão exemplificados no Capítulo 5. São métodos globais:

- Método por Abordagem Fuzzy - Abordagem de Huang [Huang et al(1995)] (página 43)
- Método por Abordagem Fuzzy - Abordagem de Yager [Yager (1979)] (página 43)
- Método por Dois Picos [Parker (1996)] (página 50)
- Método por Entropia de Pun [Pun (1981)] (página 44)

- Método por Entropia de Kapur [Kapur et al (1985)] (página 45)
- Método por Entropia de Johannsen [Johannsen et al (1982)] (página 46)
- Método por Entropia de Li-Lee [Li et al (1993)] (página 46)
- Método de Kittler-Illingworth [Sahoo et al (1988)] (página 47)
- Método de Otsu [Otsu (1979)] (página 48)
- Método por Pixels de Borda [Weszka et al (1974)] (página 49)
- Método por Porcentagem [Wilson et al (1988)] (página 49)
- Método por Seleção iterativa de Thrussel [Thrussel (1979)] (página 50)
- Método por Valor Médio [Parker (1996)] (página 50)

4.5.1 Método por Abordagem Fuzzy

Este método baseia-se na lógica Fuzzy onde é possível se determinar a probabilidade de um pixel de pertencer a uma das duas classes, ou ao objetos ou ao segundo plano. Inicialmente deve-se calcular a função de "classificação" que representa a probabilidade de um pixel pertencer a cada uma das classes. Para isso assume-se que μ_0 corresponde à média dos níveis de cinza do fundo e μ_1 à média dos níveis de cinza do primeiro plano. O pixel pertencerá à classe da qual seu nível mais se aproximar. Quanto menor a diferença entre o nível de cinza g de um pixel e o valor da média da classe a que pertence, maior será a probabilidade $\mu_g(g)$ de classificação.

A função de classificação é:

$$\mu_g(g) = \begin{cases} \frac{1}{1+|g-\mu_0|/C} & \text{se } g < t \\ \frac{1}{1+|g-\mu_1|/C} & \text{se } g \geq t \end{cases} \quad (4.1)$$

Um valor estimado para ambos μ_0 e μ_1 é necessário. Para o limiar t dado, obtém-se:

$$\begin{aligned} \text{para o fundo: } \mu_0(t) &= \frac{\sum_{g=0}^t gh(g)}{\sum_{g=0}^t h(g)} \\ \text{para os objetos: } \mu_1(t) &= \frac{\sum_{g=t+1}^{255} gh(g)}{\sum_{g=t+1}^{255} h(g)} \end{aligned}$$

onde $h(g)$ identifica o nível do histograma para determinado nível da escala de cinza. Ambos valores dependem do limiar t . O valor de C é calculado como uma constante: a diferença entre o

máximo e o mínimo nível de cinza na imagem.

Cada pixel, cujo nível de cinza é g , fará parte do fundo ou da classe de objetos que compõem o objetos dependendo da sua relação entre seu nível de cinza do pixel e o limiar t . Para um pixel de objeto onde $(g > t)$, a sua probabilidade de pertencer ao objeto é dada por $\mu_g(g)$ e deve estar situada entre os valores $1/2$ e 1 .

Os dois algoritmos apresentados aqui fazem uso do conceito de "fuzziness" que corresponde à distância entre a imagem original em níveis de cinza e a imagem binarizada correspondente (HUANG 1995). O fuzziness tem como valor mínimo 0 e como valor máximo 1. Esse algoritmos especificam que quanto menor o fuzziness mais preciso é o valor do limiar .

4.5.1.1 Método de Huang

Calculada a função de classificação, um primeiro meio de medir a fuzziness [Huang et al(1995)] é se basear em entropia que é calculada usando a função de Shannon.

$$H_f(x) = -x \log(x) - (1-x) \log(1-x) \quad (0 < x < 1) \quad (4.2)$$

O valor $E(t)$ da entropia do conjunto Fuzzy que constitui a imagem é portanto:

$$E(t) = \frac{1}{nm} \sum_g H_f[\mu_g(g)]h(g) \quad (4.3)$$

onde g representa o nível de cinza do pixel, $h(g)$ identifica o nível do histograma para determinado nível da escala de cinza, n representa o número de linhas da imagem e m representa o número de colunas da imagem.

Esta função é baseada em t porque μ_g também baseia-se em t . O valor de limiar t apropriado é obtido por minimização da fuzziness, na prática obtida pela minimização da função $E(t)$.

4.5.1.2 Método de Yager

Um outra forma para se medir o fuzziness baseia-se na idéia para um plano A onde não existem elementos em comum entre A e seu complemento A^c . O nível em que A e A^c são indistintos é a medida de fuzziness em A [Yager (1979)]. Pode ser calculada pôr esta expressão:

$$D_p(t) = [\sum_g |\mu_g(g) - \overline{\mu}_g(g)|^p]^{1/p} \quad (4.4)$$

para os níveis g , onde p é um inteiro e $\overline{\mu}_g(g) = 1 - \mu_g(g)$. O valor de p é um inteiro usado para definir a distância medida; sendo $p = 2$, igual a distância Euclidiana.

Novamente, o valor de limiar t apropriado é obtido por minimização da fuzziness, na prática obtida pela minimização da função $Dp(t)$, ou $t = \text{Min}Dp(t)$.

4.5.2 Método por Dois Picos

O método por dois picos ou cumes ainda remonta ao estudo do histograma em si [Parker (1996)]. Uma observação, que é freqüentemente feita, é que quando um limiar é obvio, acontece sempre no ponto mais baixo entre dois cumes adjacentes no histograma. Se o histograma possuir dois cumes, então o vale entre os dois cumes apresenta-se como um valor de limiar ótimo.

O processo para se descobrir dois picos ou cumes adjacentes, e então se determinar o valor de limiar, deve seguir os seguintes procedimentos. Deve-se achar o primeiro pico que se reporta como a faixa do histograma que possui o maior valor (maior número de pixels com a mesma tonalidade de cinza). O segundo pico será a faixa de maior valor de pixels localizada à direita da primeira, obtido pela multiplicação dos valores do histograma pelo quadrado da distância do primeiro pico. Assim, se o pico maior está no nível de cinza j do histograma, a seleção do segundo pico g é obtida buscando:

$$\text{Max}\{((g - j)^2 h(g) \mid (0 \leq g \leq 255))\} \quad (4.5)$$

onde $h(g)$ identifica o nível do histograma para o nível da escala de cinza g .

4.5.3 Método por Entropia de Pun

A entropia é uma medida de conteúdo de informação. Assim, em termos de teoria de informação, assumindo que existem n possíveis símbolos x (pôr exemplo, cartas ou dígitos) e que um símbolo i vai acontecer com probabilidade $p(x_i)$. Então a entropia associada é dada por:

$$H(t) = - \sum_{i=0}^n p(x_i) \log(p(x_i)) \quad (4.6)$$

Extrapolando-se o conceito acima; uma imagem pode ser vista como uma fonte de símbolos, ou níveis de cinza. A entropia associada, para uma imagem com 256 níveis de cinza, com os pixels pretos tendo sido limiarizada utilizando-se o limiar t , é obtida por:

$$H_b(t) = - \sum_{i=0}^t p_i \log(p_i) \quad (4.7)$$

De forma análoga, a entropia de pixels brancos é dada por:

$$H_w(t) = - \sum_{i=t+1}^{255} p_i \log(p_i) \quad (4.8)$$

onde $p_i = \frac{h(i)}{N}$, N : número total de pixels e $h(i)$ identifica o nível do histograma para o nível da escala de cinza i .

O algoritmo sugerido por Pun [Pun (1981)] tenta achar o limiar t que maximiza a função $H = H_b + H_w$. Pun demonstrou que basta maximizar a função $f(t)$ abaixo descrita:

$$f(t) = \frac{H_t}{H_T} \frac{\log(P_t)}{\log(\text{Max}\{p_0, p_1, \dots, p_t\})} + [1 - \frac{H_t}{H_T}] \frac{\log(1 - P_t)}{\log(\text{Max}\{p_{t+1}, p_{t+2}, \dots, p_{255}\})} \quad (4.9)$$

com

$$H_t = - \sum_{i=0}^t p_i \log(p_i) \quad e \quad H_T = - \sum_{i=0}^{255} p_i \log(p_i) \quad e \quad P_t = \sum_{i=0}^t p_i \quad (4.10)$$

onde H_t , 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 .

4.5.4 Método por Entropia de Kapur

O Método de 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 p_i de distribuição B do fundo da seguinte forma:

$$\begin{aligned} A &: \frac{p_0}{P_t}, \frac{p_1}{P_t}, \dots, \frac{p_t}{P_t} \\ B &: \frac{p_{t+1}}{(1 - P_t)}, \frac{p_{t+2}}{(1 - P_t)}, \dots, \frac{p_{255}}{(1 - P_t)} \end{aligned} \quad (4.11)$$

onde P_t representa a probabilidade cumulativa até o nível de cinza t definida na equação 4.10.

As entropias dos pixels pretos H_b e dos pixels brancos H_w das equações 4.7 e 4.8 são computadas baseando-se nas novas distribuições e são obtida por:

$$H_b(t) = - \sum_{i=0}^t \frac{p_i}{P_t} \log\left(\frac{p_i}{P_t}\right) \quad (4.12)$$

e

$$H_w(t) = - \sum_{i=t+1}^{255} \frac{p_i}{1 - P_t} \log\left(\frac{p_i}{1 - P_t}\right) \quad (4.13)$$

O valor do limiar ótimo é o valor de t que maximiza a expressão $H = H_b(t) + H_w(t)$. Para se encontrar este valor, todos limiar es de valores entre 0 e 255 devem ser testados e aquele que resultar no maior valor de H é o valor limiar escolhido.

4.5.5 Método por Entropia de Johannsen

O Método de Johannsen [Johannsen et al (1982)] baseia-se também, na divisão dos nível de cinza da imagem em duas partes a fim de minimizar a interdependência entre elas. Para se determinar o limiar devem ser calculados os valores de $S_b(t)$ e $S_w(t)$.

$$\begin{aligned} S_b(t) &= \log\left(\sum_{i=0}^t p_i\right) - \frac{1}{\sum_{i=0}^t p_i} [p_t \log(p_t) + (\sum_{i=0}^{t-1} p_i) \log(\sum_{i=0}^{t-1} p_i)] \\ S_w(t) &= \log\left(\sum_{i=t}^{255} p_i\right) - \frac{1}{\sum_{i=t}^{255} p_i} [p_t \log(p_t) + (\sum_{i=t+1}^{255} p_i) \log(\sum_{i=t+1}^{255} p_i)] \end{aligned} \quad (4.14)$$

O valor do limiar ótimo é o valor de t que minimiza a expressão $S_b(t) + S_w(t)$. Deve-se observar que os valores calculados para $S_b(t)$ e $S_w(t)$ devem ser desprezados para valores de t quando $pt = 0$.

4.5.6 Método por Entropia de Li-Lee

A entropia cruzada mede a distância teórica da informação entre duas distribuições $P = \{p_1, p_2, \dots, p_n\}$ e $Q = \{q_1, q_2, \dots, q_n\}$, ($0 \leq n \leq N$) pôr:

$$D(Q, P) = - \sum_{k=0}^N q_k \log_2 \left(\frac{q_k}{p_k} \right) \quad (4.15)$$

O método da entropia mínima cruzada pode ser visto como uma extensão do método da entropia máxima pela atribuição de estimativas iniciais para todas as probabilidades quando nenhuma informação anterior está disponível.

Trabalhos aplicando o método de entropia máxima em segmentação de imagem consideram o processo de geração de pixels como tentativas independentes. Eles usam o histograma normalizado de níveis de cinza como as distribuições de probabilidade de níveis de cinza baseados em tentativas aleatórias de pixels individuais que tem um certo nível de cinza e medem a entropia da distribuição dos pixels.

No esquema proposto por [Li et al (1993)], a imagem binarizada $g(x, y)$ será construída a partir de $f(x, y)$ da seguinte forma:

$$g(x, y) = \begin{cases} \mu_1 & \text{se } f(x, y) < t \\ \mu_2 & \text{se } f(x, y) \geq t \end{cases} \quad (4.16)$$

A imagem binarizada $g(x, y)$ é unicamente determinada a partir de $f(x, y)$ pela especificação de três parâmetros desconhecidos: t , μ_1 e μ_2 . Uma função critério η deve ser construída para possibilitar o encontro do $g(x, y)$ ótimo ou equivalentemente o conjunto ótimo de parâmetros t , μ_1 e

μ_2 que assemelha-se a $f(x, y)$ o máximo possível. Ou seja, $\eta[g(x, y)] \equiv \eta(t, \mu_1, \mu_2)$.

A função critério η é geralmente algum tipo de medida de distorção. Neste método é utilizada a entropia cruzada, que é a medida preferida para distribuições positivas e aditivas. Li-Lee definiram a função critério η da seguinte maneira:

$$\eta(t) = \sum_{g=1}^{t-1} gh(g) \log_2\left(\frac{g}{\mu_1(t)}\right) + \sum_{g=t}^{255} gh(g) \log_2\left(\frac{g}{\mu_2(t)}\right) \quad (4.17)$$

onde μ_1 e μ_2 são definidos como segue:

$$\mu_1(t) = \frac{\sum_{g=0}^{t-1} gh(g)}{\sum_{g=0}^{t-1} h(g)} \quad e \quad \mu_2(t) = \frac{\sum_{g=t}^{255} gh(g)}{\sum_{g=t}^{255} h(g)}$$

com g o nível de cinza da imagem e $h(g)$ o valor do histograma para cada nível de cinza.

O valor do limiar ótimo é então selecionado para o valor de t que minimiza a função critério $\eta(t)$.

4.5.7 Método de Kittler-Illingworth

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

$$p(g) = \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-((g-\mu_1)^2/2\sigma_1^2)} + \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-((g-\mu_2)^2/2\sigma_2^2)} \quad (4.18)$$

onde os parâmetros μ_i e σ_i^2 , com $i = 1, 2$, representam respectivamente a média e o desvio padrão de cada classe.

Tomando o logaritmo da equação 4.18, e assumindo que a probabilidade a priori de cada classe é p_i , o valor do limiar t representa a solução da equação quadrática em g :

$$\frac{(g - \mu_1)^2}{\sigma_1^2} + \log(\sigma_1) - 2\log(p_1) = \frac{(g - \mu_2)^2}{\sigma_2^2} + \log(\sigma_2) - 2\log(p_2) \quad (4.19)$$

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

$$J(t) = 1 + 2(p_1(t)\log(\sigma_1(t)) + p_2(t)\log(\sigma_2(t))) - 2(p_1(t)\log(p_1(t)) + p_2(t)\log(p_2(t))) \quad (4.20)$$

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

$$\begin{aligned} p_1(t) &= \sum_{g=0}^t h(g) & e & & p_2(t) &= \sum_{g=t+1}^{255} h(g) \\ \mu_1(t) &= \frac{\sum_{g=0}^t gh(g)}{p_1(t)} & e & & \mu_2(t) &= \frac{\sum_{g=t+1}^{255} gh(g)}{p_2(t)} \\ \sigma_1^2(t) &= \frac{\sum_{g=0}^t h(g)(g - \mu_1(t))^2}{p_1(t)} & e & & \sigma_2^2(t) &= \frac{\sum_{g=t+1}^{255} h(g)(g - \mu_2(t))^2}{p_2(t)} \end{aligned}$$

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 $J(t)$.

4.5.8 Método de Otsu

Este método baseia-se na análise discriminante [Otsu (1979)]. A operação de binarização é obtida pelo particionamento dos pixels de uma imagem em duas classes C_0 e C_1 (isto é, objetos e fundo) no nível de cinza t . Assim, $C_0 = \{0, 1, \dots, t\}$ e $C_1 = \{t + 1, t + 2, \dots, l - 1\}$. Podemos definir as características das classes C_0 e C_1 em termos de distribuições de Gauss, seja a partir de suas médias $\mu_0(t)$, $\mu_1(t)$, das variâncias intra-classes $\sigma_W^2(t)$, inter-classes $\sigma_B^2(t)$ e da variância total σ_T^2 .

Essas médias e variâncias podem ser calculadas da seguinte maneira:

$$\begin{aligned} \mu_0(t) &= \frac{\mu_t(t)}{w_0(t)} & e & & \mu_1(t) &= \frac{\mu_t - \mu_t(t)}{1 - w_0(t)} \\ \sigma_B^2(t) &= w_0(t)w_1(t)[\mu_0(t) - \mu_1(t)]^2 \\ \sigma_t^2 &= \sum_{g=0}^{255} (g - \mu_t)^2 p_g & e & & \sigma_W^2(t) &= \sigma_t^2 - \sigma_B^2(t) \end{aligned}$$

onde

$$\begin{aligned} w_0(t) &= \sum_{g=0}^t p_g & e & & w_1(t) &= \sum_{g=t+1}^{255} p_g = 1 - w_0(t) \\ \mu_t(t) &= \sum_{g=0}^t gp_g & e & & \mu_T &= \sum_{g=0}^{255} gp_g \\ p_g &= \frac{h(g)}{N} & , & & N &: \text{número total de pixels} \end{aligned}$$

e $h(g)$ o valor do histograma para cada nível de cinza g .

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

$$t^* \Rightarrow \text{Max}_{t \in [0,255]} [\eta(t)] = \text{Max}_{t \in [0,255]} \left[\frac{\sigma_B^2(t)}{\sigma_T^2} \right] \quad (4.21)$$

4.5.9 Método por Pixels de Borda

Um pixel de borda é definido como um pixel próximo do limite entre objetos e o fundo, ou entre dois objetos. Um histograma baseado nos níveis de cinza dos pixels de borda será mais regular que o histograma da imagem, devido a esses níveis dos pixels de borda serem mais relevantes que os níveis globais. Esta idéia foi usada [Weszka et al (1974)] para produzir um método de binarização baseado no Método Laplaciano digital, onde um operador não direcional de detecção de pixels de borda é utilizado. O limiar é achado computando-se o Laplaciano da imagem de entrada. Existem muitos modos de se obter o valor do Laplaciano mas o mais simples é convoluir a imagem com a seguinte máscara:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (4.22)$$

É possível obter-se um histograma da imagem original onde se deve considerar apenas os pixels obtidos pela interação da máscara. Estes pixels devem possuir um valor de Laplaciano grande, situados nos 85% percentil superior. Os pixels que têm um Laplaciano maior que 85% terão o nível de cinza mostrados no histograma, enquanto todos os outros pixels, que não compõem as bordas dos objetos, não o serão. O limiar é selecionado usando o histograma computado assim.

Usando uma aproximação melhor para o Laplaciano, os autores afirmam que obtém-se um melhor resultado. O algoritmo requer a passagem de um parâmetro externo que induz o valor de porcentagem para seleção dos valores do Laplaciano.

4.5.10 Método por Porcentagem

Este método de binarização foi um dos primeiros desenvolvidos [Wilson et al (1988)]. O método por porcentagem ou também conhecido por p-tile (escada) atua diretamente sobre o número de pixels de um histograma. Dado um histograma e a porcentagem de pixels pretos desejados, determina-se

o número de pixels pretos multiplicando-se a porcentagem pelo número total de pixels. Conta-se os pixels em níveis de cinza do histograma, começando pelo nível de cinza 0 (preto), até que a contagem seja maior ou iguale-se ao número desejado de pixels pretos. O limiar representa o último nível de cinza computado.

Para o método ser processado, há necessidade de se fornecer ao método a porcentagem desejada, que pode variar de 0 a 100

4.5.11 Método por Seleção interativa de Thrussel

O Método de Thrussel [Thrussel (1979)] ou Seleção Iterativa é um processo no qual um valor inicial arbitrário de limiar é refinado pôr consecutivas varreduras sobre a imagem. A idéia do processo é de não se utilizar o histograma da imagem, mas separá-lo nas classes objeto e fundo repetidamente, utilizando os níveis de cada classe para aprimorar o limiar. O valor inicial é calculada a partir da coleta de estatísticas sobre regiões claras e escuras. A média dos níveis de cinza para os pontos abaixo do limiar é encontrada e chamada de T_b , e a média para os pontos acima do limiar é chamada de T_o . Com estes dois parâmetros calculados, o novo limiar T_k é dado pela seguinte equação:

$$T_k = (T_b + T_o)/2 \quad (4.23)$$

Partindo da avaliação inicial do valor do limiar T_0 , a k -ésima avaliação do limiar T_k é, então, obtida pela seguinte equação:

$$T_k = \frac{\sum_{g=0}^{T_{k-1}} gh(g)}{2 \sum_{g=0}^{T_{k-1}} h(g)} + \frac{\sum_{g=T_{k-1}+1}^{255} gh(g)}{2 \sum_{g=T_{k-1}+1}^{255} h(g)} \quad (4.24)$$

onde $h(g)$ é o histograma da imagem para o nível de cinza g .

O processo repete-se até que não haja alteração de T_k , entre um passo e outro (quando $T_k = T_{k+1}$). O limiar T_k é, então, o valor ótimo.

4.5.12 Método por Valor Médio

O valor de limiar T obtido pelo método por valor médio [Parker (1996)] é a média dos valores dos níveis de cinza na imagem (I):

$$T = \frac{1}{nm} \sum_{g=0}^{255} h(g) \quad (4.25)$$

onde $h(g)$ identifica o nível do histograma para o nível da escala de cinza g , n representa o número de linhas da imagem e m representa o número de colunas da imagem.

Os resultados obtidos não podem ser de forma geral satisfatórios pois a imagem pode conter um desequilíbrio grande em termos de contraste e em termos de distribuição de níveis de cinza que o formalismo matemático não leva em consideração.

4.6 MÉTODOS LOCAIS OU ADAPTATIVOS

4.6.1 Introdução

Um modo comum para executar a binarização local é pela análise das intensidades dos níveis de cinza com a passagem de uma janela local pela imagem para a determinação dos limiares locais [Casey (1994)] [Kamel et al (1993)] [Wong (1978)].

O problema principal com qualquer técnica de binarização adaptativa é a escolha de tamanho de janela. O tamanho da janela escolhida deve ser grande bastante para garantir que o número de pixels de fundo incluídos é bastante grande para se obter uma estimativa boa de valor comum mas não tão grande como a média sobre as intensidades de fundo não uniformes. Porém, as características na imagem variam freqüentemente em tamanho tal que há problemas com o tamanho fixo de uma janela. Para resolver este problema, pode ser usada a informação domínio-dependente para conferir que os resultados da binarização dão as características esperadas (por exemplo, uma bolha grande de uma região não é esperada numa página de símbolos pequenos). Se o resultado é inesperado, então o tamanho de janela pode ser modificado e a binarização aplicada novamente. São métodos locais:

- Método de Bernsen [Bernsen (1986)] (página 51)
- Método por Médias Móveis [Parker (1996)] (página 52)
- Método por Relaxação de Peleg [Peleg (1980)] (página 53)

4.6.2 Método de Bernsen

No método de Bernsen [Bernsen (1986)], para cada pixel (x, y) , o limiar $T(x, y) = (Z_{min} + Z_{max})/2$ é utilizado, onde Z_{min} e Z_{max} são os valores dos níveis máximos e mínimos da escala de cinzas em uma vizinhança quadrada de dimensões $r \times r$ centralizada em (x, y) .

Porém, se o contraste $C(x, y)$ no pixel (x, y) encontra-se como $C(x, y) = (Z_{max} - Z_{min}) < v$, o quer dizer que o contraste é inferior a um valor preestabelecido v então a vizinhança constitui-se apenas de uma classe, objetos ou fundo. Essa decisão pode ser tomada em função da média total M_T dos níveis de cinza na imagem. Se o valor do nível de cinza do ponto (x, y) é inferior à média

total M_T , esse ponto será binarizado como preto. Caso contrário, esse ponto será binarizado como branco.

4.6.3 Método por Médias Móveis

Segundo o autor [Parker (1996)], este método obtém um limiar muito rápido por pixel, e produz boas segmentações. O método foi desenvolvido especialmente para imagens que contêm texto como, por exemplo, documentos escaneados. A premissa inicial para a aplicação do método é que a imagem original possua uma boa qualidade de iluminação.

Uma média móvel é apenas a média dos níveis de cinza dos últimos n pixels observados. A imagem pode ser tratada como se fosse um fluxo uni-dimensional de pixels e a média ou pode ser computada exatamente ou pode ser estimada por:

$$M_{i+1} = M_i - \frac{M_i}{n} + g_{i+1} \quad (4.26)$$

onde M_{i+1} é o valor estimado da média móvel no pixel $i + 1$ possuindo o nível de cinza g_{i+1} e M_i a média móvel prévia para o pixel i .

Qualquer pixel que tenha menos que a porcentagem fixa de sua média móvel é convertido para preto e, caso contrário é convertido para branco. Para evitar uma predeterminação para a conversão do pixel para branco ou preto, um método de verificação chamado boustrophedon (do grego, significando "como o arado do boi") é empregado.

Este método verifica os pixels em direção oposta em toda outra linha. Isso é, o pixel que segue o último em linha i é o último na linha $i + 1$, seguido pelo segundo último na linha $i + 1$, e assim por diante do fim para o começo de linha $i + 1$; isto é seguido pelo pixel de começo em linha $i + 2$, e assim procedendo-se até o final. Isto evita a descontinuidade ao término da linha. O processo começa com uma estimativa da média móvel; um valor de $127 \times n$ é selecionado, e isto afetará alguns poucos pixels iniciais na imagem. O valor de n utilizado remonta ao número de colunas dividido por 8.

A equação 4.26 é empregada para avaliar o valor da média móvel do próximo pixel que é imediatamente usado para binarizar da seguinte maneira:

$$T = \begin{cases} 0 & \text{se } g_i < \frac{M_i}{n} (1 - \frac{porc}{100}) \\ 255 & \text{caso contrário} \end{cases} \quad (4.27)$$

onde T representa o pixel binarizado e $porc$ representa um valor fixo de porcentagem, que o autor definiu com $porc = 15$.

4.6.4 Método por Relaxação de Peleg

O Método da Relaxação é um processo iterativo [Peleg (1980)]. Para o problema específico da binarização de uma imagem, os limiares para uma iteração k são computados em função da mesma vizinhança na iteração prévia $k - 1$ seguido o seguinte esquema:

- Inicialmente, deve-se partir de uma hipótese para a segmentação da imagem
- Para cada pixel, é executada uma modificação da segmentação e de um valor de confiança baseados na localização regional do pixel. A modificação é baseada na vizinhança dos oito pixels que rodeiam o pixel em exame.
- A repetição do processo deve ser feita até a segmentação total estar completa. Isto acontece quando não mais são vistas mudanças adicionais em passos sucessivos.

As estimativas de confiança possuem a aparência das probabilidades. Um modo para se achar uma classificação inicial, é usar o nível de cinza médio como um valor padrão. Segundo o autor, um pixel com valor maior que o valor médio tem uma probabilidade de ser branco quando a distância relativa de seu nível de cinza está na proporção dos 3/4 do intervalo total dos níveis de cinza. Para um pixel com valor menor que o médio, utiliza-se o 1/4 do intervalo dos níveis de cinza.

Assim, uma possibilidade para a classificação inicial é dada pela equação de Rosenfeld:

$$\begin{aligned} \text{para um pixel maior que a média: } p_i^0 &= \frac{1}{2} \left[1 + \frac{g_i - \mu}{max - \mu} \right] \\ \text{para um pixel menor que a média: } q_i^0 &= \frac{1}{2} \left[1 + \frac{\mu - g_i}{\mu - min} \right] \end{aligned} \quad (4.28)$$

onde μ , max , min e g_i são respectivamente o valor médio, o maior valor de nível de cinza, o menor valor de nível de cinza, e o nível de cinza para o pixel i .

O valor p_i^0 é a probabilidade inicial que o pixel i é branco, e q_i^0 representa a probabilidade inicial que o pixel i é preto. O índice 0 se refere à iteração que é inicialmente zero.

O problema principal resume-se em; dado que as probabilidades iniciais de que o pixel e sua vizinhança possa ser branco ou preto, como refinar as probabilidades? Idealmente estas probabilidades deveriam se reverter a 1 ou 0, fornecendo uma segmentação final correta. O que é necessário se obter é alguma medida de compatibilidade que pode ser usada para decidir se uma classificação particular é razoável ou não. Por exemplo, se um pixel é preto e toda sua vizinhança

é composta de pixels brancos, pareceria mais provável que o pixel preto deva ser modificado assumindo a cor branca. A compatibilidade deste pixel com seus vizinhos é baixa, sugestionando uma mudança.

A compatibilidade é calculada pela função $C(i, c_1, j, c_2) = [I(i, c_1) - I(j, c_2)]/256$ que resulta em um valor entre -1 e 1 , de quanto o pixel i é compatível (que é pertencente à classe c_1) com o pixel j (que é pertencente à classe c_2) na imagem inicial I . Para a questão da binarização existem apenas duas classes, preto ou branco, e para uma vizinhança pequena os pixels i e j serão adjacente entre si. Não é possível saber o que esta função deveria ser com certeza, porque depende das probabilidades achadas na imagem limiarizada final, o que não é desconhecido. Porém, um implementação simples teria $C = 1$ quando $c_1 = c_2$, e $C = -1$ caso contrário; isto é, os pixels serão compatíveis.

Uma vez que é possível se ter duas classes possíveis para qualquer pixel, a média dessas poderia ser usada como uma compatibilidade global entre quaisquer dois pixels:

$$Q_{ij} = C(i, c_1, j, \text{branco})p_j + C(i, c_1, j, \text{preto})q_j \quad (4.29)$$

A compatibilidade de uma região ao redor do pixel i pode ser definida como a compatibilidade comum de todos os oito pixels vizinhos:

$$Q_i(c_1) = 1/8 \sum_{j \in N} Q_{ij} = 1/8 \sum_{j \in N} C(i, c_1, j, \text{branco})p_j + C(i, c_1, j, \text{preto})q_j \quad (4.30)$$

para a vizinhança N de um pixel centralizada em i . Este será o incremento para p_i a cada vez que as probabilidades são atualizadas. Porém, assegurar que os valores de p_i e q_i permanecem positivos, deve-se somar 1 ao valor de Q_i . Assim, os valores devem ser normalizados sobre a região. A seguinte equação fornece a normalização:

$$p_i^{k+1} = \frac{p_i^k(1 + Q_i^k)}{p_i^k(1 + Q_i^k(\text{branco})) + q_i^k(1 + Q_i^k(\text{preto}))} \quad (4.31)$$

onde o sobrescrito reflete o número da iteração. Uma expressão semelhante calcula os valores de q_i .

Cada repetição do processo de relaxamento envolve todos os pixels da imagem atualizando os valores de p_i e q_i . Uma vez que um valor de p_i se modifica para 0 ou 1, ele permanece desta forma; assim, a classificação inicial é muito importante ao sucesso do método. De fato, os valores dos pixels atuais nunca são examinados depois da classificação inicial estar completa; todas as

outras iterações são executadas baseadas nas probabilidades.

No presente Capítulo foram comentadas as metodologias de segmentação por binarização utilizadas neste trabalho. No próximo Capítulo serão discutidas as estratégias de avaliação de qualidade adotadas e os resultados obtidos por ensaio.

AVALIAÇÃO ADOTADA E

Capítulo 5

ESTRATÉGIA DE AVALIAÇÃO DE QUALIDADE ADOTADA E RESULTADOS

5.1 INTRODUÇÃO

Neste capítulo serão descritos os procedimentos de aplicação dos métodos de binarização sobre os cheques bancários brasileiros assim como os resultados obtidos. Será também descrita a estratégia adotada para avaliar a qualidade desses resultados a partir de um *OCR* (do inglês Optical Character Recognition) e serão apresentadas as tabelas das taxas de reconhecimento obtidas por banco e por algoritmo. É bom ressaltar que, para uma avaliação rigorosa, as imagens submetidas aos processos de binarização não foram previamente filtradas.

5.2 QUALIDADE DA IMAGEM

A definição de qualidade de uma imagem abrange áreas muito diversas. Para uma imagem fotográfica, características como cor, contraste, granulação, sensibilidade e poder de resolução do filme fotossensível são alguns dos fatores determinantes. [Langford (1984)] define a qualidade de uma imagem fotográfica como aquela que além de observar uma composição fotográfica adequada, obedecer às técnicas de processamento e laboratório, agrada visualmente aos olhos. No âmbito de uma imagem digital, estes preceitos básicos da fotografia, como foco, contraste, equilíbrio de cores, etc., são aliados valiosos para se estabelecer um critério para a definição da qualidade de uma imagem.

Porém, como ainda descrito em [Langford (1984)] , o fator mais poderoso para a definição da qualidade reside na interpretação humana do resultado. Isto torna extremamente difícil estabelecer padrões específicos para a determinação dessa "qualidade" em uma imagem. No âmbito dos cheques bancários brasileiros, pode-se partir do pressuposto que a imagem segmentada do canal de intensidade (sistema *HSI*) apresenta uma boa qualidade em termos de contraste de níveis de cinza, foco e granulação (dada pela quantidade de DPI's da imagem digitalizada). Já no tocante à imagem binarizada, a qualidade restringe-se à possibilidade de leitura dos caracteres obtidos após a operação, uma vez que os fatores de contraste, cor, granulação, etc., são eliminados da imagem. Isto pode ser visto na figura 5.1 onde é apresentado um cheque bancário binarizado por dois métodos diferentes.

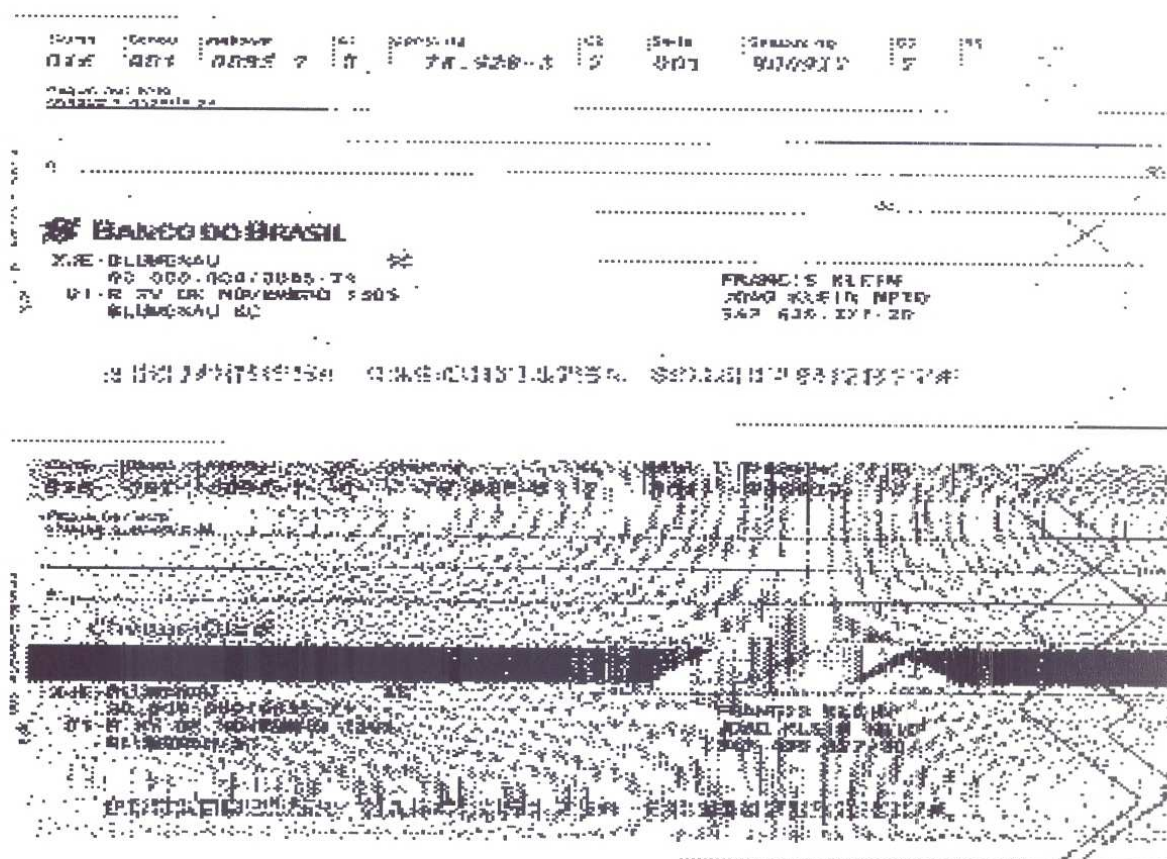


Figura 5.1: Cheque bancário binarizado por dois métodos diferentes

A qualidade da imagem binarizada de um cheque bancário fica associada a dois componentes

básicos:

- O processo de segmentação por binarização em si: a extração das informações relevantes fica subordinada ao método utilizado, pois alguns métodos não produzem o resultado esperado (algumas informações não são totalmente eliminadas);
- O fator de legibilidade dos caracteres: o ato da leitura fica subordinada à facilidade de reconhecer os caracteres segmentados.

Em relação a esse segundo ponto, [Trie et al (1995)] aconselha o estabelecimento de dois critérios para a validação de resultados de imagens obtidas por binarização:

- O primeiro proveniente da avaliação pela visão humana: o ser humano considera que a imagem binária é de boa qualidade quando ele consegue ler o que está escrito. Porém esse critério se torna difícil de ser empregado por três razões: o grande volume a ser processado, a subjetividade de discernimento do ser humano e principalmente a impossibilidade de quantificação do reconhecimento;
- O segundo proveniente da avaliação computacional: reconhecer caracteres faz parte da área de Reconhecimento de Padrões, onde muitas abordagens já foram desenvolvidas. As grandes vantagens na aplicação desse critério residem no fato de não introduzir o fator subjetividade de discernimento, na possibilidade de automatização do processo e principalmente na possibilidade de quantificar o reconhecimento.

5.3 METODOLOGIA

Muitas abordagens de reconhecimento já foram desenvolvidas. Devido a não dispor de nosso próprio método de reconhecimento e pela necessidade de um sistema que requer imagens binárias como dados de entrada, concluiu-se que o sistema mais propício seria um sistema ótico de reconhecimento de caracteres conhecido por *OCR*. Dentro do âmbito comercial e pela facilidade de obtenção, foi utilizado o Calera Word Scan versão 4.2 para a validação dos resultados da segmentação por binarização.

Devido às grandes variações de padrões encontrados na base de dados, devemos reconhecer que a escolha da metodologia adotada para a apresentação dos resultados foi complexa. Segue abaixo uma lista dos itens que foram levantados para justificar a abordagem adotada:

- | | | | | | | | | | |
|-------|-------|---------|----|----------|----|-------|-----------|----|----|
| Comp. | Banco | Agência | CI | Conta ng | C2 | Serie | Cheque ng | CA | Br |
| 009 | 001 | 0009-4 | 5 | 522.0605 | | 100 | 008106 | 2 | |

Pague por este
cheque a quantia de

R\$

ou à sua ordem

de 19

BANCO DO BRASIL

XE4 CURITIBA-CENTRO PR
00.000.000/0009-49
06 MOD FUNC1 7 AND ALA A
PCA TIRADENTES

3069640902120

59

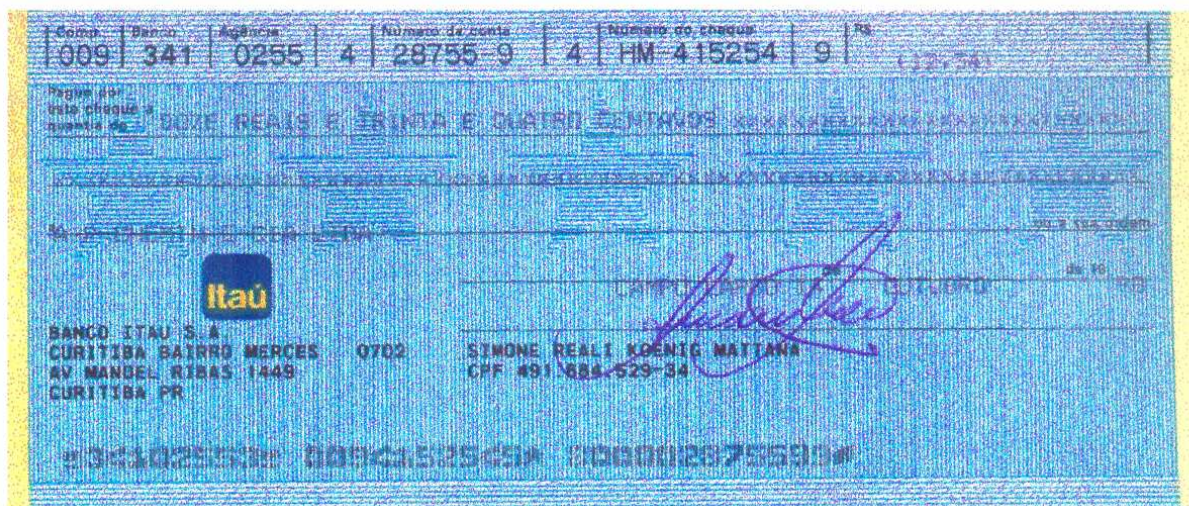


Figura 5.3: Cheque bancário do Banco Banestado S.A.

Baseando nos itens anteriores, a metodologia adotada para a apresentação dos resultados será ensaiada em um cheque de média representatividade, preenchido à máquina sem erros de posicionamento, com fundo artístico complexo e sem características peculiares de fontes (como por exemplo no caso do Banco *HSBC*). OS resultados serão visualizados a partir de um cheque do Banco Itaú S.A. (figura 5.4).

A metodologia empregada consiste em submeter a imagem binarizada do cheque ao *OCR* e em calcular o número de dígitos e caracteres reconhecidos. Esse cálculo foi realizado de forma manual, verificando caractere a caractere o resultado do *OCR* comparativamente com a leitura humana do cheque original. Para explicar com clareza as informações que foram realmente computadas, há necessidade de se fazer um levantamento das informações contidas num cheque. As informações relevantes e suscetíveis de serem reconhecidas por um *OCR*, chamadas informação tipo *R*, são:

- dados de identificação dos itens do cabeçalho ("comp.", "banco", "agência", "C1" ou "DV", "número da conta" ou "conta", "C2", "série", "número do cheque" ou "cheque No", "C3", "R\$");
- dados de identificação do extenso ("pague por este cheque a quantia de", "a", "ou a sua ordem");
- dados de identificação do local e data ("de", "de 19" ou "de 199");



a)



b)

Figura 5.4: Cheque do Banco Itaú S.A.: a) colorido b) canal de Intensidade

- dados de identificação da agência;
- dados de identificação do cliente;
- caracteres de segurança usados para delimitar campos de preenchimento e adicionados na hora de preencher o cheque (por exemplo "(", "#", "*").

Além disso, existem informações importantes que não podem ser reconhecidas por um *OCR* de forma correta, chamadas informação tipo *NR*, como:

- separadores dos dados de identificação dos itens do cabeçalho que vão de simples linhas verticais ou horizontais até retângulos;
- linhas guias de preenchimento;
- logotipo do banco;
- linha MICR CMC-7 (código de barra);
- linhas de cruzamento indicando depósito.

Como explicado em [Trie et al (1995)], diferenciou-se o reconhecimento correto do incorreto da seguinte maneira:

- a informação tipo *R* corretamente reconhecida;
- a informação tipo *R* incorretamente reconhecida;
- a informação tipo *NR* corretamente reconhecida;
- a informação tipo *NR* incorretamente reconhecida;
- a informação tipo *NR* reconhecida como uma informação tipo *R*.

Para cada cheque submetido ao *OCR*, foram computados somente os dois primeiros itens. Chamando o número total de itens tipo *R*, o número de itens tipo *R* corretamente reconhecidos e o número de itens tipo *R* incorretamente reconhecidos como Num_{Total} , Num_{Rec} e Num_{IRec} respectivamente, obtém-se a simples relação $Num_{Rec} = Num_{Total} - Num_{IRec}$.

Na metodologia, foram avaliados três taxas de reconhecimento:

- a taxa de cada cheque, TX_{CH} , calculada na forma de porcentagem com a seguinte regra:

$$TX_{CH} = 100 \times Num_{Rec} / Num_{Total};$$

- a taxa média por banco versus algoritmo, TX_{MBA} , calculada com a seguinte regra:

$$TX_{MBA} = \sum \frac{TX_{CH}}{Num_{TotalChequesdoBanco}};$$

- a taxa global por algoritmo, TX_{GA} , calculada com a seguinte regra:

$$TX_{MBA} = \sum (TX_{MBA} \times \frac{Num_{TotalChequesdoBanco}}{Num_{TotalCheques}}).$$

5.4 RESULTADOS DOS MÉTODOS GLOBAIS

5.4.1 Método por Valor Médio

Para o exemplo do cheque bancário do Banco Itaú S.A., o valor de limiar obtido é 183 (figura 5.5).

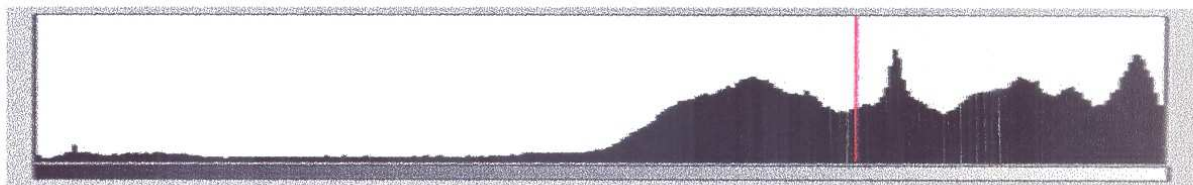


Figura 5.5: Visualização no histograma do valor de limiar obtido pelo método por valor médio.

Esse método não conseguiu eliminar o fundo artístico. Pelo contrário, os resultados obtidos deixam a impressão que esse fundo foi realçado. Visualmente falando, os dígitos e os caracteres confundem-se com esse fundo (figura 5.6).

A figura 5.7 ilustra o resultado do reconhecimento por *OCR* do cabeçalho do cheque bancário binarizado pelo método por valor médio.

Devido a péssima qualidade dos resultados da binarização, as taxas de reconhecimento por *OCR* por banco versus método por valor médio foram inferiores a 12% para cheques não preenchidos e inferiores a 2% para cheques preenchidos a máquina.

Os cheques que possuem fundos artísticos mais claros tiveram um resultado de reconhecimento um pouco melhor como evidenciado pelo cheque sem preenchimento do Banco América do Sul S.A. (taxa de 12%) que possui um fundo verde esmaecido. O fundo não chegou a ser muito realçado facilitando o reconhecimento pelo *OCR*. Os resultados obtidos nos demais cheques foram muito ruins não permitindo o reconhecimento de muitos caracteres como no caso do cheque preenchido do Banco Real S.A que teve uma taxa de reconhecimento inferior a 1%.

5.4.2 Métodos por Porcentagem e por Pixels de Borda

Esses métodos não foram testados devido à passagem obrigatória de parâmetros externos.

5.4.3 Método por Dois Picos

Para o exemplo do cheque bancário do Banco Itaú S.A., o valor de limiar obtido é 21 (figura 5.8).

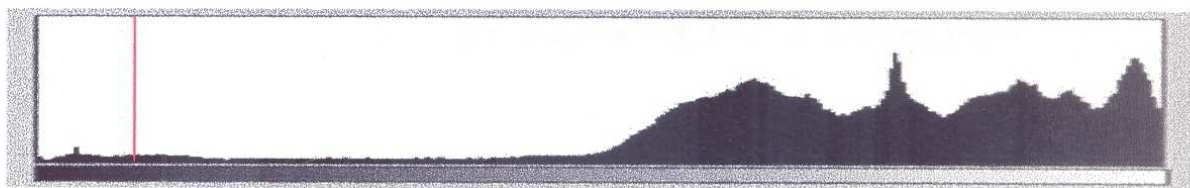


Figura 5.8: Visualização no histograma do valor de limiar obtido pelo método por dois picos.

Esse método eliminou totalmente o fundo artístico como também uma grande parte dos dados relevantes. Devido ao baixo valor de limiar obtido, a quase totalidade das informações são perdidas (figura 5.9).

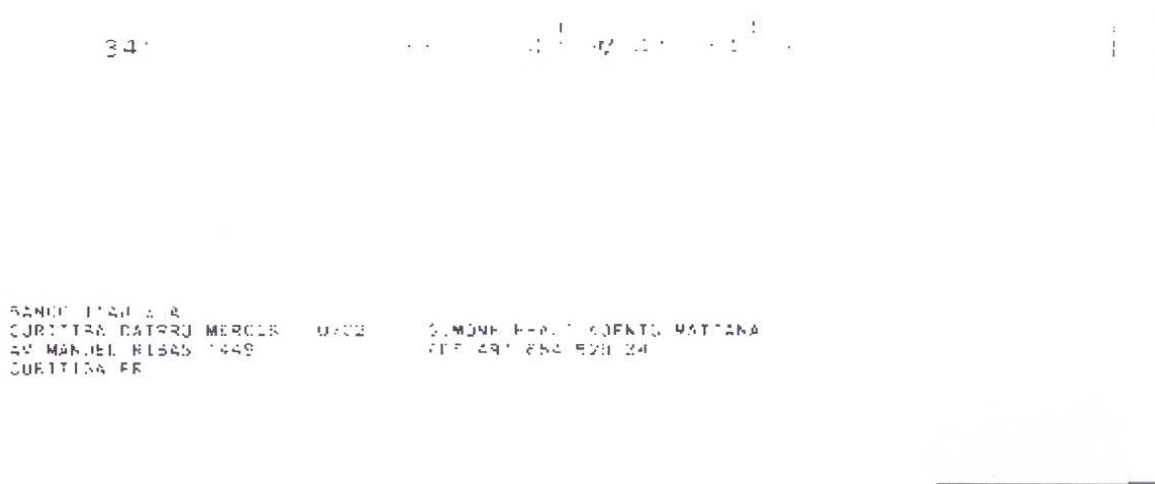


Figura 5.9: Resultado da binarização do cheque bancário pelo método por dois picos.

A figura 5.10 ilustra o resultado do reconhecimento por *OCR* do cabeçalho do cheque bancário binarizado pelo método por dois picos.

a)

Ban-~nd~|_*~st~%44~~/ it'~ ihl ~ ~f- =~== ~ ~: ~ ~-at-- :
 00~| 3=i 0 9 -=~| 4_~| yk 7d443ss ~ ~D \$ 4- HM~ 41 3R4 | 9 | ~.
 . ~ -. # = ~-- ~ -- ih - ~ ~ ~ 1, ~ w ~ lk 1 ~ i: 2 -affid## \$ ~--

b)

Figura 5.10: Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por dois picos; b) do reconhecimento por *OCR* desse cabeçalho binarizado.

Devido aos valores de limiar obtidos, pode-se notar que cheques que possuem as impressões de dados com cores mais fortes são binarizados com um valor um pouco mais alto, possibilitando a recuperação das informações pelo *OCR*. Porém, os valores das taxas de reconhecimento não são aceitáveis. As taxas de reconhecimento por *OCR* por banco versus método por dois picos foram inferiores a 8% para cheques não preenchidos e inferiores a 4% para cheques preenchidos a máquina.

5.4.4 Método de Kittler-Illingworth

Para o exemplo do cheque bancário do Banco Itaú S.A., o valor de limiar obtido é 188 (figura 5.11).

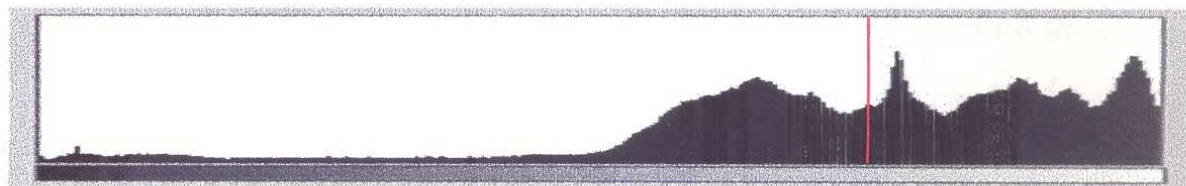


Figura 5.11: Visualização no histograma do valor de limiar obtido pelo método de Kittler-Illingworth.

Esse método não conseguiu eliminar o fundo artístico. Pelo contrário, os resultados obtidos



Figura 5.12: Resultado da binarização do cheque bancário pelo método de Kittler-Illingworth.



a)

Ban-~d~|_*~==~t=~~#~/ithy~ihl ~ ~f- =~-== -g ;g ~ ~ -at-- :
 ~h~|3=i O9 -=~~| %_~|2k7Ss ~ ~D \$ 4- HZ~ 4x 152354 | 8 |R ~,
 . ~ -, #7 ~-- ~ --ih - ~ !!!!! ~ ~!@1, ~ w ~ lha 1~ i: 1 - iid~+~| --|

b)

Figura 5.13: Resultados: a) da binarização do cabeçalho do cheque bancário pelo método de Kittler-Illingworth; b) do reconhecimento por OCR desse cabeçalho binarizado.

deixam a impressão que esse fundo foi realçado. Visualmente falando, os dígitos e os caracteres confundem-se com esse fundo (figura 5.12).

A figura 5.13 ilustra o resultado do reconhecimento por *OCR* do cabeçalho do cheque bancário binarizado pelo método de Kittler-Illingworth.

Os resultados das taxas de reconhecimento obtidos foram muito ruins (inferiores a 3% para cheques não preenchidos e inferiores a 2% para cheques preenchidos a máquina) não permitindo o eficiente reconhecimento pelo *OCR*. "Evidencia-se" o cheque sem preenchimento do banco América do Sul S.A devido ao seu fundo claro.

5.4.5 Método de Otsu

Para o exemplo do cheque bancário do Banco Itaú S.A., o valor de limiar obtido é 179 (figura 5.14).

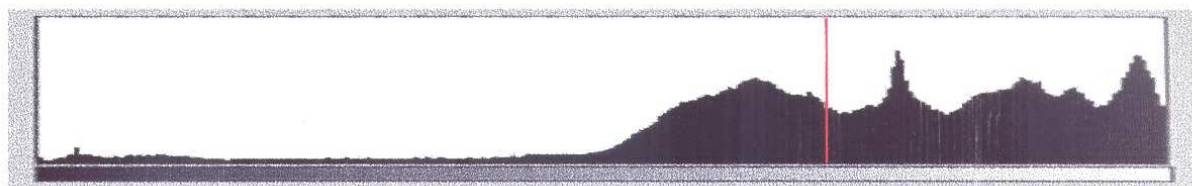


Figura 5.14: Visualização no histograma do valor de limiar obtido pelo método de Otsu.

Esse método conseguiu eliminar o fundo artístico de forma quase completa. Visualmente falando, os dígitos e os caracteres forma bem segmentados sem perdas perceptíveis (figura 5.15).

A figura 5.16 ilustra o resultado do reconhecimento por *OCR* do cabeçalho do cheque bancário binarizado pelo método de Otsu.

Nas tabelas 5.1 e 5.2 estão agrupados os resultados das taxas de reconhecimento (em %) por *OCR* por banco versus método de Otsu respectivamente para cheques não preenchidos e preenchidos a máquina.

Os resultados da taxa de reconhecimento obtidos foram muito bons permitindo o eficiente reconhecimento pelo *OCR*. Houve alguma perda de reconhecimento nos cheques preenchidos devido a erros de posicionamento dos caracteres no momento do preenchimento. Outro fator a ser observado nos mesmo grupo de cheques diz respeito a característica do preenchimento, algumas vezes esmaecida pelo desgaste da fita de impressão na máquina de preenchimento, outras por borrões de tinta da própria fita.

Banco	Cheque não preenchido (%)
Banco Bamerindus SA	79
Banco Banestado	88
Banco América do Sul	92
Banco do Brasil	94
Banco Bradesco	84
Caixa Econômica federal	89
Banco HSBC Bamerindus SA	71
Banco Itaú	82
Banco Meridional	91
Banco Real	90
Unibanco	86

Tabela 5.1: Taxas de reconhecimento (em %) por *OCR* por banco versus método de Otsu para cheques não preenchidos.

Banco	Cheque preenchido (%)
Banco Bamerindus SA	66
Banco Banestado	68
Banco América do Sul	70
Banco do Brasil	74
Banco Bandeirantes	65
Banco Bradesco	74
Caixa Econômica federal	65
Citybank SA	79
Banco HSBC Bamerindus SA	53
Banco Itaú	59
Banco Real	62
Banco Meridional	67
Unibanco	57

Tabela 5.2: Taxas de reconhecimento (em %) por *OCR* por banco versus método de Otsu para cheques preenchidos a máquina.

5.4.6 Método por Fuzzy Set - abordagem de Huang

Para o exemplo do cheque bancário do Banco Itaú S.A., o valor de limiar obtido é 201 (figura 5.17).

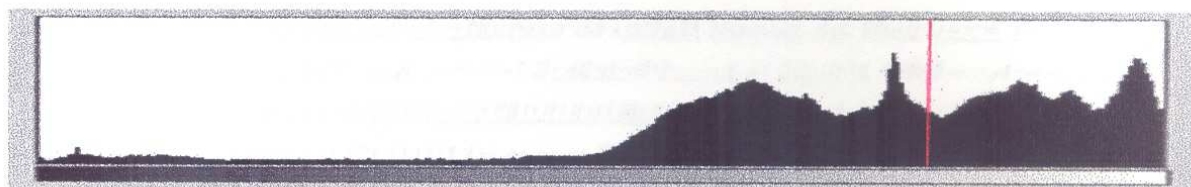


Figura 5.17: Visualização no histograma do valor de limiar obtido pelo método por Fuzzy Set - abordagem de Huang.

Esse método se comportou de forma bem irregular dependendo do fundo artístico e de suas cores. Para cheques de certos bancos (Caixa Econômica Federal, Banco HSBC por exemplo), a eliminação foi visualmente total. Entretanto, para cheques de fundo azul, o método não conseguiu eliminar o fundo artístico. Pelo contrário, esse fundo foi realçado (figura 5.18).

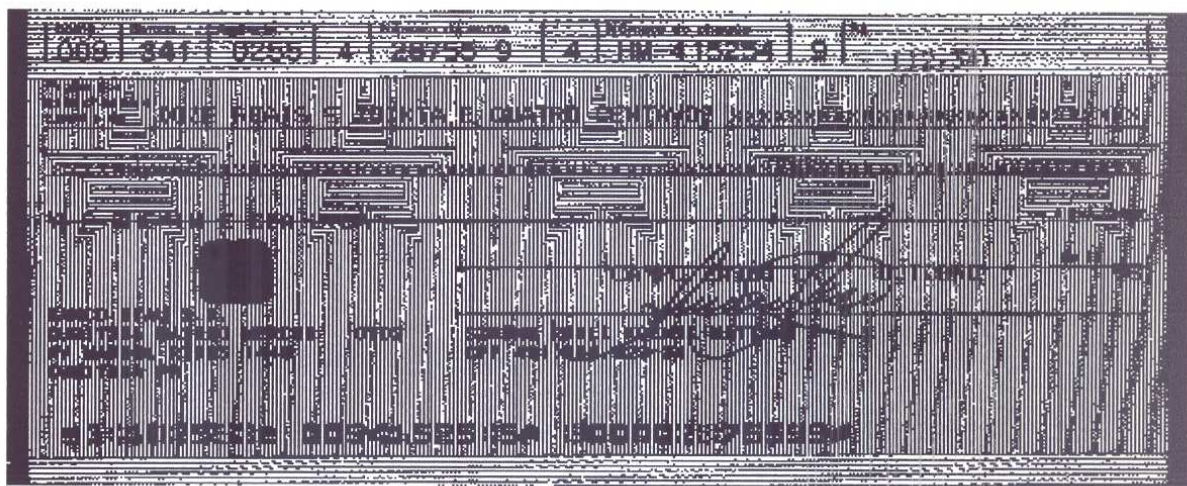


Figura 5.18: Resultado da binarização do cheque bancário pelo método por Fuzzy Set - abordagem de Huang.

A figura 5.19 ilustra o resultado do reconhecimento por *OCR* do cabeçalho do cheque bancário binarizado pelo método por Fuzzy Set - abordagem de Huang.

Os resultados das taxas de reconhecimento obtidos foram muito ruins (inferiores a 2% para

Camp	Banco	Agência	4	Número da conta	4	Número do cheque	9	R\$
009	341	0255		28755-9		HM-415254		10/04

Pague por
esta cheque a
quantia de

DOZE REAIS E TRINTA E QUATRO CENTAVOS R\$ 12,34

ou a sua ordem

17 de 19

Itaú

BANCO ITAÚ S.A.
CURITIBA BAIRRO MERCES 0702
AV MANOEL RIBAS 1449
CURITIBA PR

SIMONE REALI KOENIG MATTANA
CPF 491.684.529-34

3410025532 009045254534 8600028755910

Figura 5.21: Resultado da binarização do cheque bancário pelo método por Fuzzy Set - abordagem de Yager.

Camp	Banco	Agência	4	Número da conta	4	Número do cheque	9	R\$
009	341	0255		28755-9		HM-415254		10/04

Pague por
esta cheque a
quantia de

DOZE REAIS E TRINTA E QUATRO CENTAVOS R\$ 12,34

a)

1009 ~ 341 1 0255 14 1 22755-9
1 4 1 MM-415254 19 1 ; ~ +
I Pai ~ p0s esta cteque a quant1 @ de D==~P+x E |
t3 ~ ~ # 9 J CE ! # \$ \$ av OS ~ ! ~ @ # # ~ ~ ~ ~ ~

b)

Figura 5.22: Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por Fuzzy Set - abordagem de Yager; b) do reconhecimento por OCR desse cabeçalho binarizado.

Banco	Cheque não preenchido (%)
Banco Bamerindus SA	86
Banco Banestado	89
Banco América do Sul	91
Banco do Brasil	90
Banco Bradesco	88
Caixa Econômica federal	86
Banco HSBC Bamerindus SA	58
Banco Itaú	77
Banco Meridional	83
Banco Real	90
Unibanco	89

Tabela 5.3: Taxas de reconhecimento (em %) por *OCR* por banco versus método por Fuzzy Set - abordagem de Yager para cheques não preenchidos.

Banco	Cheque preenchido (%)
Banco Bamerindus SA	65
Banco Banestado	62
Banco América do Sul	61
Banco do Brasil	61
Banco Bandeirantes	68
Banco Bradesco	67
Caixa Econômica federal	55
Citybank SA	64
Banco HSBC Bamerindus SA	42
Banco Itaú	65
Banco Real	57
Banco Meridional	59
Unibanco	57

Tabela 5.4: Taxas de reconhecimento (em %) por *OCR* por banco versus método por Fuzzy Set - abordagem de Yager para cheques preenchidos a máquina.

por *OCR* por banco versus método por Fuzzy Set - abordagem de Yager respectivamente para cheques não preenchidos e preenchidos a máquina.

Os resultados da taxa de reconhecimento obtidos foram muito bons permitindo o eficiente reconhecimento pelo *OCR*. Como no método de Otsu, houve alguma perda de reconhecimento nos cheques preenchidos devido a erros de posicionamento dos caracteres e as características de preenchimento já citadas.

5.4.8 Método por Seleção iterativa de Thrussel

Para o exemplo do cheque bancário do Banco Itaú S.A., o valor de limiar obtido é 129 (figura 5.23).

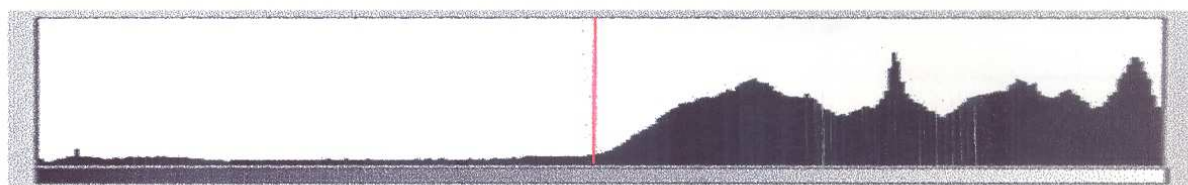


Figura 5.23: Visualização no histograma do valor de limiar obtido pelo método por seleção iterativa de Thrussel.

Esse método conseguiu eliminar quase totalmente o fundo artístico. Visualmente falando, os dígitos e os caracteres são bem legíveis, porém existe um aspecto ruidoso do tipo ruído "pimenta" (figura 5.24).

A figura 5.25 ilustra o resultado do reconhecimento por *OCR* do cabeçalho do cheque bancário binarizado pelo método por seleção iterativa de Thrussel.

Nas tabelas 5.5 e 5.6 estão agrupados os resultados das taxas de reconhecimento (em %) por *OCR* por banco versus método por seleção iterativa de Thrussel respectivamente para cheques não preenchidos e preenchidos a máquina.

Os resultados da taxa de reconhecimento obtidos foram muito bons permitindo o eficiente reconhecimento pelo *OCR*. Ainda como nos métodos de Otsu e de Yager, houve alguma perda de reconhecimento nos cheques preenchidos devido a erros de posicionamento dos caracteres e as características de preenchimento já citadas.

Banco	Cheque não preenchido (%)
Banco Bamerindus SA	84
Banco Banestado	90
Banco América do Sul	88
Banco do Brasil	87
Banco Bradesco	86
Caixa Econômica federal	87
Banco HSBC Bamerindus SA	54
Banco Itaú	87
Banco Meridional	79
Banco Real	88
Unibanco	83

Tabela 5.5: Taxas de reconhecimento (em %) por *OCR* por banco versus método por seleção iterativa de Thrussel para cheques não preenchidos.

Banco	Cheque preenchido (%)
Banco Bamerindus SA	62
Banco Banestado	67
Banco América do Sul	65
Banco do Brasil	63
Banco Bandeirantes	71
Banco Bradesco	61
Caixa Econômica federal	59
Citybank SA	66
Banco HSBC Bamerindus SA	49
Banco Itaú	67
Banco Real	62
Banco Meridional	64
Unibanco	66

Tabela 5.6: Taxas de reconhecimento (em %) por *OCR* por banco versus método por seleção iterativa de Thrussel para cheques preenchidos a máquina.

5.4.9 Método por Entropia de Pun

Para o exemplo do cheque bancário do Banco Itaú S.A., o valor de limiar obtido é 235 (figura 5.26).

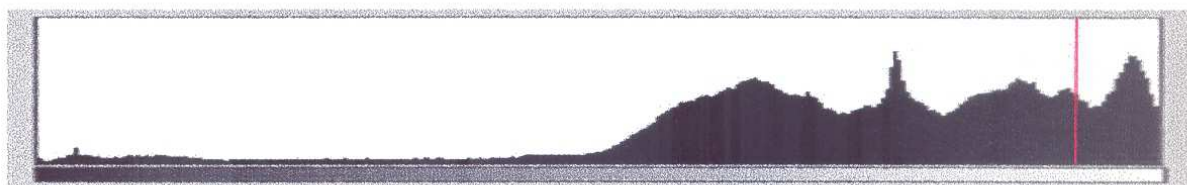


Figura 5.26: Visualização no histograma do valor de limiar obtido pelo método por entropia de Pun.

Esse método não conseguiu eliminar o fundo artístico. Pelo contrário, os resultados obtidos deixam a impressão que esse fundo foi realçado. Visualmente falando, os dígitos e os caracteres confundem-se com esse fundo (figura 5.27).

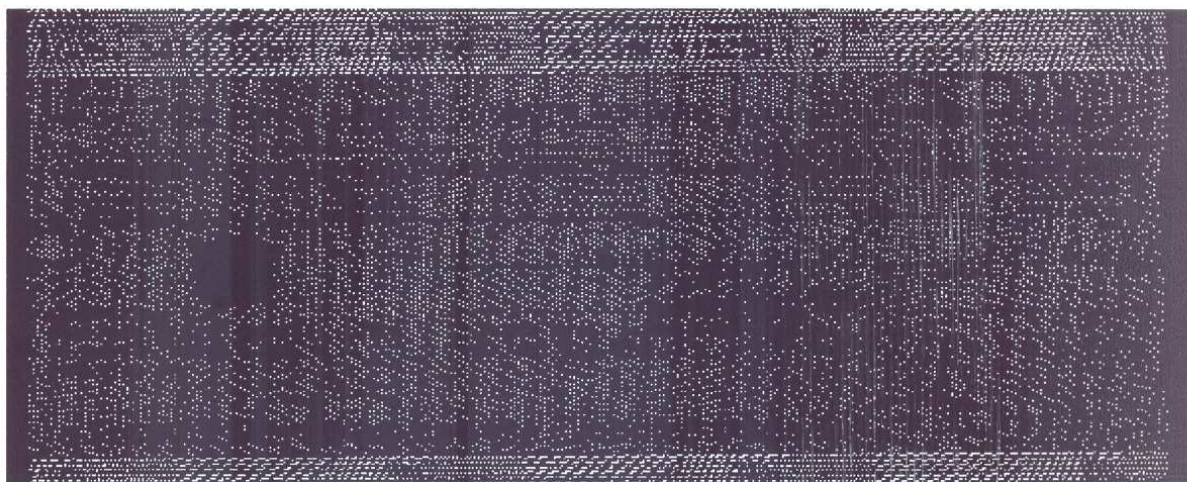


Figura 5.27: Resultado da binarização do cheque bancário pelo método por entropia de Pun.

A figura 5.28 ilustra o resultado do reconhecimento por *OCR* do cabeçalho do cheque bancário binarizado pelo método por entropia de Pun.

Devido aos péssimos resultados da binarização em termos de qualidade, as taxas de reconhecimento por *OCR* por banco versus método por entropia de Pun respectivamente para cheques não preenchidos e preenchidos a máquina forma inferiores a 1%.

Banco	Cheque não preenchido (%)
Banco Bamerindus SA	59
Banco Banestado	52
Banco América do Sul	40
Banco do Brasil	44
Banco Bradesco	21
Caixa Econômica federal	35
Banco HSBC Bamerindus SA	19
Banco Itaú	17
Banco Meridional	29
Banco Real	27
Unibanco	26

Tabela 5.7: Taxas de reconhecimento (em %) por *OCR* por banco versus método por entropia de Kapur para cheques não preenchidos.

Banco	Cheque preenchido (%)
Banco Bamerindus SA	41
Banco Banestado	50
Banco América do Sul	40
Banco do Brasil	41
Banco Bandeirantes	43
Banco Bradesco	34
Caixa Econômica federal	31
Citybank SA	47
Banco HSBC Bamerindus SA	21
Banco Itaú	30
Banco Real	25
Banco Meridional	27
Unibanco	26

Tabela 5.8: Taxas de reconhecimento (em %) por *OCR* por banco versus método por entropia de Kapur para cheques preenchidos a máquina.

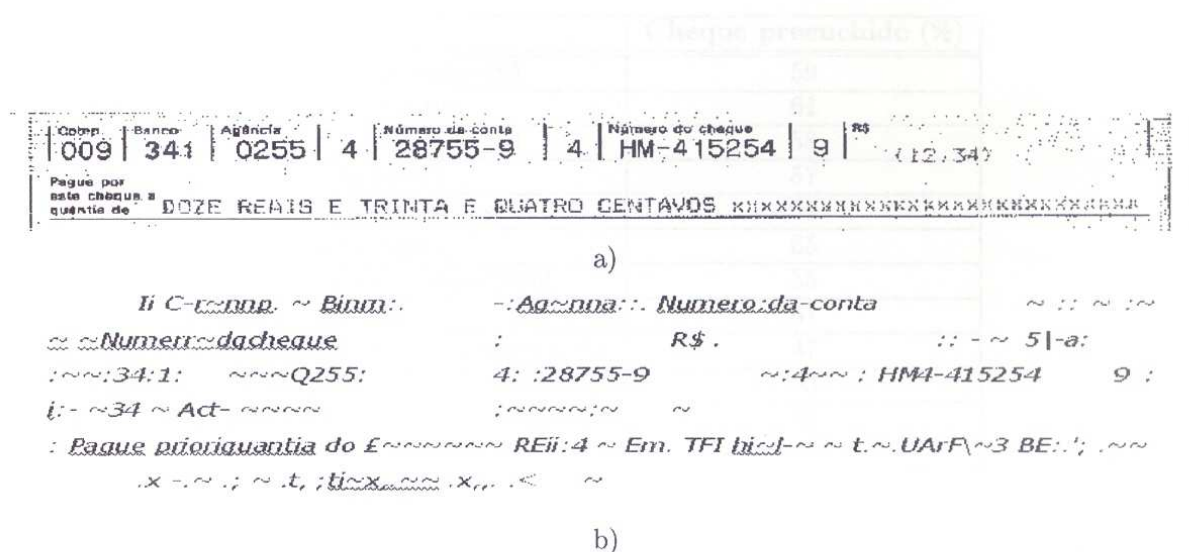


Figura 5.34: Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por entropia de Johannsen; b) do reconhecimento por *OCR* desse cabeçalho binarizado.

Banco	Cheque não preenchido (%)
Banco Bamerindus SA	82
Banco Banestado	89
Banco América do Sul	84
Banco do Brasil	83
Banco Bradesco	83
Caixa Econômica federal	81
Banco HSBC Bamerindus SA	58
Banco Itaú	81
Banco Meridional	77
Banco Real	84
Unibanco	82

Tabela 5.9: Taxas de reconhecimento (em %) por *OCR* por banco versus método por entropia de Johannsen para cheques não preenchidos.

Banco	Cheque preenchido (%)
Banco Bamerindus SA	59
Banco Banestado	61
Banco América do Sul	58
Banco do Brasil	57
Banco Bandeirantes	69
Banco Bradesco	65
Caixa Econômica federal	55
Citybank SA	61
Banco HSBC Bamerindus SA	47
Banco Itaú	63
Banco Real	62
Banco Meridional	60
Unibanco	67

Tabela 5.10: Taxas de reconhecimento (em %) por *OCR* por banco versus método por entropia de Johannsen para cheques preenchidos a máquina.

consequência disso, vários dígitos e caracteres desgastaram-se prejudicando muito a leitura (figura 5.33).

A figura 5.34 ilustra o resultado do reconhecimento por *OCR* do cabeçalho do cheque bancário binarizado pelo método por entropia de Johannsen.

Nas tabelas 5.9 e 5.10 estão agrupados os resultados das taxas de reconhecimento (em %) por *OCR* por banco versus método por entropia de Kapur respectivamente para cheques não preenchidos e preenchidos a máquina.

Os resultados da taxa de reconhecimento obtidos foram muito bons permitindo o reconhecimento pelo *OCR*. Como já citado, as taxas de acerto para cheques preenchidos deve-se ao erro de posicionamento dos caracteres e anomalias no preenchimento por borrões de tinta.

5.4.12 Método por Entropia de Li-Lee

Para o exemplo do cheque bancário do Banco Itaú S.A., o valor de limiar obtido é 104 (figura 5.35).

Esse método teve um muito bom desempenho, eliminando o fundo artístico, sem prejudicar as informações relevantes, inclusive os dígitos e os caracteres (figura 5.36).

A figura 5.37 ilustra o resultado do reconhecimento por *OCR* do cabeçalho do cheque bancário binarizado pelo método por entropia de Li-Lee.

Nas tabelas 5.11 e 5.12 estão agrupados os resultados das taxas de reconhecimento (em

Banco	Cheque não preenchido (%)
Banco Bamerindus SA	78
Banco Banestado	87
Banco América do Sul	93
Banco do Brasil	91
Banco Bradesco	86
Caixa Econômica federal	88
Banco HSBC Bamerindus SA	62
Banco Itaú	76
Banco Meridional	87
Banco Real	92
Unibanco	88

Tabela 5.11: Taxas de reconhecimento (em %) por *OCR* por banco versus método por entropia de Li-Lee para cheques não preenchidos.

Banco	Cheque preenchido (%)
Banco Bamerindus SA	63
Banco Banestado	67
Banco América do Sul	69
Banco do Brasil	71
Banco Bandeirantes	66
Banco Bradesco	72
Caixa Econômica federal	63
Citybank SA	64
Banco HSBC Bamerindus SA	51
Banco Itaú	60
Banco Real	65
Banco Meridional	62
Unibanco	64

Tabela 5.12: Taxas de reconhecimento (em %) por *OCR* por banco versus método por entropia de Li-Lee para cheques preenchidos a máquina.

%) por *OCR* por banco versus método por entropia de Li-Lee respectivamente para cheques não preenchidos e preenchidos a máquina.

Os resultados da taxa de reconhecimento obtidos foram muito bons permitindo o reconhecimento pelo *OCR*. Como já citado, as taxas de acerto para cheques preenchidos deve-se ao erro de posicionamento dos caracteres e anomalias no preenchimento por borrões de tinta.

5.5 RESULTADOS DOS MÉTODOS LOCAIS ADAPTATIVOS

5.5.1 Método de Bernsen

Esse método forneceu resultados muito bons mesmo se, as vezes, não eliminou totalmente a faixa de segurança. Ele não prejudicou as informações relevantes, inclusive os dígitos e os caracteres (figura 5.38).



Figura 5.38: Resultado da binarização do cheque bancário pelo método de Bernsen.

A figura 5.39 ilustra o resultado do reconhecimento por *OCR* do cabeçalho do cheque bancário binarizado pelo método de Bernsen.

Nas tabelas 5.13 e 5.14 estão agrupados os resultados das taxas de reconhecimento (em %) por *OCR* por banco versus método de Bernsen respectivamente para cheques não preenchidos e preenchidos a máquina.

Os resultados das taxas de reconhecimento obtidos foram muito bons permitindo o eficiente reconhecimento pelo *OCR*. Observou-se algumas perdas decorrentes das características de

Comp.	Banco	Agência		Número da conta		Número do cheque		R\$
009	341	0255	4	28755-9	4	HM-415254	9	(12,34)

Pague por este cheque a quantia de DOZE REAIS E TRINTA E QUATRO CENTAVOS XXXXXXXXXXXXXXXXXXXXXXXXXXXX

a)

I. .C-omp; I Banco . IAgência: ::: I~ I. Numern:da-conta: ~ ~ ~ ~ I I ~ ~
 INumero do~chequ~ ~ I ~ I R\$ ~ ~ ~ ~ ~
 -009 I;~341 ~ I~~;Q255 ~ I~ 4~~: ~28755-9~~~I~4 ~: ~HM-4152~54~;9 ~ ~ ~ ~ ~ ~ ~ ~ ~
 (12.34) ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ I ~ ~ ~ ~ ~
 ~ Pague por ~ ~ ~ ~ ~ ~ i::: ~ :este~~chequ~~a~~qua ntia de ~ OZE REAl:'S trit-A e qurrio
 CENt~~s xxxx,xxxxxxxxxx~x~~;xxx

b)

Figura 5.39: Resultados: a) da binarização do cabeçalho do cheque bancário pelo método de Bernsen; b) do reconhecimento por OCR desse cabeçalho binarizado.

Banco	Cheque não preenchido (%)
Banco Bamerindus SA	91
Banco Banestado	93
Banco América do Sul	96
Banco do Brasil	97
Banco Bradesco	87
Caixa Econômica federal	95
Banco HSBC Bamerindus SA	79
Banco Itaú	96
Banco Meridional	93
Banco Real	94
Unibanco	91

Tabela 5.13: Taxas de reconhecimento (em %) por OCR por banco versus método de Bernsen para cheques não preenchidos.

Banco	Cheque preenchido (%)
Banco Bamerindus SA	81
Banco Banestado	84
Banco América do Sul	85
Banco do Brasil	87
Banco Bandeirantes	85
Banco Bradesco	86
Caixa Econômica federal	83
Citybank SA	85
Banco HSBC Bamerindus SA	79
Banco Itaú	87
Banco Real	89
Banco Meridional	84
Unibanco	86

Tabela 5.14: Taxas de reconhecimento (em %) por *OCR* por banco versus método de Bernsen para cheques preenchidos a máquina.

preenchimento como já citadas.

5.5.2 Método por Médias Móveis

Esse método se comportou de forma irregular, as vezes eliminando o fundo artístico, outras vezes mantendo um aspecto ruidoso do tipo ruído "pimenta", outras vezes ainda eliminando linhas e até informações relevantes. Dígitos e caracteres sofreram desgastem prejudicando a leitura (figura 5.40).

A figura 5.41 ilustra o resultado do reconhecimento por *OCR* do cabeçalho do cheque bancário binarizado pelo método por médias móveis.

Nas tabelas 5.15 e 5.16 estão agrupados os resultados das taxas de reconhecimento (em %) por *OCR* por banco versus método por médias móveis respectivamente para cheques não preenchidos e preenchidos a máquina.

Os resultados das taxas de reconhecimento obtidos foram bons permitindo o reconhecimento pelo *OCR*. Devido ao ruído produzido e ao desgaste dos caracteres as taxas de reconhecimento mostraram-se baixas para este método.

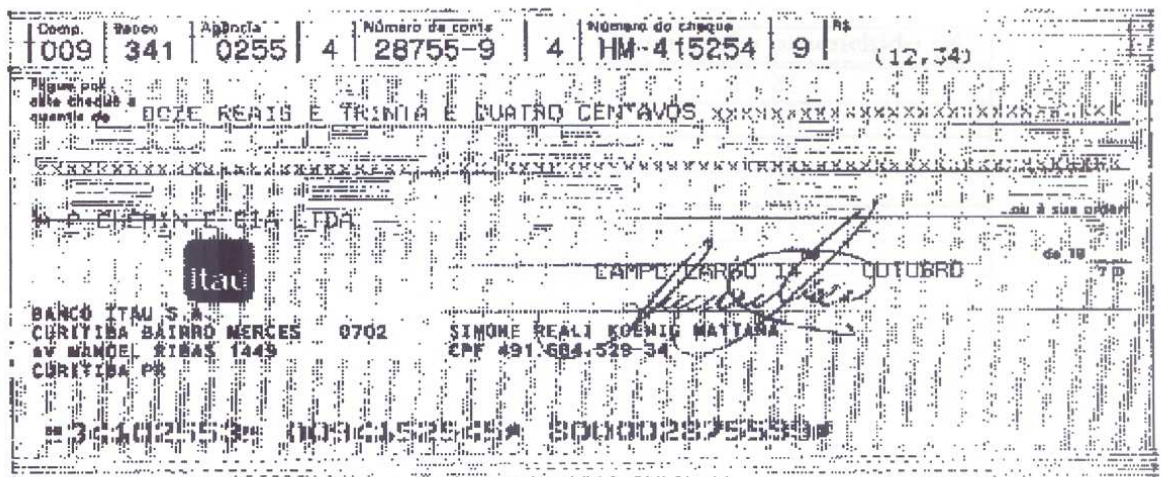


Figura 5.40: Resultado da binarização do cheque bancário pelo método por médias móveis.



a)

Or ii iiii j~ iim: iiii I~ t~ I~ ~t ~ ~ IS
;t0091 3411 02rs 1 41 Z3755-9 1 4 1 Ht-4~52541 91 t32~
HI
Q~Zs Re;4lb ,E I' -i ~ :: COCOS hi; it t;

b)

Figura 5.41: Resultados: a) da binarização do cabeçalho do cheque bancário pelo método por médias móveis; b) do reconhecimento por OCR desse cabeçalho binarizado.

Banco	Cheque não preenchido (%)
Banco Bamerindus SA	31
Banco Banestado	37
Banco América do Sul	42
Banco do Brasil	47
Banco Bradesco	22
Caixa Econômica federal	37
Banco HSBC Bamerindus SA	16
Banco Itaú	19
Banco Meridional	24
Banco Real	28
Unibanco	23

Tabela 5.15: Taxas de reconhecimento (em %) por *OCR* por banco versus método por médias móveis para cheques não preenchidos.

Banco	Cheque preenchido (%)
Banco Bamerindus SA	34
Banco Banestado	42
Banco América do Sul	40
Banco do Brasil	38
Banco Bandeirantes	32
Banco Bradesco	34
Caixa Econômica federal	41
Citybank SA	37
Banco HSBC Bamerindus SA	24
Banco Itaú	38
Banco Real	36
Banco Meridional	34
Unibanco	37

Tabela 5.16: Taxas de reconhecimento (em %) por *OCR* por banco versus método por médias móveis para cheques preenchidos a máquina.

5.5.3 Método por Relaxação

Esse método foi descartado devido aos péssimos resultados obtidos. Apesar da eliminação do fundo artístico, de forma singular, o processo espessou os dígitos e os caracteres, os tornando incompreensíveis (figura 5.42).

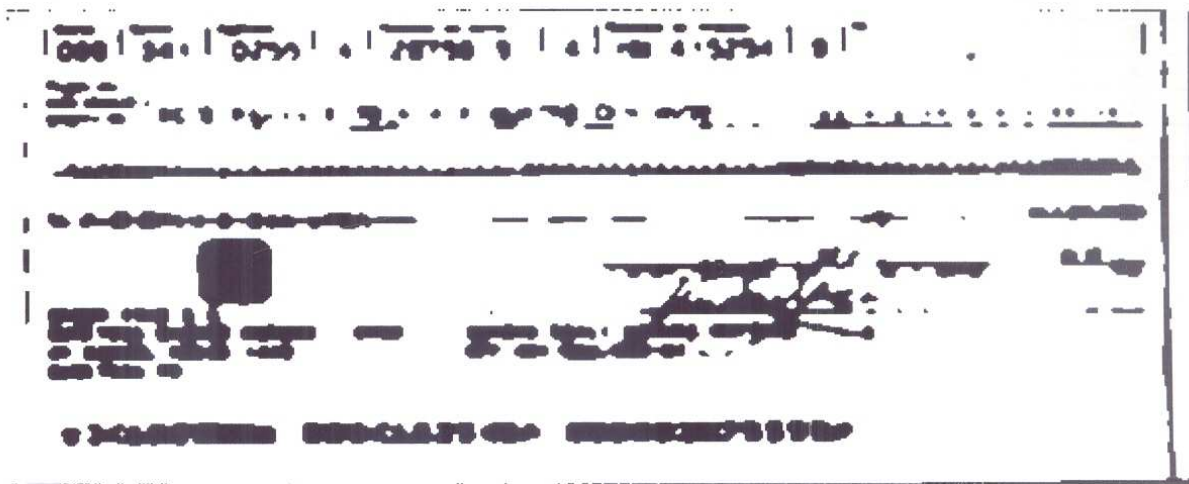


Figura 5.42: Resultado da binarização do cheque bancário pelo método por relaxação.

5.6 TAXAS DE RECONHECIMENTO POR BANCO VERSUS MÉTODOS GLOBAIS E LOCAIS

Nas tabelas 5.17 e 5.18 estão apresentadas (em %) as taxas TX_{MBA} médias por banco versus algoritmo, para os cheques não preenchidos e todos os algoritmos.

Nas tabelas 5.19 e 5.20 estão apresentadas (em %) as taxas TX_{MBA} médias por banco versus algoritmo, para os cheques preenchidos a máquina e todos os algoritmos.

5.7 TAXAS GLOBAIS DE RECONHECIMENTO

Na tabela 5.21 estão reunidas (em %) as taxas TX_{GA} globais de reconhecimento por *OCR* para todos os algoritmos de binarização em ordem decrescente.

Método	Valor Médio	2 Picos	Kittler	Otsu	Huang	Yager	Thrussel
Banco							
Banco Bamerindus SA	< 12	< 8	< 3	79	< 2	86	84
Banco Banestado	< 12	< 8	< 3	88	< 2	89	90
Banco América do Sul	12	8	3	92	< 2	91	88
Banco do Brasil	< 12	< 8	< 3	94	< 2	90	87
Banco Bradesco	< 12	< 8	3	84	< 2	88	86
Caixa Econômica federal	< 12	< 8	< 3	89	< 2	86	87
Banco HSBC Bamerindus SA	< 12	< 8	< 3	71	< 2	58	54
Banco Itaú	< 12	< 8	< 3	82	< 2	77	87
Banco Meridional	< 12	< 8	< 3	91	< 2	83	79
Banco Real	< 12	< 8	< 3	90	< 2	90	88
Unibanco	< 12	< 8	< 3	86	< 2	89	83

Tabela 5.17: Taxas de reconhecimento por *OCR* por banco versus métodos globais e locais para cheques não preenchidos.

Método	Pun	Kapur	Johannsen	Li-Lee	Bernsen	Médias
Banco						
Banco Bamerindus SA	< 10	59	82	78	91	31
Banco Banestado	< 10	52	89	87	93	37
Banco América do Sul	< 10	40	84	93	96	42
Banco do Brasil	< 10	44	83	91	97	47
Banco Bradesco	< 10	21	83	86	87	22
Caixa Econômica federal	< 10	35	81	88	95	37
Banco HSBC Bamerindus SA	< 10	19	58	62	79	16
Banco Itaú	< 10	17	81	76	96	19
Banco Meridional	< 10	29	77	87	93	24
Banco Real	< 10	27	84	92	94	28
Unibanco	< 10	26	82	88	91	23

Tabela 5.18: Taxas de reconhecimento (em %) por *OCR* por banco versus métodos globais e locais para cheques não preenchidos.

Método	Valor Médio	2 Picos	Kittler	Otsu	Huang	Yager	Thrussel
Banco							
Banco Bamerindus SA	< 2	< 4	< 2	66	< 2	65	62
Banco Banestado	2	4	< 2	68	< 2	62	67
Banco América do Sul	2	4	< 2	70	< 2	61	65
Banco do Brasil	< 2	< 4	< 2	74	< 2	61	63
Banco Bandeirantes	2	< 4	< 2	65	< 2	68	71
Banco Bradesco	< 2	4	< 2	74	< 2	67	61
Caixa Econômica federal	2	< 4	< 2	65	< 2	55	59
Citybank SA	< 2	< 4	< 2	79	< 2	64	66
Banco HSBC Bamerindus SA	< 2	< 4	< 2	53	< 2	42	49
Banco Itaú	< 2	< 4	< 2	59	< 2	65	67
Banco Real	< 2	< 4	< 2	62	< 2	57	62
Banco Meridional	2	< 4	< 2	67	< 2	59	64
Unibanco	< 2	< 4	< 2	65	< 2	57	66

Tabela 5.19: Taxas de reconhecimento por *OCR* por banco versus métodos globais e locais para cheques preenchidos a máquina.

Método	Pun	Kapur	Johannsen	Li-Lee	Bernsen	Médias
Banco						
Banco Bamerindus SA	< 1	41	59	63	81	34
Banco Banestado	< 1	50	61	67	84	42
Banco América do Sul	< 1	40	58	69	85	40
Banco do Brasil	< 1	41	57	71	87	38
Banco Bandeirantes	< 1	43	69	66	85	32
Banco Bradesco	< 1	34	65	72	86	34
Caixa Econômica federal	< 1	31	55	63	83	41
Citybank SA	< 1	47	61	64	85	37
Banco HSBC Bamerindus SA	< 1	21	47	51	79	24
Banco Itaú	< 1	30	63	60	87	38
Banco Real	< 1	25	62	65	89	36
Banco Meridional	< 1	27	60	62	84	34
Unibanco	< 1	26	67	64	86	37

Tabela 5.20: Taxas de reconhecimento (em %) por *OCR* por banco versus métodos globais e locais para cheques preenchidos a máquina.

Método de binarização	Algoritmo	Cheque não preenchido (%)	Cheque preenchido a máquina. (%)
Técnica global	Dois picos	< 4	< 3
	Huang	< 1	< 1
	Johannsen	82.86	59.37
	Kapur	33.03	35.28
	Kittler-Illingworth	< 2	< 2
	Li and Lee	85.38	64.09
	Otsu	86.12	66.1
	Pun	< 1	< 1
	Thrussel	85.74	62.29
	Yager	86.60	59.55
	Valor médio	< 4	< 1
Técnica local	Bernsen	92.79	84.03
	Médias móveis	31.38	36.13

Tabela 5.21: Taxas globais de reconhecimento (em %) por *OCR* de todos os métodos de segmentação por binarização.

5.8 CONCLUSÕES SOBRE AS TAXAS FINAIS DE RECONHECIMENTO

Os resultados das taxas de reconhecimento obtidos foram muito bons para certos algoritmos permitindo o eficiente reconhecimento do conteúdo pelo *OCR*. Ainda que existam perdas pelo preenchimento dos cheques, as taxas mostram-se mais elevadas para os cheques sem preenchimento, como era de se esperar.

Algoritmos como o Método por Valor Médio, por Dois Picos, de Huang, de Kittler apresentaram resultados muito ruins, pois além de reforçar o fundo artístico ou o suavizar de forma excessiva - o valor do limiar obtido (muito baixo ou muito alto) inviabiliza o reconhecimento pelo *OCR* pelo baixo ou alto contraste da imagem binarizada - torna incompreensíveis os caracteres impressos ou de preenchimento.

Métodos por Porcentagem e por Pixels de Borda foram descartados devido a passagem de parâmetros externos, pois em um método automatizado não é possível se manter um parâmetro comum para todos os cheques devido a grande variação de resultados obtidos.

O Método por Relaxação fornece bons resultados em termos da extração do fundo artístico, porém espessa de tal forma os caracteres de impressão/preenchimento que torna inviável qualquer operação de reconhecimento pôr *OCR*.

Na literatura, o Método Global mais citado é o de Otsu. No caso específico de cheques bancários, deve-se destacar que os melhores resultados pôr Métodos Globais não foram obtidos pelo Método de Otsu (86.12 %) como era de se esperar mas pelo Método Fuzzy da abordagem de Yager com 86.60 % para cheques não preenchidos. No caso de cheques preenchidos a máquina, o Método de Otsu voltou a liderar com 67,87 % seguido do Método por entropia de Li-Lee com 64.09 %. Já, pôr Métodos Locais Adaptativos, o Método de Bersen forneceu 92.79 % e 84.03 % respectivamente.

No presente Capítulo foram mostrados os resultados obtidos com a aplicação das metodologias de segmentação por binarização aplicadas à base de dados. No próximo Capítulo serão discutidas as conclusões e perspectivas futuras.

Capítulo 6

CONCLUSÕES E PERSPECTIVAS

O processamento de documentos, no campo da Informática, é um assunto apaixonante. A possibilidade de se informatizar alguns processos manuais onde pode haver economia de tempo e diminuição de trabalhos manuais é um grande desafio.

Como foi ressaltado neste trabalho, um dos documentos mais utilizados no Brasil é o cheque bancário como forma de pagamento em transações financeiras. A automatização da cadeia de compensação de um cheque seria muito mais eficiente se a informática estivesse presente em todo o processo. Para isso, o primeiro passo é a binarização da imagem do cheque para o seu pós-processamento.

Propunha-se um estudo sobre os Métodos de Segmentação por Binarização para cheques bancários visando a análise de qualidade final de reconhecimento da imagens obtidas. Objetivava-se ainda, quantificar a qualidade das imagens processadas pôr meio de reconhecimento pôr *OCR*. Mostrou-se que isso é viável, partindo-se de uma imagem colorida, segmentando-a em canais *HSI* e submetendo-a a Métodos de Binarização Locais e Globais, após o que, executou-se a operação de reconhecimento pôr *OCR* quantificando os resultados em tabelas. Todos os métodos ensaiados suprimiram de alguma forma, informações - de forma desejável ou não - afetando a imagem final. Algumas das técnicas de binarização se mostraram eficientes, tanto para cheques sem preenchimento como para cheques com preenchimento mecânico. Foi possível, assim, verificar que o fator "qualidade" é imprescindível para qualquer informação que se deseja extrair do documento após a sua binarização. Além do método de avaliação visual desta "qualidade", o método computacional,

baseado em um *OCR* foi utilizado com sucesso. O grande impecílio desta técnica de avaliação reside na não possibilidade de um *OCR* ler caracteres manuscritos. Desta forma, ainda não há como se executar uma avaliação de qualidade em imagens binarizadas com caracteres manuscritos, além da avaliação visual. Na figura 6.1 são ilustrados os resultados da segmentação por binarização de dois cheques preenchidos com caracteres manuscritos (dos bancos Itaú e Baneste) e a partir das técnicas que mais se destacaram do capítulo anterior 5. Podemos constatar que, apesar das excelentes taxas de reconhecimento obtidos, a qualidade dos resultados situa-se aquém do que era de se esperar no caso de cheques preenchidos com caracteres manuscritos,

As perspectivas futuras podem revelar-se promissoras em termos de resultados no caso genérico de segmentação por binarização de imagens diversas. Isto pode consistir em ensaiar outros métodos de binarização, sobretudo Métodos Locais Adaptativos como de Eikvil, de Mardia/Hainsworth, de Niblack, de Taxt, de Yanowitz entre outros, que não foram analisados pela falta de tempo necessária e pela dificuldade de obtenção de seus procedimentos matemáticos para a implementação. No caso específico de imagens de cheques bancários, a figura 6.1 mostra que as técnicas existentes devem ser ainda aprimoradas. A segmentação multimodal parece ser uma outra abordagem muito promissora. Isto porque um documento do tipo cheque bancário não possuem dois, mas três ou mais níveis de intensidade bem diferentes, todos com uma certa relevância. O fato de segmentar por binarização bimodal elimina, no caso de cheques bancários, muitas dessas informações relevantes. A segmentação multimodal é uma abordagem que já foi introduzida por [Otsu (1979)] mas ainda poucos trabalhos tem sido focalizados nesse tipo de segmentação.

021 021 0230 5 3.541.175 4 000440 5 R\$ 150,00 #

Pague por este cheque a quantia de Quento e cinquenta reais

JACOBS FAGUN

BANESTES

PAR-SERVIDOR PUBLICO
RUA PADRE RIBESIO PINTO, 38 TEL. 325-3664
PRAIA DO SUA-VITORIA-ES

CLIENTE 03/95

Vitoria, 19 de janeiro de 1998

RITA DE CASSIA FERREIRA
CPF 899.910.877-53

a)

021 021 0230 5 3.541.175 4 000440 5 R\$ 150,00 #

Pague por este cheque a quantia de Quento e cinquenta reais

JACOBS FAGUN

BANESTES

PAR-SERVIDOR PUBLICO
RUA PADRE RIBESIO PINTO, 38 TEL. 325-3664
PRAIA DO SUA-VITORIA-ES

CLIENTE 03/95

Vitoria, 19 de janeiro de 1998

RITA DE CASSIA FERREIRA
CPF 899.910.877-53

b)

021 021 0230 5 3.541.175 4 000440 5 R\$ 150,00 #

Pague por este cheque a quantia de Quento e cinquenta reais

JACOBS FAGUN

BANESTES

PAR-SERVIDOR PUBLICO
RUA PADRE RIBESIO PINTO, 38 TEL. 325-3664
PRAIA DO SUA-VITORIA-ES

CLIENTE 03/95

Vitoria, 19 de janeiro de 1998

RITA DE CASSIA FERREIRA
CPF 899.910.877-53

c)

021 021 0230 5 3.541.175 4 000440 5 R\$ 150,00 #

Pague por este cheque a quantia de Quento e cinquenta reais

JACOBS FAGUN

BANESTES

PAR-SERVIDOR PUBLICO
RUA PADRE RIBESIO PINTO, 38 TEL. 325-3664
PRAIA DO SUA-VITORIA-ES

CLIENTE 03/95

Vitoria, 19 de janeiro de 1998

RITA DE CASSIA FERREIRA
CPF 899.910.877-53

d)

Figura 6.1: Resultados da segmentação por binarização: a) cheque bancário do banco Baneste preenchido com caracteres manuscritos, b) método de Yager, c) método de Otsu, d) método de Bernsen

Referências

- [BCB (1983)] Banco Central do Brasil, 1983, *Livro das Normas e Instruções*.
- [Bernsen (1986)] Bernsen, J., oct. 1986, Dynamic Thresholding of gray-level images, *Proc. Eighth Int'l Conf. Conf. on Pattern Recognition*, Paris, France (1251-1255).
- [Casey (1994)] Casey, E., Wong, S., Adaptative Thresholding Metodology for Images Based on Intensity Gradient Information, *Computer Vision, Graphics and Image Processing*, 51, pp 238-247, 1994
- [Gonzales et al (1992)] Gonzales, R., Woods, R. E. *Digital Image Processing*. Addison-Wesley Publishing Company, Inc. 3a.Ed, USA, 1992.
- [O ´Gorman et al (1995)] O ´Gorman L., Kasturi, R. capítulo 2, "Pixel-level processing" em Document Image Analysis, *IEEE Computer Society Press*, Los Alamitos, California, USA, 1995.
- [Haralick et al(1992)] Haralick, R.M., Shapiro,L.G. Computer and robot vision, *Addison-Wesley Publishing Company, Inc. USA*, 1992.
- [Huang et al(1995)] Huang L.K, and Wang M.J.J., 1995, Image Thresholding by Minimizing the Measures of Fuzziness, *Pattern Recognition*, v.28, no 1,(41-51).
- [Jähne (1995)] Jähne, B., Digital Image Processing - Concepts, Algorithms and Scientific Applications, 3nd ed. *Spring-Verlang Berlin Heidelberg*, 1995.
- [Johannsen et al (1982)] Johannsen G. and Bille J., A Threshold Selection Method using Information Measures. *Proceedings, 6th Int. Conf. Pattern Recognition, Munich, Germany*. pp.140-143, 1982.
- [Kamel et al (1993)] Kamel. G. H., Zhao, J., Analyzing Oriented Patterns, *IEEE Computer Society Press*, 1993.

- [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*. v.29, pp.273-285, 1985.
- [Koerich (1997)] Koerich. A. L., *Uma Proposta para o Processamento de Imagens de Cheques Bancários: Extração e Armazenamento de Informações*, Campinas, Dissertação (Mestrado), Universidade Estadual de Campinas, São Paulo, Fev. 1997
- [Langford (1984)] Langford, M. J., *Fotografia Básica - 4ª Edição*, Ed. Martins Fontes, São Paulo, 1984.
- [Li et al (1993)] Li C.H. and Lee C.K., 1993, Minimum Cross Entropy Thresholding, *Pattern Recognition*, v.26, no 4,(616-626).
- [Mitchell et al (1989)] Mitchell, C.H., Gillies, L., *Fundamentals of Digital Image Processing*, Prentice-Hall Information and System Sciencies Series, Englewood Cliffs, NJ, 1989.
- [Morita (1998)] Morita M. E., *Estudo para a melhoria da correção da inclinação da linha de base de palavras manuscrita*, Curitiba, Dissertação (Mestrado), Pós-Graduação em Engenharia Elétrica e Informática Industrial, Centro Federal de Educação Tecnológica do Paraná, 1998.
- [Ohta et al (1980)] Ohta, Y., Kanade, T., Sakai, T. Color Information for Region Segmentation. *Computer Graphics and Image Processing*. v.13, pp.222-241, 1980.
- [Otsu (1979)] Otsu N., A Threshold Selection Method from Gray-level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*. v. SMC 9, no 1, pp.62-66, 1979.
- [Pacheco (1998)] Pacheco, J.L.Jr., email: "jpacheco@banco-central.gov".
- [Parker (1996)] Parker J.R., 1996, Algorithms for Image Processing and Computer Vision, *John Wiley and Sons*, (145-148).
- [Peleg (1980)] Peleg S., A New Probabilistic Relaxation Scheme. *IEEE Trans. Pattern Analysis and Machine Intelligence*. PAMI-2, pp.362-369, 1980.
- [Pun (1981)] Pun T., 1981, Entropic Thresholding, the new Approach, *Computer Graphics and Image Processing*, v.16, (210-239).
- [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*. v.41, pp.233-260, 1988.

- [Thrussel (1979)] Thrussel, H.J., Picture Threshold using an Iterative Selection Method. *IEEE Transactions on Systems, Man and Cybernetics*. v.SMC 9, no.5, pp.311, 1979.
- [Trie et al (1995)] Trier, O.D. and Jain A.K., Goal-Directed Evaluation of Binarization Methods. *IEEE Trans. Pattern Analysis and Machine Intelligence*. vol.17, no.12, pp.1191-1201, december 1995.
- [Umbaugh et al (1993)] Umbaugh, S.E. et al., Automatic Color Segmentation Algorithms with Application to Skin Tumor Feature Identification. *IEEE Enginnering in Medicine and Biology*. pp.75-82, September, 1993.
- [Weber (1997)] Weber A. *Avaliação da proliferação celular baseada em PCNA-CICLINA por análise morfológica aplicada a linfomas malignos não-Hodgkin*, Curitiba, Dissertação (Mestrado), Pós-Graduação em Engenharia Elétrica e Informática Industrial, Centro Federal de Educação Tecnológica do Paraná, 1997.
- [Weszka et al (1974)] , Weszka J.S., Nagel R.N., Rosenfeld A., A Threshold Selection Technique. *IEEE Trans. Computer*. v.C-23, pp.1322-1326, 1974.
- [Wilson et al (1988)] Wilson, R. and Spann M., Image Segmentation and Uncertainly. *New York: John Wiley & Sons Inc.*. 1988.
- [Wong (1978)] Wong, K., Image Segmentation by texture. *IEEE Computer Society Press*, 1978
- [Yager (1979)] Yager R.R., 1979, On the Measures of Fuzziness and Negation.Part.1: Membership in the Unit Interval, *Int Journal of Gen. Sys*, v.5, (221-229).