

**EDERSON MARCOS SGARBI**

**SEGMENTAÇÃO DO CONTEÚDO E  
ESTIMATIVA DO FUNDO POR  
MORFOLOGIA MATEMÁTICA EM  
COR DA PRIMEIRA BÍBLIA DE  
GUTENBERG**

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

Curitiba  
2013

**EDERSON MARCOS SGARBI**

**SEGMENTAÇÃO DO CONTEÚDO E  
ESTIMATIVA DO FUNDO POR  
MORFOLOGIA MATEMÁTICA EM  
COR DA PRIMEIRA BÍBLIA DE  
GUTENBERG**

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

Área de Concentração: Ciência da Computação

Orientador: Prof. Dr. Jacques Facon

Curitiba  
2013

Dados da Catalogação na Publicação  
Pontifícia Universidade Católica do Paraná  
Sistema Integrado de Bibliotecas – SIBI/PUCPR  
Biblioteca Central

S142s Sgarbi, Ederson Marcos  
2013 Segmentação do conteúdo e estimativa do fundo por morfologia matemática em cor da primeira bíblia de Gutenberg / Ederson Marcos Sgarbi ; orientador, Jacques Facon. – 2013.  
xx, 240 f. : il. ; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Paraná,  
Curitiba, 2013  
Bibliografia: f. 234-240

1. Bíblia – Conservação e restauração. 2. Sistema Binário (Matemática).  
3. Informática. I. Facon, Jacques. II. Pontifícia Universidade Católica do  
Paraná. Programa de Pós-Graduação em Informática. III. Título.

CDD 20. ed. – 004



Pontifícia Universidade Católica do Paraná  
ATA DE DEFESA DE TESE DE DOUTORADO  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA  
ÁREA DE CONCENTRAÇÃO: CIÊNCIA DA COMPUTAÇÃO

**DEFESA DE TESE DE DOUTORADO Nº 021/2013**

Aos 25 dias de Novembro de 2013 realizou-se a sessão pública de Defesa da Tese de Doutorado intitulada "**Segmentação do Conteúdo e Estimativa do Fundo por Morfologia Matemática em Cor da Primeira Bíblia de Gutenberg**" apresentada pelo aluno **Ederson Marcos Sgarbi** como requisito parcial para a obtenção do título de Doutor em Informática, perante uma Banca Examinadora composta pelos seguintes membros:

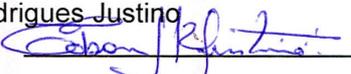
Prof. Dr. Jacques Facon  
PUCPR (Orientador)

  
\_\_\_\_\_  
(assinatura) Aprovado  
(aprov/reprov.)

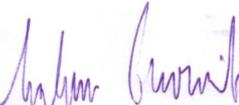
Profª. Drª. Cinthia Obladen de Almendra Freitas  
PUCPR

  
\_\_\_\_\_  
APROVADO

Prof. Dr. Edson José Rodrigues Justino  
PUCPR

  
\_\_\_\_\_  
APROVADO

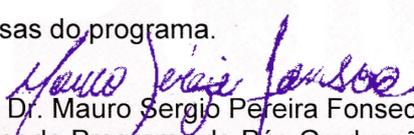
Prof. Dr. Sidnei Paciornik  
PUC-RJ

  
\_\_\_\_\_  
APROVADO

Profª. Drª. Leila Maria Vriesmann  
UEPG

  
\_\_\_\_\_  
APROVADO

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

  
Prof. Dr. Mauro Sergio Pereira Fonseca  
Diretor do Programa de Pós-Graduação em Informática



À minha esposa, Juliana, que me incentivou muito para vencer esta etapa. Soube ter paciência e maturidade durante os dias que permaneci em Curitiba. Às duas princesas da minha vida, Antônia e Yolanda, que me dão muitas alegrias e também muito trabalho. Aos meus pais, Maria Alice e José Carlos, exemplos de vida, pela minha formação, educação e pelo incansável incentivo. Ao meu irmão Sérgio, pelo apoio, incentivo, e ao meu irmão Edilson e sua esposa Sandra, que sempre me acolheram muito bem em sua casa em Curitiba.

# Agradecimentos

A Deus, por estar me proporcionando este momento tão importante na minha vida. Sem Ele não sou nada e com Ele sou tudo.

Ao prof. Jacques Facon, meu orientador, pelos ensinamentos transmitidos, incentivo, paciência e amizade durante o Doutorado.

Aos membros da banca examinadora, prof. Dr. Sidnei Paciornik, prof. Dr<sup>a</sup> Cinthia Obladen de Almendra Freitas, prof. Dr. Edson José Rodrigues Justino e prof. Dr<sup>a</sup> Leila Maria Vriesmann.

Aos professores e funcionários do PPGIA, meus sinceros agradecimentos pelo apoio, compreensão e colaboração.

Aos professores e colegas de viagem a Curitiba: Daniela Guilhermino, Luiz Fernando Legore, Ricardo Castanho e Cristiane Yanase (*in memoriam*), grandes companheiros e amigos durante este período, sempre um animando o outro para jamais desistir.

Ao prof. Wellington Aparecido Della Mura, grande amigo e colega de trabalho, que me ajudou muito nas dificuldades encontradas para otimização da implementação.

Ao aluno Victor Garcia Ronci, que me ajudou na elaboração das imagens sintéticas.

A prof. Alba Codato de Mello Martins, que me ajudou na correção do texto.

À Coordenação do Curso de Sistemas de Informação - Prof. Biluka - e Ciência da Computação - Prof. Merlin - da UENP, que disponibilizaram-me tempo de estudos para que eu terminasse este trabalho.

A todos os professores e funcionários do Centro de Ciências Tecnológicas da UENP, que sempre me incentivaram para esta conquista.

À UENP e Fundação Araucária, pelo apoio a esta pesquisa.

À Pontifícia Universidade Católica do Paraná - PUCPR - que forneceu espaço e tecnologias para o desenvolvimento deste trabalho.

A todas as pessoas aqui não mencionadas, porém eternamente lembradas e que estiveram ao meu lado colaborando de alguma forma para a realização deste trabalho o meu MUITO OBRIGADO!

# Sumário

<b>Agradecimentos</b>	iii
<b>Sumário</b>	iv
<b>Lista de Figuras</b>	ix
<b>Lista de Tabelas</b>	xvii
<b>Lista de Símbolos</b>	xviii
<b>Resumo</b>	xix
<b>Abstract</b>	xx
<b>Capítulo 1</b>	
<b>Introdução</b>	1
1.1 Objetivos Gerais . . . . .	2
1.1.1 Objetivos Específicos . . . . .	2
1.2 Justificativa . . . . .	4
1.3 Contribuição . . . . .	5
1.4 Estrutura do Trabalho . . . . .	5
<b>Capítulo 2</b>	
<b>História dos Livros Antigos</b>	7
2.1 História dos Livros antigos antes da invenção de Gutenberg . . . . .	8
2.2 História da obra de Gutenberg . . . . .	10
2.2.1 Bíblias antigas existentes de Gutenberg . . . . .	13
2.2.2 Base de Dados da Bíblia de Gutenberg - Museu Mainz - Alemanha	14
2.3 Pesquisas sobre a Bíblia de Gutenberg . . . . .	15
2.3.1 Projeto HUMI . . . . .	15
2.3.2 Projeto Debora . . . . .	16
2.4 Considerações Finais da História dos Livros Antigos . . . . .	18

## Capítulo 3

<b>Fundamentação Teórica</b>	19
3.1 Percepção Humana da cor . . . . .	19
3.2 Espaço de Cor . . . . .	20
3.2.0.1 Modelo de cor RGB . . . . .	22
3.2.0.2 Modelos de cor HSI, HSL e HSV . . . . .	23
3.2.0.3 Modelos de cor YUV, YIQ e YCrCb . . . . .	29
3.3 Morfologia Matemática . . . . .	31
3.3.1 Morfologia Matemática Binária . . . . .	32
3.3.1.1 Operadores Morfológicos Binários . . . . .	33
3.3.1.2 Morfologia Matemática Soft . . . . .	34
3.3.2 Morfologia Matemática em Níveis de Cinza . . . . .	35
3.3.2.1 Operadores Morfológicos em Níveis de Cinza . . . . .	35
3.3.3 Morfologia Matemática em cor . . . . .	38
3.3.3.1 Ordenações . . . . .	38
3.3.3.2 Ordem dos canais . . . . .	44
3.3.3.3 Operadores Morfológicos em Cor . . . . .	44
3.3.4 Reconstrução . . . . .	46
3.4 Binarização . . . . .	47
3.4.1 Limiarização Global . . . . .	47
3.4.1.1 Método de Limiarização por Entropia Johannsen . . . . .	48
3.4.1.2 Limiarização por Entropia de Li-Lee (Mínima Cruzada) . . . . .	48
3.4.1.3 Limiarização MCC . . . . .	49
3.4.2 Limiarização Local . . . . .	50
3.4.2.1 Limiarização de Sauvola Otimizado . . . . .	50
3.5 Técnica do Perfil de Projeção . . . . .	52
3.6 Técnica de rotulação . . . . .	52
3.7 Técnica de Bounding Box . . . . .	53
3.8 Avaliação da eficiência dos trabalhos . . . . .	54
3.8.1 Geração de imagens <b>ground truth</b> . . . . .	54
3.8.2 Geração de imagens sintéticas . . . . .	55
3.8.3 Métricas de Avaliação . . . . .	57
3.9 Considerações Finais da Fundamentação Teórica . . . . .	59

## Capítulo 4

<b>Estado da Arte</b>	60
4.1 Restauração de Documentos Antigos . . . . .	60
4.1.1 Remoção de ruídos por filtros em documentos de imagens binárias .	60
4.1.2 Restauração de livros antigos afetados por manchas avermelhadas e papéis amarelados . . . . .	61
4.1.2.1 Restauração de livros antigos afetados por alterações mecânicas e químicas . . . . .	63
4.2 Segmentação de documentos antigos . . . . .	66
4.2.1 Segmentação baseada na morfologia matemática para identificar e extrair conteúdos de livros antigos . . . . .	66
4.2.2 Caracterização de documentos antigos baseados na análise de tex- tura das páginas . . . . .	70
4.2.3 Geração de uma versão sintética de documentos antigos . . . . .	71
4.2.4 Segmentação de imagens coloridas para a separação do fundo e do primeiro plano . . . . .	73
4.2.5 Segmentação de conteúdos de documentos antigos baseada em regiões: texto impresso de máquina, manuscritos e desenhos . . . . .	74
4.2.6 Segmentação de conteúdo de documentos antigos de diversas coleções	76
4.2.7 Separação do fundo e do primeiro plano de documentos coloridos de baixa qualidade degradados com o tempo . . . . .	78
4.2.8 Avaliação de desempenho de seis algoritmos de segmentação . . . .	79
4.2.9 Segmentação das letras contidas dentro de capitulares . . . . .	83
4.2.10 Detecção de textura baseada em ferramentas de caracterização de livros antigos . . . . .	85
4.2.11 Segmentação de texto em imagens coloridas . . . . .	87
4.2.12 Análise Reflexiva dos trabalhos de segmentação de documentos antigos	89
4.3 Aplicações baseadas na morfologia matemática em cor . . . . .	90
4.3.1 Morfologia Matemática FUZZY em processamento de imagens co- loridas . . . . .	90
4.3.2 Trabalhos relacionados com a Ordenação Lexicográfica . . . . .	92
4.3.3 Morfologia Matemática no espaço de cor HSL . . . . .	92
4.3.4 Algoritmo para Análise de Imagem de retina . . . . .	97
4.3.5 Remoção de brilhos excessivos em imagens coloridas . . . . .	98
4.3.6 Estudo Comparativo de Ordenação dos espaços de cor para morfo- logia matemática . . . . .	100
4.3.7 Análise Reflexiva dos trabalhos relacionados com a morfologia em cor	103

## Capítulo 5

<b>Método Proposto</b>	105
5.1 Introdução do Método Proposto . . . . .	105
5.1.1 Escolha do espaço de cor . . . . .	108
5.1.2 Ordenação para o espaço de cor . . . . .	108
5.2 Descrição do Método Proposto . . . . .	108
5.2.1 Nível 1 - Processo 1: Estimativa do fundo e Processo 2: Segmentação do conteúdo . . . . .	108
5.2.1.1 Filtragem Lexicográfica - Operação morfológica reconstrução	109
5.2.1.2 Nível 1 - Processo 3: Segmentação do conteúdo colorido . . . . .	113
5.2.2 Nível 2 - Processo 4: Segmentação do conteúdo do primeiro plano . . . . .	122
5.2.2.1 Diferenciação dos dois grupos . . . . .	123
5.2.2.2 Aplicação da operação de abertura binária . . . . .	124
5.2.2.3 Abordagem morfológica por abertura binária . . . . .	124
5.2.2.4 Técnica de perfil de projeção . . . . .	125
5.2.3 Nível 3 - Segmentação do conteúdo do Grupo 1: figuras, títulos, numeração de páginas e manuscritos . . . . .	128
5.2.3.1 Segmentação de figuras, títulos, manuscritos e numeração de páginas . . . . .	128
5.2.3.2 Técnica de Rotulação . . . . .	132
5.2.3.3 Técnica de Bounding Box . . . . .	133
5.2.3.4 Separação das figuras, títulos, número de páginas e manuscritos . . . . .	133
5.3 Avaliação do método proposto . . . . .	136
5.3.1 Geração das imagens <b>ground truth</b> . . . . .	136
5.3.2 Avaliação da classificação baseado nas métricas . . . . .	142
5.4 Considerações . . . . .	142

## Capítulo 6

<b>Resultados Experimentais</b>	143
6.1 Base de Dados - Bíblia de Gutenberg . . . . .	143
6.2 Experimentos . . . . .	146
6.3 Experimentos do Nível 1 . . . . .	147
6.3.1 Processo 1: Estimativa do fundo . . . . .	147
6.3.2 Processo 2: Segmentação do conteúdo do primeiro plano . . . . .	166
6.3.3 Processo 3: Segmentação do conteúdo colorido . . . . .	170

6.3.4	Resultados finais dos 03 processos do nível 1 . . . . .	175
6.3.5	Avaliação do Nível 1 e seus 03 processos . . . . .	178
6.3.5.1	Avaliação do Processo 1: Estimativa do fundo homogêneo	194
6.3.5.2	Avaliação do Processo 2: Segmentação do conteúdo do primeiro plano . . . . .	194
6.3.5.3	Avaliação do Processo 3: Segmentação do conteúdo colorido	203
6.4	Experimentos do Nível 2 . . . . .	212
6.4.1	Processo 4: Segmentação do conteúdo do primeiro plano . . . . .	212
6.4.1.1	Busca automática das regiões . . . . .	214
6.4.1.2	Limiar de corte para linhas e colunas . . . . .	214
6.4.1.3	Considerações do processo 4: Segmentação do grupo 1 e 2	217
6.5	Experimentos do Nível 3 . . . . .	217
6.5.1	Processo 5: Segmentação do Grupo 1 . . . . .	218
6.5.1.1	Aplicação da operação morfológica reconstrução em cor . .	218
6.5.1.2	Aplicação da operação abertura binária . . . . .	219
6.5.1.3	Aplicação da técnica de rotulação dos pixels do Grupo 1 .	219
6.5.1.4	Classificação dos pixels rotulados . . . . .	220
6.5.1.5	Avaliação do Nível 3 - Processo 5: Segmentação do Grupo 1 . . . . .	222
6.5.1.6	Considerações dos testes do Nível 3 - Processo 5 . . . . .	228
6.6	Considerações Finais dos Experimentos . . . . .	229

## **Capítulo 7**

<b>Conclusão</b>	231
7.1 Trabalhos Futuros . . . . .	233
<b>Referências Bibliográficas</b>	234

## Lista de Figuras

Figura 1.1	Exemplar da Bíblia no Museu de Mainz - Alemanha. (Fonte: (MAINZ, 2008)) . . . . .	2
Figura 1.2	Exemplar da Bíblia no Museu da Universidade do Texas. (Fonte: University of Texas) . . . . .	2
Figura 1.3	Página da Bíblia de Gutenberg.(Fonte:Base de Dados de Mainz) . .	3
Figura 2.1	Exemplos dos tipos de escrita dos povos antigos (Fonte: (CALDEIRA, 2002)). . . . .	8
Figura 2.2	Exemplos do Livro de Kells (Fonte: (HEITLINGER, 2006)). . . . .	9
Figura 2.3	Exemplos de páginas de livros de vários séculos.(Fonte: (CALDEIRA, 2002)) . . . . .	10
Figura 2.4	Foto de Johannes Gutenberg. (Fonte: (MAINZ, 2008)) . . . . .	11
Figura 2.5	Invenção de Gutenberg - prensa impressão. (Fonte: (MAINZ, 2008))	11
Figura 2.6	Elementos da Página da Bíblia de Gutenberg. (Fonte: (MAINZ, 2008)) . . . . .	12
Figura 2.7	Exemplo de uma página da Bíblia de Gutenberg. (Fonte: (MAINZ, 2008)) . . . . .	13
Figura 2.8	Páginas da Bíblia de Gutenberg. (Fonte: (MAINZ, 2008)) . . . . .	14
Figura 2.9	Resultado da correção de inclinação das páginas da Bíblia. (Fonte: (OZAWA; MASSAKI, 2001)) . . . . .	15
Figura 2.10	Resultado do processo de fusão de páginas.(Fonte: (OZAWA; MASSAKI, 2001)) . . . . .	16
Figura 2.11	Exemplos de documentos antigos Digital Scriptorium. (Fonte: <i>http : //sunsite.berkeley.edu/Scriptorium</i> ) . . . . .	18
Figura 3.1	Ilustração do olho humano.(Fonte: <i>http://profs.ccems.pt/PaulaFrota/olho.htm</i> )	

Figura 3.2	Ilustração da variação da saturação do matiz vermelho.(Fonte: (WITTE, 2006)) . . . . .	21
Figura 3.3	Ilustração da luminosidade com variação do fundo.(Fonte: (WITTE, 2006)) . . . . .	22
Figura 3.4	Exemplo do cubo RGB.(Fonte: adaptado de <a href="http://www.cin.ufpe.br/marcelow.html">http://www.cin.ufpe.br/marcelow.html</a> ) . . . . .	23
Figura 3.5	Exemplo do Cone HSI. (Fonte: (KUMAR, 2006)) . . . . .	24
Figura 3.6	Conversão RGB para HSI. (Fonte: (WITTE, 2006)) . . . . .	26
Figura 3.7	Exemplo do modelo HSL. . . . .	27
Figura 3.8	Exemplo de dois elementos estruturantes. . . . .	32
Figura 3.9	Exemplo dos operadores soft erosão e dilatação. . . . .	35
Figura 3.10	Estratégia marginal. . . . .	39
Figura 3.11	Exemplo da operação de dilatação de uma imagem colorida com a ordenação marginal. (Fonte: (CALIXTO, 2005)) . . . . .	40
Figura 3.12	Histograma das imagens: (a) representando a imagem original e (b) a imagem reconstruída após a dilatação. (Fonte: (CALIXTO, 2005)) . . . . .	40
Figura 3.13	Estratégia vetorial. . . . .	41
Figura 3.14	Exemplo da imagem reconstruída em cor. . . . .	47
Figura 3.15	Ilustração da aplicação do perfil de projeção horizontal e vertical. . . . .	52
Figura 3.16	Ilustração da rotulação ideal dos pixels de uma imagem da Bíblia. . . . .	53
Figura 3.17	Exemplo de dois conjuntos rotulados e etiquetados com as áreas retangulares. . . . .	54
Figura 3.18	Exemplo de imagem <i>ground truth</i> da segmentação do texto da imagem(a). . . . .	55
Figura 3.19	Exemplo da união das imagens fundo e primeiro plano. . . . .	56
Figura 4.1	Exemplo do filtro MDMF e IGSF (PING; LIHUI; ALEX, 2000). . . . .	61
Figura 4.2	Exemplo da restauração do fundo (RAMPONI, 2005). . . . .	62
Figura 4.3	Exemplo da restauração de uma imagem com manchas avermelhadas (RAMPONI, 2005). . . . .	63
Figura 4.4	Exemplo da restauração de rachaduras e riscos de uma imagem (ARCIDIACONO; PORTUESE; SCIFO, 2005). . . . .	63
Figura 4.5	Exemplo da restauração de uma imagem com papel rasgado (ARCIDIACONO; PORTUESE; SCIFO, 2005). . . . .	64
Figura 4.6	Exemplo da restauração de uma imagem amarelada (ARCIDIACONO; PORTUESE; SCIFO, 2005). . . . .	65

Figura 4.7 Exemplo da restauração de uma imagem com mancha avermelhada (ARCIDIACONO; PORTUESE; SCIFO, 2005). . . . .	65
Figura 4.8 Exemplo da restauração de uma imagem com manchas de água (ARCIDIACONO; PORTUESE; SCIFO, 2005). . . . .	66
Figura 4.9 Fluxograma do método de segmentação do conteúdo das páginas antigas (GRANADO; PINA; MUGE, 2001). . . . .	67
Figura 4.10 Nível 1: Segmentação do texto e não texto (GRANADO; PINA; MUGE, 2001). . . . .	68
Figura 4.11 Nível 2: Segmentação das figuras e não figuras da imagem (GRANADO; PINA; MUGE, 2001). . . . .	68
Figura 4.12 Nível 3: Segmentação de listras e letras iluminadas (GRANADO; PINA; MUGE, 2001). . . . .	69
Figura 4.13 Nível 2: Segmentação das anotações manuscritas e do texto central impresso (GRANADO; PINA; MUGE, 2001). . . . .	70
Figura 4.14 Exemplo baseado no cálculo das direções dos pixels das imagens. (JOURNET; RAMEL; MULLOT, 2008) . . . . .	71
Figura 4.15 Exemplo da metodologia. (MELLO; LINS, 2002) . . . . .	72
Figura 4.16 Resultado da aplicação da metodologia de Mello e Lins (2002) . . .	73
Figura 4.17 Exemplo de um manuscrito com o fundo irregular e o método de segmentação aplicado à imagem (BOUSSELLAA; ZAHOUR; ALIMI, 2008). . .	74
Figura 4.18 Resultado da classificação de duas imagens. Fonte: (BAIRD, 2007). . .	76
Figura 4.19 Um exemplo de uma imagem do documento do conjunto de testes (BAIRD; MOLL, 2008). . . . .	77
Figura 4.20 Dois Exemplos da segmentação de Manuscritos históricos (GARAIN; PAQUET; HEUTTE, 2005). . . . .	79
Figura 4.21 Exemplo da segmentação dos seis algoritmos analisados. (SHAFAIT; KEYSERS; BREUEL, 2007) . . . . .	82
Figura 4.22 Exemplo da decomposição de uma imagem capitular em três camadas (COUSTATY, 2010). . . . .	84
Figura 4.23 Exemplo da extração da letra de um capitular (COUSTATY, 2010). . .	85
Figura 4.24 Etapas do processo de detecção de textura em classes (JOURNET, 2006). . . . .	86
Figura 4.25 Exemplo da segmentação da área de caracter pelo algoritmo K-means (HASE, 2004). . . . .	87
Figura 4.26 Exemplo da extração das áreas de caracteres com 05, 10 e 30 cores (HASE, 2004). . . . .	88

Figura 4.27 Resultado da comparação entre os algoritmos de popularidade e o método proposto (HASE, 2004). . . . .	89
Figura 4.28 Ilustração da dilatação e erosão Fuzzy por Popov (2007). . . . .	91
Figura 4.29 Resultado da erosão e dilatação com ordem LSH. (Fonte: (HANBURY; SERRA, 2001)) . . . . .	93
Figura 4.30 Resultado da erosão e dilatação com ordem SLH. (Fonte: (HANBURY; SERRA, 2001)) . . . . .	94
Figura 4.31 Resultado da erosão e dilatação com H somente e ponderado.(Fonte: (HANBURY; SERRA, 2001)) . . . . .	96
Figura 4.32 Reconstrução geodésica de imagens coloridas. (Fonte: (ORTIZ, 2002)) . . . . .	98
Figura 4.33 Resultado da remoção de brilho excessivo. (Fonte: Ortiz, Torres e Gil (2005)) . . . . .	100
Figura 4.34 Resultado da erosão e dilatação baseado na ordenação lexicográfica RGB. (Fonte: Puente (2006)) . . . . .	101
Figura 4.35 Resultado da Erosão: (a) Lexicográfico HSI; (b) Lexicográfico SHI; (c) Lexicográfico ISH e (d) Lexicográfico HSI com $H_{ref} = 0^\circ$ . (Fonte: Puente (2006)) . . . . .	102
Figura 4.36 Resultado da Dilatação: (a) Lexicográfico HSI; (b) Lexicográfico SHI; (c) Lexicográfico ISH e (d) Lexicográfico HSI com $H_{ref} = 0^\circ$ . (Fonte: Puente (2006)) . . . . .	103
Figura 5.1 Fluxograma do método proposto de segmentação do conteúdo das páginas da Bíblia impressa por Gutenberg. . . . .	106
Figura 5.2 Exemplo de uma página da Bíblia colorida e em níveis de cinza. . . . .	107
Figura 5.3 Exemplo de recorte de uma página da Bíblia colorida e em níveis de cinza. . . . .	107
Figura 5.4 Fluxograma da restauração do fundo e extração do conteúdo do primeiro plano de documentos antigos. . . . .	109
Figura 5.5 Exemplo da imagem da Bíblia, da marcadora e da máscara. . . . .	110
Figura 5.6 Exemplo do fundo reconstruído. . . . .	110
Figura 5.7 Exemplo do fundo reconstruído com TopHat. . . . .	111
Figura 5.8 Exemplo do fundo reconstruído com a tonalidade original dos elementos do primeiro plano. . . . .	112
Figura 5.9 Exemplo do fundo reconstruído branco e estimado. . . . .	113
Figura 5.10 Exemplo da identificação do conteúdo colorido. . . . .	114

Figura 5.11 Exemplo de identificação das pequenas letras vermelhas. . . . .	115
Figura 5.12 Fluxograma da segmentação texto colorido. . . . .	116
Figura 5.13 Resultado da operação de reconstrução com ordem SHI. . . . .	117
Figura 5.14 Resultados da filtragem matemática binária soft com $m$ variando de 2 a 4. . . . .	118
Figura 5.15 Resultados da filtragem matemática binária soft com $m$ variando de 5 a 8. . . . .	119
Figura 5.16 Resultados da segmentação com $m$ variando de 2 a 4. . . . .	120
Figura 5.17 Resultados da segmentação com $m$ variando de 5 a 8. . . . .	121
Figura 5.18 Exemplo grupo 1 e 2. . . . .	122
Figura 5.19 Fluxograma da segmentação em grupos 1 e 2. . . . .	123
Figura 5.20 Exemplo da segmentação ideal do conteúdo do primeiro plano de uma página da Bíblia. . . . .	125
Figura 5.21 Ilustração da segmentação ideal dos dois grupos de conteúdos do primeiro plano. . . . .	126
Figura 5.22 Ilustração da aplicação do perfil de projeção horizontal e vertical para segmentação ideal das colunas da imagem. . . . .	126
Figura 5.23 Ilustração da intersecção ideal dos histogramas dos perfis de projeções.	127
Figura 5.24 Ilustração da segmentação ideal do conteúdo do grupo 1 e do grupo 2.	128
Figura 5.25 Ilustração do conteúdo do grupo 1: título, figura, numeração de páginas e manuscritos. . . . .	129
Figura 5.26 Fluxograma da segmentação das figuras,títulos, número de páginas e manuscritos. . . . .	130
Figura 5.27 Ilustração dos pixels originais da reconstrução do grupo 1. . . . .	131
Figura 5.28 Exemplo do grupo 1 rotulada ideal. . . . .	132
Figura 5.29 Exemplo das bounding box rotuladas para classificação. . . . .	133
Figura 5.30 Ilustração da localização dos títulos, número de páginas, manuscri- tos e figuras. . . . .	134
Figura 5.31 Ilustração da segmentação ideal do conteúdo do grupo 1: Título, número de páginas, manuscritos e figura. . . . .	135
Figura 5.32 Exemplo da geração de uma imagem sintética. . . . .	137
Figura 5.33 Exemplo da geração de uma imagem sintética e sua imagem <i>ground truth</i> . . . . .	138
Figura 5.34 Ilustração da geração da imagem <i>ground truth</i> para análise quanti- tativa. . . . .	139
Figura 5.35 Exemplo da geração de uma imagem <i>ground truth</i> dos Grupos 1 e 2.	140

Figura 5.36	Exemplo da geração da imagem <i>ground truth</i> para o grupo 1. . . .	141
Figura 6.1	Exemplo do Título de imagem da Bíblia no formato JPEG. . . . .	144
Figura 6.2	Exemplos das categorias das páginas da Bíblia. . . . .	145
Figura 6.3	Resultado da reconstrução do fundo com uso de 03 canais HSI. . .	148
Figura 6.4	Resultado da reconstrução do fundo com uso de 03 canais HSI com H variando. . . . .	149
Figura 6.5	Resultado da reconstrução do fundo com uso de 02 canais HSI. . . .	150
Figura 6.6	Resultado da reconstrução do fundo com uso de somente 01 canal HSI. . . . .	151
Figura 6.7	Resultado da reconstrução do fundo com uso de 03 canais HSL. . .	152
Figura 6.8	Resultado da reconstrução do fundo com uso de 03 canais HSL com H variando. . . . .	153
Figura 6.9	Resultado da reconstrução do fundo com uso de 02 canais HSL. . .	154
Figura 6.10	Resultado da reconstrução do fundo com uso de somente 01 canal HSL. . . . .	155
Figura 6.11	Resultado da reconstrução do fundo com uso de 03 canais YCrCb. . .	156
Figura 6.12	Resultado da reconstrução do fundo com uso de 02 canais YCrCb. .	157
Figura 6.13	Resultado da reconstrução do fundo com uso de 01 canal YCrCb. .	158
Figura 6.14	Resultado da reconstrução do fundo com uso dos 03 canais YIQ. . .	159
Figura 6.15	Resultado da reconstrução do fundo com uso de 02 canais YIQ. . .	160
Figura 6.16	Resultado da reconstrução do fundo com uso de 01 canal YIQ. . . .	161
Figura 6.17	Exemplo da diferença entre as reconstruções. . . . .	162
Figura 6.18	Resultado da estimativa do fundo baseada nas imagens recons- truídas do fundo da imagem 029r1l. . . . .	163
Figura 6.19	Resultado da estimativa do fundo baseada nas imagens recons- truídas do fundo da imagem 084r1l. . . . .	164
Figura 6.20	Exemplos da restauração do fundo com espaço de cor YCrCb. . . .	166
Figura 6.21	Resultados da operação Top Hat. . . . .	167
Figura 6.22	Testes com as técnicas de binarização. . . . .	168
Figura 6.23	Testes com as Limiarizações. . . . .	169
Figura 6.24	Resultado da extração do conteúdo do primeiro plano com pixels originais. . . . .	170
Figura 6.25	Resultados da Filtragem matemática binária Soft com $m$ variando de 2 a 4. . . . .	172

Figura 6.26 Resultados da Filtragem matemática binária Soft com $m$ variando de 5 a 8. . . . .	173
Figura 6.27 Resultado do conteúdo colorido segmentado pelo processo 3. . . . .	174
Figura 6.28 Resultado do fundo estimado com conteúdo do primeiro plano. . . . .	175
Figura 6.29 Resultado da extração do conteúdo do primeiro plano. . . . .	176
Figura 6.30 Resultado do conteúdo colorido pelo processo 3. . . . .	177
Figura 6.31 Imagens com diferentes fundos. . . . .	179
Figura 6.32 Imagens com diferentes fundos. . . . .	180
Figura 6.33 Fundos reconstruídos para geração das imagens sintéticas. . . . .	181
Figura 6.34 Fundos reconstruídos para geração das imagens sintéticas. . . . .	182
Figura 6.35 Imagem sintética do primeiro plano. . . . .	184
Figura 6.36 Imagens sintéticas do primeiro plano. . . . .	185
Figura 6.37 Ilustração da junção dos pixels de duas imagens. . . . .	187
Figura 6.38 Ilustração da junção dos pixels de uma imagem sintética com ampliação. . . . .	188
Figura 6.39 Geração das imagens sintéticas com variação da opacidade. . . . .	189
Figura 6.40 Exemplos de uma imagem sintética com 04 fundos diferentes. . . . .	190
Figura 6.41 Exemplos de 04 imagens sintéticas. . . . .	192
Figura 6.42 Exemplos de imagens sintéticas e ideais. . . . .	193
Figura 6.43 Ilustração da segmentação do conteúdo do primeiro plano de uma imagem sintética. . . . .	195
Figura 6.44 Ilustração do resultado da comparação das limiarizações para a imagem sintética. . . . .	196
Figura 6.45 Conteúdo do primeiro plano utilizado para os testes com os 08 fundos. . . . .	199
Figura 6.46 Imagens para avaliação quantitativa do processo 3. . . . .	204
Figura 6.47 Resultado da limiarização das imagens para avaliação. . . . .	205
Figura 6.48 Resultado da operação de reconstrução da imagem ideal. . . . .	206
Figura 6.49 Resultado da limiarização das imagens. . . . .	207
Figura 6.50 Ilustração do resultado da aplicação da técnica de perfil de projeção horizontal e vertical em uma imagem segmentada da Bíblia pelo nível 1. . . . .	213
Figura 6.51 Ilustração do resultado da intersecção entre os histogramas horizontal e vertical. . . . .	214
Figura 6.52 Ilustração do resultado da subtração da intersecção e da imagem limiarizada por Sauvola. . . . .	215
Figura 6.53 Ilustração do resultado da operação de dilatação da imagem resultante da intersecção dos histogramas. . . . .	216

Figura 6.54	Ilustração do resultado da aplicação do filtro para eliminar os ruídos.	217
Figura 6.55	Ilustração da reconstrução em cor do Grupo 1.	218
Figura 6.56	Ilustração da operação abertura binário.	219
Figura 6.57	Resultado da Imagem Rotulada.	220
Figura 6.58	Classificação dos rótulos.	221
Figura 6.59	Ilustração dos pixels originais classificados.	222
Figura 6.60	Ilustração da classificação do conjunto Título.	223
Figura 6.61	Ilustração da classificação do conjunto Manuscrito.	224
Figura 6.62	Ilustração da imagem sem conexão.	225
Figura 6.63	Ilustração da imagem com conexão.	226
Figura 6.64	Exemplos de 03 figuras segmentadas pelo processo 5.	227
Figura 6.65	Exemplos de letras capitulares pequenas.	228

## Lista de Tabelas

Tabela 4.1	Resultados das técnicas de pré-processamento. . . . .	61
Tabela 4.2	Resultado da classificação. . . . .	75
Tabela 4.3	Resultado da classificação. . . . .	81
Tabela 4.4	Análise Reflexiva dos trabalhos relacionados com ordenações lexicográficas. . . . .	103
Tabela 6.1	Resultado da avaliação das limiarizações pelas métricas. . . . .	197
Tabela 6.2	Resultado da comparação das técnicas de limiarizações com a variação da opacidade. . . . .	198
Tabela 6.3	Resultado com variação dos fundos com opacidade 70 % . . . . .	200
Tabela 6.4	Resultado com variação dos fundos com Opacidade 80 % . . . . .	200
Tabela 6.5	Resultado com variação dos fundos com Opacidade 90 % . . . . .	201
Tabela 6.6	Resultado com variação dos fundos com Opacidade 100 % . . . . .	201
Tabela 6.7	Resultado com variação da opacidade nas 160 imagens testadas para o processo 2. . . . .	202
Tabela 6.8	Resultado do Processo 3 - Limiarização direta da imagem ideal. . .	205
Tabela 6.9	Resultado do Processo 3 - Limiarização após a reconstrução da imagem ideal. . . . .	208
Tabela 6.10	Resultado com a variação da opacidade para as 160 imagens testadas para o processo 3 - Com Limiarização Direta. . . . .	209
Tabela 6.11	Resultado com a variação da opacidade para as 160 imagens testadas para o processo 3 - Com Limiarização após Reconstrução. . . . .	210
Tabela 6.12	Resultado da avaliação do processo 3 para imagens com opacidade 100 %. . . . .	210

## Lista de Símbolos

$\varepsilon$	erosão
$\delta$	dilatação
$\gamma$	abertura
$\phi$	fechamento
$\mathit{tophat}_\gamma^B(f)$	top hat abertura
$\mathit{tophat}_\phi^B(f)$	top hat fechamento
$\rho(f)$	reconstrução
$\mathit{inf}$	mínimo
$\mathit{sup}$	máximo
$\mathit{wedge}$	mínimo entre vários valores
$\mathit{vee}$	máximo entre vários valores
$D(H, H_{ref})$	distância entre $H$ e $H_{ref}$
$ME$	erro de classificação
$RAE$	erro relativo da área do primeiro plano
$P$	precisão
$A$	acurácia
$R$	revocação
$ER$	erro
$FM$	medida F
$NRM$	a taxa negativa
$T^+$	verdadeiro positivo
$T^-$	verdadeiro negativo
$F^+$	falso positivo
$F^-$	falso negativo

# Resumo

Segmentar documentos antigos é uma tarefa árdua que envolve a aquisição e análise de conteúdo específico. Este trabalho apresenta uma abordagem de segmentação do conteúdo e estimativa do fundo das páginas da Bíblia impressa por Gutenberg, empregando ferramentas de morfologia matemática em cor, em cinza e binária e algoritmos de limiarização. O uso das operações da morfologia matemática em cor exige a escolha adequada do espaço de cor, do tipo de ordenação e da ordem dos canais. Neste estudo, a ordenação lexicográfica em conjunto com vários espaços de cor descorrelacionados, baseados na percepção humana da cor (HSI, HSL, YCrCb, YIQ), foram usados e testados. A abordagem proposta é dividida em 05 etapas, permitindo a estimativa do fundo e a segmentação do conteúdo do primeiro plano. O processo de segmentação consiste em dividir em duas classes as informações do primeiro plano, visando separá-las em grupos específicos como: Títulos, Figuras, Manuscritos, Numeração de Páginas e colunas centrais impressas. Foram realizados testes com imagens reais e imagens sintéticas das páginas da Bíblia impressa por Gutenberg. As imagens sintéticas foram geradas juntamente com as imagens ground truth para a validação pelas métricas de avaliação. Os resultados apresentados mostram-se promissores no que diz respeito à estimativa do fundo, à segmentação do conteúdo do primeiro plano e à segmentação das letras pequenas coloridas.

**Palavras-chave:** 1 - Segmentação; 2 - Restauração; 3 - Morfologia em cor; 4- Documentos Antigos;

# Abstract

Old documents segmentation is an arduous task that involves the acquisition and analysis of specific contents. This thesis presents an approach to estimate the background and segment texts of the 1st Bible pages printed by Gutenberg. This approach is performed by using color, gray and binary mathematical morphological tools and thresholding algorithms. The use of color mathematical morphological operators requires the choice of the appropriate color space, the kind of morphological order and the channel order. In this study, the lexicographical ordering and uncorrelated channel color spaces, based on human color perception ( HSI , HSL , YCrCb , YIQ ), were used and tested. The approach is divided into 05 steps, allowing the background estimation and foreground content segmentation. The process of segmentation consists in dividing the foreground in two classes to separate specific groups as: Titles, Figures, Manuscripts, Page Numbering and central printed columns. Tests were performed with real images and synthetic images of the 1st Bible pages printed by Gutenberg.. The synthetic images were generated with the images ground truth to validate the evaluation metrics. The results showed this approach promising to estimate the background, to segment complex foreground contents and small colored letters.

**Keywords:** 1- Restauration; 2 - Segmentation; 3 - Morphology in color; 4-Old Documents.

# Capítulo 1

## Introdução

Com a melhoria da tecnologia de impressão a partir do século XV, houve um crescimento na publicação de documentos, na elaboração e confecção de livros impressos. A maioria das obras literárias já desapareceu e o que foi salvo, pode ter sofrido com armazenamentos inadequados e desgaste do tempo. Para evitar a destruição e os desgastes provenientes do manuseio, as obras literárias antigas são, muitas vezes, inacessíveis a pesquisadores e ao público.

Os livros antigos devem ser preservados e disponibilizados, pois fazem parte do patrimônio histórico da humanidade. Para se consultar essas obras sem danificá-las e degradá-las, elas devem ser disponibilizadas em formato de mídia digital. A intenção de criar versões digitais de documentos deve ser enriquecida com a possibilidade de acesso de uma forma interativa, permitindo a sua consulta com, por exemplo, a presença ou ausência de certas características gráficas ou palavras-chaves. Os documentos antigos apresentam-se em várias categorias, gêneros, línguas e escritas, cada qual possuindo suas particularidades, como desenhos, manuscritos, letras impressas, letras capitulares, letras coloridas e espaços entre linhas. Por isso há uma série de tentativas para manter, organizar, segmentar e restaurar documentos antigos impressos. Para atingir esses objetivos é necessário, em primeiro lugar, identificar os componentes que constituem as páginas dos livros antigos, como por exemplo a existência de desenhos, texto impresso, texto manuscrito, títulos e outros. Na última década, várias técnicas de análise de imagens digitais têm sido amplamente utilizadas, mostrando uma possibilidade de criar procedimentos automáticos para processar grandes quantidades de dados de livros antigos.

A Bíblia impressa por Johannes Gutenberg (MAINZ, 1900) é o símbolo-chave de um momento de transição da história humana e exemplo de um documento antigo impresso com suas particularidades. A invenção da imprensa provocou uma revolução: a propagação do conhecimento para todos. Para Gutenberg, o Projeto Bíblia foi a obra de

sua vida. Não somente pelo significado deste livro, mas também porque a sua extensão foi considerável. A obra de dois volumes compreende 1.282 páginas com 42 linhas cada - daí provém a abreviação B-42 para a Bíblia de Gutenberg - e aproximadamente 3 milhões de caracteres. Esta impressão da Bíblia integra o Antigo e o Novo Testamento. Durante 3 anos, de 1452 até 1455, Gutenberg trabalhou com 20 colaboradores na obra. Do ponto de vista econômico foram custos consideráveis. Dos 180 exemplares da edição B-42, 150 deles foram impressos em papel de manufatura Italiana e 30 em pergaminho, existem hoje apenas 49 exemplares. As Bíblias foram impressas na oficina de Gutenberg, em Mainz, Alemanha. A Figura 1.1 ilustra um exemplar da Bíblia disponível no Museu de Mainz - Alemanha e a Figura 1.2 um exemplar na Biblioteca da Universidade do Texas, USA. A Figura 1.3 ilustra uma página da Bíblia do Museu de Mainz - Alemanha.



Figura 1.1: Exemplar da Bíblia no Museu de Mainz - Alemanha. (Fonte: (MAINZ, 2008))



Figura 1.2: Exemplar da Bíblia no Museu da Universidade do Texas. (Fonte: University of Texas)

## 1.1 Objetivos Gerais

Nesta pesquisa, o objetivo principal é a segmentação do conteúdo semântico da Bíblia impressa por Gutenberg com a utilização da morfologia matemática binária, cinza e em cor, com o auxílio da técnica de limiarização, para auxiliar os historiadores nas suas pesquisas. A segmentação consiste em separar o texto tradicional dos desenhos complexos e coloridos, dos títulos, das numerações de páginas, do verso e das anotações manuscritas.

### 1.1.1 Objetivos Específicos

A estimativa do fundo e a segmentação do conteúdo semântico das páginas da Bíblia impressa por Gutenberg incluem:

Equare historiam et nascentis ecclesie  
 infantiam regere: sed si nouerim? scripro-  
 ran eorum lucam esse medicinam. cuius  
 laus est i euangelio. animaduertim?  
 pariter omnia uerba illi? anime lan-  
 guentis esse medicinam. **Jacobus**. pre-  
 lohanes. iudas. septem epistolas edi-  
 derunt tam misticas q̄ succindas. et  
 breues pariter ac longas: breues in uer-  
 bis. longas in sententijs: ut rarus sit  
 qui nō in earū recusat ledione. **Apo-**  
**calipsis** iohannis tot habet sacramen-  
 ta quor uerba. **Paru** dixi: et pro meo  
 uoluminis laus omnis inferior est.  
**In** uerbis singulis. multiplices latent  
 intelligentie. **Oro** te fratere carissime  
 in hoc uicere ista meditari: nil ali-  
 ud nosse. nichilq; aliud querere: nōne  
 uiderur tibi iam hic in terris regni ce-  
 lestis habitaculū? **Nolo** ut offendam  
 in scripturis sanctis. simplicitate et sibi  
 uilitate uerborū: que uel uicio interceptū  
 uel de industria sic prolata sunt. ut ru-  
 dicam contumeliam facillime insuecēt:  
 et in una eademq; sententia aliter do-  
 ctus. aliter sentietur indoctus. **Nō** sum  
 tam peccans et hebes ut hec me nol-  
 se pollicear: et eorum fructus carere i ter-  
 ra. quorū radices in celo fixæ sunt: sed  
 uelle fatore: sedenti me profecto: magi-  
 strum renuens. comitem spondeo. **Pe-**  
**terenti** datur: pulsam aperitur: quecōs  
 inuenit. **Dilcans** i ceteris: quorū scētia  
 nobis p̄fueret in celo. **Abiis** te ma-  
 nibs recipiā: et ut inepte aliquid ac de  
 hec magore tumiditate effundā. quicqd  
 quecōs tecum scire conabor. **C. viij.**

**H**abes hic amanuallium cui scan-  
 dulum. qui litteras tuas michi  
 gratia duplicauit: referens honestatē  
 morū tuorū. et tempū seculi. fidem ami-  
 citie. amorem xpi. **Nam** prudentiā et

eloquiū ueritatē etiā absq; illo ipsa  
 epistola p̄ferebat. **Festina** queso te: et  
 herēti i salo nauicade funem magis p̄-  
 scinde q̄ solue. **Auaro** renūciatūrus  
 seculo bene potest uendere. que cōtem-  
 pte ut uendat. **Quicquid** in sumptū  
 de tuo tuleris: pro lucro computa. **Ali-**  
**quis** didum est. **Auaro** dicit tam qd  
 habet: q̄ quod non habet. **Credenti**  
 totus mūdus diuinarū est: **Infidelis**  
 autē etiā obulo indiget. **Hic** uiuam?  
 quasi nil habentes et omnia possiden-  
 tes. **Uidus** atq; uestitus diuitie xpi  
 norū. **Si** habes in potestate rem tuā  
 uende: si non habes proice. **Tolle** mi  
 tunicam: et palliū. relinquendū est.  
**Nisi** scilicet tu semper recalcinans et  
 diem de die rathēs cante et pedreptum  
 tuas possessiones uendidicis: non  
 habet xps unde alat pauperes suos.  
**Totū** deo dedit: qui se obtulit. **Aposto-**  
**li** tantū nauem et rema reliquit: ui-  
 dua duo rea mitte in gazophilatū:  
 et p̄ferur ceteri diuitijs. **Facile** cōcipit  
 oia: qui se longū cogitat esse moriturū.

**Incipit prologus in pentateuchi mori.**  
**F**ideli mei desideratas  
 accepi litteras. qui quodā  
 prelagio futurorū. cum  
 danielle sorci? est nomen:  
 obsecrans. ut translationem in latinā  
 linguam de hebreo sermone pentate-  
 uicū nostrorū auribus tradere. **Pe-**  
**riculosum** opus certe. et obsecratorū  
 meorū lacrimibus patens: qui me affe-  
 runt i septuaginta interpretum singilla-  
 rionem. noua p̄ uerbis audire. ita  
 ingenū q̄i uim probantes: cum ego  
 sepissime testat? sum me pro uisi por-  
 one in tabernaculo dei offere q̄ possim:  
 nec opes aliorū aliorū pauprate sed a-  
 ri. **Quo** ut autem uigenti me studiū

Figura 1.3: Página da Bíblia de Gutenberg. (Fonte: Base de Dados de Mainz)

- Estudar os conceitos de morfologia matemática em cor, as ordenações e os espaços de cor envolvidos;
- Implementar e comparar, no âmbito da segmentação, a eficácia de várias morfologias em cor numa base de imagens coloridas complexas;
- Segmentar o fundo e o primeiro plano (separar em duas classes);
- Estimar o fundo deteriorado com o tempo;
- Extrair informação sobre a localização e tamanho (largura e altura) das colunas;
- Extrair informação sobre possíveis títulos nas páginas (localização, tamanho, altura e cores);
- Extrair informação da existência de trechos manuscritos (localização, tamanho e inclinação);
- Diferenciar os grafismos (desenhos complexos);
- Detectar somente o texto colorido;
- Comparar a eficácia dos resultados.

## 1.2 Justificativa

Um historiador, especialista na área de documentos antigos, atualmente encontra dificuldade em localizar e estudar, de forma rápida e hábil, o conteúdo complexo de documentos antigos, isto devido às suas particularidades e disponibilização. Uma ferramenta computacional para segmentar o texto impresso, os manuscritos, os títulos e os grafismos, trará benefícios ao estudo de documentos antigos. A Bíblia impressa por Gutenberg é um marco histórico, sendo o primeiro livro impresso e, por ser um livro antigo, apresenta vários desafios:

- Texto escrito em latim, com fonte gótica, em preto e também colorido, na maioria pré-impresso mas também podendo ser manuscrito em alguns trechos pequenos.
- Desenhos complexos coloridos.
- Manchas e desgastes
- Vazamento de tinta do reto (informações do primeiro plano) para o verso.

- Estrutura bem definida, sem inclinação e sem distorções.

A área de processamento de documentos antigos ainda carece de metodologias no âmbito de cor. Propor uma metodologia de segmentação inovadora, baseada em morfologia matemática em cor, de um documento tão antigo e tão complexo quanto a Bíblia impressa por Gutenberg, trará avanços significativos às áreas de segmentação de documentos antigos e de morfologia matemática em cor, ainda pouco divulgada e aplicada.

A morfologia matemática em cor representa atualmente uma nova linha de pesquisa desafiadora da área, possibilitando trabalhar e explorar características geométricas com informações coloridas em vários espaços em cor. Aplicando-a às páginas da Bíblia, será possível avaliar a sua eficácia para segmentar texto preto e colorido, títulos, manuscritos e grafismos.

Como a morfologia matemática explora características geométricas, será possível segmentar as páginas da Bíblia, quanto às colunas impressas, títulos, manuscritos e grafismos.

### **1.3 Contribuição**

O estudo aprofundado de vários mecanismos de morfologia matemática em cor (ordenações lexicográficas diversas e outros espaços de cores) permite avaliar visual e numericamente o potencial dessas novas técnicas e sua eficiência para imagens complexas. Outra contribuição importante da pesquisa é estender os mecanismos morfológicos conhecidos para imagens em níveis de cinza, adaptando-os para novos mecanismos morfológicos em cores. Os resultados do método proposto de segmentação e de estimativa do fundo poderão abrir novas frentes de pesquisa no que diz respeito ao processamento de documentos coloridos, antigos e recentes. Os elementos segmentados de um texto antigo formarão uma base de dados que poderá ser estudado por especialistas da área, como os historiadores, com mais facilidade e rapidez.

### **1.4 Estrutura do Trabalho**

O Capítulo 2 apresenta a História dos livros antigos, o Capítulo 3 a fundamentação teórica dos documentos antigos, das imagens antigas, da Base de dados da Bíblia de Gutenberg, da morfologia matemática binária, em níveis de cinza e em cor. Também fundamenta os algoritmos de limiarização, a técnica do perfil de projeção e as formas de avaliação dos resultados. O capítulo 4 apresenta o estado da arte, com a apresentação

de abordagens de segmentação e de restauração de documentos antigos, mecanismos de avaliação da eficiência da segmentação, a geração de imagens *ground truth* e as aplicações baseadas na morfologia em cor. O Capítulo 5 apresenta o método proposto, composto por 05 processos divididos em 03 níveis. No Capítulo 6 estão os experimentos realizados. O capítulo 7 expõe as conclusões e os trabalhos futuros.

## Capítulo 2

### História dos Livros Antigos

Há diversos livros antigos que podem ser classificados de acordo com a evolução do ser humano. As informações eram escritas em diferentes tipos de materiais e de forma manual, pelos escribas, até a invenção de Johann Gutenberg. Este capítulo descreve os livros antigos antes e após a invenção de Gutemberg. Para tanto, é preciso definir o que é um documento e um documento histórico.

Um documento (do latim *documentum*, derivado de *docere* "ensinar, demonstrar") é qualquer meio, sobretudo gráfico, que comprove a existência de um fato, a exatidão ou a verdade de uma afirmação. É o registro de uma informação independentemente da natureza.

Um documento histórico é todo material produzido em um determinado período, que possa auxiliar o historiador em sua análise. Pode-se constituir desde documentos produzidos por governos ou entidades (públicas e privadas), até mesmo objetos como: imagens, textos, pinturas, esculturas, canções, etc (DEBORA, 2000). Estes documentos são as fontes históricas de contextos, fatos, teses, que fundamentaram algum momento, sendo considerados documentos antigos, mas existem também documentos recentes considerados como históricos. Estes tipos de documentos sofrem muito com a degradação, ocasionada pelo armazenamento inadequado, pelo manuseio repetitivo sem cuidados. O acesso a estes documentos é outro desafio, devido à permissão restrita em manuseá-los, pelo fato de serem raros e também pela distância de onde estão localizados, dificultando muito a pesquisa. Os catálogos de coleções de livros antigos pertencentes a bibliotecas são geralmente compostos de ficheros específicos, escritos por especialistas de acordo com as características de cada um dos elementos da coleção, e as obras são tão frágeis e preciosas que requerem um tratamento muito cuidadoso. Portanto, digitalizar tais arquivos e coleções é uma operação mais complexa do que a digitalização de obras modernas (DEBORA, 2000).

## 2.1 História dos Livros antigos antes da invenção de Gutenberg

Nesta seção consideram-se livros antigos aqueles produzidos antes da invenção de Gutenberg, anteriores ao século XV. Caldeira (2002), no seu estudo, explica que os livros antigos tinham como objetivo registrar e difundir as informações e evoluíram de acordo com os povos:

- Sumérios: guardavam as informações em tijolos de barro.
- Indianos: guardavam as informações em folhas de palmeiras.
- Maias e Astecas: escreviam as informações (os livros) em um material macio existente entre a casca das árvores e a madeira.
- Romanos: guardavam as informações em tábuas de madeira cobertas com cera.
- Egípcios: desenvolveram a tecnologia do papiro.

A Figura 2.1 ilustra a evolução dos tipos de materiais utilizados para a escrita dos povos citados por Caldeira (2002), desde os materiais de tijolos de barros até o pergaminho.



Figura 2.1: Exemplos dos tipos de escrita dos povos antigos (Fonte: (CALDEIRA, 2002)).

Caldeira explica ainda que, após o surgimento do papiro, apareceu o pergaminho, feito da pele de carneiro. No processo de evolução, surgiu o papel, na China, no início do século II, sendo considerado o principal meio para divulgação das informações e conhecimento humano. Dados históricos mostram que este material foi difundido no ano de 1150, com a instalação da primeira fábrica de papel na cidade de Játiva, Espanha.

Vários livros importantes foram escritos antes da invenção de Gutenberg. O Livro Irlandês, de Kells (HEITLINGER, 2006), é uma das obras mais importantes de livros manuscritos iluminados da Idade Média, com figuras ornamentais, feito por monges celtas por volta do ano de 800, no estilo conhecido por arte insular. O livro de Kells é escrito em latim e contém os quatro Evangelhos do Novo Testamento, além de notas preliminares e explicativas, com ilustrações coloridas e está disponível na Biblioteca do Trinity College de Dublin, República da Irlanda, sob a referência MS A. I.. A Figura 2.2 mostra exemplos de páginas do Livro de Kells.



Figura 2.2: Exemplos do Livro de Kells (Fonte: (HEITLINGER, 2006)).

Existem inúmeras imagens de diversos livros antigos antes da invenção de Gutenberg. A Figura 2.3 ilustra imagens do século VII, IX, XIII e XIV. Nota-se pelas imagens que os desenhos e palavras escritas eram feitos de forma manual, por escribas, demonstrando uma diferença entre os tamanhos das letras manuscritas em uma mesma imagem. As imagens da Figura 2.3 ilustram um fundo deteriorado com o tempo.



Figura 2.3: Exemplos de páginas de livros de vários séculos.(Fonte: (CALDEIRA, 2002))

## 2.2 História da obra de Gutenberg

No século XV, o alemão Johannes Gensfleisch zur Laden Zum Gutenberg, que nasceu por volta de 1398 e morreu em 03 de fevereiro de 1468, revolucionou o modo de impressão para a Europa com a invenção da impressão mecânica de tipo móvel, considerado um marco importante no período moderno. Gutenberg foi o primeiro europeu a usar tipos móveis de impressão, por volta de 1439. Entre as suas muitas contribuições para a impressão estão: a invenção de um processo para a produção em massa de tipo móvel e a utilização do óleo-base de tinta. A invenção de Gutenberg permitiu a produção em massa de livros impressos, sendo considerável economicamente a produção e o custo dos livros.

Com o uso do tipo móvel de impressão houve uma melhora acentuada no uso do manuscrito, que era o método existente de produção de livros na Europa. A tecnologia de impressão de Gutenberg se espalhou rapidamente por toda a Europa e depois para o mundo.

A Figura 2.4 ilustra uma foto de Johannes Gutenberg e a Figura 2.5 ilustra a sua invenção, exposta no museu de Mainz - Alemanha.



Figura 2.4: Foto de Johannes Gutenberg. (Fonte: (MAINZ, 2008))



Figura 2.5: Invenção de Gutenberg - prensa impressão. (Fonte: (MAINZ, 2008))

Sua principal obra, a Bíblia de Gutenberg (também conhecida como a Bíblia de 42 linhas), foi aclamada por sua alta qualidade estética e técnica. Os elementos que compõem um documento antigo variam de acordo com a época, com o estilo, com desenhos, escritas a mão, escritas impressas, títulos e outros elementos. Os elementos que compõem as páginas da Bíblia de Gutenberg são:

- Títulos;
- Desenhos;
- Letras capitulares coloridas;
- Escritas impressas pretas e coloridas;
- Manuscritos;
- Paginação.

A Figura 2.6 ilustra um recorte de cada um dos elementos das páginas da Bíblia de Gutenberg.

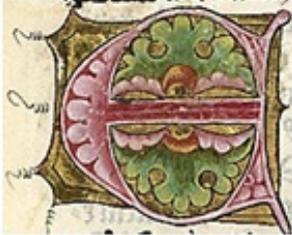
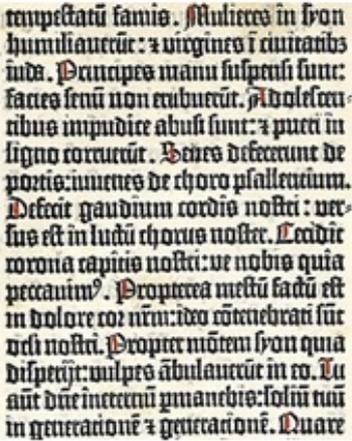
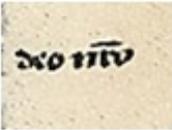
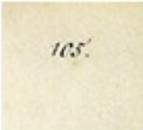
	
a) Título	b) Letra Capitalar
	
c) Escritas impressas e coloridas	d) Título
	
e) Manuscrito	f) Paginação

Figura 2.6: Elementos da Página da Bíblia de Gutenberg. (Fonte: (MAINZ, 2008))

A Figura 2.7 mostra um exemplo dos elementos que compõem uma página da Bíblia de Gutenberg: as 2 colunas de texto, os caracteres pretos e coloridos, as letras capitulares, os títulos, a numeração da página, o fundo desgastado, a dobra na parte inferior da imagem. Existe um problema de digitalização nesta imagem, pois aparece o fundo do scanner, fundo que, obviamente, não faz parte das páginas da Bíblia, mas que será levado em consideração na hora de processar as imagens digitais.

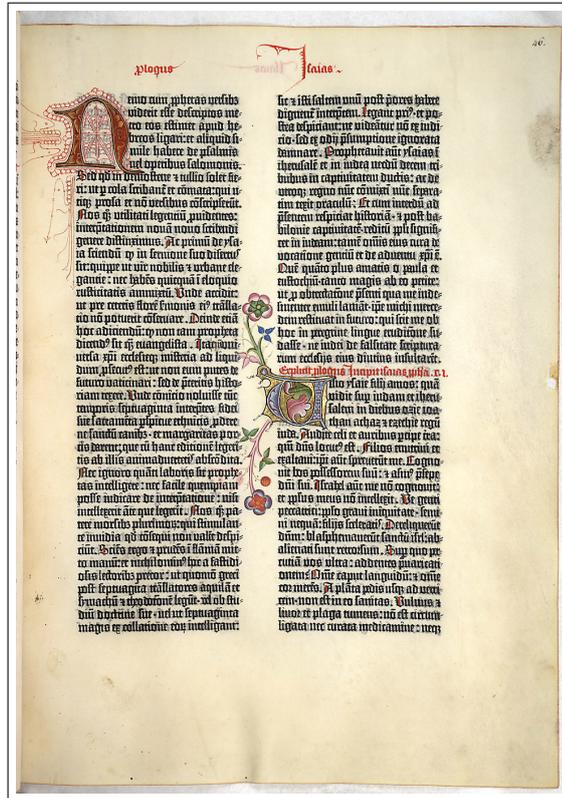


Figura 2.7: Exemplo de uma página da Bíblia de Gutenberg. (Fonte: (MAINZ, 2008))

### 2.2.1 Bíblias antigas existentes de Gutenberg

Das 180 Bíblias impressas pelo processo de Gutenberg, segundo informações do Museu do Texas (HARRY, 2001), hoje existem aproximadamente 48 exemplares, localizados em diversos países e que são listados pelo Museu. Neste trabalho iremos utilizar a base de dados adquirida do Museu de Gutenberg, em Mainz, na Alemanha.

### 2.2.2 Base de Dados da Bíblia de Gutenberg - Museu Mainz - Alemanha

A base de dados da Bíblia impressa por Gutenberg, adquirida do museu de Gutenberg, em Mainz - Alemanha (MAINZ, 2008), é composta por 1282 páginas digitalizadas. As imagens possuem dimensões de 965x1390 pixels no formato JPEG, com resolução de 72 dpi e 24 bits. Como pode ser visto na Figura 2.8, as páginas da Bíblia possuem 02 colunas de texto, caracteres pretos e coloridos, letras capitulares, títulos, numeração de página, fundo desgastado, vazamento de tinta e o problema de digitalização, já citado anteriormente, que faz aparecer o fundo do *scanner*.



Figura 2.8: Páginas da Bíblia de Gutenberg. (Fonte: (MAINZ, 2008))

## 2.3 Pesquisas sobre a Bíblia de Gutenberg

### 2.3.1 Projeto HUMI

O Projeto HUMI (WU, 2003), instituído pela Universidade Keio - Tóquio - Japão, trabalha a restauração de livros antigos, com a aplicação de tecnologias de processamento de imagens digitais, para o arquivamento de livros raros orientado a bibliotecas digitais. No momento da digitalização, as páginas não são totalmente abertas, tomando-se os devidos cuidados para não deteriorar o livro antigo. A página digitalizada pode, portanto, aparecer inclinada, deformada ou em pedaços. Os autores propõem métodos de processamento de imagens para corrigir a inclinação da página, para conectar imagens parciais (páginas digitalizadas pela metade, fazendo a união das partes) e para representar as ilustrações das páginas. Para os exemplos da utilização dos métodos foram utilizadas páginas da Bíblia de Gutenberg. O museu de Gutenberg em Mainz - Alemanha, digitalizou as páginas da Bíblia utilizando o Projeto HUMI. Para o método de representação das ilustrações foi utilizado o espaço de cor RGB. Os autores não mencionaram a quantidade de páginas de livros antigos testadas, não divulgaram o tempo de processamento, nem as taxas de acertos. A Figura 2.9 (a) mostra o exemplo da imagem original da Bíblia de Gutenberg e (b) o resultado do ajustamento da imagem.

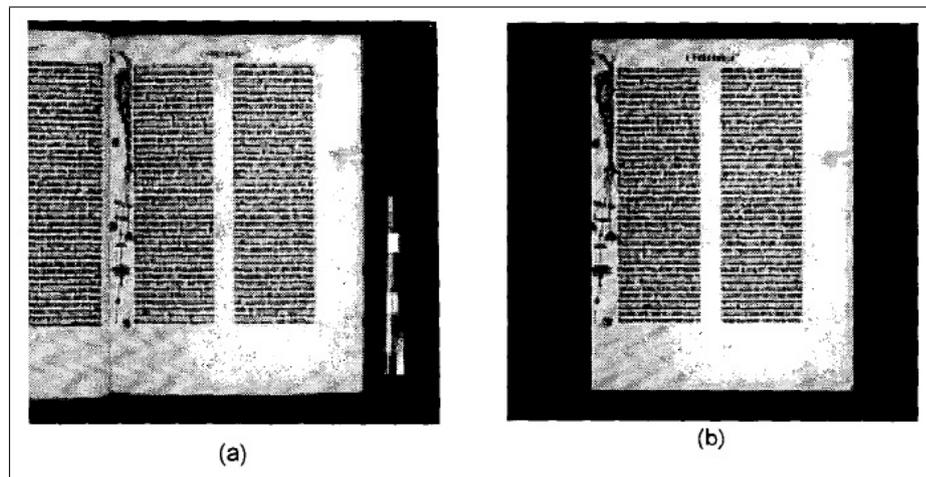


Figura 2.9: Resultado da correção de inclinação das páginas da Bíblia. (Fonte: (OZAWA; MASSAKI, 2001))

A Figura 2.10 ilustra o resultado do processo de união de partes de uma mesma página. As imagens (a) e (b) são partes de uma página e a imagem (c), o resultado da união das duas páginas.

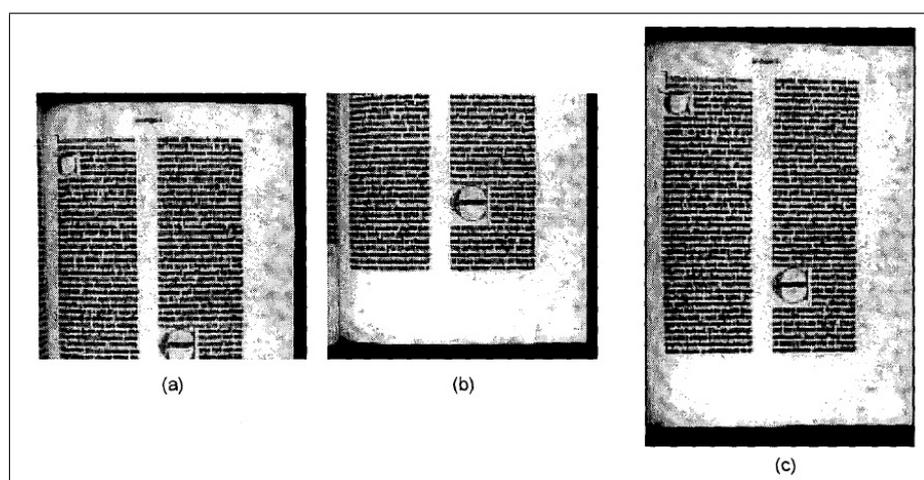


Figura 2.10: Resultado do processo de fusão de páginas. (Fonte: (OZAWA; MASSAKI, 2001))

### 2.3.2 Projeto Debora

O Projeto europeu denominado DEBORA (DEBORA, 2000) significa o acesso a livros do Renascimento, tendo como principal objetivo permitir o acesso remoto e colaborativo a livros do século XVI, sem a necessidade de contato direto com as bibliotecas em que os originais estão localizados. Disponibiliza-se para os leitores do mundo todo livros raros, que só poderiam ser consultados por um pequeno número de especialistas e estudiosos nas coleções das bibliotecas próprias. Foi escolhido o século XVI, não o XV, devido à existência de recursos mais abundantes no primeiro século de impressão e porque foi nessa época que o livro impresso assumiu sua forma moderna. Este projeto é parte do movimento para modernizar tecnologicamente bibliotecas. Desde o final dos anos oitenta, as bibliotecas tinham como objetivo tornar o acesso aos livros antigos através de redes. Numerosas iniciativas públicas e privadas já foram estudadas para o desenvolvimento de uma biblioteca digital e alguns dos problemas culturais, técnicos, organizacionais, sócio-econômicos levantados já foram resolvidos.

O projeto Debora abrange as bibliotecas de:

- Roma (Itália), com arquivos do período renascentista.
- Lisboa (Portugal), com o patrimônio histórico dos grandes navegadores.
- Lyon (França) com um acervo grande dos livros impressos do século XVI.

Os laboratórios de pesquisa de desenvolvimento do Projeto Debora foram os seguintes:

- Departamento de Ciência da Computação da Universidade de Lancaster (Inglaterra).
- Laboratório ERSICO da Universidade de Lyon (França).
- Laboratório de formas renascentista e visão de INSA de Lyon (França).
- Escola Nacional de Ciências da Informação de Villeurbanne (França).

A digitalização de acervos arquivísticos responde a uma necessidade real em centros de documentação e bibliotecas, e se justifica pela redução de custos e a criação de novos serviços prestados aos usuários. No entanto, essa escolha de disponibilizar as coleções digitalizadas tem consequências importantes, não só em níveis técnico e financeiro, mas também em termos de seu impacto sobre as práticas culturais e evolutivas. Assim, é importante considerar com cuidado os riscos e restrições desta escolha.

A digitalização abrange três principais campos de aplicação:

1. Preservação do patrimônio;
2. Acesso remoto às coleções;
3. Demanda de produção de multimídia de livros.

Este projeto relata que as coleções antigas são muito frágeis e preciosas, que requerem um tratamento especial, muito cuidadoso e a digitalização de tais documentos é uma operação muito complexa. É anunciado ainda nesse trabalho que os documentos antigos abaixo já foram digitalizados e estão disponíveis on-line.

1. Gutenberg: o volume 1 da Bíblia de Gutenberg foi digitalizado pela Universidade de Keio, Tóquio, Japão, de um exemplar da Edição Latina de 1455.
2. Cervantes 2001: um site dedicado às obras de Cervantes, incluindo a biblioteca digital com as suas obras, localizado no servidor da Universidade do Texas (HARRY, 2001).
3. O Scriptorium Digital é um banco de dados com imagem de manuscritos medievais e renascentistas disponíveis para o ensino e pesquisa. Há imagens de diversos locais do mundo, como por exemplo da Universidade da Califórnia, Universidade de Missouri, Universidade do Estado de São Francisco, Biblioteca Pública de Nova York, Biblioteca Hungtnton de San Marino e outras. Alguns exemplos são mostrados na Figura 2.11.

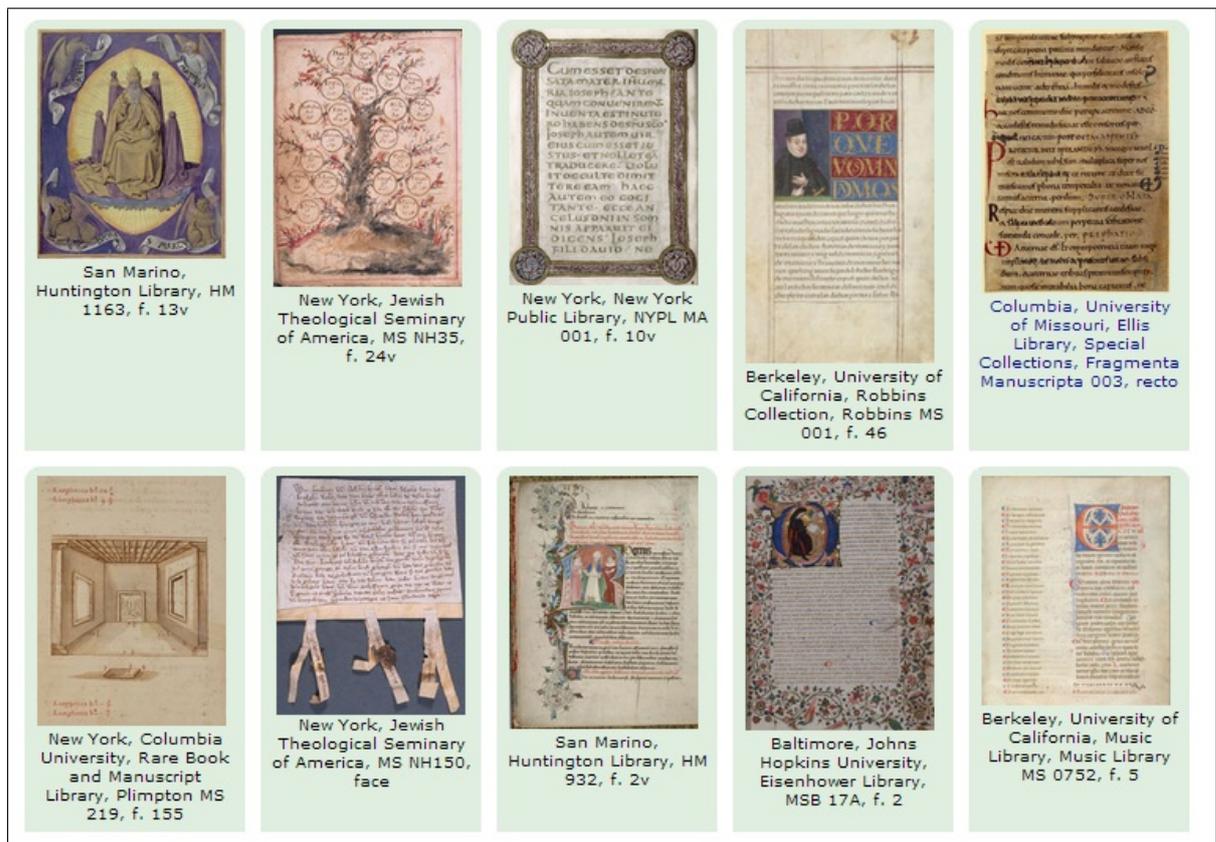


Figura 2.11: Exemplos de documentos antigos Digital Scriptorium. (Fonte: <http://sunsite.berkeley.edu/Scriptorium>)

De acordo com o Projeto Debora (2000), a produção de documentos digitalizados depende de vários critérios que devem ser levados em consideração antes de se iniciar um trabalho para as bibliotecas digitais: I) a periodicidade; II) a habilidade técnica disponível; III) o financiamento, se a biblioteca dispõe de meios financeiros; IV) a natureza dos documentos, alguns são fáceis de manusear, outros têm valor de patrimônio alto e são frágeis para serem deslocados.

## 2.4 Considerações Finais da História dos Livros Antigos

O capítulo 2 apresentou a história dos livros antigos, descrevendo o que é um documento antigo e histórico, a história da obra de Gutenberg e os trabalhos realizados com a Bíblia de Gutenberg. O Capítulo 3 apresenta a fundamentação teórica deste projeto.

## Capítulo 3

### Fundamentação Teórica

Neste capítulo, são apresentados os conceitos teóricos necessários para a realização deste trabalho. O capítulo está organizado na seguinte forma: na Seção 3.2, são descritos os conceitos dos espaços de cor com suas formulações; na Seção 3.3, são apresentados os conceitos da morfologia matemática; na 3.4 os algoritmos de limiarizações; em 3.5 apresenta-se a técnica do perfil de projeção. Na Seção 3.6 é apresentada a técnica de rotulação; em 3.7, a técnica do bounding box e na Seção 3.8 são apresentadas as formas de avaliação.

#### 3.1 Percepção Humana da cor

A percepção da cor pelo olho humano é realizado pela interpretação da luz incidente. A Figura 3.1 ilustra o esboço do olho humano.

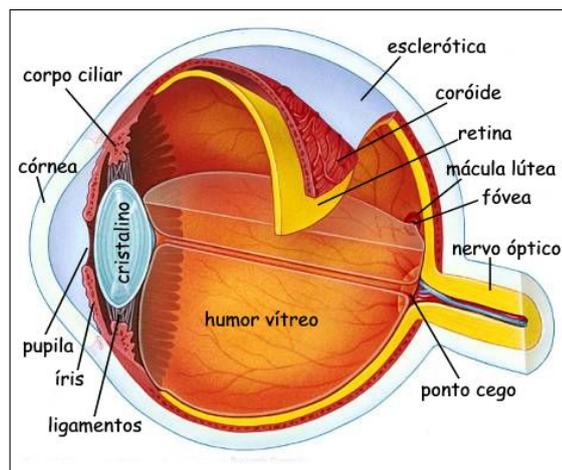


Figura 3.1: Ilustração do olho humano. (Fonte: <http://profs.ccems.pt/PaulaFrota/olho.htm>)

Witte (2006) descreve a forma de interpretação pelo olho humano. Toda luz que atinge as partes do olho sensíveis à luz tem que passar pela pupila, que é uma abertura, cujo tamanho é controlado pelo diafragma, um músculo circular. De acordo com o tamanho da pupila, mais ou menos luz pode entrar no olho. Contraíndo e relaxando o diafragma, o olho pode se adaptar a diferentes condições de luz, como dia ou noite. A córnea é uma camada protetora transparente que cobre tanto a pupila quanto a íris, sendo a primeira lente do sistema de visão. A lente interna, a segunda lente do olho, é um corpo transparente capaz de contrair e relaxar. A forma que a lente interna adquire as informações, pode permitir que o olho se adapte a uma grande gama de distâncias. Nesta lente interna também se concentram radiações de luz de entrada, ou seja, toda luz proveniente do mesmo ponto físico do espaço é reunido (focalizado) num único ponto sobre a retina. A retina é composta de várias camadas, com tarefas específicas, como por exemplo, receber luz ou convertê-la em sinais elétricos. A fóvea é chamada de centro da retina. Esta pequena área é responsável pela visão central mais nítida. A retina contém dois tipos de receptores sensíveis à luz, também chamados de fotorreceptores, os cones e bastonetes, que traduzem a luz recebida em sinais elétricos. Os sinais elétricos produzidos pelos cones e bastonetes são, então, enviados ao longo do nervo óptico a uma parte de trás do cérebro para posterior processamento.

Pode-se concluir que a resposta à pergunta "o que é cor?" é subjetiva. As cores são criadas em nosso cérebro como uma reação à luz. Cor é como o olho e o cérebro humano observam diferentes comprimentos de onda de luz.

## 3.2 Espaço de Cor

Um espaço de cor é um modelo matemático utilizado para representar uma cor por meio da combinação de componentes, sendo possível especificar, criar e visualizar cores. A percepção de cores pelo ser humano é um fenômeno fisio-psicológico que ainda não é completamente compreendido. Entretanto, a natureza física das cores pode ser expressa numa base formal suportada por resultados experimentais e teóricos (GONZAGA, 2008). Existem vários modelos de cores, como por exemplo, RGB, HSI, HSV, YCrCB, Luv, Lav, XYZ. Em especial para a área de processamento de imagens, são mais utilizados os espaços de cores RGB e HSI. O modelo RGB (*Red, Green, Blue*) tem como propósito principal a reprodução de cores em dispositivos eletrônicos como monitores de TV, computador, datashows, scanners, câmeras digitais, assim como na fotografia tradicional. O modelo HSI (H - matiz, S - saturação e I - intensidade ou brilho) é utilizado em sistemas de visão artificial, pois suas características tornam-se uma ferramenta útil para o desenvolvimento

de algoritmos de processamento de imagens fortemente baseados no modelo de percepção de cor pelo ser humano (GONZAGA, 2008).

Witte (2006) destaca que nem todas as cores visíveis pelo ser humano podem ser obtidas pelo modelo matemático RGB, mesmo existindo diferentes espaços de cor com base no modelo RGB. Espaços de cor RGB bem utilizados são sRGB (standard RGB) e Adobe RGB. A sRGB foi desenvolvida pela Hewlett-Packard e Microsoft Corporation usando uma definição de cores independente, de dispositivo simples e robusto para trabalhar com cor nos sistemas operacionais, drivers de dispositivo e na Internet. O espaço de cor Adobe RGB foi concebido pela Adobe Systems para abranger a maioria das cores realizáveis em impressoras, mas utilizando cores primárias RGB sobre o monitor do computador. Para o estabelecimento dos espaços de cor (HSV, HSI e HSL) é preciso definir alguns conceitos, descritos por Witte (2006):

**Matiz** - é a tonalidade da cor, a característica que define e distingue uma cor. Vermelho, verde e azul são matizes de cor, são considerados cores puras. Preto, branco e os valores de cinza são chamados de "cores neutras", pois não possuem tonalidade.

**Saturação** - é a característica (parâmetro) que especifica a qualidade de um matiz de cor pelo seu grau de mesclagem com a cor branca. No modelo de cor HSI, a saturação é a proporção de quantidade de cor em relação à cor cinza médio. Quanto menos cinza na composição da cor, mais saturada ela é. Cores puras são completamente saturadas. Cores que não estão saturadas possuem valores de cinza. A Figura 3.2 ilustra a variação da saturação para o matiz vermelho.

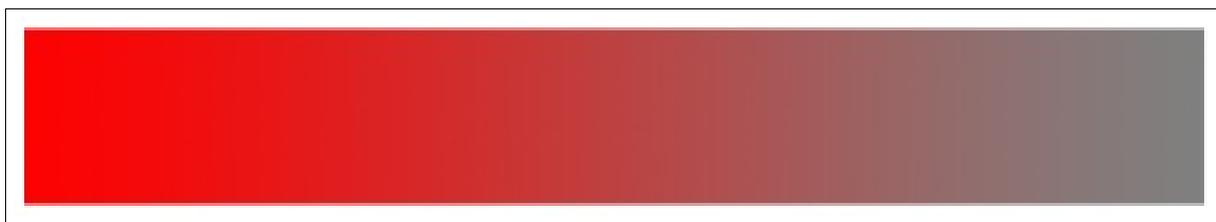


Figura 3.2: Ilustração da variação da saturação do matiz vermelho.(Fonte: (WITTE, 2006))

**Brilho** - é definido como o modo que a luz é refletida pela superfície, é uma sensação visual humana que exhibe uma área com mais ou menos luz, parecendo ser mais ou menos clara.

O brilho de um objeto depende da maneira que é iluminado, ou seja, quanto mais luz e quanto mais intensa ela for, mais brilhante será a cor. A luminosidade pode ser definida como a intensidade de uma área em relação ao brilho de uma superfície

branca igualmente iluminada. Brilho refere-se à percepção absoluta da quantidade de luz do elemento cor de interesse, enquanto a luminosidade pode ser vista como o brilho relativo. Somente cores relacionadas apresentam luminosidade. A luminosidade do fundo escolhido para ser usado pode causar uma diferença no objeto observado, chamado de efeito de contraste luminoso, e é ilustrado na Figura 3.3. Cada pequeno quadrado cinza tem o mesmo brilho físico (intensidade), isto é, os nossos olhos recebem exatamente a mesma quantidade de luz de cada uma das imagens. Espera-se que todos eles possuam o mesmo valor de cinza, que os nossos olhos observem como o mesmo valor de cinza. Mas o quadrado cinza, no fundo cinza escuro, parece ser muito mais brilhante do que o quadrado cinza no fundo cinza claro: quanto mais escuro o fundo, o pequeno quadrado cinza parece ser mais luminoso.

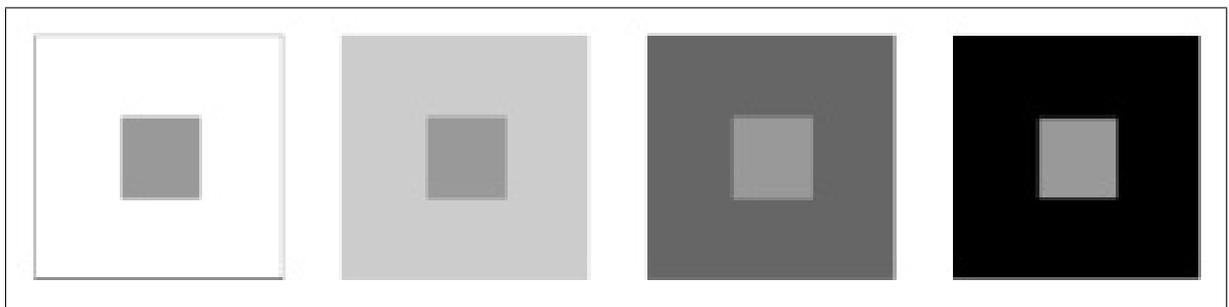


Figura 3.3: Ilustração da luminosidade com variação do fundo.(Fonte: (WITTE, 2006))

Com as definições apresentadas, serão descritos os modelos de cor RGB, HSI, YCrCb e YIQ.

### 3.2.0.1 Modelo de cor RGB

O espaço de cor RGB é um modelo no qual cada cor aparece nos seus componentes espectrais primários, vermelho, verde e azul, baseado na síntese aditiva, sendo possível representar uma cor com a adição das três cores primárias. Os componentes do modelo RGB podem variar entre 0 e 255, definindo-se um cubo de cor, sendo cada ponto representante de uma cor, em que o valor (0,0,0) corresponde ao preto, e o valor (255,255,255) corresponde ao branco. Vermelho, verde e azul são as cores primárias, e as outras cores são as secundárias, como a cor ciano, magenta e amarelo (GONZALEZ, 2000). A Figura 3.4 representa o cubo com as representações das cores no espaço RGB.

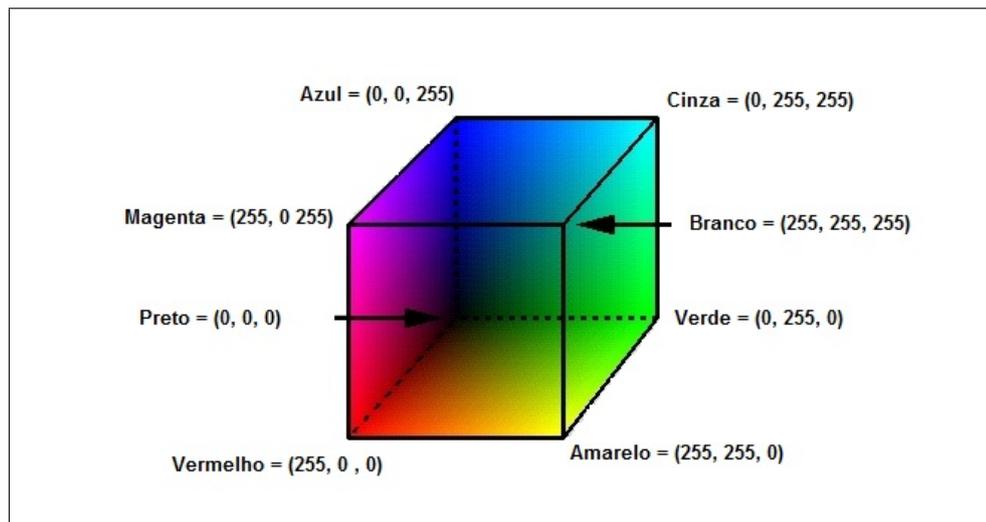


Figura 3.4: Exemplo do cubo RGB.(Fonte: adaptado de <http://www.cin.ufpe.br/marcelow.html>)

### 3.2.0.2 Modelos de cor HSI, HSL e HSV

Esses modelos usam eixos equivalentes em sua representação de cor. No modelo HSV, a cor é definida pelos componentes matiz, saturação e valor; no HSI, utilizam os componentes matiz, saturação e intensidade; no modelo HSL, utilizam matiz, saturação e luminância. O canal matiz varia de 0 a  $2\pi$ . A saturação indica a pureza da cor, variando de 0 a 1, em que a saturação igual a 0 indica que a cor não está saturada, e com valor igual a 1 indica que a cor está completamente saturada. O canal intensidade da cor de um modelo representa o brilho/claridade/luminosidade da cor, variando de 0 a 1, sendo a intensidade máxima igual a 1, indicando o branco puro e a intensidade mínima como preto puro. Os modelos HSV, HSL e HSI diferem nas fórmulas para os valores de intensidade. Por exemplo, cores totalmente saturadas com diferentes matizes têm o mesmo valor de  $V = 1$  em HSV e a mesma luminância  $L = 0,5$  em HSL e  $I=1$  para HSI. No entanto, isso nem sempre é verdade na percepção humana, por exemplo, amarelo completamente saturado é sempre mais leve que o azul totalmente saturado, ou seja, visualmente a cor amarela parece mais clara que o azul. No modelo de cor HSL encontramos mais brilho, uma cor de luminosidade mais intensa com valor  $L = 0,5$  (metade do máximo). Os modelos HSV, HSL e HSI são usados em aplicações de computação gráfica e processamento de imagens.

A seguir será descrito os modelos de cor HSI e HSL com as formulações de conversão do espaço de cor RGB para ambos os espaços.

### a) Espaço de cor HSI

O modelo HSI representa uma cor em torno dos valores de matiz ( $H$  - *hue*), saturação ( $S$  - *saturation*) e intensidade ( $I$  - *intensity*). O espaço de cor HSI é muito importante e atrativo para aplicações de processamento de imagens. No modelo RGB existe uma grande correlação entre seus canais, pois a alteração em apenas um canal reflete significativamente na imagem como um todo. O modelo HSI apresenta canais mais descorrelacionados, sendo mais apropriado para processar canais separadamente. Além disso, a cor é representada de uma forma que corresponde à forma da percepção pelo olho humano. Basicamente, o modelo HSI distribui as cores em dois cones conforme a Figura 3.5:

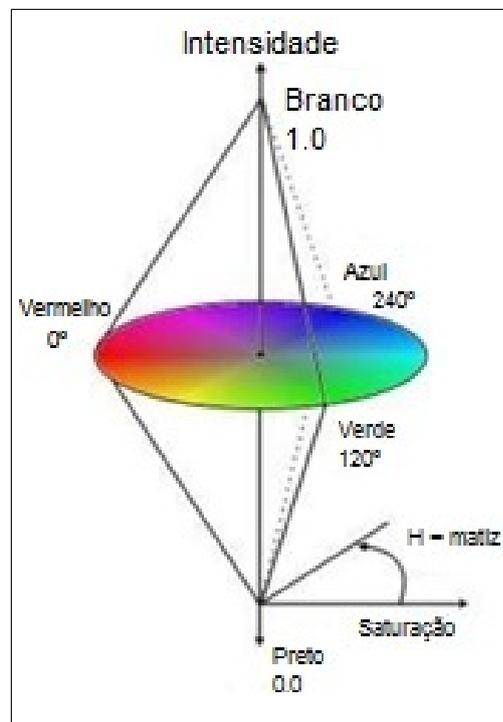


Figura 3.5: Exemplo do Cone HSI. (Fonte: (KUMAR, 2006))

#### Conversão RGB para HSI

Pode-se usar a formulação matemática seguinte para converter a imagem do espaço de cor RGB para HSI, (KUMAR, 2006).

1. Normalizar cada valor RGB no intervalo [0,1]:

$$r = \frac{R}{R + G + B} \quad g = \frac{G}{R + G + B} \quad b = \frac{B}{R + G + B} \quad (3.1)$$

2. Cada componente HSI normalizado é obtido pelas seguintes regras:

- Se  $b \leq g$  (componente b normalizado menor ou igual ao valor g normalizado)

$$h = \cos^{-1} \left( \frac{0,5 * [(r - g) + (r - b)]}{[(r - g)^2 + (r - b) * (g - b)]^{\frac{1}{2}}} \right) \quad (3.2)$$

- Se  $b > g$  (componente b normalizado for maior que g normalizado, isto significa que é um ângulo maior que  $180^\circ$  e menor  $360^\circ$ ).

$$h = 2\pi - \cos^{-1} \left( \frac{0,5 * [(r - g) + (r - b)]}{[(r - g)^2 + (r - b) * (g - b)]^{\frac{1}{2}}} \right)$$

$$h \in [0, 2\pi]$$

$$s = 1 - 3 * \min(r, g, b) \quad s \in [0, 1]$$

$$i = (((R + G + B) \div 3) \div 255) \quad i \in [0, 1] \quad (3.3)$$

### Conversão de HSI para RGB

Para realizar a conversão de volta, de HSI para RGB, faz-se da seguinte forma (KUMAR, 2006):

$$x = i * (1 - s) \quad s \in [0, 1], \quad i \in [0, 1]$$

$$y = i * \left[ 1 + \frac{s \cdot \cos(h)}{\cos(\frac{\pi}{3} - h)} \right] \quad s \in [0, 1], \quad i \in [0, 1], \quad h \in [0, 2\pi]$$

$$z = 3.1 - (x + y) \quad i \in [0, 1] \quad (3.4)$$

- Se  $h < \frac{2\pi}{3}$  (h for menor que  $120^\circ$ )

$$r = y, \quad g = z, \quad b = x \quad (3.5)$$

- Se  $\frac{2\pi}{3} \leq h < \frac{4\pi}{3}$  (h for maior que  $120^\circ$  e menor que  $240^\circ$ )

$$h = h - \frac{2\pi}{3}; \quad r = x, \quad g = y, \quad b = z \quad (3.6)$$

- Se  $\frac{4\pi}{3} \leq h < 2\pi$  (h for maior que  $240^\circ$  e menor que  $360^\circ$ )

$$h = h - \frac{4\pi}{3}; \quad r = z, \quad g = x, \quad b = y \quad (3.7)$$

Os valores de r,g e b obtidos pertencem ao intervalo [0,1]. Para melhor visualização da imagem resultante, os valores deverão ser normalizados no intervalo [0,255], conforme a equação a seguir:

$$R = 255.r \quad G = 255.g \quad B = 255.b \quad (3.8)$$

A Figura 3.6 ilustra a conversão de uma imagem no espaço RGB para o espaço HSI, e os canais H,S e I separadamente, normalizados entre [0,255].

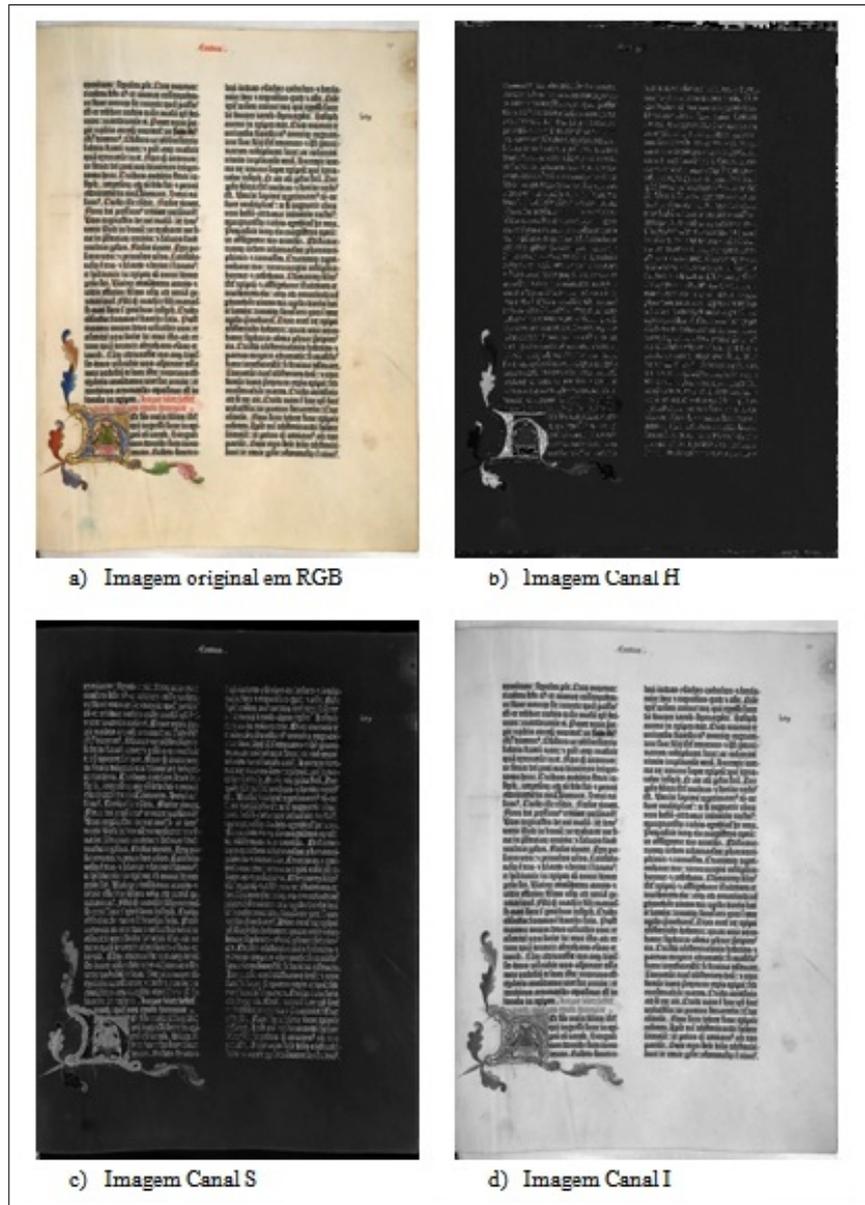


Figura 3.6: Conversão RGB para HSI. (Fonte: (WITTE, 2006))

## b) Espaço de cor HSL

HSL significa (Hue-Saturation-Luminance) em que H é o matiz, S a saturação e L a luminância. É o espaço de cor mais utilizado por fotógrafos com equipamento digital. A Figura 3.7 é uma representação do espaço de cor HSL. O canal H é um valor em ângulos variando de  $0^\circ$  a  $360^\circ$  e os canais S e L variam de 0 a 1.

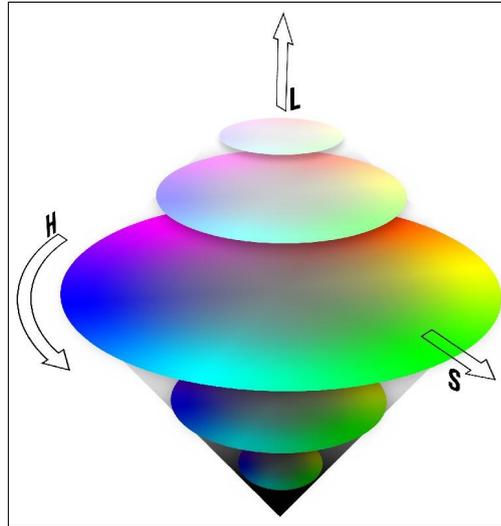


Figura 3.7: Exemplo do modelo HSL.

### Conversão RGB para HSL

Os valores de RGB são normalizados para o intervalo 0 e 1. A luminosidade  $L$  e a saturação  $S$  estão normalizados entre 0 e 1. O matiz  $H$  varia entre  $[0, 360]$ .

Para realizar a conversão do modelo RGB para o modelo HSL é preciso encontrar o máximo e o mínimo de R,G e B para cada pixel, ou seja, para cada pixel calcula-se:

$$Max = \text{maximo}(R, G, B) \text{ e } Min = \text{minimo}(R, G, B) \quad (3.9)$$

Com essas definições preliminares, as formulações a seguir demonstram como obter a conversão do espaço de cor RGB para HSL e de HSL para RGB.

$$H = \begin{cases} 60 * \frac{(G-B)}{(Max-Min)}, & \text{se } Max = R \\ 60 * \frac{(B-R)}{(Max-Min)} + 120, & \text{se } Max = G \\ 60 * \frac{(R-G)}{(Max-Min)} + 240, & \text{se } Max = B \end{cases} \quad (3.10)$$

$$S = \begin{cases} 0, & \text{se } Max = Min \\ \frac{(Max-Min)}{(Max+Min)}, & \text{se } 0 < L < 0.5 \\ \frac{(Max-Min)}{2-(Max-min)}, & \text{se } L > 0.5 \end{cases} \quad (3.11)$$

$$L = \left\{ \frac{(Max + Min)}{2} \right. \quad (3.12)$$

### Conversão HSL para RGB

Para realizar a conversão de HSL para RGB é preciso verificar se o canal de cor H está no intervalo  $0^\circ$  a  $360^\circ$ , incluindo  $0^\circ$ , mas não  $360^\circ$ . Os canais L e S devem estar no intervalo  $[0.0, 1.0]$ . Após a verificação procede-se da seguinte maneira;

- Se o valor do canal S for igual 0, os três canais (RGB) recebem o valor de L, sendo o canal H irrelevante nesta situação.
- Se o valor do canal S for diferente de 0, procede-se da seguinte forma:

$$Q = \begin{cases} L * (1.0 + S), & \text{se } L < 0.5 \\ L + S - (L * S), & \text{se } L \geq 0.5 \end{cases} \quad P = 2.0 * L - Q$$

$$H_k = \frac{H}{360} (\text{Normaliza } H \text{ entre } 0 \text{ e } 1)$$

$$T_R = H_k + \frac{1}{3}$$

$$T_G = H_k$$

$$T_B = H_k - \frac{1}{3} \quad (3.13)$$

- Se  $T_c < 0$

$$T_c = T_c + 1.0, \quad \text{para cada } c = R, G, B \quad (3.14)$$

- Se  $T_c > 1$

$$T_c = T_c - 1.0, \quad \text{para cada } c = R, G, B \quad (3.15)$$

- Para cada cor  $c=R,G,B$  faça:

$$c = \begin{cases} P + ((Q - P) * 6.0 * T_c), & \text{se } T_c < \frac{1}{6} \\ Q, & \text{se } \frac{1}{6} \leq T_c < \frac{1}{2} \\ P + ((Q - P) * (\frac{2}{3} - T_c) * 6.0), & \text{se } \frac{1}{2} \leq T_c < \frac{2}{3} \end{cases} \quad (3.16)$$

### 3.2.0.3 Modelos de cor YUV, YIQ e YCrCb

Chamados de modelos de cor de transmissão de televisão, os modelos YUV e YIQ, são utilizados em televisão em cores, sendo uma recodificação do modelo RGB, e também para a codificação de imagens coloridas em televisão e em vídeo de cor, e muito útil para a transmissão de sinais de televisão para a TV preto e branco. O modelo de cores YCbCr é independente de sistemas de codificação para sinais de TV e é usado para a representação de imagens de TV em sistemas digitais.(WITTE, 2006). Estes modelos de cor estão baseadas no fato de que o olho humano é mais sensível para as mudanças no brilho do que para mudanças na tonalidade ou na saturação. O componente Y contém informação sobre o brilho da imagem, enquanto que os outros componentes contêm as informações sobre a cor(tonalidade e saturação). Por ser o olho humano mais sensível a mudanças no brilho, esses modelos utilizam um intervalo maior para o componente Y de que para os outros dois canais. O componente Y contém todas as informações necessárias para a televisão preto e branco, isto é, o componente Y dá a imagem em escala de cinza de uma imagem a cores. (WITTE, 2006).

A seguir serão descritas as formulações para conversão dos espaços de cor YCrCb e YIQ.

#### a) Espaço YCrCb

O Espaço de cor YCrCb é da família dos espaços utilizados em vídeos e fotografias digitais. O canal Y corresponde à luminância (informa os tons de cinza ou brilho da imagem), enquanto os canais Cr e Cb correspondem à crominância, sendo que Cr corresponde ao canal vermelho (*Chroma red*) e o Cb ao canal azul (*Chroma blue*). A luminância neste formato representa sua forma relativa, que demonstra a densidade de intensidade luminosa de uma fonte de luz. A crominância carrega a informação sobre a cor, separada do canal que representa a luminância. Esse espaço

de cor possui a vantagem de ser mais perceptível, devido ao fato de que, assim como nossa visão, ela trabalha separadamente o brilho e a cor. (SHAHBAHRAMI; JUURLINK; VASSILIADI, 2006) As formulações da conversão do espaço RGB para YCrCb e vice-versa, são apresentadas nas Equações 3.17 e 3.18 (SHAHBAHRAMI; JUURLINK; VASSILIADI, 2006).

### Conversão RGB para YCrCb

A formulação matemática para conversão do espaço RGB para YCrCb é descrita pela equação 3.17.

$$\begin{cases} Y = (0,257 * R) + (0,504 * G) + (0,098 * B) + 16 \\ Cr = (0,439 * R) - (0,368 * G) - (0,071 * B) + 128 \\ Cb = (0,148 * R) - (0,291 * G) + (0,439 * B) + 128 \end{cases} \quad (3.17)$$

### Conversão YCrCb para RGB

A formulação matemática de retorno do espaço YCrCb para RGB é descrita pela Equação 3.18.

$$\begin{cases} R = 1,164.(Y - 16) + 1,596 * (Cr - 128) \\ G = 1,164.(Y - 16) - 0,813.(Cr - 128) - 0,391.(Cb - 128) \\ B = 1,164.(Y - 16) + 2,018.(Cb - 128) \end{cases} \quad (3.18)$$

## b) Espaço YIQ

No espaço de cores YIQ, a cromaticidade da imagem é representada pelos componentes  $I$  e  $Q$ , enquanto o componente  $Y$  representa a luminosidade. É usado na transmissão comercial de televisão colorida. Trata-se basicamente de uma recodificação de RGB, para eficiência da transmissão e para manutenção da compatibilidade com os padrões monocromáticos de televisão. Isso é possível através do componente  $Y$ , o qual fornece todas as informações de vídeo necessárias para um aparelho de televisão monocromático (SILVA; COSTA, 2006). A equação 3.19 apresenta a formulação do espaço de cor RGB para YIQ.

### Conversão RGB para YIQ

$$\begin{cases} Y & 0.299 & 0.587 & 0.114 & R \\ I & = & 0.596 & -0.274 & -0.322 & G \\ Q & & 0.211 & -0.523 & 0.312 & B \end{cases} \quad (3.19)$$

### Conversão YIQ para RGB

A equação 3.20 apresenta a formulação do espaço de cor YIQ para RGB.

$$\begin{array}{l} R \\ G \\ B \end{array} = \begin{array}{llll} 1.000 & 0.956 & 0.621 & Y \\ 1.000 & -0.272 & -0.647 & I \\ 1.000 & -1.106 & 1.703 & Q \end{array} \quad (3.20)$$

## 3.3 Morfologia Matemática

A morfologia matemática explora características geométricas das imagens. É uma teoria usada para análise, filtragem e segmentação da imagem (GRANADO; PINA; MUGE, 2001). Foi criada em meados de 1960 por Georges Matheron e Jean Serra da *École des Mines de Paris*, com o objetivo de quantificar as estruturas de acordo com sua geometria. Sua evolução teórica tem gerado ferramentas poderosas não só para lidar com a geometria das estruturas, mas também com a análise e os procedimentos de modelagem e simulação.

O princípio básico da morfologia matemática, consiste em extrair informações relativas à geometria e à topologia do conteúdo de uma imagem. O objetivo de estudo do pesquisador é quantificar as estruturas geométricas da imagem a partir de um conjunto perfeitamente definido e conhecido (forma e tamanho) chamado de: *Elemento Estruturante*. O resultado dessa transformação permite gerar novas imagens e avaliar o conteúdo das mesmas através dos operadores de filtragem, segmentação, detecção de bordas, esqueletização, afinamento, análise de formas, compressão e *etc.* (FACON, 1996).

A morfologia matemática, originalmente desenvolvida para imagens binárias, foi estendida para imagens em níveis de cinza e as pesquisas atuais se concentram em sua extensão para imagens coloridas. Os pilares desta teoria são as duas operações básicas, de erosão e de dilatação, a partir das quais, por composição, é possível realizar muitos outros operadores poderosos. A morfologia matemática se apoia no princípio da análise da estrutura geométrica das imagens (FACON, 1996). A Figura 3.8 apresenta dois exemplos de elementos estruturantes, sendo a imagem (a) em forma de cruz com conectividade 4, e em (b) em forma de quadrado com conectividade 8.

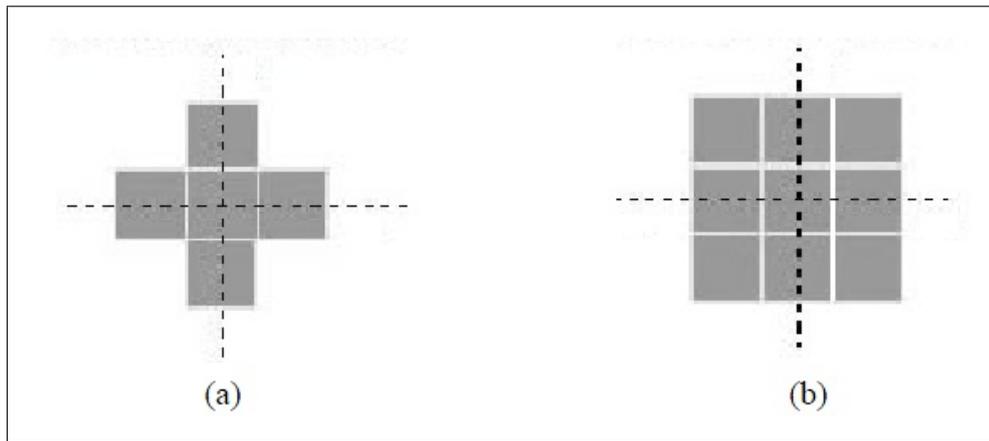


Figura 3.8: Exemplo de dois elementos estruturantes.

As vantagens da morfologia matemática estão aqui relacionadas.

- Simplicidade de implementação e aproveitamento das implementações existentes envolvendo morfologia binária e cinza, destacando-a das outras técnicas de processamento de imagens (FACON, 1996).
- Capacidade de preservar o formato dos objetos da imagem, quaisquer que sejam seus tamanhos, através dos filtros morfológicos, enquanto que filtros lineares têm como desvantagens deteriorar a nitidez dos objetos (LEDDA; PHILIPS, 2002).
- Aspecto intuitivo aplicado em processamento de imagens, uma vez que trabalha diretamente sobre o domínio espacial (OROZ, 2005). Os elementos estruturantes são considerados como 'tijolos fundamentais' para desempenhar o mesmo papel das frequências no caso de processamento de sinais.

### 3.3.1 Morfologia Matemática Binária

Os operadores morfológicos binários são definidos em imagens com pixels pretos e brancos, somente. Um objeto é considerado ser um conjunto matemático de pixels pretos, sendo cada pixel identificado pelos seus índices de linha e coluna, chamado de um ponto no espaço bidimensional. A seguir serão descritos os operadores morfológicos binários.

### 3.3.1.1 Operadores Morfológicos Binários

- **Erosão:** Serra (1982) define a operação de erosão binária, utilizando uma imagem  $f$  pelo elemento estruturante  $b$ , da seguinte maneira:

$$\varepsilon^B(f) = \{x \in \xi : B_x \subset f\} \quad (3.21)$$

em que:

$B_x$  representa o elemento estruturante  $B$  transladado na posição  $x$ .

$\xi$  representa o conjunto não vazio.

Por erosão binária, os efeitos obtidos são: a) Diminuir conjuntos, desconectá-los e eventualmente eliminá-los; b) Aumentar e abrir cavidades.

- **Dilatação:** a operação de dilatação binária é definida por Serra (1982), utilizando uma imagem  $f$  pelo elemento estruturante  $B$  da seguinte maneira:

$$\delta^B(f) = \{x \in f : B_x \cap f \neq 0\} \quad (3.22)$$

em que:

$B_x$  representa o elemento estruturante  $B$  transladado na posição  $x$ .

Por dilatação binária, os efeitos obtidos são: a) Aumentar os conjuntos e eventualmente conectá-los, caso o tamanho do elemento estruturante for maior que o espaço entre eles; b) Diminuir e preencher cavidades.

- **Abertura:** a operação abertura binária consiste em erodir e depois dilatar o resultado da erosão, servindo para eliminar pequenos ruídos, sem modificar o tamanho dos outros conjuntos. Serra (1982) a define:

$$\gamma^B(f) = \delta^B(\varepsilon^{\tilde{B}}(f)) \quad (3.23)$$

em que:

$\tilde{B}$  representa o transposto de  $B$  obtido por simetria central pela origem  $o$  do sistema de referência.

- **Fechamento:** a operação de fechamento binário consiste em dilatar e depois erodir o resultado da dilatação, preenchendo cavidades sem modificar o tamanho dos outros conjuntos. Serra (1982) apresenta sua equação assim:

$$\phi^B(f) = \varepsilon^B(\delta^{\tilde{B}}(f)) \quad (3.24)$$

em que:

$\tilde{B}$  representa o transposto de B obtido por simetria central pela origem  $o$  do sistema de referência.

### 3.3.1.2 Morfologia Matemática Soft

A morfologia soft é uma abordagem alternativa mais tolerante a ruídos, possui muitas características desejáveis presentes nos operadores morfológicos tradicionais (PRESTON, 1983), (WILSON, 1989) e (KOSKINEN; ASTOLA; NEUVO, 1991). Zmuda e Tamburino (1996) apresenta dois novos algoritmos que processam eficientemente os operadores *soft* (erosão e dilatação), são definidos os domínios no qual cada algoritmo domina em termos de eficiência. Os operadores *soft* podem ser configurados através de um parâmetro  $m$  da seguinte forma:

- Na erosão, um *pixel*  $(x,y)$  preto continuará sendo preto na imagem resultado sempre que o número de pixels pretos vizinhos conforme o formato do elemento estruturante em  $(x,y)$  na imagem origem for maior ou igual a  $m$ .
- Na dilatação, um *pixel*  $(x,y)$  branco na imagem origem será pintado de preto na imagem resultado apenas quando o número de pixels pretos vizinhos conforme o formato do elemento estruturante aplicado sobre o ponto  $(x,y)$  na imagem origem for maior ou igual que  $m$ .

Os operadores morfológicos *soft* servem para eliminar ou engordar objetos na imagem.

A Figura 3.9 apresenta os resultados à medida que varia o argumento  $m$  em uma aplicação de um operador básico *soft* utilizando o elemento estruturante Cruz.

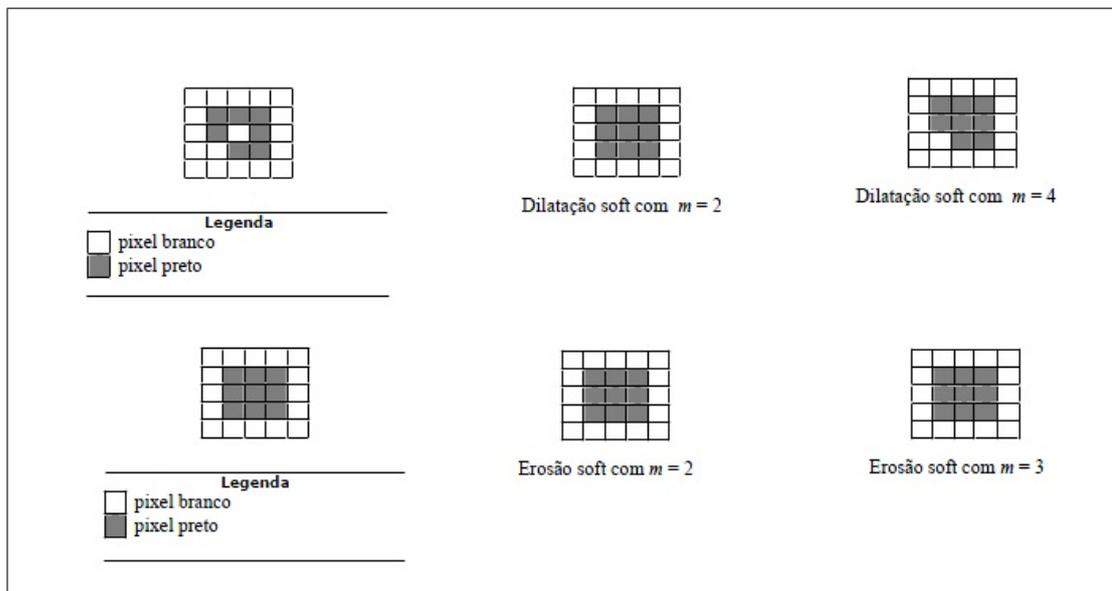


Figura 3.9: Exemplo dos operadores soft erosão e dilatação.

### 3.3.2 Morfologia Matemática em Níveis de Cinza

Nesta etapa serão descritos os operadores morfológicos em níveis de cinza.

#### 3.3.2.1 Operadores Morfológicos em Níveis de Cinza

- **Erosão em níveis de cinza:** é definida por uma imagem  $f$  e por um elemento estruturante  $B$  e pode ser descrita pela equação 3.25 (FACON, 1996).

$$\varepsilon^B(f(x)) = \wedge \{f(x) - B(x - y) : y \in E\} \quad (3.25)$$

em que:

- $\wedge$  representa o mínimo entre vários valores.
- $E$  representa o domínio do elemento estruturante.

A erosão é uma transformação não comutativa que apresenta várias propriedades interessantes como:

- Diminuir as partículas;
- Eliminar conjuntos de tonalidade clara, de tamanho inferior ao tamanho do elemento estruturante;

- Aumentar os buracos;
  - Separar conjuntos claros próximos;
  - Ressaltar as partes escuras da imagem que ficam mais '*gordas*'.
- **Dilatação em níveis de cinza:** é definida por meio de um imagem  $f$  por um elemento estruturante  $B$  e pode ser descrita pela equação 3.26 (FACON, 1996).

$$\delta^B(f(x)) = \vee \{f(y) + B(x - y) : y \in E\} \quad (3.26)$$

sendo que:

- $\vee$  representa o máximo entre vários valores.
- $E$  representa o domínio do elemento estruturante.

A dilatação é uma transformação comutativa, cujos efeitos interessantes são:

- Engordar partículas;
  - Eliminar conjuntos de tonalidade escura, de tamanho inferior ao tamanho do elemento estruturante;
  - Preencher pequenos buracos;
  - Conectar conjuntos claros próximos.
- **Abertura em níveis de cinza:** Consiste em erodir um conjunto  $f$  por  $B$  e depois dilatar esse conjunto erodido pelo mesmo elemento estruturante. A equação 3.27 descreve a formulação de acordo com Facon (1996).

$$\gamma^B(f) = \delta^B(\varepsilon^B(f)) \quad (3.27)$$

em que

- $\delta$  é a dilatação em níveis de cinza;
- $\varepsilon$  é a erosão em níveis de cinza.

O comportamento da abertura resulta em:

- Separar picos próximos (pico é considerado o ponto máximo de ocorrência de pixels no histograma);
- Eliminar os picos inferiores, em tamanho, ao elemento estruturante;

- Conservar vales afastados;
  - Emendar vales próximos;
  - As entidades restantes (são os pixels que não estão nos vales e nos picos), após abertura, ficam quase idênticas;
  - A imagem aberta é mais regular que a imagem original;
  - A imagem aberta é menos rica em detalhes que a imagem original.
- **Fechamento em níveis de cinza:** consiste na operação dual da operação de abertura em níveis de cinza, podendo ser definida como o fechamento de um sinal  $f$  por um elemento estruturante  $B$ , conforme a equação 3.28 (FACON, 1996).

$$\phi^B(f) = \varepsilon^B(\delta^B(f)) \quad (3.28)$$

O comportamento do fechamento em níveis de cinza consiste em:

- Separar vales próximos;
  - Eliminar vales inferiores;
  - Conservar picos afastados;
  - Emendar picos próximos;
  - As entidades restantes, após fechamento, ficam quase idênticas;
  - A imagem fechada é mais regular que a imagem original;
  - A imagem fechada é menos rica em detalhes que a imagem original.
- **Transformação TopHat em níveis de cinza:** Esta transformação detecta os picos e/ou vales e pode ser definida por Abertura ou Fechamento e estão descritas a seguir.
    - **TopHat por Abertura:** utilizado para detecção de picos. Consiste em se usar a combinação entre uma imagem original e a imagem correspondente aberta, podendo ser descrita conforme a equação 3.29.

$$tophat_{\gamma}^B(f) = f - \gamma^B(f) \quad (3.29)$$

O resultado da operação abertura fica abaixo do sinal original, portanto a transformação é sempre positiva. Com o uso de um elemento estruturante

adequado, o processo de abertura permite a eliminação dos picos. Fazer a diferença, entre os sinais da imagem original e da imagem com o resultado da abertura, permite tirar o ruído, ressaltando as informações dos picos.

- **Tophat por fechamento:** utilizado para detecção de vales. Consiste em usar uma combinação entre uma imagem original e a imagem correspondente fechada. A equação 3.30 descreve a técnica.

$$\text{tophat}_{\phi}^B(f) = \phi^B(f) - f \quad (3.30)$$

Durante o fechamento, o vale a ser detectado desaparece. Dessa forma, durante a subtração, esse vale reaparece.

### 3.3.3 Morfologia Matemática em cor

Na morfologia em níveis de cinza anteriormente apresentada, os operadores baseiam-se nos conceitos de mínimo e máximo de níveis de cinza. Conceitos que não apresentam nenhuma dificuldade teórica e/ou matemática. Por exemplo, procurar o mínimo ou o máximo entre dois níveis de cinza, 100 e 110 não apresenta nenhuma dificuldade computacional. Mas como definir o mínimo ou o máximo entre, por exemplo, a cor verde e a cor azul? Essa pergunta representa o desafio matemático e teórico que os pesquisadores tentam solucionar na morfologia em cor.

A dificuldade é definir uma ordem para as imagens coloridas, ou seja, um caminho único para ordenar as cores. A morfologia matemática em cor representa atualmente uma nova linha de pesquisa desafiadora para a área. Os desafios matemáticos e teóricos residem na escolha apropriada do espaço de cor e na definição da ordenação das cores.

Os dados multivariáveis podem ser ordenados por várias técnicas (TRAHANIAS; VENETSANOPOULOS, 1992), entre as quais a marginal e a vetorial. O modelo de cor adotado para o processamento é importante, uma vez que imagens coloridas podem ser representadas por vários modelos de cores (SANGWINE; HORNE, 1998). A seguir serão descritas as ordenações.

#### 3.3.3.1 Ordenações

Os operadores morfológicos básicos de erosão e dilatação para a morfologia em cor baseiam-se nos conceitos de mínimo e máximo. Assim, é importante definir uma ordenação apropriada dos dados (cores), no espaço de cor selecionado. Os dados dos

canais são multivariados e existe uma série de maneiras diferentes para ordená-los. A seguir serão apresentadas as ordenações descritas por Aptoula e Lefèvre (2007).

### 1. Ordenação Marginal

Esta abordagem consiste em avaliar cada um dos componentes de cor separadamente, realizando a ordenação de cada canal e posterior combinação dos canais para a geração do resultado. A Figura 3.10 apresenta a estratégia marginal por meio dos canais RGB (LAMBERT; CHANUSSOT, 2000), (APTOULA; LEFÈVRE, 2007) e (CALIXTO, 2005).

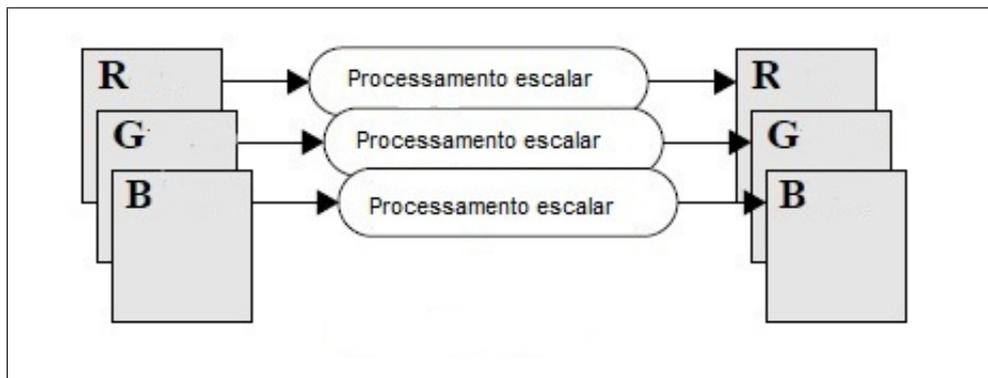


Figura 3.10: Estratégia marginal.

Esta abordagem não é totalmente satisfatória por diferentes razões:

- (a) Não leva em conta a correlação inter-componente (entre os canais);
- (b) Se a imagem tem  $N$  componentes diferentes,  $N$  processamentos são necessários, levando a consumir tempo de operação;
- (c) O processamento pode produzir cores falsas, alterando o equilíbrio de cor da imagem, que é uma desvantagem para várias aplicações, como filtros morfológicos.

O problema desta abordagem é a possibilidade de gerar uma cor que não faz parte da imagem original, após a aplicação de um operador morfológico. Calixto (2005) apresenta um exemplo da utilização da morfologia matemática em cor com imagens no espaço de cor RGB. Ele ilustra um exemplo da aplicação da dilatação utilizando o conceito da ordenação marginal no espaço de cor R,G,B. Neste exemplo, aplicam-se as 03 dilatações, em cada um dos componentes da imagem, que são recombinados

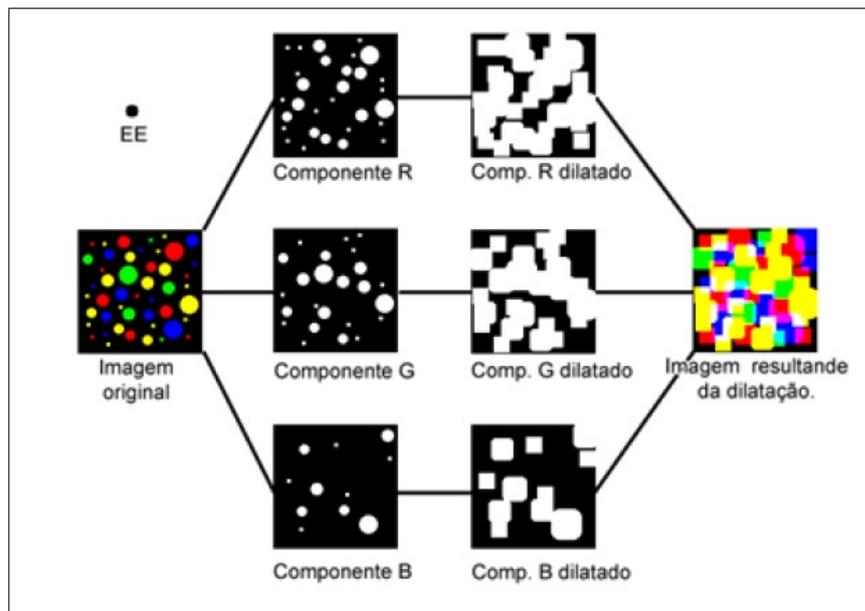


Figura 3.11: Exemplo da operação de dilatação de uma imagem colorida com a ordenação marginal. (Fonte: (CALIXTO, 2005))

para gerar a imagem dilatada final. A Figura 3.11 ilustra as distorções geradas pela aplicação direta da dilatação nos canais R,G,B com ordenação marginal.

A Figura 3.12 apresenta os histogramas da imagem original apresentada pela Figura 3.11 e do resultado da dilatação da imagem original utilizando a ordenação marginal.

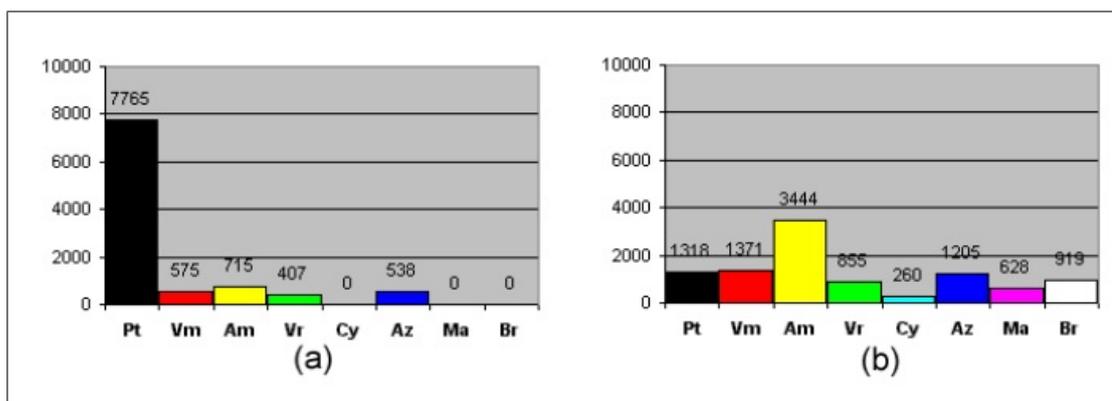


Figura 3.12: Histograma das imagens: (a) representando a imagem original e (b) a imagem reconstruída após a dilatação. (Fonte: (CALIXTO, 2005))

Nas Figuras 3.11 e 3.12, é possível perceber que aplicar diretamente a dilatação nos canais R, G e B, com ordenação marginal, foram geradas cores novas ("cores

falsas”) que não existiam na imagem original, como a cor branca. A geração de cores indesejáveis é um dos principais problemas do uso da ordenação marginal.

## 2. Ordenação Vetorial

Definido como uma relação de ordem dos canais de um espaço de cor, consiste na utilização de uma abordagem puramente vetorial, conforme a Figura 3.13, sendo realizado o processamento vetorial para todos os canais do espaço de cor. Esta estratégia requer a definição da ordenação do vetor para encontrar o máximo e o mínimo para qualquer imagem com  $N$  dimensões. É uma forma de ordenar o espaço, sendo um ponto de diferenciação na construção de uma morfologia matemática colorida, garantindo a não geração de cores falsas. A figura 3.13 mostra a estratégia vetorial em ordenar os canais de cor RGB.

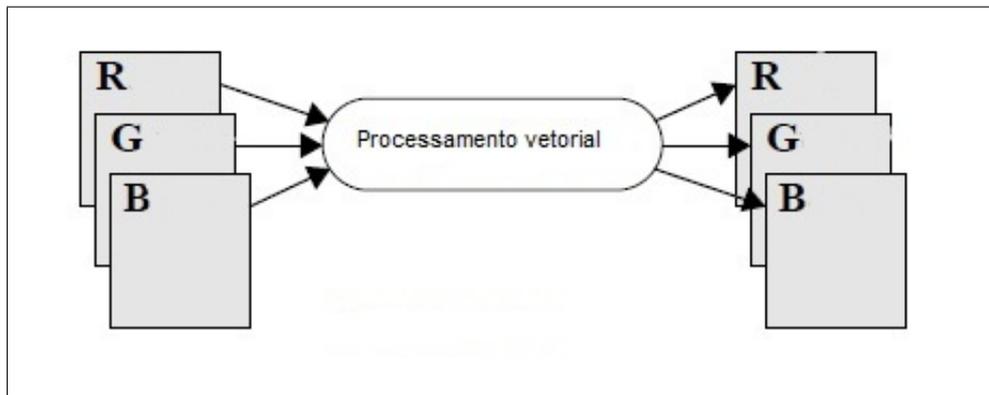


Figura 3.13: Estratégia vetorial.

A ordenação vetorial pode ser classificada em mais de um grupo conforme Aptoula e Lefèvre (2007) e Barnet (1976).

- **Ordenação Condicional:** os vetores são ordenados por meio de seus componentes, selecionados sequencialmente de acordo com diferentes condições. A ordenação lexicográfica é um exemplo muito conhecido da ordenação condicional empregando potencialmente todos os componentes disponíveis dos vetores de dados, como na equação 3.31.

$$\forall v, v' \in \mathbb{R}^n, v \leq v' \forall i \in 1, \dots, n, v_i \leq v'_i \quad (3.31)$$

Também pode-se utilizar para comparação apenas um subconjunto, ou seja, um dos três canais na comparação (HARDIE; ARCE, 1991). A ordenação con-

dicional é indicada para os casos em que se pode estabelecer uma prioridade, ou seja, uma importância entre os canais da imagem. A ordenação condicional subdivide-se em: pré-ordenação condicional total e ordenação condicional total.

**Pré-ordenação total condicional:** consiste em restringir o processo de ordenação a apenas um ou mais canais dos vetores de dados. Este tipo é adequado a situações em que determinados canais são mais privilegiados que outros. Por exemplo, no caso de empregar somente o primeiro canal do espaço de cor. É apresentado pela equação 3.32:

$$v \leq v' \Leftrightarrow v_1 \leq v' \quad (3.32)$$

O principal problema é a escolha adequada do componente a ser privilegiado, pois a ordenação do vetor, apenas utilizando um canal (componente) de cor é justificável apenas se o dado for suficiente para a aplicação. De forma geral, esta abordagem atribui importância demais para o canal selecionado, desconsiderando os outros canais. O problema está em casos em que todos os canais são importantes, ficando claro que a utilização desta abordagem não se torna eficaz.

**Ordenação Total condicional** Do ponto de vista teórico, há várias vantagens, dentre as quais a utilização. Em primeiro lugar, utiliza-se a totalidade dos canais, que são considerados vetores de preservação. É por isso que a maioria das tentativas de estender os operadores morfológicos baseiam-se em ordenações totais. Em particular, a ordenação lexicográfica (ordenação condicional), juntamente com suas variantes, está entre as opções mais implementadas. A priorização de um canal implica o uso de espaços de cor decorrelacionados, ou seja, é preciso converter a imagem do espaço de cor RGB para HSI ou YCrCr ou YIQ, por exemplo, pois, não há como definir prioridades para os canais RGB. Para os espaços decorrelacionados pode-se definir prioridades para determinados canais. A ordenação lexicográfica é considerada uma ordenação total condicional e será descrita com mais detalhes a seguir.

**Ordenação Lexicográfica** A ordem lexicográfica é uma forma de organizar as palavras através de uma técnica de ordenação vetorial, que se refere à ordem na qual são organizadas as palavras nos dicionários. Aptoula e Lefèvre (2007) relatam que esta técnica foi introduzida por Angulo (2005) e é considerada a ordenação total mais utilizada pela morfologia matemática em cor. A

ordem condicional, conhecida por utilizar potencialmente todos os canais disponíveis nos vetores de dados, é mais adequada para situações em que existe a ordem de importância entre os canais disponíveis: a ordem é decidida por um canal (letra), seguida de um segundo canal e depois de um terceiro canal. Se a primeira comparação dos componentes satisfazer-se não se precisa realizar a segunda comparação, e assim por diante.

O algoritmo de ordenação lexicográfica requer uma ordenação interna em cada componente e outra ordenação entre os componentes (PITAS; VENETSANOPOULOS, 1992) e (ORTIZ, 2000). Os três canais nunca serão recombina-dos/separados, ou seja, é realizado a ordenação da cor com seus três canais juntos, isto exclui a o aparecimento de cores falsas. A equação 3.33 ilustra a ordenação lexicográfica genérica, para ordenar os vetores com três canais de cor.

$$(a_1, b_1, c_1) < (a_2, b_2, c_2) \text{ se } \left\{ \begin{array}{l} a_1 < a_2 \\ \text{ou} \\ a_1 = a_2 \text{ e } b_1 < b_2 \\ \text{ou} \\ a_1 = a_1 \text{ e } b_1 = b_2 \text{ e } c_1 < c_2 \end{array} \right. \quad (3.33)$$

Um exemplo do funcionamento do algoritmo da ordenação lexicográfica é dado a seguir.

- (a) Escolhe-se o espaço de cor (descorrelacionado), obtém-se os três canais.
- (b) Escolhe-se a ordem em que os canais serão processados.
- (c) Primeiramente ordenam-se os pixels utilizando a janela do elemento estruturante, baseado no primeiro canal escolhido na ordenação.
- (d) A ordenação baseada no segundo canal será processada caso, na procura do mínimo/máximo, o valor do pixel central do primeiro canal for igual ao mínimo/máximo do vetor ordenado. No segundo canal, essa ordenação será somente realizada para os pixels que possuírem o mesmo valor do mínimo/máximo do primeiro canal.
- (e) A ordenação baseada no terceiro canal será processada da mesma forma que para o segundo canal, caso permaneça o empate.

Aptoula e Lefèvre (2007) fornecem a descrição matemática pela equação 3.34.

$$\forall v, v' \in \mathbb{R}^n, v \leq v' \forall i \in 1, \dots, n, v_i \leq v'_i \quad (3.34)$$

- Ordenação Parcial: é a ordenação dos dados em grupos de equivalência com relação à classificação e ordem. Os dados parciais são geralmente geométricos na natureza e possuem inter-relação entre os componentes (canais). Esse tipo de ordenação também pode conduzir à desvantagem de múltiplos máximos. A formulação é descrita pela equação 3.35.

$$v \leq v' \Leftrightarrow v_1 \leq v'_1 \quad (3.35)$$

- Ordenação Reduzida: a ordenação dos vetores é reduzida a valores escalares, que são classificados de acordo com sua ordem escalar natural. Esta ordenação consiste em classificar através da distância e da projeção. Por exemplo: Uma ordenação reduzida em  $R_n$  pode consistir na definição da transformação  $h : R_n \rightarrow R$  e, em seguida na ordenação do vetor de  $R_n$  com a respectiva ordem escalar da sua projeção  $R$  por  $h$  definido pela equação 3.36.

$$\forall v, v' \in \mathbb{R}^n, v \leq v' \Leftrightarrow h(v) \leq h(v') \quad (3.36)$$

Uma vantagem adicional da ordenação reduzida, descrita por Aptoula e Lefèvre (2007), reside no fato de que a escolha adequada da transformação  $h$ , pode-se atribuir prioridade igual a todos os componentes de cor, ao contrário da ordenação condicional. A desvantagem é que este procedimento pode conduzir à existência de mais de um mínimo/máximo e, assim, introduzir ambiguidade ao resultado dos dados.

### 3.3.3.2 Ordem dos canais

Após a escolha de um espaço de cor e do tipo de ordenação a ser utilizado, é preciso escolher o ordem dos canais do espaço de cor a ser processado pelos operadores morfológicos. Por exemplo: a ordenação lexicográfica há muito utilizada e possui várias formas de se ordenar os canais de um espaço de cor. Para o espaço de cor HSI, têm-se as seguintes ordenações lexicográficas possíveis: HSI, ISH, IHS, SHI, SIH e HSI.

### 3.3.3.3 Operadores Morfológicos em Cor

Nesta seção são descritos os operadores morfológicos em cor: erosão, dilatação, abertura, fechamento, tophat abertura e tophat fechamento.

- **Erosão em cor:** A operação de erosão colorida é definida como a seleção do valor da cor da imagem com menor distância absoluta (diferença de cor), ou seja, a erosão de uma imagem  $f$  por um elemento estruturante  $B$  é definida conforme a equação 3.37.

$$\varepsilon^B(f(x)) = \wedge\{f(y) - B(x - y) : y \in E\} \quad (3.37)$$

- **Dilatação em cor:** A operação de dilatação colorida é a seleção do valor da cor da imagem com maior distância absoluta, ou seja, a dilatação de uma imagem  $f$  por um elemento estruturante  $B$  é definida pela equação 3.38.

$$\delta^B(f(x)) = \vee\{f(y) + B(x - y) : y \in E\} \quad (3.38)$$

- **Abertura em cor:** A abertura de uma imagem  $f$  por um elemento estruturante  $B$  consiste em erodir o conjunto  $f$  para depois dilatar esse novo conjunto, sendo descrita pela equação 3.39.

$$\gamma_B(f)(x, y) = \delta_B(\varepsilon_B(f)(x, y)) \quad (3.39)$$

- **Fechamento em cor:** O fechamento de uma imagem  $f$  por um elemento estruturante  $B$  consiste em dilatar o conjunto  $f$  para depois erodir esse novo conjunto, sendo descrita pela equação 3.40.

$$\phi_B(f)(x, y) = \varepsilon_B(\delta_B(f)(x, y)) \quad (3.40)$$

- **Top Hat Abertura em cor:** O tophat por abertura é definido como a subtração da operação de abertura na imagem  $f$  da imagem original  $f$  e pode ser descrita pela equação 3.41.

$$TopHat_{\gamma}^B(f)(x, y) = f - \gamma_B(f)(x, y) \quad (3.41)$$

- **Top Hat Fechamento em cor:** O tophat por fechamento é definido como a subtração da imagem original  $f$  da operação de fechamento na imagem  $f$  e pode ser

descrita pela Equação 3.42.

$$TopHat_{\phi}^B(f)(x, y) = \phi_B(f)(x, y) - f \quad (3.42)$$

### 3.3.4 Reconstrução

A reconstrução é uma técnica que permite recuperar informações importantes a partir de pequenas dados de um conjunto. Facon (1996), a reconstrução é descrita da seguinte forma: dado um subconjunto  $S$  da imagem  $f$ , imaginemos ter um "pedacinho"  $Z$  de  $S$ , onde  $S$  é marcado por  $Z$ , sendo, portanto,  $Z$  o marcador de  $S$ , e  $S$  é a máscara, assim recupera-se o conjunto  $S$  a partir desse marcador  $Z$ . A reconstrução trabalha com duas imagens: marcadora e máscara (modelo). Após uma dilatação geodésica na imagem marcadora e realiza-se uma comparação dos pixels dilatados com os pixels da imagem máscara, conforme a equação 3.43, baseada na função da reconstrução em cor  $\rho$  da dilatação geodésica  $\delta_{...}$ , i.e.,

$$\rho(f) = \lim_{n \rightarrow +\infty} \underbrace{\delta_g(\delta_g(\dots \delta_g(f)))}_n \text{ com } \delta_g(f) = \delta(f) \wedge g \quad (3.43)$$

A dilatação geodésica em cor  $\delta_g(f)$  é baseada na restrição da dilatação em cor da imagem marcadora  $f$  pela imagem máscara  $g$ . Sendo:

- $f$ , a imagem marcadora (imagem que será reconstruída);
- $g$ , a imagem máscara (imagem modelo).

Para realizar o processo da reconstrução é preciso seguir as três etapas abaixo:

- Primeira: precisa-se da imagem marcadora e da imagem máscara (modelo).
- Segunda: aplica-se a dilatação geodésica na imagem marcadora.
- Terceira: realiza-se a comparação dos pixels correspondentes entre a imagem marcadora e a imagem máscara, buscando o ínfimo.

A reconstrução em cor obedece aos mesmos princípios da reconstrução em níveis de cinza.

A Figura 3.14 apresenta um exemplo da reconstrução geodésica, sendo a imagem (a) a máscara, a (b) a marcadora e a imagem(c) o resultado da reconstrução. A imagem (b) apresenta o resultado da erosão com 20 iterações, da imagem (a). A imagem (b) possui sementes (marcadores), as bolas coloridas remanescentes. A imagem (c) representa a reconstrução dessas bolas coloridas.

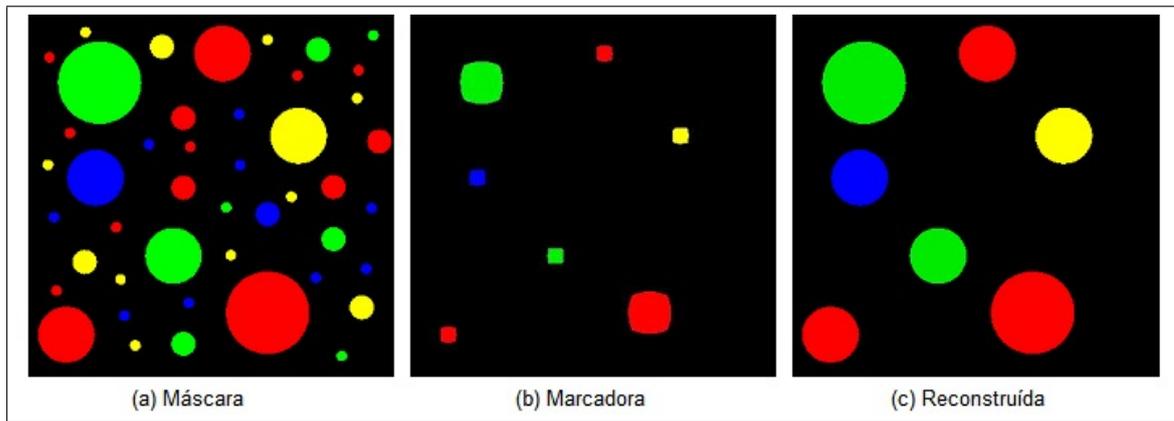


Figura 3.14: Exemplo da imagem reconstruída em cor.

## 3.4 Binarização

A limiarização é também conhecida como binarização, que consiste em segmentar as informações do fundo da imagem daquelas do primeiro plano. O objetivo é rotular o fundo com um valor, e as informações do primeiro plano com outro valor, normalmente, branco e preto, gerando uma imagem binária. Existem várias abordagens para aplicação da limiarização, de acordo com Facon (2005). Este trabalho descreve algumas técnicas de limiarização globais e locais.

### 3.4.1 Limiarização Global

A limiarização global consiste em segmentar o conteúdo de uma imagem através de um limiar único para toda a imagem. A maneira mais direta de encontrá-lo, de forma automática, é através da análise do histograma de níveis de cinza da imagem. No caso das informações do primeiro plano serem bem distintas do fundo, o histograma fornece claramente dois vales, um representando o fundo e o outro com as informações do primeiro plano. A limiarização global fundamenta-se em escolher no histograma um valor  $T$ , tal como para cada pixel  $P(x,y)$  (FACON, 2005):

$$P(x,y) = \begin{cases} \in \text{primeiro plano} & \text{se o valor de } P(x,y) > T \\ \in \text{fundo} & \text{caso contrario} \end{cases} \quad (3.44)$$

O limiar  $T$  fica em algum lugar do vale, sendo selecionado o valor do limiar para binarização da imagem, classificando-o em preto e branco. A limiarização global possui a

vantagem de ser um processo simples e precisa de pouca memória para processamento. As desvantagens são que: 1- Quando as imagens não possuem um fundo e primeiro plano bem distintos, um valor de limiar único não permite uma boa limiarização; 2- Imagens com partes gráficas possuem mais de um vale, ou seja, pequenos vales, tornando-se difícil encontrar o vale entre estes picos.

### 3.4.1.1 Método de Limiarização por Entropia Johannsen

O Método de Johannsen fundamenta-se na Entropia, que é uma medida de conteúdo de informação. A função entropia é dada por:  $E(x) = -x \log(x)$ , sendo a imagem formada por níveis de cinza (JOHANNSEN; BILLE, 1982). O processo de entropia, para uma imagem com 256 níveis de cinza, com pixels pretos, é limiarizada utilizando-se o valor de corte  $t$ , dado por  $S_b(t)$ :

$$S_b(t) = \log\left(\sum_{i=0}^t p_i\right) + \frac{1}{\sum_{i=0}^t p_i} \left[ E(p_t) + E\left(\sum_{i=0}^{t-1} p_i\right) \right] \quad (3.45)$$

De forma análoga, a entropia de pixels brancos é dada por  $S_w(t)$ :

$$S_w(t) = \log\left(\sum_{i=t}^{255} p_i\right) + \frac{1}{\sum_{i=t}^{255} p_i} \left[ E(p_t) + E\left(\sum_{i=t+1}^{255} p_i\right) \right] \quad (3.46)$$

No algoritmo de Johannsen (JOHANNSEN; BILLE, 1982), o valor ótimo do limiar é dado por  $t$  que minimiza a soma  $S_b(t) + S_w(t)$ .

### 3.4.1.2 Limiarização por Entropia de Li-Lee (Mínima Cruzada)

O método de entropia, chamado de mínima cruzada (LI; LEE, 1993), calcula a distância teórica da informação entre duas distribuições,  $P = p_1, p_2, \dots, p_n$  e  $Q = q_1, q_2, \dots, q_n$  dado por:

$$D(Q, P) = \sum_{k=1}^N q_k \log_2 \frac{q_k}{p_k} \quad (3.47)$$

A seleção do limiar é calculada através da minimização da entropia cruzada entre a imagem original e a limiarizada. O método pode ser visto como uma extensão do método da entropia máxima, calculado pela atribuição de estimativas iniciais para todas as probabilidades, quando nenhuma informação anterior está disponível (FACON, 1996). Para determinar a imagem binarizada é preciso determinar:

$$\begin{aligned} \eta[g(x, y)] &\equiv \eta(t, \mu_1, \mu_2) = \\ \eta(t) &= \sum_{g=1}^{t-1} gh(g) \log_2\left(\frac{g}{\mu_1(t)}\right) + \sum_{g=t}^{255} gh(g) \log_2\left(\frac{g}{\mu_2(t)}\right) \end{aligned} \quad (3.48)$$

com  $\mu_1$  e  $\mu_2$  definidos da seguinte forma:

$$\mu_1(t) = \frac{\sum_{g=0}^{t-1} gh(g)}{\sum_{g=0}^{t-1} h(g)} \quad e \quad \mu_2(t) = \frac{\sum_{g=t}^{255} gh(g)}{\sum_{g=t}^{255} h(g)} \quad (3.49)$$

com  $g$  sendo 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  $\eta(t)$ .

### 3.4.1.3 Limiarização MCC

A limiarização MCC (YEN; CHANG, 1995) está fundamentada no Critério de Correlação Máxima, considerando a discrepância entre a imagem original e a limiarizada e também no número de bits requeridos para representar a imagem binarizada. A idéia básica consiste em escolher o limiar baseado na correlação máxima do objeto e do fundo da imagem. Considerando uma imagem  $f(x, y)$ , de tamanho  $N \times N$  pixels representados por  $m$  níveis de cinza, seja o conjunto de níveis de cinza  $G_m = 0, 1, \dots, (m - 1)$  e  $f_i, i \in G_m$  como sendo as frequências dos níveis de cinza da imagem  $f$ . A probabilidade do nível de cinza  $i$  na imagem  $f$  pode ser calculada da seguinte forma:

$$p_i = \frac{f_i}{N \times N} \quad i \in G_m \quad (3.50)$$

Assim, obtém-se a distribuição  $\frac{p_i}{i} \in G_m$ . Para um dado valor de nível de cinza  $s$ , verifica-se se  $\sum_{i=0}^{s-1} p_i < s < 1$ , caso positivo, as duas distribuições seguintes podem ser derivadas desta distribuição após a normalização, sendo dada da seguinte maneira:

$$\begin{aligned}
A &\equiv \left\{ \frac{p_0}{P(s)}, \frac{p_1}{P(s)}, \dots, \frac{p_{s-1}}{P(s)} \right\} \\
B &\equiv \left\{ \frac{p_0}{1-P(s)}, \frac{p_{s-1}}{1-P(s)}, \dots, \frac{p_{m-1}}{1-P(s)} \right\} \\
\text{com } P(s) &= \sum_{i=0}^{s-1} p_i
\end{aligned} \tag{3.51}$$

Com as definições apresentadas acima, a quantidade total de correlação fornecida pelas distribuições A e B é dada por;

$$\begin{aligned}
TC(s) &= C_A(s) + C_B(s) = -\ln[G(s)xG'(s)] + 2\ln[P(s)x(1-P(s))] \\
\text{com } G(s) &= \sum_{i=0}^{s-1} p_i^2 \quad e \quad G'(s) = \sum_{i=s}^{m-1} p_i^2
\end{aligned} \tag{3.52}$$

Assim, o Critério de Correlação Máxima (MCC) implica em determinar o limiar ( $s^*$ ) que maximiza  $TC(s)$ :

$$TC(s^*) = TC(s) \quad \text{sendo } s \in G_m \tag{3.53}$$

### 3.4.2 Limiarização Local

É a definição de vários limiares para regiões diferentes em uma mesma imagem, ou seja, definir mais de um limiar para a imagem. Nas técnicas de limiarização local, o objetivo é calcular um limiar  $t(x, y)$  para cada pixel da seguinte maneira:

$$o(x, y) = \begin{cases} 0 & \text{se } g(x, y) \leq t(x, y) \\ 255 & \text{para os outros casos} \end{cases} \tag{3.54}$$

#### 3.4.2.1 Limiarização de Sauvola Otimizado

O algoritmo de limiarização de Sauvola (SHAFAIT; KEYSERS; BREUEL, 2007) é uma modificação do algoritmo de Niblack para imagens com problemas de iluminação. A

formulação do limiar é descrita da seguinte forma:

$$t(x, y) = m(x, y) \left[ 1 + k \left( \frac{s(x, y)}{R} \right) - 1 \right] \quad (3.55)$$

A localização dos pixels vizinhos é definida por uma janela de tamanho fixo. Na técnica de Sauvola, o limiar  $t(x, y)$  é calculado pela média  $m(x, y)$  e o desvio padrão  $s(x, y)$  das intensidades dos pixels da janela  $w \times w$  centrada no pixel  $(x, y)$ . Para determinar o limiar  $t(x, y)$  é preciso definir o valor de  $R$ , que é o valor máximo do desvio padrão, sendo utilizado o valor  $R=128$  para um documento em escala de cinza, e  $K$  como parâmetro que aceita valores positivos na faixa  $[0.2, 0.5]$ .

Para otimizar o algoritmo de Sauvola pode-se utilizar imagens integrais como método (SHAFAIT; KEYSERS; BREUEL, 2008). Define-se que uma imagem integral  $i$  de qualquer imagem de entrada  $g$ , é dada por uma imagem cuja intensidade de um determinado pixel seja igual à soma de todos os pixels acima e à esquerda dele, da seguinte maneira:

$$I(x, y) = \sum_{i=0}^x \sum_{j=0}^y g(i, j) \quad (3.56)$$

Com a imagem integral, precisa-se calcular a média local  $m(x, y)$  utilizando um tamanho de janela  $w$ , as coordenadas do pixel  $x, y$ , e realizar simplesmente duas operações de adição e uma subtração, descritas pela equação 3.57 :

$$m(x, y) = \frac{\left( \begin{array}{l} I(x + \frac{w}{2}, y + \frac{w}{2}) + \\ I(x - \frac{w}{2}, y - \frac{w}{2}) - \\ I(x + \frac{w}{2}, y - \frac{w}{2}) - \\ I(x - \frac{w}{2}, y + \frac{w}{2}) \end{array} \right)}{w^2} \quad (3.57)$$

Similarmente, considerando o cálculo da variância local, temos:

$$s^2(x, y) = \frac{1}{w^2} \sum_{i=x-\frac{w}{2}}^{x+\frac{w}{2}} \sum_{j=y-\frac{w}{2}}^{y+\frac{w}{2}} g^2(i, j) - m^2(x, y) \quad (3.58)$$

Uma vez calculada a imagem integral das intensidades dos pixels e os quadrados das intensidades dos pixels, pode-se calcular as médias locais e variâncias, independentemente do tamanho da janela.

### 3.5 Técnica do Perfil de Projeção

A técnica de perfil de projeção assume que uma imagem é formada por regiões de texto, organizadas em linhas paralelas, sendo possível determinar aquelas regiões que possuem uma grande concentração de pixels na mesma direção, vertical ou horizontal (MASCARO; CAVALCANTI; MELLO, 2010). Esta técnica consegue detectar os pixels de outros elementos como figuras, tabelas ou ruído.

A Figura 3.15 ilustra o exemplo da aplicação do perfil de projeção horizontal e vertical (MASCARO; CAVALCANTI; MELLO, 2010).

- A imagem (a) é o resultado da limiarização do conteúdo do primeiro plano de uma página da Bíblia.
- A imagem (b) é o resultado da aplicação do perfil de projeção horizontal.
- A imagem (c) é o resultado da aplicação do perfil de projeção vertical.

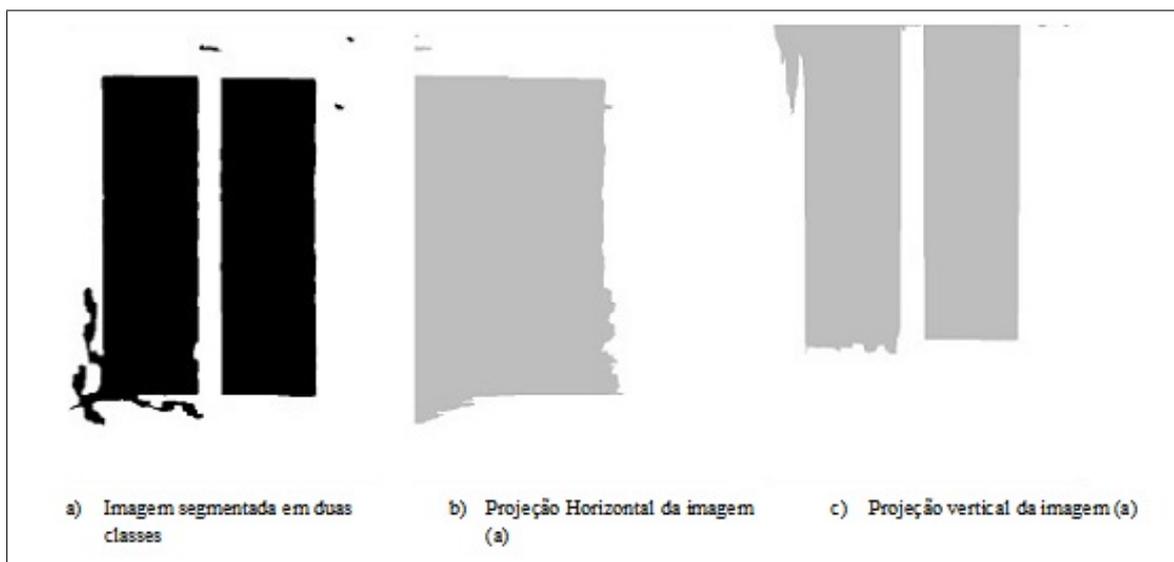


Figura 3.15: Ilustração da aplicação do perfil de projeção horizontal e vertical.

Com os perfis de projeção horizontal e vertical, aplicações podem-se utilizar destas informações para determinar as áreas de texto de interesse.

### 3.6 Técnica de rotulação

A rotulação consiste em etiquetar todos os pixels de uma determinada região da imagem binarizada com um valor de rótulo  $L$ , representado por um valor numérico (GOSE;

JOHNSONBAUGH; JOST, 1996) . A Figura 3.16 ilustra o exemplo de rotulação de um conjunto de pixels de uma imagem da Bíblia.

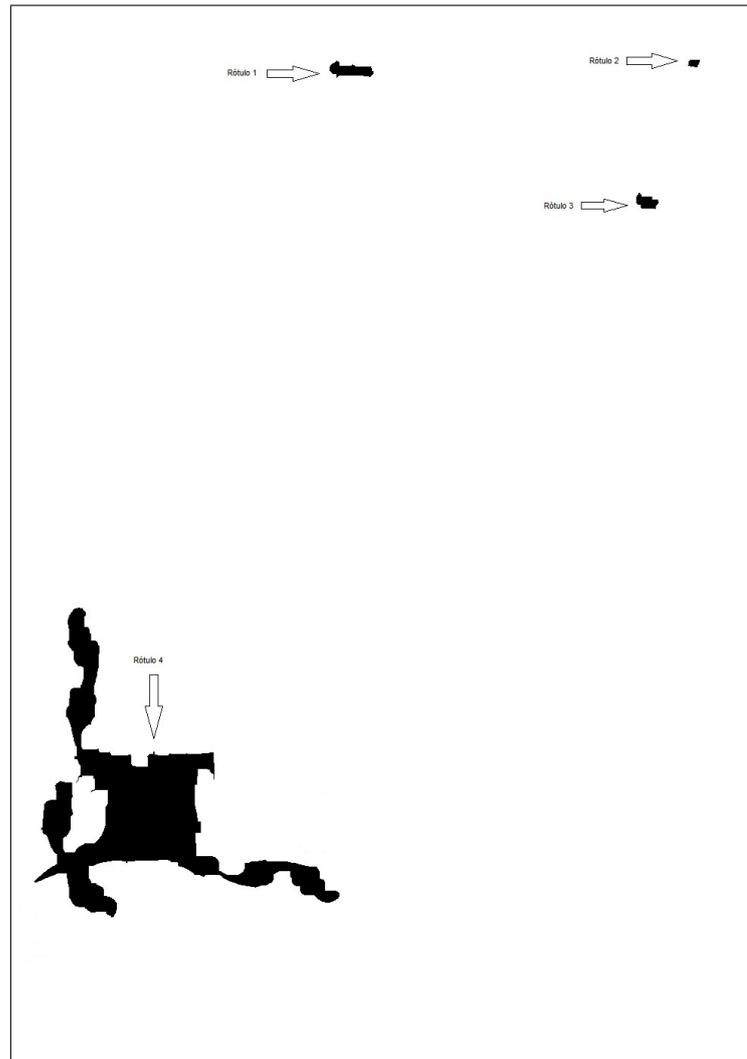


Figura 3.16: Ilustração da rotulação ideal dos pixels de uma imagem da Bíblia.

### 3.7 Técnica de Bounding Box

A técnica de bounding box consiste em gerar, para cada rotulação, uma área retangular que delimita a altura e largura desta, contendo os pixels da rotulação, que servirá para posterior classificação. A figura 3.17 ilustra o exemplo de duas áreas retangulares em dois conjuntos de uma imagem.

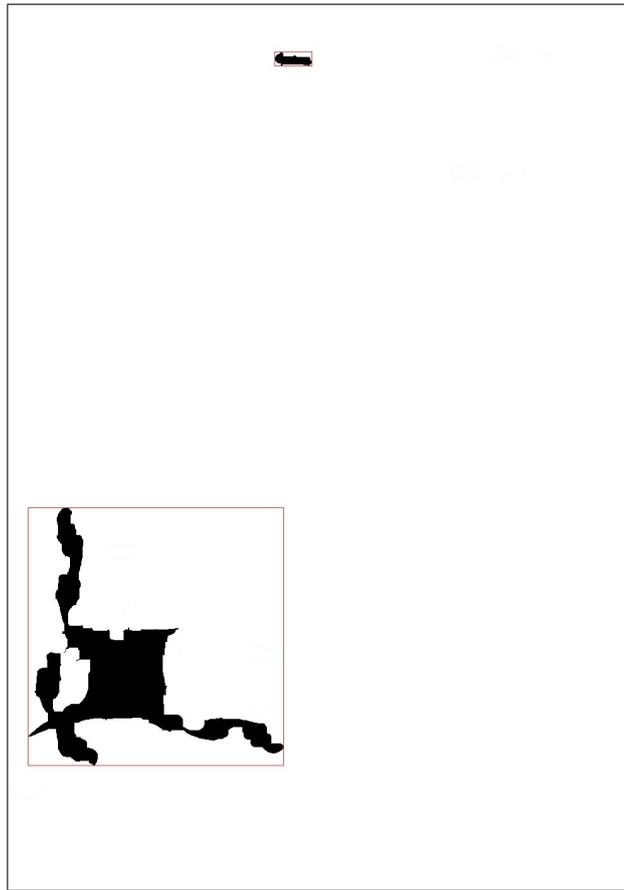


Figura 3.17: Exemplo de dois conjuntos rotulados e etiquetados com as áreas retangulares.

## 3.8 Avaliação da eficiência dos trabalhos

A avaliação da eficiência pode ser medida por meio da comparação dos resultados obtidos com as imagens *ground truth* através das métricas. Nesta seção estão descritas a geração de imagens *ground truth* e as métricas que podem ser utilizadas.

### 3.8.1 Geração de imagens *ground truth*

A imagem *ground truth* é uma imagem de referência que apresenta o resultado da segmentação ideal/desejada de qualquer algoritmo de segmentação. A geração das imagens segmentadas *ground truth* é de suma importância para se poder avaliar a segmentação de forma quantitativa, sendo possível realizar uma avaliação pixel a pixel do resultado. A geração das imagens *ground truth* pode ser realizada de forma manual, semi-automática ou automática. A Figura 3.18 mostra, na imagem (a), uma página da Bíblia e em (b), a imagem *ground truth* da Bíblia gerada de forma manual, em duas classes: classe 1(o fundo); e classe 2 (o conteúdo com o título, desenhos, manuscritos e colunas).



Figura 3.18: Exemplo de imagem *ground truth* da segmentação do texto da imagem(a).

As imagens *ground truth* podem ser geradas manualmente por um 'Zoner' (profissional da área de imagens) (BAIRD, 2007), por retângulos isométricos, com a desvantagem de alguns conteúdos manuscritos não serem detectados por esses retângulos.

### 3.8.2 Geração de imagens sintéticas

Uma base de imagens artificiais foi gerada unindo uma imagem somente com o fundo (imagem de fundo) com uma imagem com as informações do primeiro plano. Aqui está descrito como foram criadas essas imagens.

- Geração de um fundo: pode-se utilizar uma imagem da Bíblia de Gutenberg que possua somente o fundo, ou gerar uma imagem de fundo estimado das imagens da Bíblia.
- Geração de uma imagem com informações do primeiro plano: foram geradas imagens com informações do primeiro plano, com texto verdadeiro da Bíblia traduzido em latim, com fonte gótica e com desenhos parecidos com as páginas da Bíblia de Gutenberg.

- Geração da imagem artificial: é a união da imagem com o fundo e a imagem com o conteúdo do primeiro plano gerada sinteticamente.
- Foram geradas imagens sintéticas com variação da opacidade em 70, 80, 90 e 100% para mostrar a influência do fundo e do primeiro plano.

A Figura 3.19 ilustra o exemplo da união da imagem fundo com a imagem do primeiro plano.

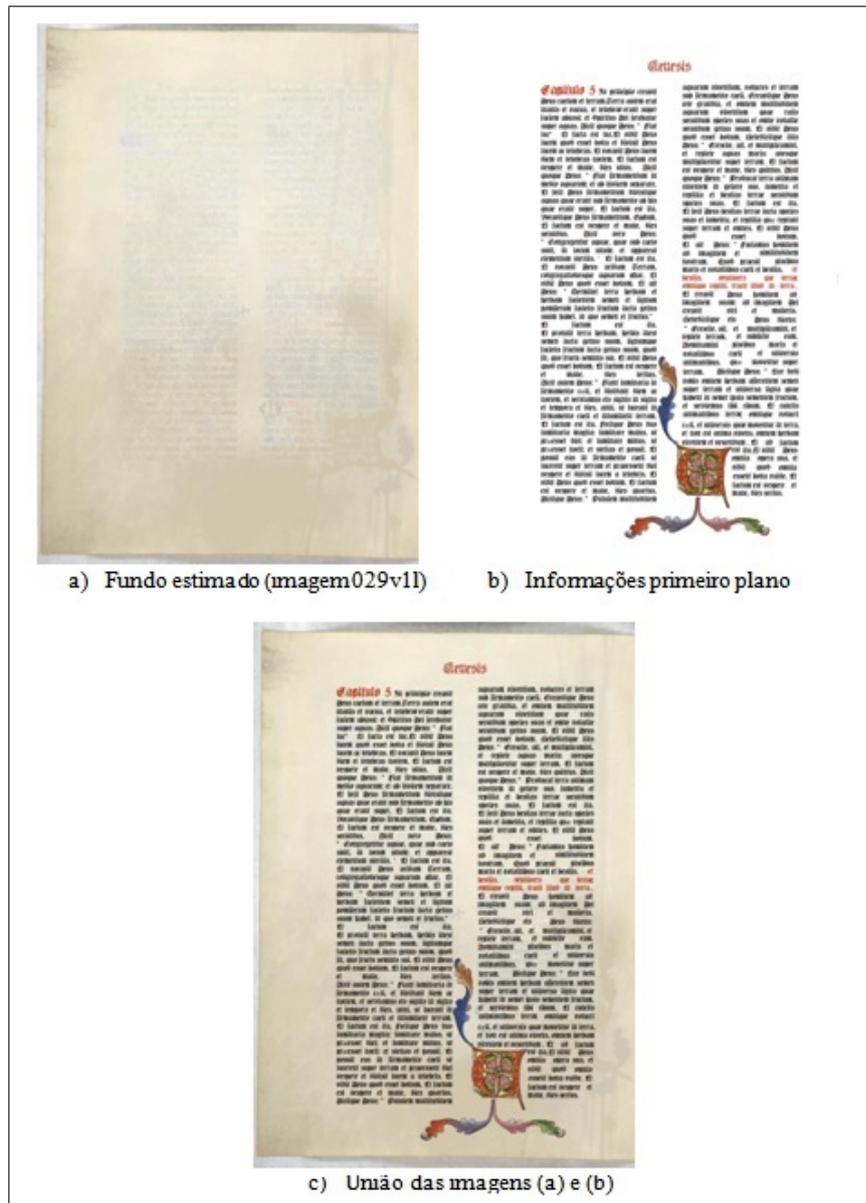


Figura 3.19: Exemplo da união das imagens fundo e primeiro plano.

No momento da união o fundo vai interferir com as informações do primeiro plano. Isto é tratado com a geração de várias imagens com diferentes valores de opacidade,

variando de 0% a 100 %. A fórmula a seguir ilustra a geração da opacidade (LINS; SILVA, 2007).

$$interf(i, j) = \alpha * paper(i, j) + (1 - \alpha) * ink(i, j) \quad (3.59)$$

em que  $paper(i, j)$  representa a intensidade da tinta das informações do primeiro plano e  $ink(i, j)$  representa a intensidade da tinta do fundo.

### 3.8.3 Métricas de Avaliação

As métricas são utilizadas para medir a eficiência e avaliar o desempenho do método de segmentação proposto, inclusive as páginas da Bíblia de Gutenberg. Todos os pixels segmentados podem ser avaliados de três maneiras (BAIRD, 2007).

1. Precisão por pixel: é a fração de todos os pixels do documento, que são classificados corretamente de acordo com as áreas zoneadas. É uma medida objetiva e quantitativa, ainda que um pouco arbitrária devido à variedade de formas com que os conteúdos podem ser zoneados, embora alguns conteúdos manuscritos muitas vezes não podem ser descritos por zonas retangulares. No entanto, esta métrica provê uma simples generalização de como a classificação está sendo executada.
2. Análise de precisão por página: consiste na análise dos conjuntos especificados de uma página. Por exemplo: analisar a quantidade de desenhos que foram segmentados em uma página, que ela contenha pelo menos 70 % dos pixels dos desenhos, classificando os desenhos como segmentados ou não.
3. Qualidade da segmentação subjetiva: esta análise consiste em classificar, de forma expressa, a segmentação em "bom", "justo" e "ruim".

Para a utilização das métricas deve-se usar a estratégia precisão por pixel. Abaixo estão descritas as métricas mais utilizadas para avaliação em morfologia.

- A métrica que representa a taxa de erro de classificação ( $ME$ ) (SEZGIN; SANKUR, 2001), retorna 0 em caso de perfeita segmentação e pode ser descrita pela equação 3.60.

$$ME = 1 - \frac{|B_g \cap B_s| + |F_g \cap F_s|}{|B_g| + |F_g|} \quad (3.60)$$

na qual  $(B_g, F_g)$  e  $(B_s, F_s)$  representam o fundo e o primeiro plano da segmentação da imagem *ground truth* e da segmentação obtida pelo método proposto, respectivamente, sendo  $|X|$  a cardinalidade da imagem X.

- A métrica que representa o erro relativo da área do primeiro plano ( $RAE$ ) (SEZGIN; SANKUR, 2001), retorna 0 em caso de segmentação perfeita, e pode ser descrita pela equação 3.61.

$$RAE = \begin{cases} \frac{|F_g| - |F_s|}{|F_g|} \text{ if } |F_g| > |F_s| \\ ou \\ \frac{|F_s| - |F_g|}{|F_s|} \text{ if } |F_g| < |F_s| \end{cases} \quad (3.61)$$

em que  $F_g$  e  $F_s$  representam a segmentação do primeiro plano da imagem verdade e a segmentação obtida pelo método proposto, respectivamente, sendo  $|X|$  a cardinalidade da imagem  $X$ .

- A métrica que representa a precisão ( $P$ ) (BIMBO, 1999), retorna 1 em caso de segmentação perfeita e pode ser descrita pela equação 3.62.

$$P = \frac{T^+}{T^+ + F^+} \quad (3.62)$$

- A métrica que representa a Acurácia ( $A$ ) (BIMBO, 1999), retorna 1 em caso de segmentação perfeita, é definida como quanto se consegue chegar até próximo a meta, e pode ser descrita pela equação 3.63.

$$A = \frac{T^+ + T^-}{T^+ + T^- + F^+ + F^-} \quad (3.63)$$

- A métrica que representa a revocação ( $R$ ) (BIMBO, 1999), retorna 1 em caso de perfeita segmentação e pode ser descrita pela equação 3.64.

$$R = \frac{T^+}{T^+ + F^-} \quad (3.64)$$

- A métrica que representa a taxa de erro ( $ER$ ) (BIMBO, 1999), retorna 0 em caso de segmentação perfeita e pode ser descrita pela equação 3.65.

$$ER = \frac{F^+}{T^+ + F^-} \quad (3.65)$$

- A métrica que representa a medida F ( $FM$ ) (GATOS; NITIROGIANNIS; PRATIKAKIS, 2009), retorna 1 em caso de segmentação perfeita e pode ser descrita pela equação 3.66.

$$FM = \frac{2 * P * R}{P + R} \quad (3.66)$$

- A métrica que representa a taxa negativa  $NRM$  (GATOS; NITIROGIANNIS; PRATIKAKIS, 2009), retorna 0 em caso de segmentação perfeita e pode ser descrita pela equação 3.67.

$$NRM = \frac{NR_{FN} + NR_{FP}}{2} \text{ onde } NR_{FN} = \frac{F^-}{T^+ + F^-}, NR_{FP} = \frac{F^+}{T^- + F^+}; \quad (3.67)$$

Esclarecimentos finais:

- A métrica do verdadeiro positivo ( $T^+$ ) representa os pixels verdadeiros segmentados (detectados) pelo método proposto, de acordo com a classe determinada.
- A métrica do verdadeiro negativo ( $T^-$ ) representa os pixels verdadeiros segmentados (detectados) pelo método proposto como sendo de outra classe.
- A métrica do falso positivo ( $F^+$ ) representa os pixels detectados de outras classes como sendo da classe desejada.
- A métrica do falso negativo ( $F^-$ ) representa os pixels detectados de classe desejada como sendo pixels de outras classes.

### 3.9 Considerações Finais da Fundamentação Teórica

Foram apresentados os requisitos necessários sobre da fundamentação teórica para implementação deste projeto. O próximo capítulo apresenta os trabalhos relacionados envolvendo restauração e segmentação de documentos antigos.

# Capítulo 4

## Estado da Arte

Neste capítulo serão apresentados trabalhos relacionados a métodos de restauração, segmentação de livros antigos, avaliação da eficiência dos trabalhos por meio de métricas e geração de imagens *ground truth*.

### 4.1 Restauração de Documentos Antigos

Esta seção apresenta as abordagens sobre a restauração de livros antigos afetados por ruídos, manchas avermelhadas, papéis amarelados e rasgados.

#### 4.1.1 Remoção de ruídos por filtros em documentos de imagens binárias

Ping, Lihui e Alex (2000) apresentam o processo de remoção de ruídos irregulares em documentos de imagens binárias por meio de filtros. A remoção de ruído aumenta a capacidade de compreensão e análise de documentos e melhora a taxa de reconhecimento de caracteres por ferramentas OCR. Esse trabalho apresenta dois algoritmos de filtragem para remoção de ruídos: 1) O MDMF (Filtro morfológico direcional modificado), utilizado para remoção de pequenos ruídos espalhados na imagem, é baseado nos operadores morfológicos, erosão e dilatação, e consiste em 04 operações de erosão seguida de uma dilatação; 2) O IGSF (Filtro da estrutura geométrica das imagens), utilizado para remover ruídos maiores, tem sua base nas informações geométricas do curso (traçado) dos caracteres. Os autores citados mostraram pelos experimentos que a utilização do filtro MDMF funcionou de forma satisfatória para eliminar ruídos cujos diâmetros eram menores que o menor traço dos caracteres. No entanto, esse algoritmo mostrou-se inadequado para o processamento de ruídos feitos pelo homem, tais como: retângulo, linhas e rabiscos. O software de digitalização "OCR TextBridge" usado nos experimentos reconheceu apenas

94 de 170 caracteres quando nenhum filtro foi utilizado. Com o uso do filtro da mediana, foram reconhecidos 119 caracteres e com o filtro MDMF foram reconhecidos 137 caracteres. Com o filtro IGSF, após a filtragem MDMF, foram reconhecidos 151 caracteres. As taxas de reconhecimento de caracteres de várias técnicas em pré-processamento estão descritas na tabela 4.1.

Tabela 4.1: Resultados das técnicas de pré-processamento.

Técnicas	Caracteres de Jornais	Caracteres Manuscritos
Sem processamento	55,30 %	51,10 %
Filtro Mediana	70,00 %	65,30 %
Filtro MDMF	80,60 %	77,20 %
Filtro MDMF + IGSF	88,80 %	89,50 %

A Figura 4.1 ilustra o uso dos filtros MDMF e IGSF.

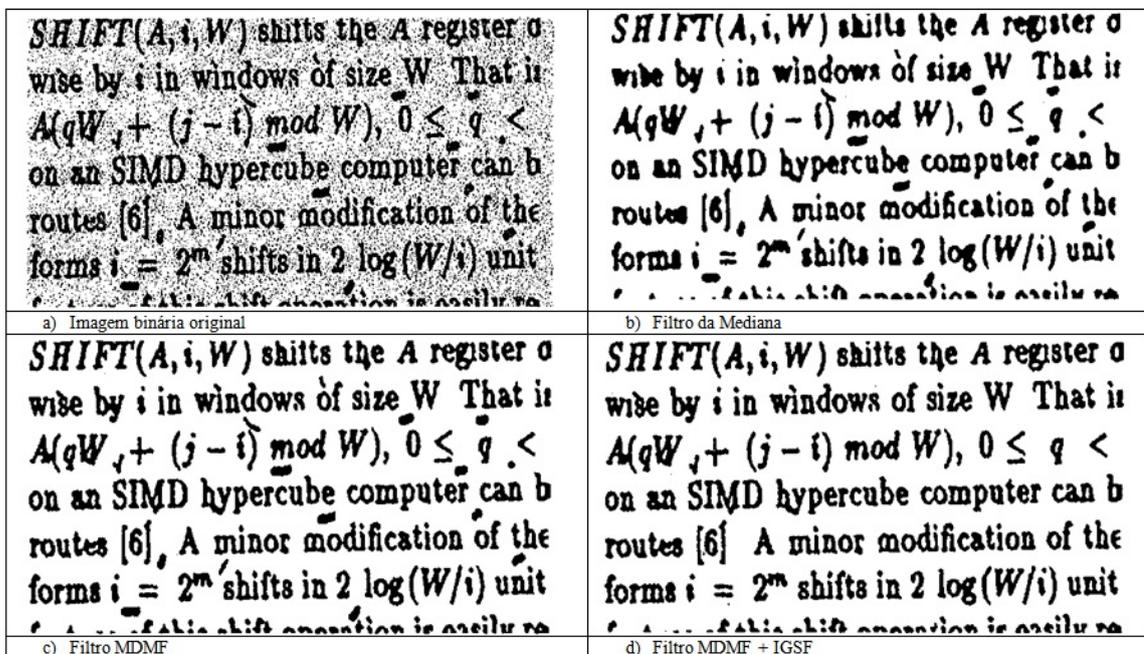


Figura 4.1: Exemplo do filtro MDMF e IGSF (PING; LIHUI; ALEX, 2000).

#### 4.1.2 Restauração de livros antigos afetados por manchas avermelhadas e papéis amarelados

Ramponi (2005) apresenta uma nova técnica automática para a restauração de papéis antigos, dos séculos XIV a XVIII, afetados por manchas avermelhadas e pelo amarelado do papel causado pelo processo de deterioração. A imagem de entrada, no espaço

de cor RGB, é convertida para o espaço de cor YCbCr. O processo aprimorou o aspecto do fundo do papel, tornando-o mais homogêneo, corrigindo a descoloração do fundo e reduzindo o amarelado do papel. O algoritmo dividiu o histograma de luminosidade em duas partes: texto e fundo, processo baseado no algoritmo de limiarização de Otsu para a seleção do limiar de corte, para binarizar o fundo e o primeiro plano. A correção de cores do fundo foi baseada nas informações derivadas da luminância (canal Y), sendo modificados os valores de crominância dos pixels que pertencem ao fundo, ficando os pixels do primeiro plano inalterados. Os valores dos pixels do fundo foram baseados na média das crominâncias dos pixels do fundo. A Figura 4.2 apresenta um exemplo da restauração do fundo de uma imagem, deixando o fundo mais homogêneo.

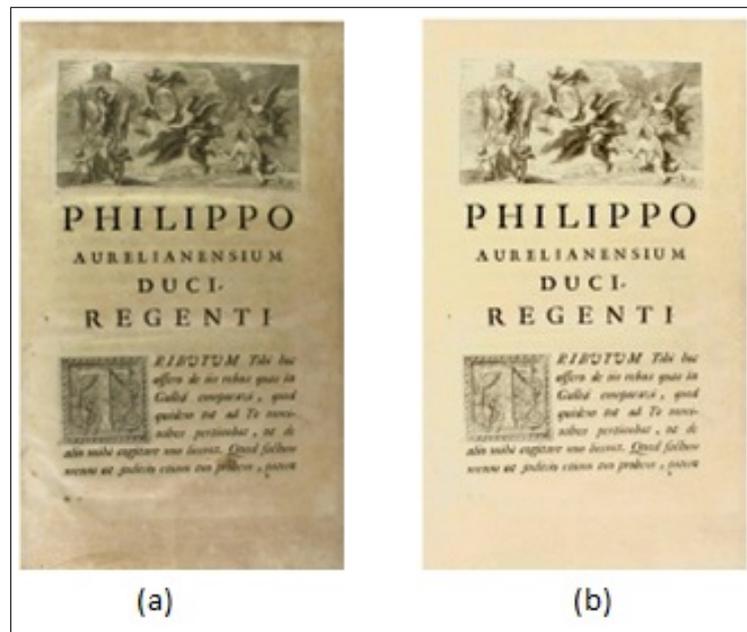


Figura 4.2: Exemplo da restauração do fundo (RAMPONI, 2005).

Para a restauração das manchas avermelhadas nas imagens, o método analisa e identifica as manchas pelo valor de crominância  $Cr$  dos pixels, em que os pixels que representam as manchas possuem valores maiores em  $Cr$ . Assim, para identificar os pixels, é preciso determinar um valor máximo de  $Cr$  inserido pelo usuário. A Figura 4.3, ilustra a restauração de um documento com manchas avermelhadas, sendo a imagem (a) original com as manchas, (b) o resultado da restauração das manchas avermelhadas, e a imagem (c) a restauração do fundo.

Os documentos originais utilizados nos experimentos pertencem à Biblioteca Re-dentoristi em S. Maria della Consolazione, em Veneza, Itália. Neste trabalho não foram divulgadas as taxas e a quantidade de imagens testadas.

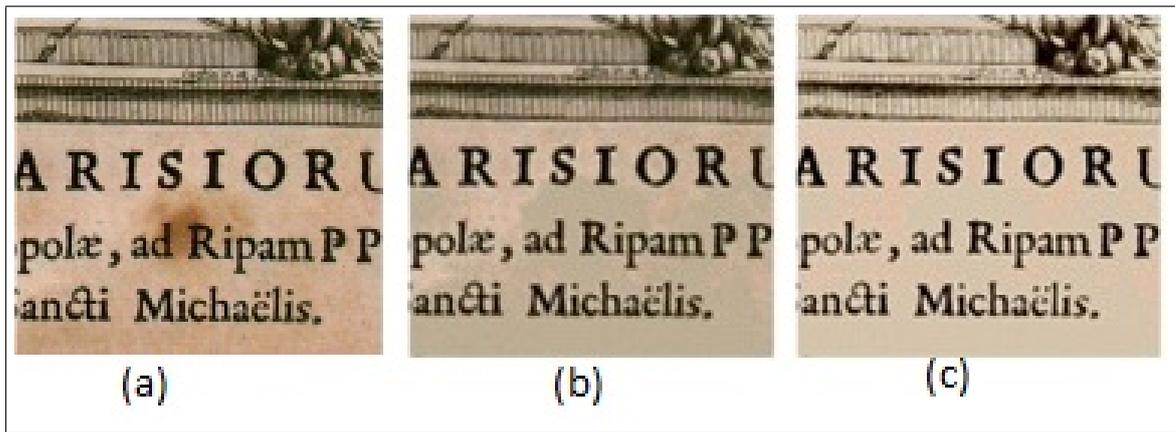


Figura 4.3: Exemplo da restauração de uma imagem com manchas avermelhadas (RAM-PONI, 2005).

#### 4.1.2.1 Restauração de livros antigos afetados por alterações mecânicas e químicas

Arcidiacono, Portuese e Scifo (2005) abordam o tema da restauração digital de livros antigos deteriorados por alterações mecânicas e químicas, tais como: papel rasgado, riscos, manchas avermelhadas, amarelado do papel e borrão de água. O objetivo é realizar a remoção das manchas, reconstruindo as áreas danificadas para melhorar o reconhecimento de textos por processos de OCR. Os defeitos de rachaduras e riscos podem ser detectados por operadores de borda e reparados pela técnica de *inpainting*, permitindo completar a falta de informação em imagens, seja por falha ou por remoção intencional de alguma área ou objeto (BERTALMIO, 2000). A Figura 4.4 mostra a imagem com rachaduras e a recuperação da imagem.

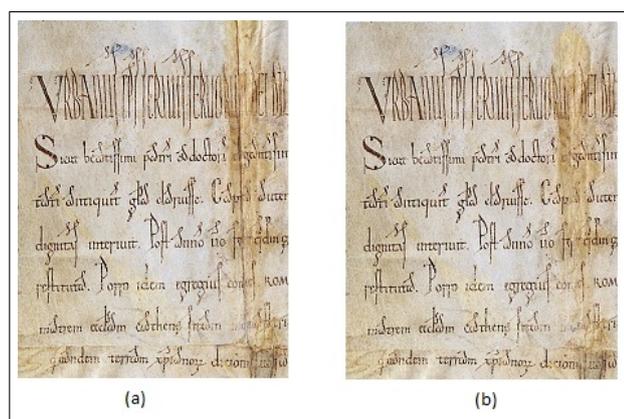


Figura 4.4: Exemplo da restauração de rachaduras e riscos de uma imagem (ARCIDIA-CONO; PORTUESE; SCIFO, 2005).

### a) Correção de rasgos

O papel rasgado pode comprometer drasticamente o documento porque danifica uma área de interesse do documento. A detecção do papel rasgado pode ser feita com o uso de uma cor uniforme predeterminada do fundo durante a digitalização do papel. A reconstrução das partes rasgadas pode ser realizada pelo algoritmo de *inpaintin* (BERTALMIO, 2000). A Figura 4.5 mostra a imagem com papel rasgado (a) e a imagem restaurada em (b).

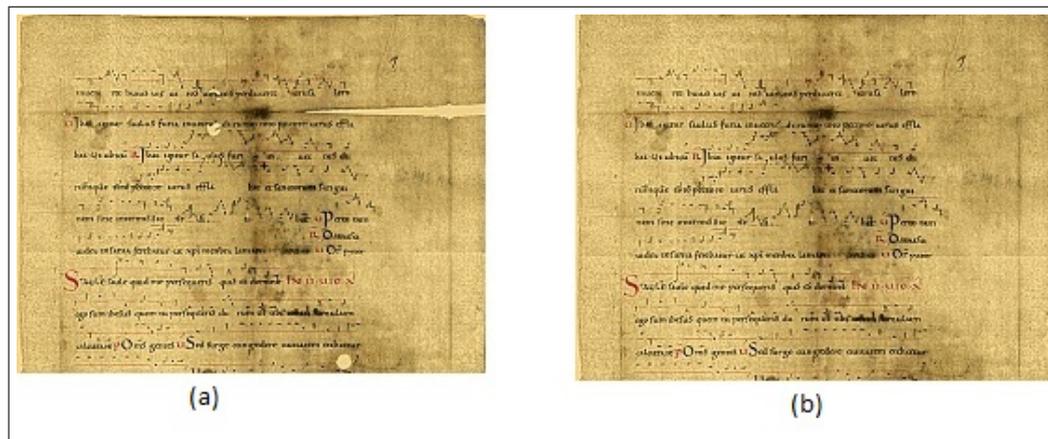


Figura 4.5: Exemplo da restauração de uma imagem com papel rasgado (ARCIDIACONO; PORTUESE; SCIFO, 2005).

### b) Correção da oxidação do papel

A alteração química, causada pela oxidação do papel, é um processo natural que afeta todo papel velho, sendo caracterizado pela cor amarelada. A remoção deste efeito deixa o fundo mais homogêneo, melhorando a legibilidade do documento para os algoritmos de reconhecimento de caracteres. O algoritmo tenta separar o texto do fundo por meio do particionamento do histograma de luminância. A restauração ocorre pelo ajuste da cor do pixel que pertence ao fundo com a sua mediana. A Figura 4.6 mostra um exemplo de restauração, sendo a imagem (a) com amarelado do papel e (b) a restauração.

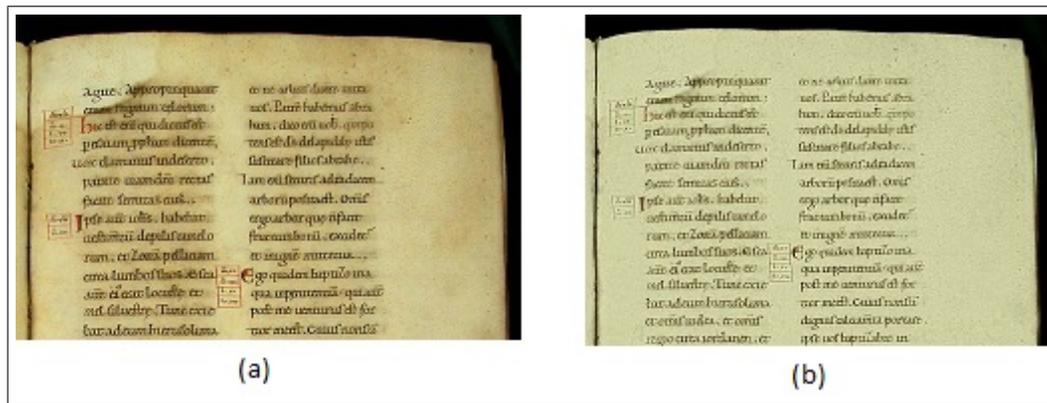


Figura 4.6: Exemplo da restauração de uma imagem amarelada (ARCIDIACONO; PORTUESE; SCIFO, 2005).

### c) Correção das manchas avermelhadas

As manchas avermelhadas do papel podem ter diversas origens, como por exemplo o crescimento de fungos ocorrido pela oxidação de metais, como a utilização de ferro em tintas velhas. As manchas avermelhadas podem ser detectadas por meio da análise da matriz de crominância Cr no espaço de cor YCrCb. As matrizes de crominância Cr e Cb são ajustadas para corrigir a cor de áreas avermelhadas. A Figura 4.7 mostra um exemplo de imagem com uma área avermelhada.

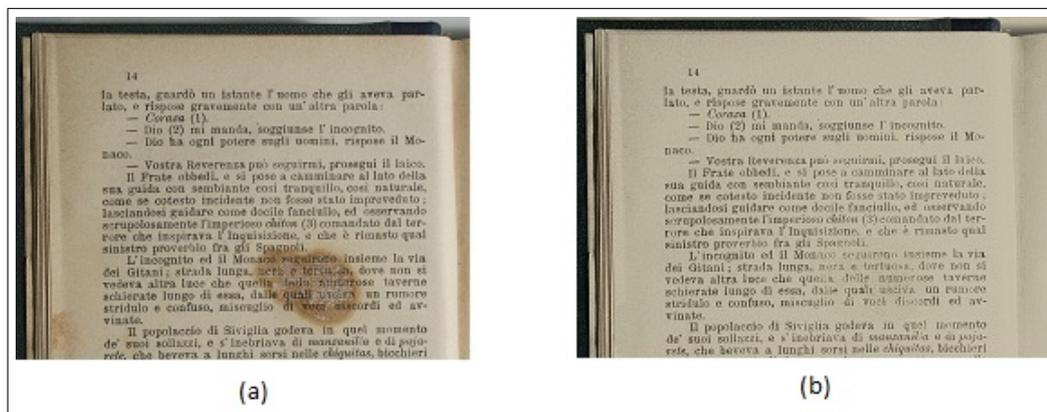


Figura 4.7: Exemplo da restauração de uma imagem com mancha avermelhada (ARCIDIACONO; PORTUESE; SCIFO, 2005).

#### d) Correção de manchas de água

O borrão d'água é outro tipo de alteração química, geralmente proveniente de ambientes adversos em que o livro foi conservado. É caracterizado por sua forma irregular (delimitada por uma borda visível) e se apresentam com cor mais escura que o fundo. As manchas de água podem ser detectadas através de uma sequência de reconstruções morfológicas e a restauração é realizada por uma equação descrevendo a área danificada em função da área original (modelo adição/multiplicação). A Figura 4.8 mostra um exemplo de uma imagem com manchas de água (a) e a imagem restaurada (b).

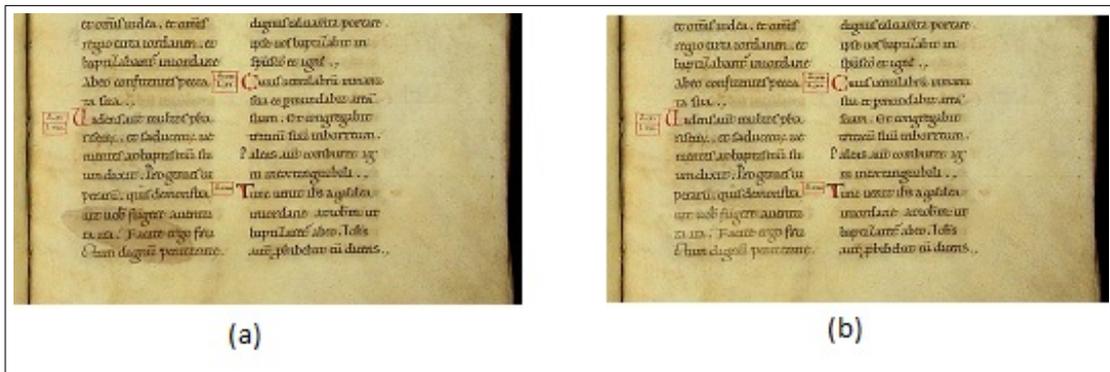


Figura 4.8: Exemplo da restauração de uma imagem com manchas de água (ARCIDIACONO; PORTUESE; SCIFO, 2005).

## 4.2 Segmentação de documentos antigos

A segmentação de documentos antigos permite o estudo mais aprofundado de cada grupo semântico do texto, como por exemplo, títulos, manuscritos, desenhos etc. Nesta seção são descritas várias abordagens de segmentação de documentos coloridos e binários.

### 4.2.1 Segmentação baseada na morfologia matemática para identificar e extrair conteúdos de livros antigos

Os autores Granado, Pina e Muge (2001) apresentam uma metodologia de segmentação baseada na morfologia matemática, para imagens binárias, para identificar e extrair conteúdos de livros antigos impressos no ano de 1554. Foram previamente classificados em cinco grupos diferentes: letras capitulares, listras, figuras, anotações e texto impresso, sendo caracterizados por suas particularidades geométricas. Para realizar a

segmentação foram utilizados os operadores morfológicos: erosão, dilatação, abertura e fechamento. A metodologia deste trabalho é composta por 03 níveis descritos a seguir e está representada pela Figura 4.9.

- Nível 1: Separação do conjunto texto e não-texto;
- Nível 2: Separação de figuras e não figuras do conjunto considerado não-texto e separação de anotações e do texto central.
- Nível 3: Separação dos elementos que não são figuras como as listras e letras iluminadas.

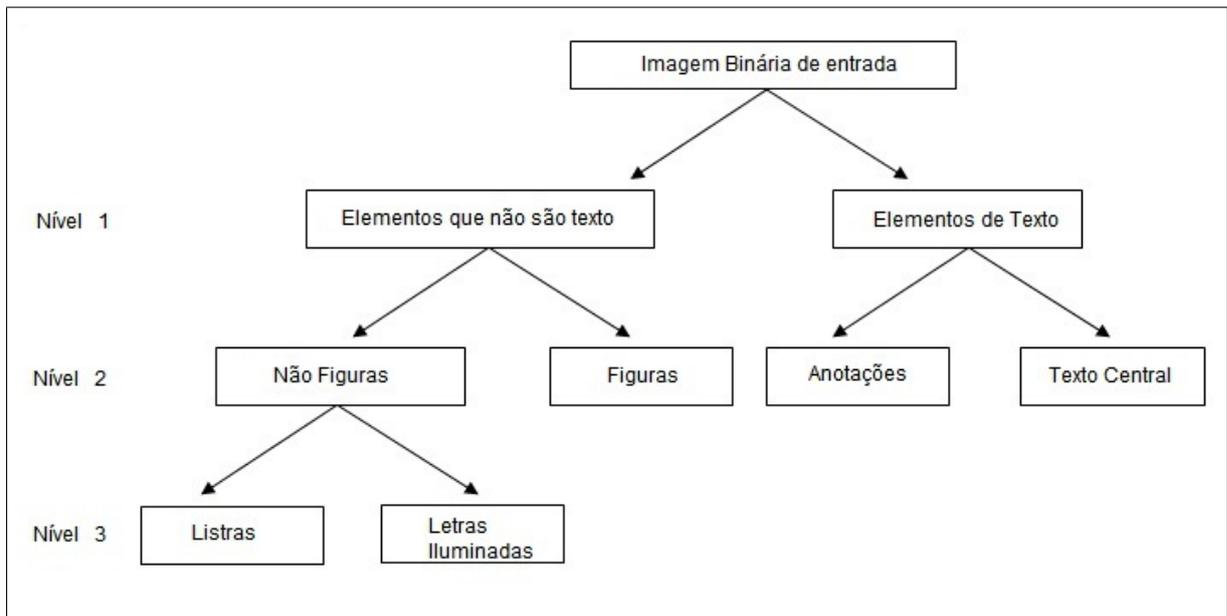


Figura 4.9: Fluxograma do método de segmentação do conteúdo das páginas antigas (GRANADO; PINA; MUGE, 2001).

Para o nível 1 é realizada a operação morfológica fechamento, seguida das operações de abertura e reconstrução. A Figura 4.10 ilustra as operações realizadas, sendo as imagens (a) original, (b) o resultado da operação fechamento, (c) o resultado da abertura, (d) a imagem reconstruída, (e) a recuperação dos pixels originais e (f) o resultado da intersecção entre as imagens (e) e (a).

No nível 2 foi realizada a operação fechamento, seguida das operações de reconstrução, subtração e intersecção. A Figura 4.11 ilustra os resultados para este nível, sendo a imagem (a) original, (b) o resultado do fechamento, (c) o resultado da erosão, (d) o resultado da reconstrução, (e) apresenta a recuperação dos pixels originais e (f) o resultado dos elementos que não são figuras.

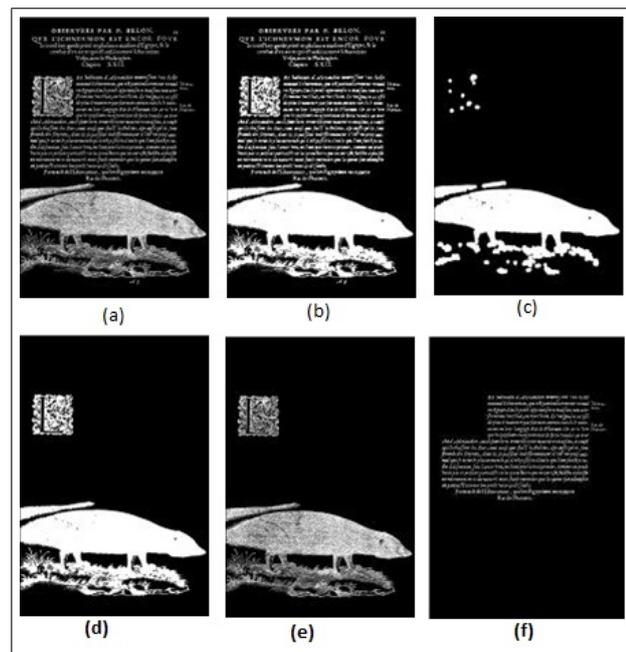


Figura 4.10: Nível 1: Segmentação do texto e não texto (GRANADO; PINA; MUGE, 2001).

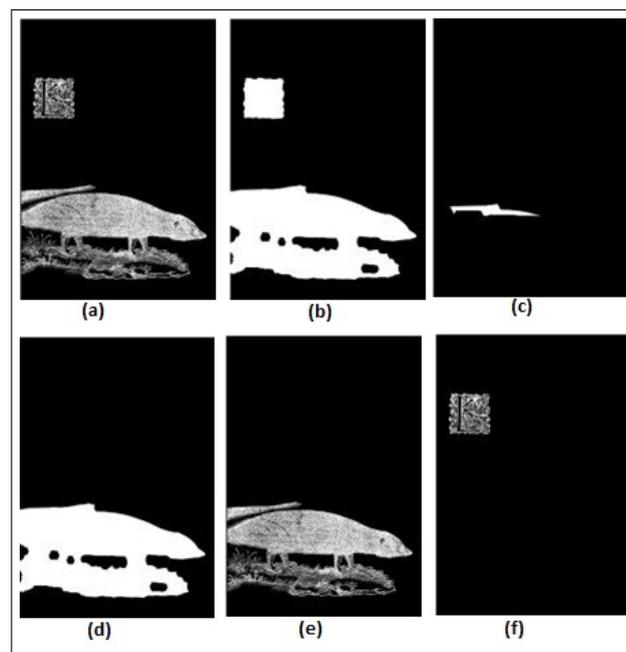


Figura 4.11: Nível 2: Segmentação das figuras e não figuras da imagem (GRANADO; PINA; MUGE, 2001).

No nível 3 foram realizadas as operações de fechamento, abertura, erosão, reconstrução, intersecção e subtração. A Figura 4.12 ilustra os resultados para este nível, sendo a imagem (a) de entrada, (b) resultado da operação de fechamento e abertura, (c) a erosão,

(d) a reconstrução, (e) a recuperação dos pixels originais que representam os desenhos e (f) a subtração com as letras iluminadas.

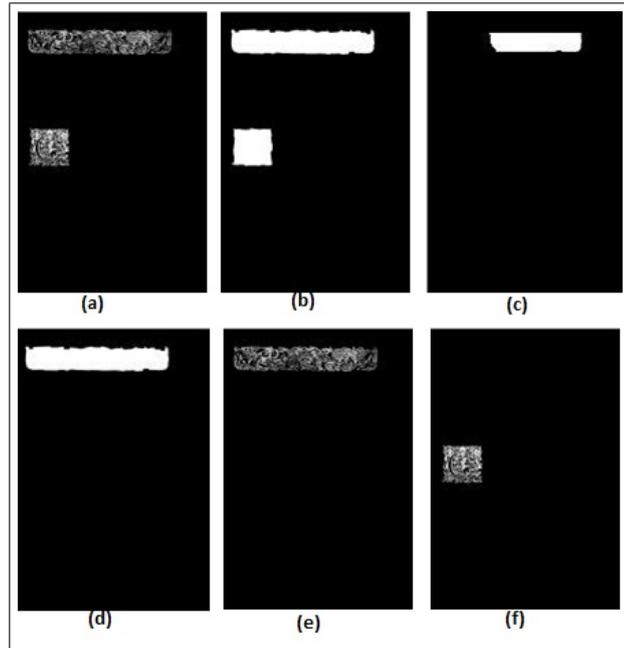


Figura 4.12: Nível 3: Segmentação de listras e letras iluminadas (GRANADO; PINA; MUGE, 2001).

A Figura 4.13 ilustra a separação de anotações e do texto central com base nas operações de fechamento, abertura, erosão, reconstrução, intersecção e subtração, sendo a imagem (b) o resultado da operação fechamento, (c) o resultado da abertura, (d) a reconstrução, (e) o resultado da intersecção das imagens (a) e (d), e a imagem (f) o resultado da subtração entre as imagens (e) e (a).

Os testes foram realizados em páginas de um livro antigo, de 1554, "Les observations de plusieurs singularités et choses mémorables trouvées en Grèce", em Asie, em Judée de Pierre de Belon, com 451 páginas. Os resultados apresentados foram:

- Nível 1 (separação do texto e não-texto,) a taxa de classificação foi de 97,78 % .
- Nível 2 (separação das figuras e elementos que não eram figuras), a taxa de sucesso foi de 90,71 % . Para a separação das anotações e do texto central, a taxa de classificação foi de 74,51 % .
- Nível 3 (separação das letras iluminadas e listras), a taxa de acerto foi de 100 % .

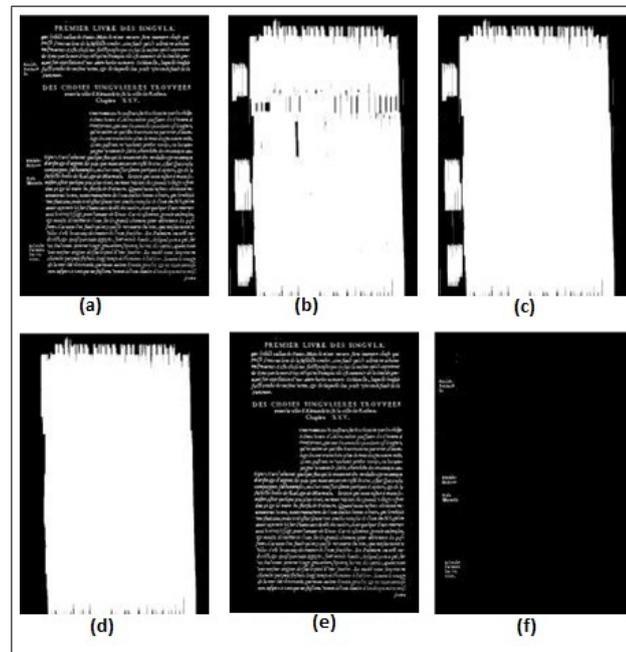


Figura 4.13: Nível 2: Segmentação das anotações manuscritas e do texto central impresso (GRANADO; PINA; MUGE, 2001).

#### 4.2.2 Caracterização de documentos antigos baseados na análise de textura das páginas

Journet, Ramel e Mullot (2008) propuseram um método de caracterização de imagens de documentos antigos baseado na análise da textura das páginas, com a ajuda de um estudo de multi-resolução das texturas contidas nas imagens. A originalidade da proposta é baseada em não segmentar ou extrair informações por meio das estruturas dos documentos analisados. Ao extrair informações ligadas às frequências e às orientações das diferentes partes de uma imagem, é possível extrair e comparar os elementos do conteúdo. A segmentação de imagens de documentos é realizada por meio da detecção das áreas com características específicas de orientações e frequências, sendo utilizado conhecimento, *a priori*, sobre formas geométricas ou estruturas do documento. Para os testes foram utilizadas imagens em níveis de cinza, de 9 obras diferentes de documentos antigos, num total de 400 páginas. Cada teste começou com a aplicação do algoritmo de classificação em três grupos (texto, desenhos e fundo). A taxa de classificação dos desenhos foi de 83% e dos textos de 92%. O tempo de processamento não foi divulgado. A Figura 4.14 ilustra as direções dos pixels com base na circunferência, calculando-se as direções dos pixels das imagens, sendo a imagem (a) com direção na vertical, (b) horizontal, (c) com quatro linhas (d) com texto distorcido.

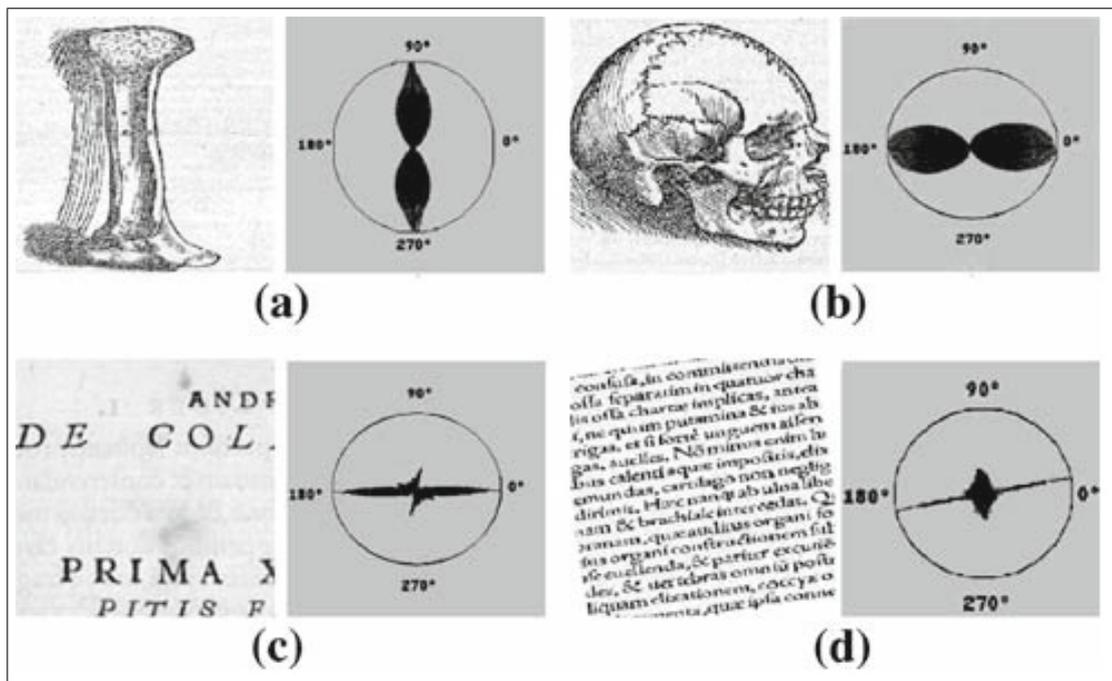


Figura 4.14: Exemplo baseado no cálculo das direções dos pixels das imagens. (JOURNET; RAMEL; MULLOT, 2008)

#### 4.2.3 Geração de uma versão sintética de documentos antigos

Os autores Mello e Lins (2002) apresentam uma abordagem sobre geração de documentos históricos por composição, parte do projeto Nabuco da Universidade Federal de Pernambuco. A base de dados é composta de quase 6.500 documentos do final do século XIX, do legado de Joaquim Nabuco, totalizando mais de 30.000 páginas. Os documentos foram digitalizados em cores verdadeiras, com resolução de 200 dpi e armazenados no formato JPEG, com 1 % de perda, e cada documento atinge, em média, 380 Kb. O trabalho apresenta um sistema eficiente para indexação, armazenamento e transmissão por rede de documentos históricos, por meio da geração de uma versão sintética de imagens por segmentação da textura e do texto (tinta). Os documentos são decompostos por suas características, como: textura do papel, cores, partes manuscritas, partes datilografadas, desenhos e outras. A Figura 4.15 ilustra o método proposto (MELLO; LINS, 2002).

Uma característica importante desta metodologia é que ela deve ser aplicada unicamente em textos datilografados. Após a segmentação do texto tinta e da textura do papel por um algoritmo baseado em entropia, as imagens são processadas separadamente para se extrair suas características principais. A parte da tinta é processada por uma ferramenta OCR para gerar um texto como entrada para uma base de dados de fonte. Para a formação da imagem sintética, é preciso fazer a remontagem do documento uti-

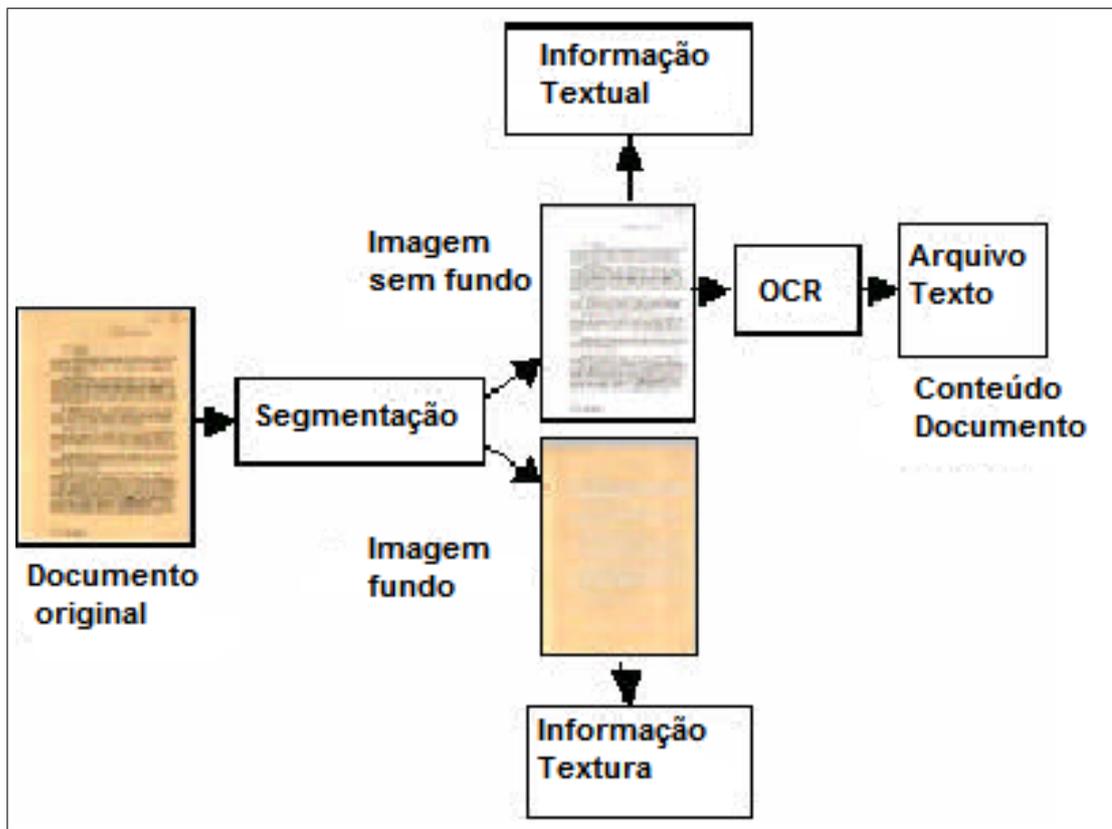


Figura 4.15: Exemplo da metodologia. (MELLO; LINS, 2002)

lizando as informações geradas pelo sistema OCR e pelas informações da textura, sendo estas últimas usadas para criar uma imagem em branco (igual a uma folha de papel), visualmente semelhante à folha original na época. A saída do sistema OCR gera uma imagem de texto com configurações padrões como: margens, tamanho dos espaços em branco e espaço entrelinha. Os outros dados são extraídos da imagem original como: cor e matiz. Assim, é gerada a imagem sintética composta (junção das informações da saída do OCR e da textura). O sistema foi aplicado a um conjunto de 30 imagens de documentos, no espaço de cor RGB, e as imagens sintéticas foram analisadas qualitativamente por meio de inspeção visual e quantitativamente pela análise de variância (ANOVA) (MILLER, 1986). O método ANOVA foi aplicado para analisar o histograma dos três canais RGB, observando-se a média, o desvio padrão, assimetria e curtose. Os autores não mencionam as taxas de acertos e erros. A Figura 4.16 ilustra o exemplo aplicado a uma imagem.

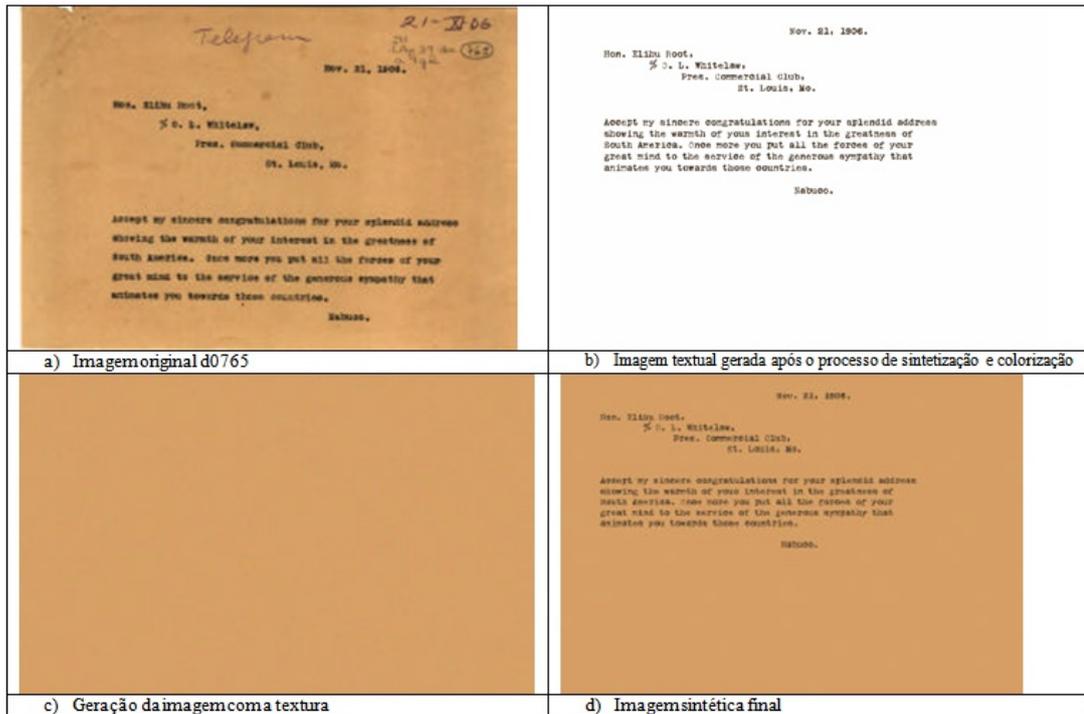


Figura 4.16: Resultado da aplicação da metodologia de Mello e Lins (2002)

#### 4.2.4 Segmentação de imagens coloridas para a separação do fundo e do primeiro plano

Boussellaa, Zahour e Alimi (2008) apresentaram uma metodologia de segmentação de imagens coloridas para a separação do primeiro plano e do fundo, em documentos manuscritos históricos árabes, usando métodos híbridos. Este método permite a melhoria da qualidade de manuscritos históricos árabes que apresentam fundo irregular e contraste baixo (provenientes do modo tradicional de produção) e os efeitos do envelhecimento e degradação. O espaço de cor escolhido foi o YIQ (Y canal de luminância, I e Q canais de crominância da cor), sendo justificada a escolha do espaço de cor pelo fato de que a visão humana é muito sensível à variação de luminosidade. Além disso, a variação na intensidade da luz, causada pelo fundo desigual dos manuscritos históricos, é capturada pelo canal Y. O algoritmo trabalha com as seguintes etapas: 1) extração do primeiro plano (texto) por meio do algoritmo de normalização de intensidade de luz; 2) ajuste do contraste por meio da correção gamma e da normalização do histograma; 3) segmentação do texto e do fundo utilizando o classificador K-means; 4) reconstrução da imagem (documento) por meio da superposição de primeiro plano e a média de fundo no espaço de cores RGB com suavização. As técnicas foram testadas em um conjunto de documentos históricos

manuscritos árabes provenientes da Biblioteca Nacional da Tunísia. As imagens foram digitalizadas em 300 dpi e salvas como TIF sem compressão. A Figura 4.17 mostra os resultados alcançados da segmentação, sendo a imagem (a) de entrada, (b) a segmentação do primeiro plano, a (c) o fundo e (d) a imagem restaurada. Não foram divulgadas a quantidade de imagens testadas nem as taxas de acerto.

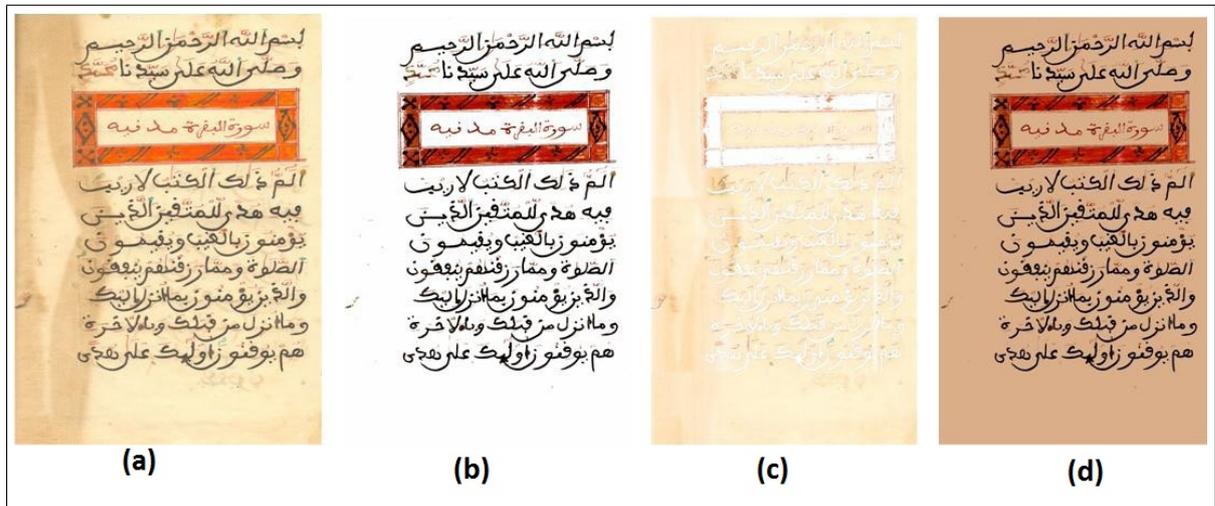


Figura 4.17: Exemplo de um manuscrito com o fundo irregular e o método de segmentação aplicado à imagem (BOUSSELLAA; ZAHOUR; ALIMI, 2008).

#### 4.2.5 Segmentação de conteúdos de documentos antigos baseada em regiões: texto impresso de máquina, manuscritos e desenhos

Baird (2007) apresenta um método de segmentação de conteúdos de documentos antigos baseado em regiões, como: texto impresso de máquina (MP), manuscritos (HW), fotografias (PH), linha arte (LA), notação matemática (MT), mapas (MA), desenhos de engenharia (ED), desenhos de química (CD), ruídos (JK) e espaços em branco (BL). Os tipos de documentos aceitos podem ser em níveis de cinza, coloridos ou preto e branco. O espaço de cor utilizado é o HSL. As regiões da imagem foram zoneadas (delimitadas) manualmente usando-se retângulos para determinar as áreas das regiões de texto impresso, manuscrito, fotografias.

Para o projeto experimental foram utilizados: 1) os algoritmos de treinamentos automáticos (Força Bruta KNN, aproximação KNN utilizando árvores Hash k-d, árvore de classificação e regressão, localização hash sensível); 2) dois conjuntos de imagens, A com 28 imagens e B com 117. Em A, 13 imagens foram colocadas no conjunto de treinamento e o restante no conjunto de teste. Em B, 31 imagens foram colocadas no conjunto de treinamento e o restante no conjunto de teste. Os documentos antigos testados incluem

caracteres em inglês, árabe e chinês. A seleção dos testes e as páginas foram escolhidas aleatoriamente. Os dois conjuntos contêm: texto impresso à máquina, textos manuscritos, desenhos (fotografias) e branco (espaços em branco).

Para avaliar a eficiência foram utilizados os três caminhos descritos: 1) precisão por pixel; 2) análise de precisão por página e 3) qualidade da segmentação subjetiva.

Para o conjunto de testes (B), a pontuação média da precisão por pixel foi de 62,4%. A classificação foi melhor para o reconhecimento de MP e HP, apresentou algumas dificuldades com o BL, e vários problemas com HW, erroneamente 43% dos pixels HW classificados como BL. Esses problemas de classificação podem estar ligados à metodologia de zoneamento das áreas dos conteúdos das imagens.

Para a análise de precisão por página, por exemplo, a questão "Encontrar todas as imagens que contenham pelo menos 30% de impressão à máquina" pode ser respondida com a revocação de 88% e 94% de precisão. As taxas de precisão por página parecem ser menos sensíveis às arbitrariedades da metodologia de zoneamento do que as taxas de precisão por pixel.

As taxas de acertos, utilizando a matriz de confusão, para o conjunto B, contendo 178 milhões de pixels, estão descritas na Tabela 4.2, pela estratégia de classificação por precisão por pixels.

Tabela 4.2: Resultado da classificação.

<b>Categorias</b>	<b>Classificados</b>	
	<b>Quantidade Pixels</b>	<b>Classificação</b>
BL	40 milhões	22,48 %
HW	12 milhões	06,74 %
MP	83 milhões	46,62 %
PH	43 milhões	24,16 %
Total	178 milhões	100 %

A Figura 4.18 ilustra duas imagens classificadas por Baird (2007). A classificação da imagem (a) atingiu 86,72% sendo ilustrada pela imagem (b), a (c) atingiu 70,40% dos pixels, sendo representada pela imagem (d).

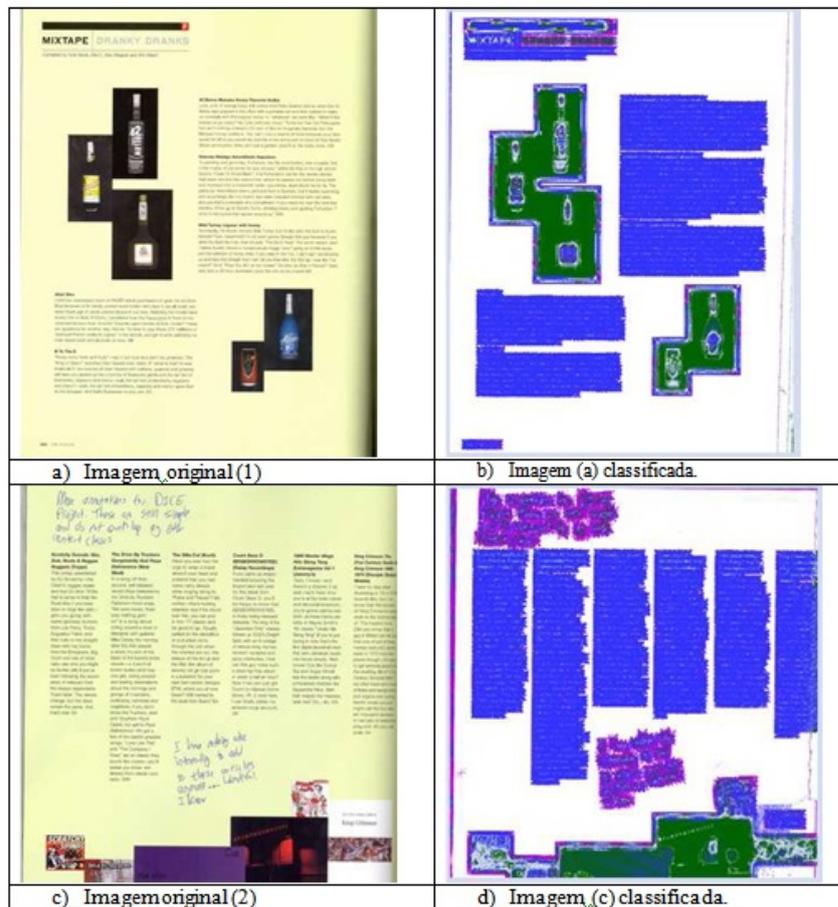


Figura 4.18: Resultado da classificação de duas imagens. Fonte: (BAIRD, 2007).

#### 4.2.6 Segmentação de conteúdo de documentos antigos de diversas coleções

Os autores Baird e Moll (2008) apresentam um método de segmentação de partes de livros antigos, realizando a extração das regiões contendo: texto manuscrito, texto impresso a máquina, fotografias e espaços em brancos. Os documentos antigos são em níveis de cinza e coloridos. A imagem é dividida em 4 classes (regiões), onde cada tipo de conteúdo é zoneado manualmente através de retângulos isométricos, criando assim as imagens de referência (*groundtruth*). A Figura 4.19 representa a imagem de entrada, com a imagem *groundtruth* e a imagem de saída, sendo que alguns conteúdos manuscritos muitas vezes não podem ser descritos por zonas retangulares. Foram estudados os algoritmos de treinamentos automáticos (Força Bruta KNN, aproximação KNN utilizando árvores Hash k-d, árvore de classificação e regressão, localização hash sensível). Para a validação dos testes, dois caminhos foram apresentados: precisão por pixel e precisão por página.

A Figura 4.19 apresenta um exemplo do conjunto de testes. A imagem (a) ilustra uma página colorida; (b) a imagem *ground truth* gerada com retângulos desenhados, por um "Zoner"<sup>1</sup>, de forma manual; a (c) é a imagem de saída da classificação, conforme Baird e Moll (2008).

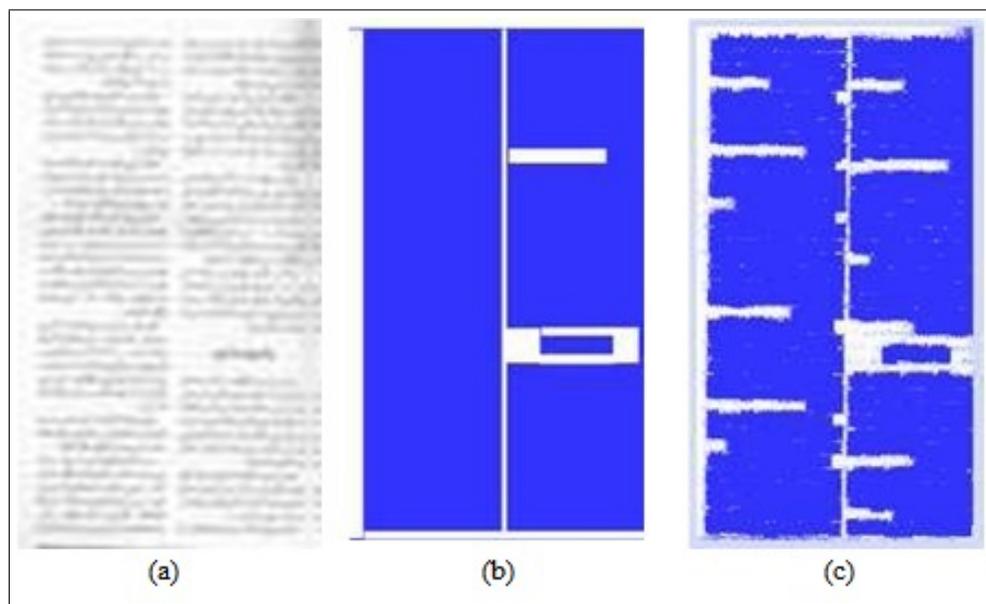


Figura 4.19: Um exemplo de uma imagem do documento do conjunto de testes (BAIRD; MOLL, 2008).

Os experimentos envolveram um conjunto (A), contendo 230 imagens, dobrando o tamanho das experiências anteriores, e um conjunto para o desenvolvimento (B), contendo quatro imagens, que foram usadas para testar a capacidade de classificação em discriminar um novo tipo de conteúdo previamente testado. Para o conjunto de dados (A), foram utilizadas 80 imagens para o conjunto de treinamento, totalizando mais de 416 milhões de pixels analisados, e 150 imagens foram utilizadas no conjunto de teste. As imagens incluem texto em inglês, árabe e chinês, representando imagens em preto e branco, níveis de cinza e colorido. A seleção de páginas de teste e de treinamento foi aleatória. Para o conjunto de dados (A), a taxa média de precisão por pixel foi de 73,7 % (para 416 milhões de pixels na formação de todo o conjunto). Os autores relataram (não divulgaram valores) que nos experimentos a taxa de classificação diminuiu quando da adição de grandes quantidades de texto impresso árabe, diminuindo consideravelmente a habilidade dos classificadores para identificar a caligrafia árabe.

<sup>1</sup>"Zoner" é um profissional da área de computação com conhecimento em processamento de imagens

#### 4.2.7 Separação do fundo e do primeiro plano de documentos coloridos de baixa qualidade degradados com o tempo

Os autores Garain, Paquet e Heutte (2005) apresentam um método de separação do fundo e do primeiro plano de imagens coloridas de baixa qualidade de documentos antigos, que sofreram degradações ligadas à iluminação irregular, ruídos e efeitos de envelhecimento. A abordagem foi automática e não exigiu interação manual, sendo aplicável a textos manuscritos e impressos. O método é bastante adaptável para capturar iluminação e fundos não uniformes, sendo possível realizar testes com imagens coloridas de baixa qualidade. O método possui as seguintes fases: 1) pré-processamento da imagem de entrada para suavizá-la com o filtro da mediana, redefinindo os pixels utilizando uma janela 3x3; 2) rotulação dos componentes conectados, capturando informações de cor e formas espaciais, considerando no mínimo 8 pixels para ser um conjunto conectado, trabalhando no espaço de cor HSV; 3) identificação dos componentes que representam o fundo mostram a sua não uniformidade. Os componentes de primeiro plano são considerados menores que os do fundo, mas alguns destes, que são menores, não são identificados; 4) divisão da imagem em diversas grades após a identificação do primeiro plano.

Os experimentos foram realizados em um conjunto de teste com 50 imagens de documentos antigos. A maioria das imagens foi digitalizada a 300 dpi (algumas estão em 150 dpi) e com tamanho médio de 3M pixels. O conjunto de teste para a parte I foi composto por 20 imagens retiradas de livros antigos e documentos históricos da Livraria Digital "DjVu". A seleção destes documentos permitiu a comparação dos resultados com os resultados da técnica de DjVu (BOTTOU; HAFFNER; HOWARD, 1998). Muitos documentos deste grupo possuem mais de 200 anos. A parte II foi formada por um conjunto de 30 documentos coloridos, digitalizados a partir de cadernos de trabalhos de vários escritores famosos dos séculos XIX e XX. A Figura 4.20 mostra dois exemplos de segmentação, sendo:

- As imagens (a) e (d) são de manuscritos de entrada;
- As imagens (b) e (e) são os resultados da segmentação pela técnica DjVu;
- As imagens (c) e (f) são os resultados da segmentação pelo método de Garain, Paquet e Heutte (2005).

A avaliação de precisão de extração do primeiro plano foi realizada em dois níveis diferentes: linhas e palavras. O resultado de avaliação foi feito manualmente, calculando-se o número de linhas e palavras da imagem, comparando-as com os extraídos da imagem

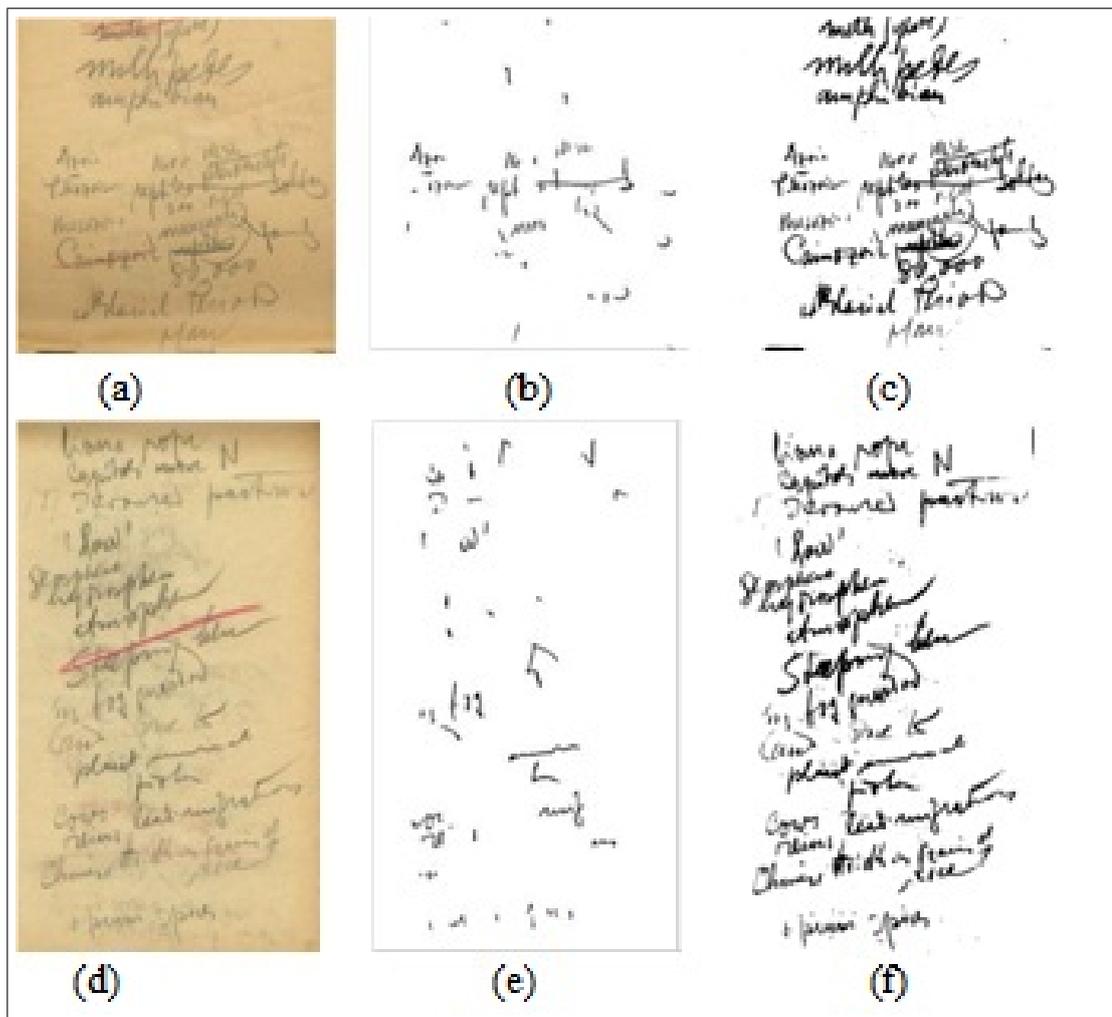


Figura 4.20: Dois Exemplos da segmentação de Manuscritos históricos (GARAIN; PAQUET; HEUTTE, 2005).

pelo método. Para a obtenção dos resultados foram utilizados 50 documentos antigos e testados para as técnicas de DjVu (BOTTOU; HAFFNER; HOWARD, 1998) e o método. Os documentos possuem 917 linhas, das quais o método de DjVu extraiu 541 linhas corretamente e o método proposto extraiu 761 linhas. Para as palavras, o conjunto de teste apresentou 59.651 palavras, sendo extraídas corretamente pela técnica de DjVu 42.934, e o método proposto 54.264.

#### 4.2.8 Avaliação de desempenho de seis algoritmos de segmentação

Shafait, Keysers e Breuel (2007) apresentaram a avaliação de desempenho e comparação de seis algoritmos de segmentação de páginas (corte xy, manchas, espaço em branco, procura de linha de texto, *docstrum*, e *Voronoi*). A medida de avaliação do de-

sempenho foi baseada na teoria dos conjuntos, sendo o bloco de texto dividido em linhas de texto utilizando a projeção horizontal. As páginas podem ser representadas por layouts classificados em Manhattan e não Manhattan. Layouts de Manhattan são definidos como disposições que podem ser decompostas em segmentos individuais por cortes verticais e horizontais e suas zonas individuais podem ser representadas por retângulos que não se sobreponham. Para layouts não Manhattan, as zonas não podem ser representadas exatamente por retângulos, mas por polígonos. Um problema comum com essas abordagens é que se precisa de um software especializado para visualizar os arquivos que representam as segmentações da página, limitando, assim, a sua portabilidade e facilidade de uso. Para contornar esse problema, foi proposta uma nova maneira de representar os segmentos das páginas em formato de imagens em cores, decompondo a imagem em segmentos homogêneos. As imagens *ground truth* são rotuladas por polígonos arbitrários, sendo a imagem dividida em diversos polígonos por meio de editores de imagens, como: GIMP ou MS-Paint. Os experimentos foram divididos em duas partes: aferição dos algoritmos com base na medida das linha de texto de precisão dos blocos segmentados e avaliação de desempenho dos algoritmos com base na pontuação vetorial. A avaliação do desempenho dos algoritmos de segmentação de páginas foi feita em um subconjunto da base de dados da Universidade de Washington (UW-III).

A base de dados é composta de:

- 1600 imagens de documentos em Inglês com layouts Manhattan, digitalizados de arquivos de jornais e revistas;
- Para os testes foram escolhidas 978 imagens;
- Somente as regiões de texto foram avaliadas, as regiões não-texto foram ignorados;
- O conjunto de dados foi dividido em 100 imagens de treinamento e 878 imagens de teste;
- Os documentos foram separados de acordo com o número de colunas, sendo 362, 449 e 67 documentos de uma, duas e três colunas, respectivamente, do conjunto de teste de 878 images.

A Tabela 4.3 ilustra os resultados para os seis algoritmos em termos de porcentagem de detecção de erro de linhas de texto.

Observou-se que os algoritmos de manchas, espaços em branco e linha de texto tiveram pior desempenho em documentos de uma coluna, em média. Este comportamento pode ser explicado pelo efeito mais forte dos blocos de ruídos que ocorrem em fotocópias de

Tabela 4.3: Resultado da classificação.

Algoritmos	Média	
	Treinamento	Teste
Corte X-Y	14,70 %	17,10 %
Manchas	13,40 %	14,20 %
Espaços em Branco	9,10 %	9,80 %
Linha Texto	8,90 %	8,50 %
Docstrum	4,30 %	6,00 %
Voronoi	4,70 %	5,50 %

imagens. O algoritmo *docstrum* não ganha precisão para documentos limpos, enquanto o algoritmo de *Voronoi* tem o melhor desempenho. Com base nos resultados experimentais e as observações, as recomendações a seguir podem ser feitas sobre a escolha do algoritmo de segmentação de página para diferentes aplicações: O algoritmo de corte XY falha na presença de ruído e tende a tomar toda a página como um único segmento, resultando em muitos erros de segmentação. Nos algoritmos de *Voronoi* e *docstrum*, ocorre a variação dos espaçamentos entre linhas, devido a diferentes tamanhos de fonte e estilos, provocando erros de segmentação. O algoritmo de manchas classifica as manchas dos blocos das linhas como sendo não texto, considerados ruídos, resultando em um grande número de erros de segmentação. Em geral os algoritmos de corte XY, *docstrum* e *Voronoi* levaram menos da metade do tempo em relação aos outros algoritmos. As recomendações abaixo devem ser consideradas no final dos resultados.

- Para documentos limpos, com pouca ou nenhuma inclinação, pode ser utilizado o algoritmo de corte XY, pois é rápido e fácil de implementar.
- Para um conjunto homogêneo de documentos, com uma quantidade variável de ruídos, os algoritmos *docstrum* e *Voronoi* são indicados.
- Para documentos que contenham tamanhos de fonte e estilos diferentes, o algoritmo procura linha de texto funciona melhor porque se baseia em modelos geométricos que são invariantes de tamanho, estilo da fonte e da resolução de digitalização.
- Para uma coleção de documentos diversos, com diferentes tamanhos de fonte e layouts, ou documentos digitalizados em diferentes resoluções, o algoritmo que procura linha de texto é uma boa escolha.
- Para os layouts não-Manhattan, texto em diferentes orientações, o algoritmo *Voronoi* pode ser uma boa escolha.

- Um problema com a avaliação dos algoritmos de segmentação é que eles dão uma segmentação única da página sem qualquer valor de confiança.

Para uma coleção de documentos homogêneos, com uma grande proporção de layouts de Manhattan, algoritmos *docstrum* e de *Voronoi* são a melhor escolha. No caso de uma coleção de documentos heterogêneos, com diferentes tamanhos de fonte, estilos e resoluções diferentes, o algoritmo que procura linhas de texto parece ser a melhor escolha. A Figura 4.21 ilustra um exemplo de imagem com a segmentação dos seis algoritmos analisados.

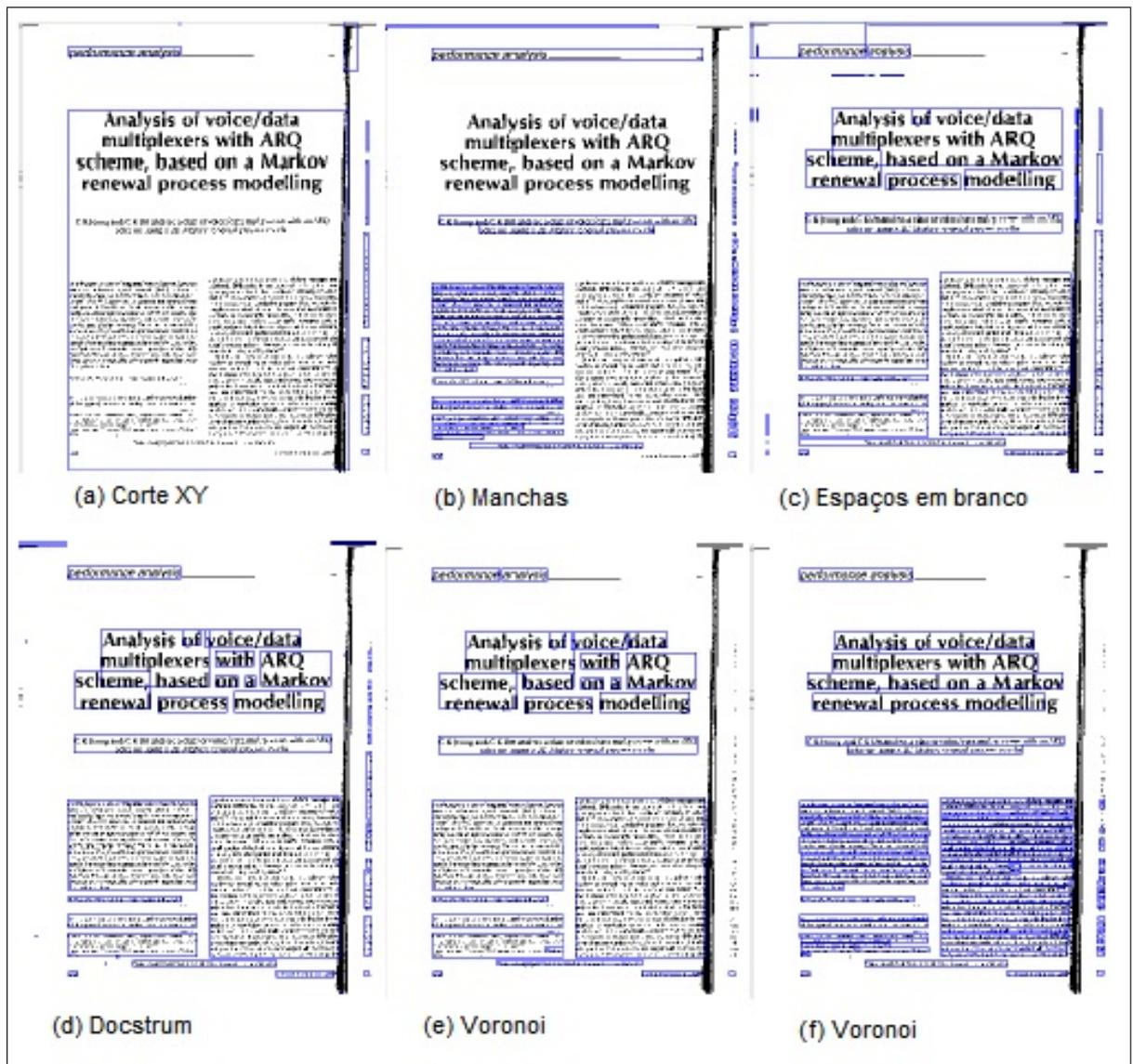


Figura 4.21: Exemplo da segmentação dos seis algoritmos analisados. (SHAFAIT; KEYSERS; BREUEL, 2007)

#### 4.2.9 Segmentação das letras contidas dentro de capitulares

Um método para extrair a letra que está dentro de um capítular (composta pela letra e pelo fundo) é apresentado por Coustaty (2010). Capitulares são imagens decorativas que inicializam um capítulo de documento antigo. As imagens dos capitulares são muito complexas e muito ricas em termos de informações e necessitam ser simplificadas. A idéia é capturar componentes puramente geométricos em uma imagem, independentemente da textura e do ruído, para extrair as formas. Após a segmentação destas imagens, é possível separar a letra e o fundo. As letras capitulares pertencem às imagens como índices. Estas imagens são compostas por dois elementos principais: a letra e o fundo. Um passo importante no processo de reconhecimento dos capitulares consiste em segmentar a letra e os elementos do plano de fundo para caracterizá-los usando uma assinatura a qual permite uma comparação simples e rápida para o processo de indexação de grandes massas de dados. As fases do método são:

1. Simplificação das imagens usando camadas;
2. Extração das formas de uma dessas camadas;
3. Seleção destas formas.

Assim, a imagem é decomposta nas três camadas descritas abaixo:

1. A camada regularizada corresponde à área da imagem que tem baixa oscilação do nível de cinza. Esta camada permite destacar a geometria, o que corresponde às formas da imagem.
2. Camada dos elementos oscilantes de uma imagem, destacando a textura dos capitulares.
3. Camada de alta oscilação, que corresponde ao ruído na imagem, em que se recupera todos os pixels que não pertencem às duas primeiras camadas.

Portanto, na camada de ruído encontram-se o texto de fundo e o envelhecimento do papel. A Figura 4.22 exemplifica a decomposição de uma imagem de capitulares em três camadas, sendo a imagem (b) a de formas, a (c) a de textura e a (d) a de ruídos.

A camada regularizada, obtida por decomposição, contém todas as formas das letras. Para extraí-las e selecionar as mais interessantes, é usada a Lei Zipf, definida por George Kingsley Zipf, que é baseada na frequência e no grau de aparecimento das letras em um texto (PARETI; VINCENT, 2006). Este método consiste em um processo de três etapas, que serão descritas a seguir.

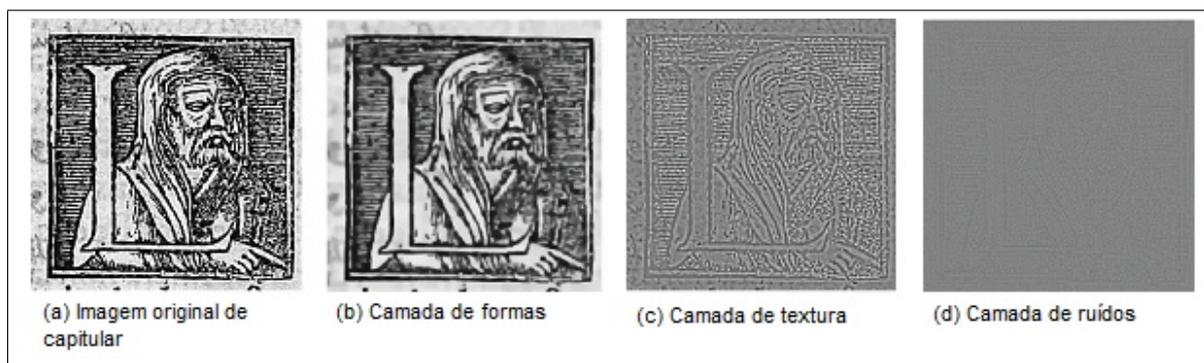


Figura 4.22: Exemplo da decomposição de uma imagem capitular em três camadas (COUSTATY, 2010).

1. Simplificação da imagem através da diminuição de seus níveis de cinza, reduzindo os padrões.
2. Busca de padrões de tamanho três por três, para obter suas frequências e suas classificações, uma vez que o número de níveis de cinza foi reduzido, uma simples contagem de cada padrão permite conhecer a sua frequência e sua classificação.
3. Classificação dos padrões em três classes de acordo com a lei da evolução da frequência em relação à sua categoria. A extração de formas é caracterizada pela procura dos componentes ligados da imagem binarizada.

Quando todos os componentes estão ligados, conforme mostra a imagem(a) da Figura 4.23, pode-se ver que as formas mais importantes têm características específicas (com base no tamanho, localização, centro de massa e excentricidade). Uma seleção dos componentes ligados de acordo com estes parâmetros permite a obtenção da região de interesse de capitulares. Um exemplo de componentes ligados pode ser visto na imagem (b) da mesma Figura. Com uma seleção precisa destes parâmetros, o componente ligado mais importante para o historiador pode ser extraído: a carta. Este pode ser obtido selecionando-se o maior componente ligado no qual o centro de massa está centrado e que não toca as fronteiras da imagem. Um exemplo da extração da letra em um capitular é apresentado pela Figura 4.23.

Para a avaliação, utilizaram-se dois sistemas comerciais OCR (FineReader e o Tesseract), a fim de implementá-la. Para os experimentos foi utilizada uma base de dados contendo 4.500 imagens. Para o conjunto de treinamento foram selecionadas 1.500 imagens, enquanto que 3.000 foram consideradas para os testes. Os resultados foram: a taxa de classificação utilizando o OCR FineReader foi de 72,8 % e para o OCR Tesseract

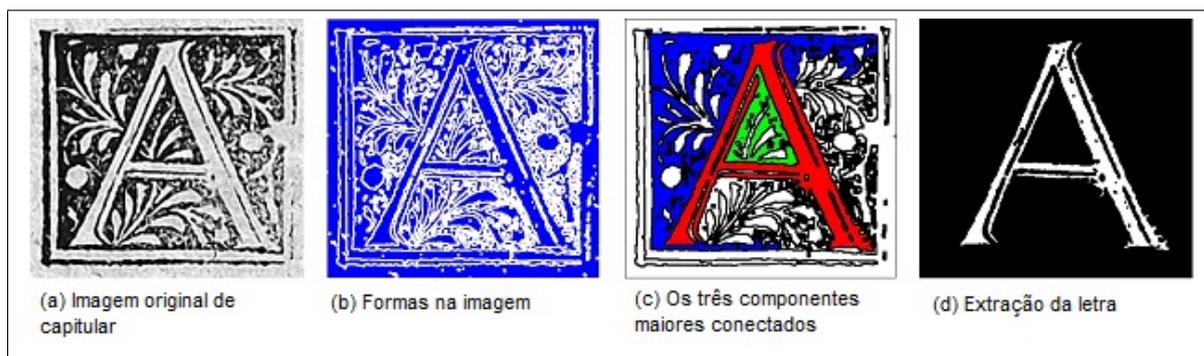


Figura 4.23: Exemplo da extração da letra de um capitular (COUSTATY, 2010).

foi de 67,9 % . Os autores relataram que os resultados foram insatisfatórios, mas muito animadores.

#### 4.2.10 Detecção de textura baseada em ferramentas de caracterização de livros antigos

Journet (2006) apresentaram um método de classificação de pixel de documentos antigos datilografados por meio da construção de mapas de multiresoluções com base em informações de textura (de orientações dos pixels, densidades do seu nível de cinza). O objetivo do sistema é realizar uma indexação robusta, que seja adaptada para grandes coleções heterogêneas de livros antigos, realizando a segmentação de forma automática das imagens de um livro, permitindo que os usuários extraiam os dados desejados usando a interface de representação. A recuperação da informação deve ser realizada sem qualquer passo de binarização, que pode causar problemas no contexto de documentos antigos. O método foi desenvolvido em duas fases:

1. Classificação automática não supervisionada, utilizando apenas as características de textura dos pixels. Esta classificação devia levar em conta os diferentes estilos de livros, diferentes tipos da qualidade de digitalização, o formato de resoluções diferentes;
2. O procedimento interativo de indexação e de semântica realizando a rotulagem das áreas homogêneas.

O trabalho apresenta principalmente a primeira etapa do processo automático. O procedimento interativo é reduzido a uma operação simples: o usuário define o número de classes de texturas que ele quer ver no início da etapa de indexação. A abordagem reside na construção de mapas de multiresoluções e análises, sendo construídos 5 mapas de

resoluções para 5 diferentes mapas de características, os quais expressam características presentes nos documentos antigos. Todos os mapas são correlacionados com base em informações de textura. Para uma imagem, podemos formar 5 mapas em 5 resoluções diferentes, formando 25 mapas que são analisados para alimentar um algoritmo de classificação. A Figura 4.24 mostra as etapas do processo, sendo a primeira a entrada com as imagens de livros antigos; a segunda a criação dos mapas de multiresoluções e a terceira é o processo interativo na qual o usuário informa a quantidade de texturas desejadas, assim é realizada a classificação dos pixels com base no mapa de dados e o resultado visual é apresentado ao usuário, por exemplo, a imagem (b) da Figura 4.24 apresenta a classificação em 3 classes e a imagem (c) em 6 classes. O problema é que os elementos principais de uma imagem (o fundo, o texto e a parte gráfica) não são definidos claramente, ficando mais difícil sua identificação. É por isso que são criadas as texturas primitivas, que levam em conta todos esses aspectos.

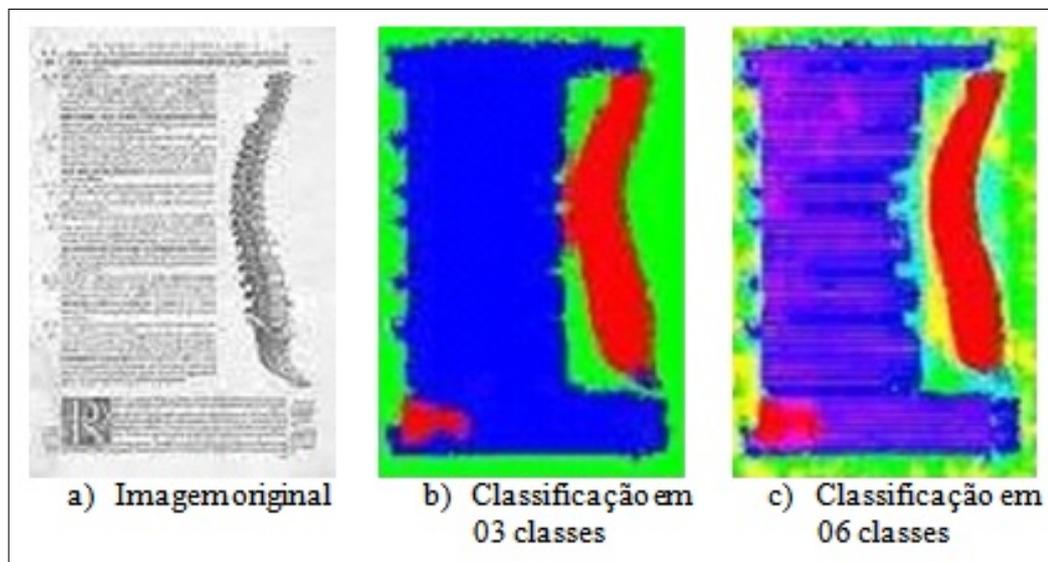


Figura 4.24: Etapas do processo de detecção de textura em classes (JOURNET, 2006).

Os resultados da classificação foram validados em três grupos usando somente características de orientação. Esses resultados têm permitido a classificação de layouts de páginas de acordo com a distribuição e quantidade de informações relacionadas a fundo, texto e gráficos. A originalidade da abordagem reside no desenvolvimento de um extrator novo e uma ferramenta de análise que é dedicada a imagens de documentos antigos sem utilizar parâmetros, limites, modelos, ou informações de estrutura. O resultado de uma análise do documento é empregado como características relevantes para a classificação de um pixel. Os autores relatam que os resultados atuais são bastante promissores, mas necessitando ainda um pouco mais de investigação para alcançar um sistema de indexação

completo para usuários não especializados.

#### 4.2.11 Segmentação de texto em imagens coloridas

Hase (2004) descreve um método de segmentação de caracteres para imagens com baixa resolução e de redução de ruídos para imagens com alta resolução, assumindo-se que o caracter é impresso e visualizado com uma única cor. O importante neste estudo é encontrar as cores dominantes em documentos coloridos, ajudando na segmentação de região ou extração de texto. O espaço de cor utilizado é o  $L^*a^*b$ , no qual  $L$  representa a intensidade e a cor é representado pela combinação de "a" e "b". Primeiramente é avaliada a uniformidade de cor em uma área de caracter, obtida a partir das cores representativas da imagem. A segmentação no espaço de cor  $L^*a^*b$  é realizada em documentos em cores, como livros, capas e cartazes. Primeiro é selecionada a área que pode ser vista como uma única cor pela inspeção do ser humano. Em seguida, procede-se à avaliação da segmentação da área de caracter pelo algoritmo K-means para o segmento  $L^*a^*b$ . A Figura 4.25 mostra o resultado de uma imagem de  $327 \times 450$  pixels em (a) e em (b) a imagem do modelo com a área "S" selecionada, ( que é a área dos caracteres), a imagem (c) representa a taxa de uniformidade de linhas sólidas e (d) a diferença mínima de cor. O modelo é uma imagem binária na qual uma área preta foi com antecedência selecionada manualmente.

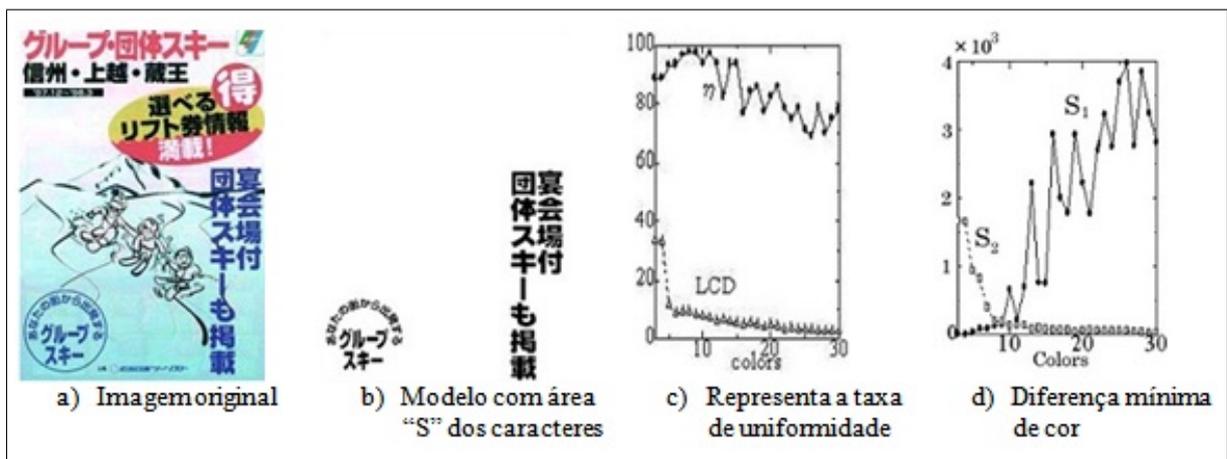


Figura 4.25: Exemplo da segmentação da área de caracter pelo algoritmo K-means (HASE, 2004).

As imagens da Figura 4.26 mostram os estados da extração das áreas de caracteres depois da segmentação da cor. A imagem (a) representa o resultado do método em 5 cores, (b) com 10 cores e (c) com 30 cores. Nota-se que as figuras não são o resultado da extração de texto, são apenas uma área de cor que domina a área de caráter mostrado nas figuras. A área de cor dominante depende do número de cores.



Figura 4.26: Exemplo da extração das áreas de caracteres com 05, 10 e 30 cores (HASE, 2004).

Nos experimentos foram utilizadas 20 imagens e 31 modelos, sendo que cada modelo corresponde a um conjunto de caracteres escolhido manualmente. Para os testes, os agrupamentos de cor foram divididos em nove grupos  $n(9)$ . Foram realizados testes para o método proposto e para o algoritmo K-emans. Para o método proposto a medida de segmentação foi de 91,73 % e para o algoritmo K-means foi de 91,24 % . Nota-se que houve pouca diferença entre os métodos, no entanto a quantidade de tempo de processamento foi de aproximadamente 01 segundo para o método proposto e poucos minutos para o K-means. A Figura 4.27 mostra o resultado da comparação entre o método proposto (b), que apresentou uma taxa de segmentação da uniformidade de cor de 97,92% e o desempenho do algoritmo de popularidade (PA) (HECKBERT, 1982), com 64,71% (c), em (a) tem-se a a imagem original.

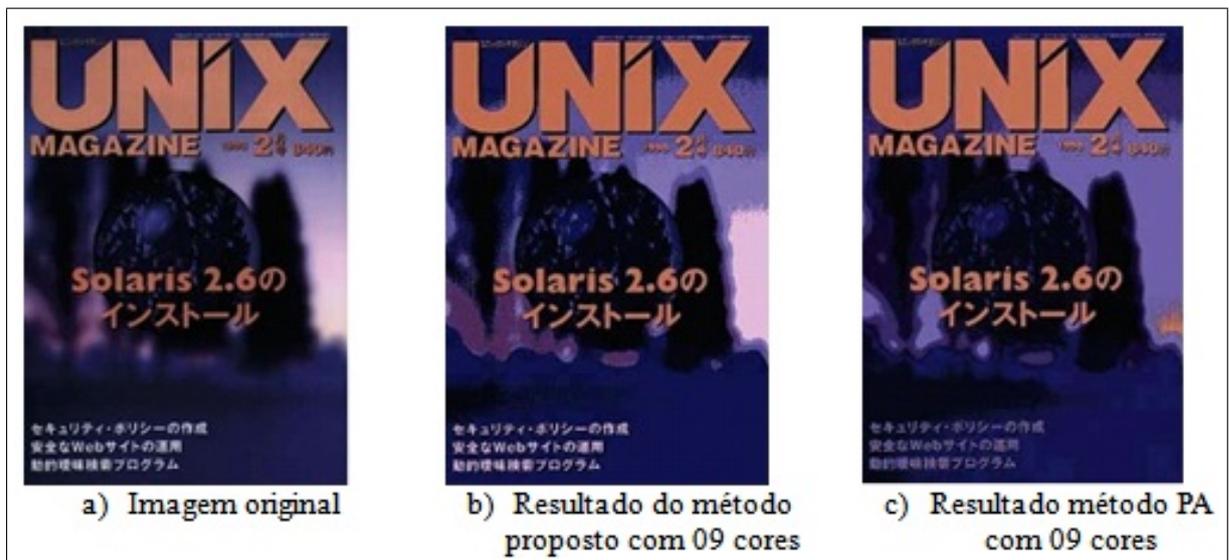


Figura 4.27: Resultado da comparação entre os algoritmos de popularidade e o método proposto (HASE, 2004).

#### 4.2.12 Análise Reflexiva dos trabalhos de segmentação de documentos antigos

Os documentos antigos podem ser segmentados de várias formas, que serão descritas nesta seção, mostrando como são processadas as imagens e quais as técnicas utilizadas. As considerações a seguir são necessárias.

- Existem abordagens que usam imagens binárias para realizar a segmentação, algumas utilizam imagens em níveis de cinza e outras as coloridas.
- Para segmentação do primeiro plano e do fundo, o trabalho de Boussellaa, Zahour e Alimi (2008) está baseado no histograma do canal intensidade do espaço de cor YIQ, com auxílio do classificador K-means e da estimação do fundo homogêneo definido pela média dos pixels do fundo no espaço RGB. O problema nesta abordagem é que a estimação da cor do fundo com base no espaço RGB leva a uma estimação de fundo diferente da cor de fundo da imagem original, isto devido à média de cada canal RGB.
- A abordagem de Granado, Pina e Muge (2001) realiza a segmentação do conteúdo de documentos antigos com base na morfologia matemática binária. Esta abordagem é somente para imagens binárias, realizando a separação do conteúdo dos documentos. Os autores não realizaram a estimação do fundo e não relataram as taxas de acertos do método proposto.

- A abordagem de Baird (2007) utiliza algoritmos de treinamentos, com imagens *ground truth* para segmentação de documentos com base em regiões, não utilizando morfologia matemática em cor.
- O trabalho de Journet, Ramel e Mullet (2008) apresenta a análise de textura das páginas por meio da detecção das áreas com características específicas de orientações e frequências.

Conclui-se que existem vários trabalhos relacionados com a segmentação de documentos antigos, mas nenhum para segmentação de imagens da Bíblia de Gutenberg baseado na morfologia matemática em cor.

### 4.3 Aplicações baseadas na morfologia matemática em cor

Nesta seção estão descritos os trabalhos relacionados com a morfologia matemática para imagens coloridas, com o uso de vários espaços de cor e suas ordenações. O espaço de cor é muito importante, uma vez que imagens coloridas podem ser representadas por vários modelos de cor. Para a definição dos operadores mínimo e máximo, é preciso determinar a ordenação a ser utilizada para os três canais de um espaço de cor, definindo as operações básicas erosão e dilatação em cor. A seguir serão descritos os trabalhos relacionados com as ordenações e os espaços utilizados até o presente momento.

#### 4.3.1 Morfologia Matemática FUZZY em processamento de imagens coloridas

Uma nova abordagem para a construção de filtros morfológicos para imagens coloridas baseados em morfologia matemática FUZZY é apresentada por Popov (2007). Essa metodologia foi desenvolvida com o intuito de deixar os operadores menos sensíveis à imprecisão da imagem, e pode ser vista simplesmente como uma alternativa à morfologia matemática para escala de cinza. Um valor fuzzy é uma probabilidade entre 0 e 1, que representa o grau de pertencimento de um ponto  $x$  em um conjunto qualquer. Sendo  $A$  um conjunto de valores fuzzy, no caso a imagem,  $B$  o elemento estruturante e  $x$  um pixel, as operações de erosão e dilatação fuzzy podem ser definidas, respectivamente, como:

$$\begin{aligned}\varepsilon_B(A)(x) &= \bigwedge_y i(B(y-x), A(y)) \\ \delta_B(A)(x) &= \bigvee_y c(B(x-y), A(y))\end{aligned}\tag{4.1}$$

A utilização de operações morfológicas para imagens coloridas exigem ordenação dos canais em três dimensões, precisando haver uma ordem das cores. O trabalho utiliza o modelo de cor YCrCb por sua simplicidade e eficiência em calcular, e os canais Cr e Cb possuem pesos iguais. A ordenação obedece à prioridade para o canal Y (luminância), seguida dos canais Cr e Cb. Assim, através dos operadores mínimo e máximo, são obtidas as erosões e dilatações fuzzy necessárias para a construção de filtros eficientes para imagens coloridas ruidosas. A equação a seguir define os operações básicas de erosão e dilatação Fuzzy, respectivamente, para imagens coloridas.

$$\begin{aligned}\varepsilon_B(X)(x) &= x^{-1} [\bigwedge_y i(B(y-x), (x(X))(y))] \\ \delta_B(X)(x) &= x^{-1} [\bigvee_y c(B(x-y), (x(X))(y))]\end{aligned}\tag{4.2}$$

A Figura 4.28 ilustra o resultado das operações dilatação e erosão fuzzy.

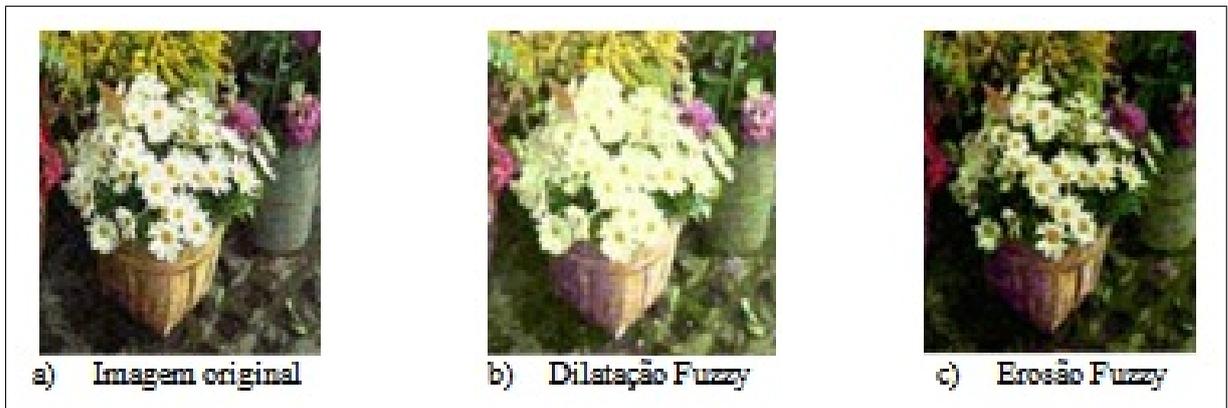


Figura 4.28: Ilustração da dilatação e erosão Fuzzy por Popov (2007).

### 4.3.2 Trabalhos relacionados com a Ordenação Lexicográfica

A fim de se utilizar as operações da morfologia matemática em cor no espaço vetorial, é preciso ordenar os vetores para se evitar a introdução de cores falsas. Isto permite que se escolha um supremo ou ínfimo, dentre os vetores de um conjunto. A ordenação lexicográfica permite ordenar os vetores conforme a ordem das palavras definidas para não gerar cores falsas. A equação 4.3 ilustra a ordenação lexicográfica genérica para ordenar os vetores com três canais de cor.

$$(a_1, b_1, c_1) < (a_2, b_2, c_2) \text{ se } \left\{ \begin{array}{l} a_1 < a_2 \\ \text{ou} \\ a_1 = a_2 \text{ e } b_1 < b_2 \\ \text{ou} \\ a_1 = a_1 \text{ e } b_1 = b_2 \text{ e } c_1 < c_2 \end{array} \right. \quad (4.3)$$

A seguir serão descritos os trabalhos que realizaram esta ordenação e os espaços de cor utilizados.

### 4.3.3 Morfologia Matemática no espaço de cor HSL

Os autores Hanbury e Serra (2001) apresentam os operadores morfológicos para o espaço de cor HSL, que é amplamente utilizado em análise de imagem porque é fisicamente intuitivo. A maioria das dificuldades da aplicação de operadores de análise de imagem neste espaço é devida ao fato de ser o componente matiz definido no círculo unitário, ou seja, não se pode dizer, entre dois valores de matiz, qual é o maior ou o menor valor.

Os operadores morfológicos erosão e dilatação são definidos pela equação 4.4, respectivamente:

$$\begin{aligned} \varepsilon f(x) &= f(y) : f(y) = \inf(f(z)), z \in B_x \\ \delta f(x) &= f(y) : f(y) = \sup(f(z)), z \in B_x \end{aligned} \quad (4.4)$$

onde:

$B_x$  indica o elemento estruturante para o ponto  $x$ .

Para ordenar os vetores, para encontrar o máximo e o mínimo, é utilizada a ordenação lexicográfica com os dois componentes, luminância e saturação, na primeira e na segunda posições para comparação. A prioridade para o canal H, na primeira posição, causa problema e é discutível, sendo sugerida uma solução na forma de uma ordem lexicográfica com matiz-saturação ponderada na primeira posição. O vetor  $c_i$  do HSL, com a ordenação lexicográfica, com prioridade para o canal L, é definido pela equação 4.5.

$$c_i > c_j \text{ se } \begin{cases} L_i > L_j \\ \text{ou} \\ L_i = L_j \text{ e } S_i < S_j \\ \text{ou} \\ L_i = L_j \text{ e } S_i = S_j \text{ e } \frac{H_i}{H_0} < \frac{H_j}{H_0} \end{cases} \quad (4.5)$$

A Figura 4.29 ilustra a erosão e a dilatação com ordenação lexicográfica em ordem LSH.

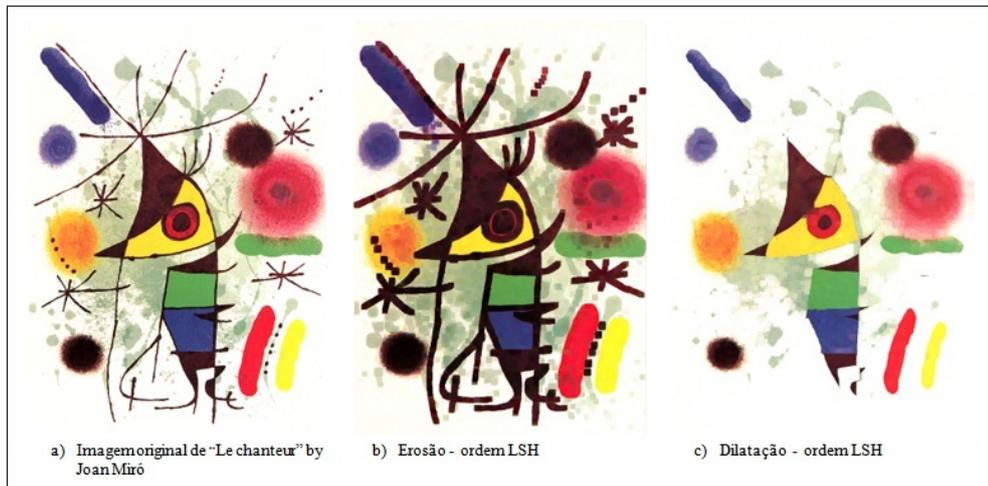


Figura 4.29: Resultado da erosão e dilatação com ordem LSH. (Fonte: (HANBURY; SERRA, 2001))

A equação 4.7 define a prioridade para o canal  $s$ , seguida para o canal  $l$  e em terceiro o  $h$ .

$$c_i > c_j \text{ se } \begin{cases} S_i > S_j \\ \text{ou} \\ S_i = S_j \text{ e } |L_i - 0.5| < |L_j - 0.5| \\ \text{ou} \\ S_i = S_j \text{ e } |L_i - 0.5| = |L_j - 0.5| \text{ e } \frac{H_i}{H_0} < \frac{H_j}{H_0} \end{cases} \quad (4.6)$$

A Figura 4.30 ilustra a erosão e a dilatação com ordenação lexicográfica em ordem SLH.

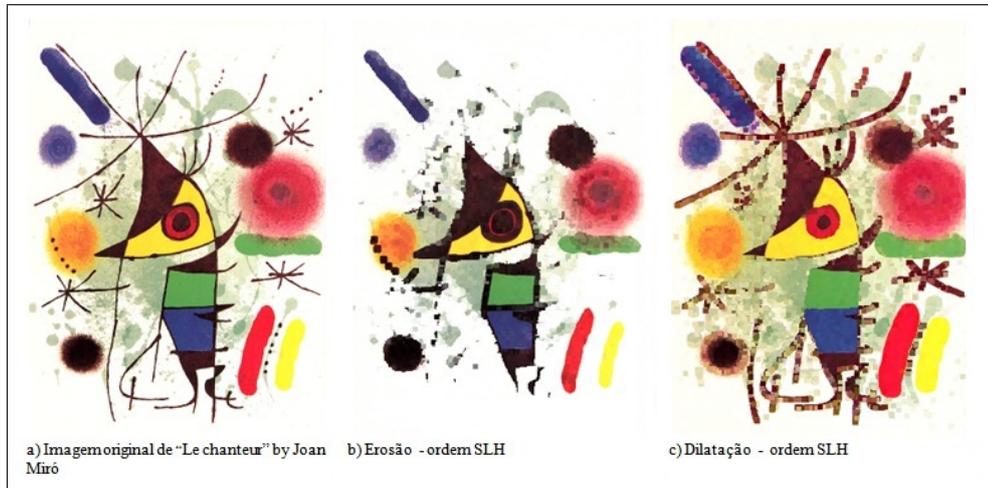


Figura 4.30: Resultado da erosão e dilatação com ordem SLH. (Fonte: (HANBURY; SERRA, 2001))

A ordenação lexicográfica com prioridade para o canal H do espaço HSL é definida conforme a equação 4.7.

$$D_i = D(H_i, H_0) = \begin{cases} |H_i - H_0| & \text{se } |H_i - H_0| \leq 180^\circ \\ 360^\circ - |H_i - H_0| & \text{se } |H_i - H_0| \geq 180^\circ \end{cases} \quad (4.7)$$

na qual  $H_0$  é o valor da matiz de origem escolhido e  $D_i$  é a distância calculada para todos os pontos do elemento estruturante. O ponto com menor distância é considerado o mínimo do elemento estruturante, e a distância maior é considerada o máximo.

Para resolver o problema da escolha somente do valor H, propõe-se uma solução com valor de saturação ponderada, dada pela equação 4.8.

$$H_i = \begin{cases} H_i - H_0 & \text{se } H_i - H_0 \geq 0 \\ 360^\circ + (H_i - H_0) & \text{se } H_i - H_0 < 0 \end{cases} \quad (4.8)$$

Para calcular o  $H'$ , a equação 4.9 .

$$H'_i = \begin{cases} \text{máximo}[H_i, 90^\circ(1 - S_i)] & \text{se } 0^\circ \leq H_i \leq 90^\circ \\ \text{mínimo}[H_i, 90^\circ(1 + S_i)] & \text{se } 90^\circ < H_i \leq 180^\circ \\ \text{máximo}[H_i, 90^\circ(3 - S_i)] & \text{se } 180^\circ < H_i \leq 270^\circ \\ \text{mínimo}[H_i, 90^\circ(3 + S_i)] & \text{se } 270^\circ < H_i < 360^\circ \end{cases} \quad (4.9)$$

A ordenação lexicográfica baseada no matiz ponderado é definida conforme equação 4.10.

$$c_i > c_j \text{ se } \begin{cases} \frac{H_i}{H_0} < \frac{H_j}{H_0} \\ \frac{H_i}{H_0} = \frac{H_j}{H_0} \text{ e } |L_i - 0.5| < |L_j - 0.5| \\ \frac{H_i}{H_0} = \frac{H_j}{H_0} \text{ e } |L_i - 0.5| < |L_j - 0.5| \text{ e } S_i > S_j \end{cases} \quad (4.10)$$

Uma desvantagem da ordenação com matiz ponderado é que, em regiões que não contêm pixels altamente saturados, com tonalidade próxima à origem selecionada, o comportamento é menos previsível. Uma vez que a ordem do vetor foi definido, permitindo a escolha de um máximo e um mínimo de um conjunto de vetores, podem-se definir os operadores de base morfológica, erosão e dilatação. A maioria das dificuldades da aplicação de operadores de análise de imagem no espaço HSL é devida ao componente matiz ser definido no círculo unitário. Agora, é interessante refletir sobre a aplicabilidade dos operadores morfológicos e das ordenações, e qual é a melhor aplicação para cada tarefa. A escolha depende, em grande parte, e das propriedades das imagens a serem processadas, e as informações que o usuário deseja extrair das imagens. Algumas aplicações possíveis são:

- Operadores com base na ordem lexicográfica com luminância na primeira posição são utilizados para preservar os contornos dos objetos na imagem. No entanto, eles só são úteis em situações em que se sabe que os objetos de interesse são mais brilhantes ou mais escuros na imagem;
- Em situações nas quais os objetos de interesse são altamente coloridos sobre um fundo cinza, os operadores com a saturação na primeira posição devem ser escolhidos.

No entanto, a escolha entre um pixel preto ou branco de baixa saturação, feita pelo ope-

rador, pode parecer arbitrária. Esse problema também se manifesta na região de textura com cores de tons diferentes, mas saturações semelhantes. Em uma aplicação em que apenas os objetos de uma cor específica é de interesse, a ordenação com o matiz ponderada é a melhor escolha. As técnicas podem ser aplicadas a outros espaços de cor com valor de matiz angular, por exemplo: HSV, HSI,  $L^*a^*b$ , precisando fazer algumas pequenas modificações para as ordens, levando-se em conta as diferentes formas dos espaços.

A Figura 4.31 ilustra o exemplo da aplicação de erosão e dilatação com H somente e ponderado.

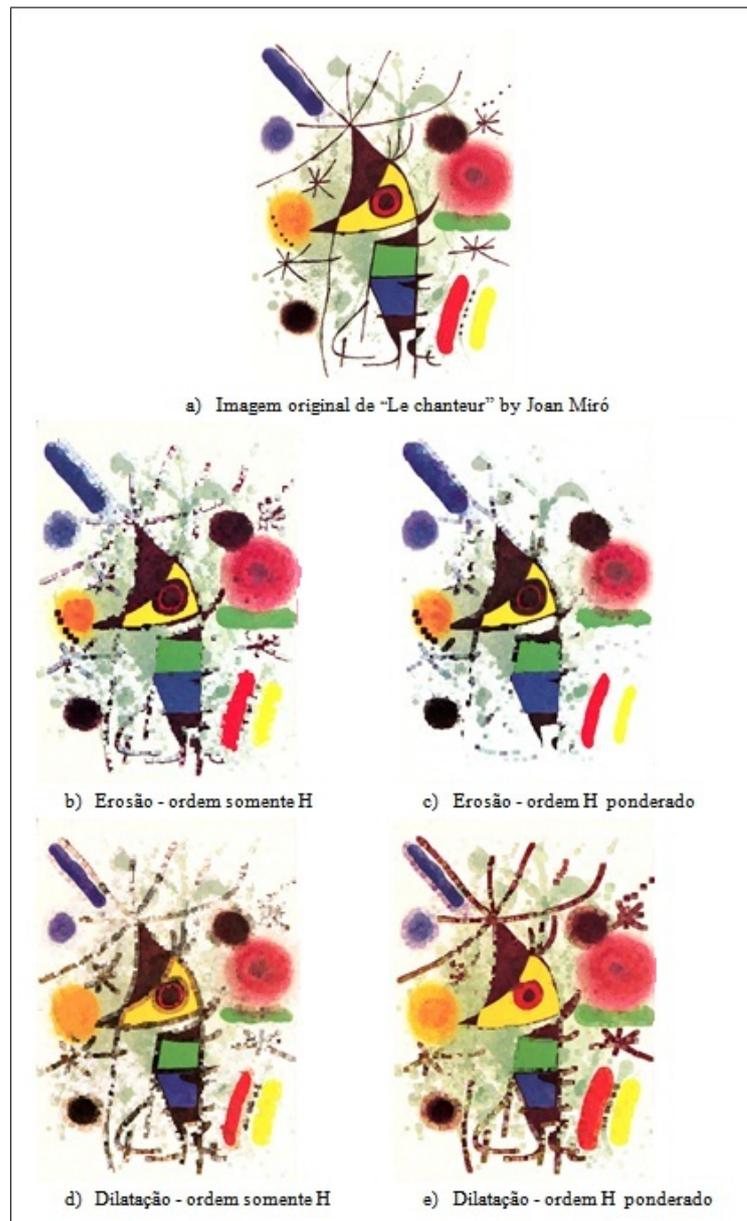


Figura 4.31: Resultado da erosão e dilatação com H somente e ponderado.(Fonte: (HAN-BURY; SERRA, 2001))

#### 4.3.4 Algoritmo para Análise de Imagem de retina

Ortiz (2002) apresenta um algoritmo para análise de imagens de retinas de macacos, com a extensão de conceitos da morfologia matemática (binária e nível de cinza) para imagens coloridas e o espaço de cor utilizado é o HSV (matiz, saturação e brilho). A preferência ou disposição dos componentes de cor depende da aplicação e da propriedade da imagem. O canal do brilho (V) é o melhor caminho para preservar contornos de objetos na imagem. Em situações nas quais se buscam objetos com cores fortes ou objetos específicos de uma cor, o melhor caminho é o matiz (hue) em primeiro lugar. A ordenação dos componentes HSV é baseada na ordenação lexicográfica. Devido à forma específica do modelo de cor HSV, surge um problema com o método de ordenação lexicográfica: a saturação e a intensidade são conjuntos totalmente ordenáveis, mas o matiz não é, além disso, não é possível ordenar cores de valores mais baixos para valores mais altos, uma vez que não faz sentido dizer que o azul é maior ou menor que o vermelho. Para solucionar este problema, Hanbury (2001) e Peteres (1997) ordenaram as cores de acordo com o valor absoluto da função de distância entre a cor da imagem ( $H_i$ ) e uma cor de referência ( $H_{ref}$ ), que é o ângulo da cor que se deseja destacar na imagem, ou seja, se o usuário está procurando objetos vermelhos na imagem, deve-se ajustar o ângulo para  $0^\circ$ . Para isso, é necessário que o usuário conheça o espectro de cores. A ordenação lexicográfica inicia pelo matiz (H), seguido pelo valor do brilho (V) e então pela saturação (S). A distância é calculada pela equação 4.7 (HANBURY; SERRA, 2001) e a equação 4.11 ilustra a ordenação lexicográfica utilizada.

$$P_i < P_j = \left\{ \begin{array}{l} d(H_i, H_{ref}) < d(H_j, H_{ref}) \\ \text{ou} \\ d(H_i, H_{ref}) = d(H_j, H_{ref}) \text{ e } V_i < V_j \\ \text{ou} \\ d(H_i, H_{ref}) = d(H_j, H_{ref}) \text{ e } V_i = V_j \text{ e } S_i < S_j \end{array} \right. \quad (4.11)$$

Com isto, são definidas as operações morfológicas básicas (erosão e dilatação) usando a ordenação lexicográfica HVS. A Figura 4.32 ilustra uma imagem de retina de macaco, sendo (a) imagem original; (b) máscara para objetos vermelhos; (c) máscara para objetos verdes; (d) reconstrução de objetos vermelhos e (e) reconstrução de objetos verdes.

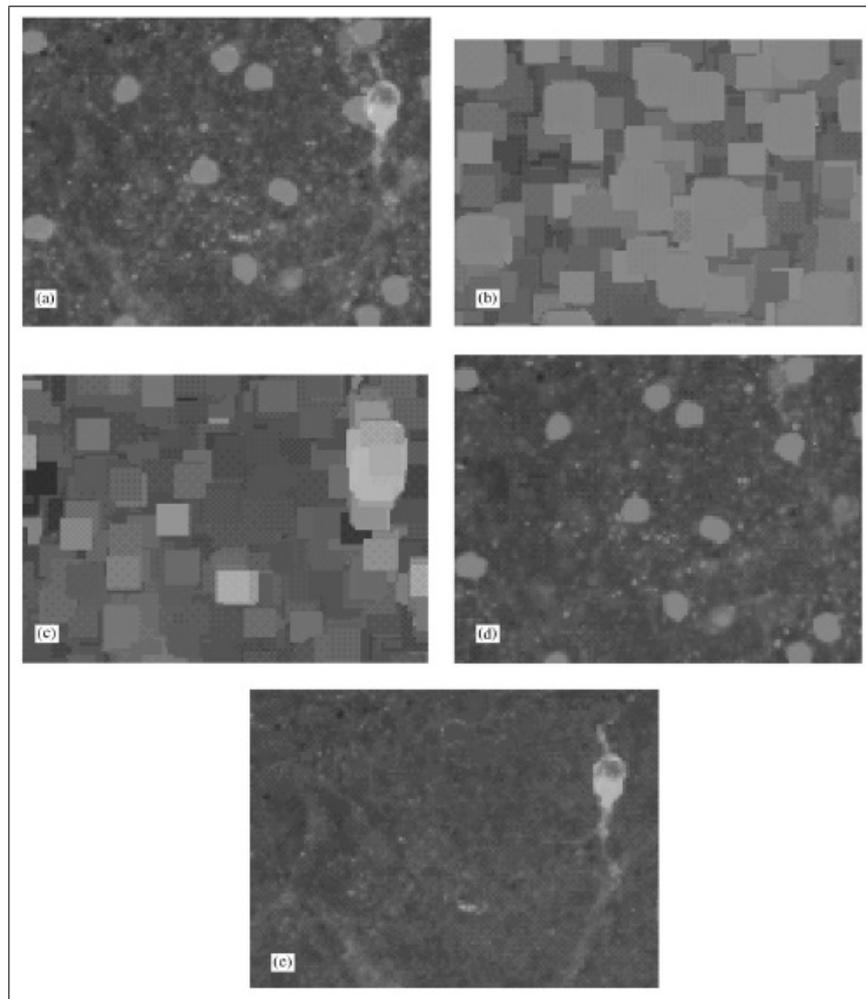


Figura 4.32: Reconstrução geodésica de imagens coloridas. (Fonte: (ORTIZ, 2002))

#### 4.3.5 Remoção de brilhos excessivos em imagens coloridas

Os autores Ortiz, Torres e Gil (2005) apresentam um estudo comparativo sobre a detecção e eliminação dos brilhos excessivos em imagens coloridas, utilizando os espaços de cores HSL e HSV, com base em coordenadas polares, que servem para detectar os brilhos em uma imagem. Os dois sistemas polares (HLS e HSV) têm algumas incoerências que impedem o uso dessas representações de cor em alguns processamentos de imagem, devido às fórmulas de conversão de RGB para os espaços de cores. Algumas instabilidades surgem na saturação dos espaços HSL e HSV para pequenas variações de valores de RGB. Além disso, a saturação das cores primárias não é visualmente agradável em HSL e HSV. Para evitar esses inconvenientes, usou-se a definição L1-normal de Serra (SERRA, 2002), que é o HSV modificado para *hsm*, onde a intensidade ( $I$ ) é modificada para ( $m$ ), sendo calculados  $m$  e  $s$ , conforme as equação 4.12.

$$\begin{aligned}
m &= \left\{ \begin{array}{l} \frac{1}{3}(r + g + b) \\ \frac{1}{2}(2r - g - b) = \frac{3}{2}(r - m) \quad \text{se } (b + r) \geq 2g \\ \text{ou} \\ \frac{1}{2}(r + g - 2b) = \frac{3}{2}(m - b) \quad \text{se } (b + r) < 2g \end{array} \right. \quad (4.12)
\end{aligned}$$

Um estudo detalhado mostrou que os brilhos aparecem em uma área da imagem e que podem ser detectados com base nos diagramas bi-dimensionais LS, VS e MS do espaço de cor modificado HSM. A detecção de brilhos excessivos serve para direcionar a reconstrução somente para os pixels que apresentem essa característica e para não afetar o restante da imagem. Para detectar as regiões de brilhos excessivos, foi utilizado o operador de morfologia matemática *tophat* por abertura em cor, com o espaço de cor HSV e com a ordenação lexicográfica VSH. A ordenação lexicográfica é baseado na equação 4.5 (HANBURY; SERRA, 2001), com mudança na comparação do valor de H, sendo realizada uma comparação simples no terceiro componente  $H1 \leq H2$ .

Para remover os brilhos excessivos realiza-se uma reconstrução morfológica em cores, com o espaço de cor HSV modificado para m, s, com a ordenação lexicográfica, com as imagens máscara e marcadora. O objetivo da reconstrução é preservar os pixels originais e alterar somente os que apresentam brilho excessivo. A imagem marcadora é obtida através de uma erosão em cor da imagem original, utilizando a ordenação lexicográfica. O trabalho de Ortiz, Torres e Gil (2005) emprega uma versão modificada do processo de reconstrução em cor, inserindo uma condição adicional na forma de uma outra imagem contendo as regiões de brilhos excessivos definindo, onde a reconstrução será realizada, sendo necessário que a imagem com regiões de brilhos excessivos seja previamente dilatada para garantir que, após o processo de reconstrução em cor, todos os brilhos excessivos sejam removidos. A Figura 4.33 apresenta o resultado da remoção de brilhos excessivos.

A eliminação do brilho foi obtida automaticamente em um tempo muito curto de processamento. Ocorreu uma redução de tempo de processamento de 50% a 80% em relação a uma reconstrução global geodésica.

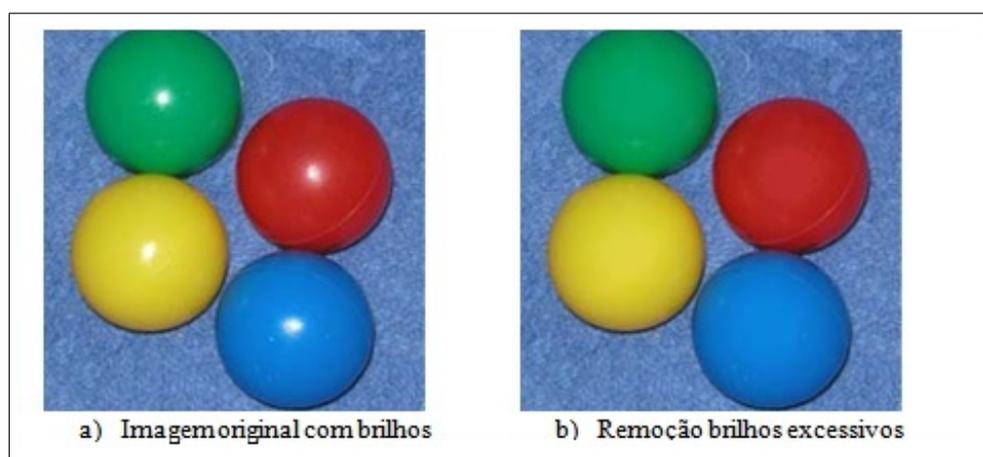


Figura 4.33: Resultado da remoção de brilho excessivo. (Fonte: Ortiz, Torres e Gil (2005))

#### 4.3.6 Estudo Comparativo de Ordenação dos espaços de cor para morfologia matemática

Em Puente (2006) são apresentadas novas ordenações lexicográficas para os espaços de cores RGB e HSI a fim de implementar os operadores morfológicos erosão e dilatação em cor. As operações morfológicas para imagens coloridas dependem da ordem estabelecida para o espaço de cor e, portanto, não possuem um significado único como em imagens binárias. Para o espaço de cor RGB estabeleceu-se a ordem lexicográfica com prioridade para os três canais igualmente, sendo comparado o canal vermelho primeiramente, e, em caso de empate, utiliza-se a comparação dos canais azul e verde da seguinte maneira: compara-se a soma dos canais azul e verde e, permanecendo o empate, compara-se o canal azul. A equação 4.13 relata a ordenação utilizada para o espaço RGB.

$$(r_1, g_1, b_1) \leq (r_2, g_2, b_2) = \begin{cases} (r_1 + g_1 + b_1) < (r_2 + g_2 + b_2) & (1) \\ \text{ou} \\ (r_1 + g_1 + b_1) = (r_2 + g_2 + b_2) \text{ e } (g_1 + b_1) < (g_2 + b_2) & (2) \\ \text{ou} \\ (r_1 + g_1 + b_1) = (r_2 + g_2 + b_2) \text{ e } (g_1 + b_1) = (g_2 + b_2) \text{ e } b_1 \leq b_2 & (3) \end{cases} \quad (4.13)$$

Os resultados obtidos com a aplicação de operadores morfológicos em imagens processadas no espaço de cor RGB, para o estabelecimento dessa ordem, são muito aceitáveis do ponto de vista intuitivo, segundo os autores. A Figura 4.34 ilustra a aplicação da erosão e dilatação, com 04 iterações e elemento estruturante  $3 \times 3$ .

O olho humano não reconhece a quantidade de cada componente vermelho, azul e verde, sem empregar atributos de luminância, ou intensidade, ou saturação. Os espaços

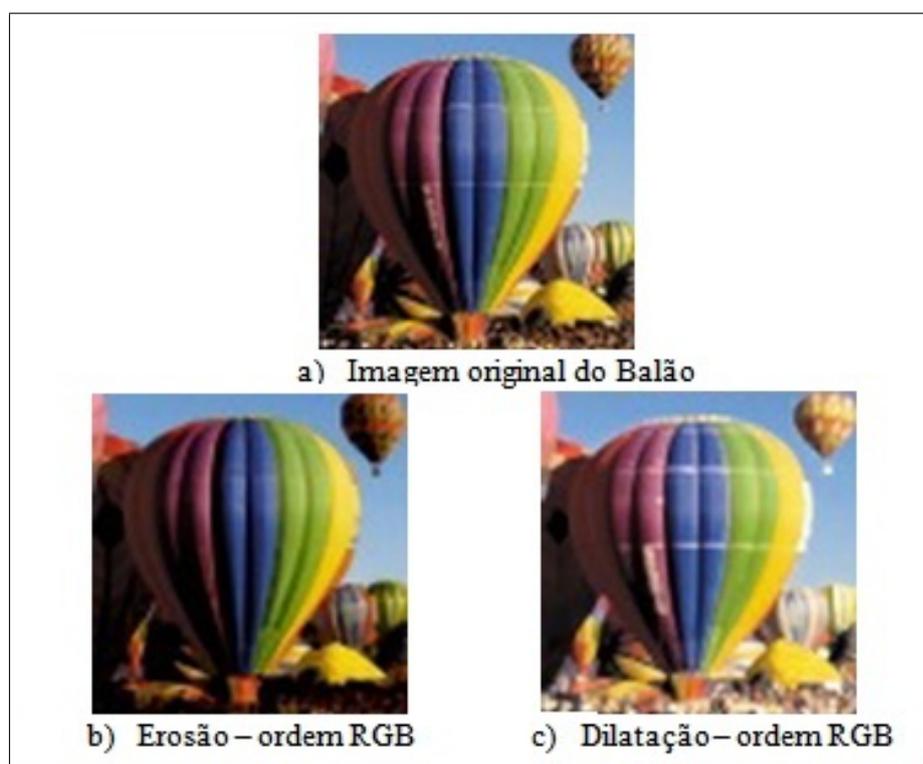


Figura 4.34: Resultado da erosão e dilatação baseado na ordenação lexicográfica RGB. (Fonte: Puente (2006))

que representam as cores com esses atributos permitem uma descrição mais intuitiva das cores. Assim, é preciso realizar a transformação do espaço de cor RGB para o HSI. Deve-se observar se a ordem natural dos componentes do espaço de cor HSI, se a intensidade e a saturação coincidem com a intuição, ou seja, se o preto é menor que o branco, e se o vermelho com baixa saturação é menor que o vermelho puro, com saturação máxima. No entanto, o componente matiz tem uma ordem distinta segundo o matiz de referência  $H_{ref}$ , isto é, a ideia intuitiva de menor e maior desaparece: Para um matiz de referência  $0^\circ$ , o mínimo corresponde à cor vermelha e ao máximo o azul, e para  $H_{ref} = 180^\circ$  o azul é o mínimo e o vermelho é o máximo. Isso acontece quando se utiliza um espaço de cor com componente angular, e isto é um fator relevante, pois os resultados podem resultar em imagens estranhas. Para estabelecer a ordem no espaço de cor HSI, fixa-se o valor de referência do matiz. O valor da saturação atua como peso de ponderação entre a intensidade e o matiz. Em uma cor com baixa saturação, o peso se dá pelo componente intensidade, e com alta saturação pelo componente matiz. A erosão de uma imagem real melhora com a ordenação lexicográfica com prioridade da intensidade, preservando os contornos. Observando os componentes das imagens, percebe-se que nas regiões com alta saturação (zonas claras) é o matiz que estabelece a ordem e em regiões com baixa

saturação (zonas escuras) é a intensidade. Os testes foram realizados somente em duas imagens, abrangendo as operações de erosão e dilatação, trocando as ordens lexicográficas e determinando o valor de referência do matiz para o espaço de cor HSI.

$$(H_1, S_1, I_1) \leq_{H_{ref}} (H_2, S_2, I_2) = \begin{cases} (1 - S_1) \times I + (S_1 \times H_1) < (1 - S_2) \times I_2 & (1) \\ \text{ou} \\ (1 - S_1) \times I + (S_1 \times H_1) = (1 - S_2) \times I_2 \text{ e } I_1 < I_2 & (2) \\ \text{ou} \\ (1 - S_1) \times I + (S_1 \times H_1) = (1 - S_2) \times I_2 \text{ e } I_1 = I_2 \text{ e } H_1 <_{H_{ref}} H_2 & (3) \\ \text{ou} \\ (1 - S_1) \times I + (S_1 \times H_1) = (1 - S_2) \times I_2 \text{ e } I_1 = I_2 \text{ e } H_1 = H_2 \text{ e } S_1 < S_2 & (4) \end{cases} \quad (4.14)$$

A Figura 4.35 ilustra a aplicação da erosão, com 04 iterações e elemento estruturante  $3 \times 3$  e suas ordenações.

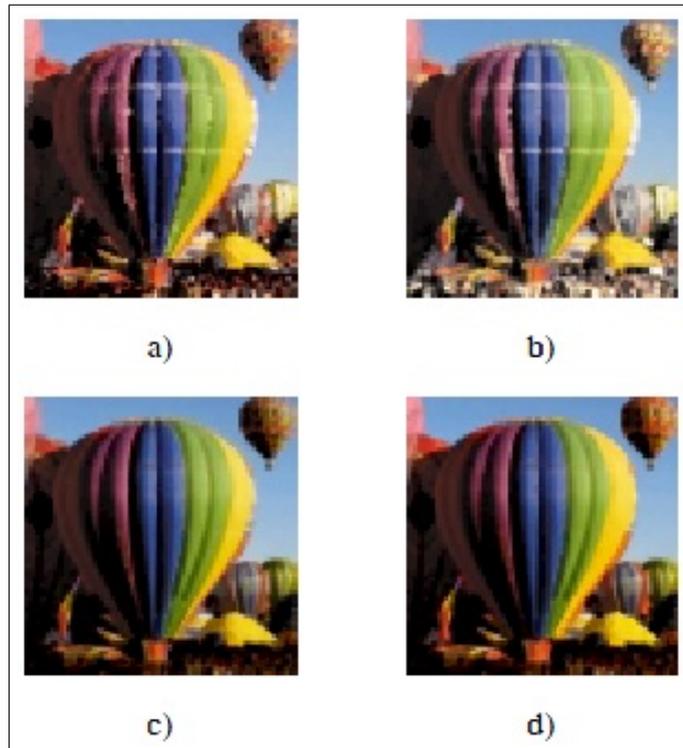


Figura 4.35: Resultado da Erosão: (a) Lexicográfico HSI; (b) Lexicográfico SHI; (c) Lexicográfico ISH e (d) Lexicográfico HSI com  $H_{ref} = 0^\circ$ . (Fonte: Puente (2006))

A Figura 4.36 ilustra a aplicação da dilatação, com 04 iterações e elemento estruturante  $3 \times 3$  e suas ordenações.

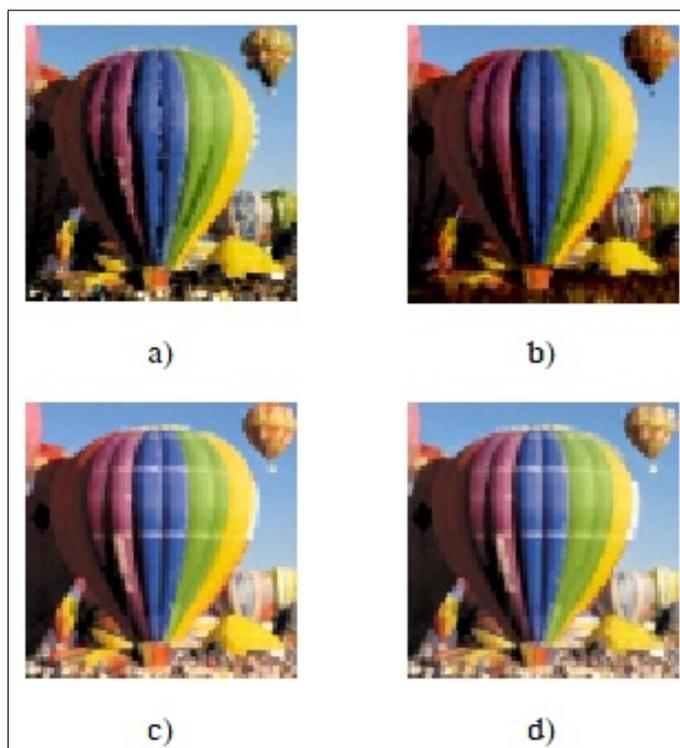


Figura 4.36: Resultado da Dilatação: (a) Lexicográfico HSI; (b) Lexicográfico SHI; (c) Lexicográfico ISH e (d) Lexicográfico HSI com  $H_{ref} = 0^\circ$ . (Fonte: Puente (2006))

#### 4.3.7 Análise Reflexiva dos trabalhos relacionados com a morfologia em cor

Nesta seção apresentaram-se os trabalhos referentes à morfologia matemática em cor, ressaltando os espaços de cor e as suas ordenações lexicográficas. A Tabela 4.4 resume os trabalhos relacionados com as ordenações lexicográficas, com os autores, espaços de cor, ordenações e descrição da estratégia do uso do canal H.

Tabela 4.4: Análise Reflexiva dos trabalhos relacionados com ordenações lexicográficas.

Autores	Espaço de Cor	Ordenações Lexicográficas	Definição H
Hanbury e Serra (2001)	HSL	LSH e SLH HLS	$\frac{H_i}{H_{ref}} \leq \frac{H_j}{H_{ref}}$ $d(H_1, H_{ref}) \leq d(H_2, H_{ref})$
Ortiz (2002)	HSV	HVS	$d(H_1, H_{ref}) \leq d(H_2, H_{ref})$
Ortiz, Torres e Gil (2005)	HSV	HSV modificado HSM	$H_1 \leq H_2$
Puente (2006)	RGB e HSI	RGB, HSI, SHI, ISH	$d(H_1, H_{ref}) \leq d(H_2, H_{ref})$

Sobre as ordenações lexicográficas algumas considerações são necessárias.

- Não há um espaço de cor único para todas as abordagens estudadas, sendo utilizados os espaços: HSL, HSV, RGB, HSI.
- Não há também uma ordenação única utilizada para morfologia matemática em cor. São utilizadas R+G+B como ordenação, VSH e HSL. Nota-se que para um mesmo espaço de cor (HSL) existem várias ordenações lexicográficas como: LSH, SLH e HLS. O uso da ordem dos canais depende da aplicação e dos objetivos esperados.
- Não há um consenso entre qual é o melhor espaço de cor ou a melhor ordenação a ser utilizado.
- Os espaços de cor YCrCb e YIQ são bem conhecidos e não são utilizados nos trabalhos relacionados. Esses dois espaços possuem a vantagem de não utilizar valores a priori para os canais, já os espaços com canal H precisam de um valor de  $H_{ref}$  a priori para o processamento na ordenação lexicográfica.

Assim, este Capítulo apresentou os diversos trabalhos relacionados e as análises reflexivas destes, será apresentado método proposto no Capítulo 5.

# Capítulo 5

## Método Proposto

### 5.1 Introdução do Método Proposto

O método proposto, inspirado na abordagem de Granada, Pina e Muge (2001) por realizar a segmentação baseada na morfologia matemática para imagens binárias, visa à estimativa do fundo e à segmentação do conteúdo das páginas da Bíblia impressa por Gutenberg, sendo composto por 3 níveis e 5 processos, que estão organizados no fluxograma da Figura 5.1.

- O nível 1 realiza os seguintes processos:
  - 1: Estimativa do fundo das imagens antigas;
  - 2: Segmentação do conteúdo do primeiro plano;
  - 3: Segmentação do conteúdo colorido.
- O nível 2 realiza o processo:
  - 4: Segmentação do conteúdo do primeiro plano em dois grupos.
- Nível 3 realiza o processo:
  - 5: Separação do conteúdo do grupo 1 (figuras, títulos, número de páginas e manuscritos).

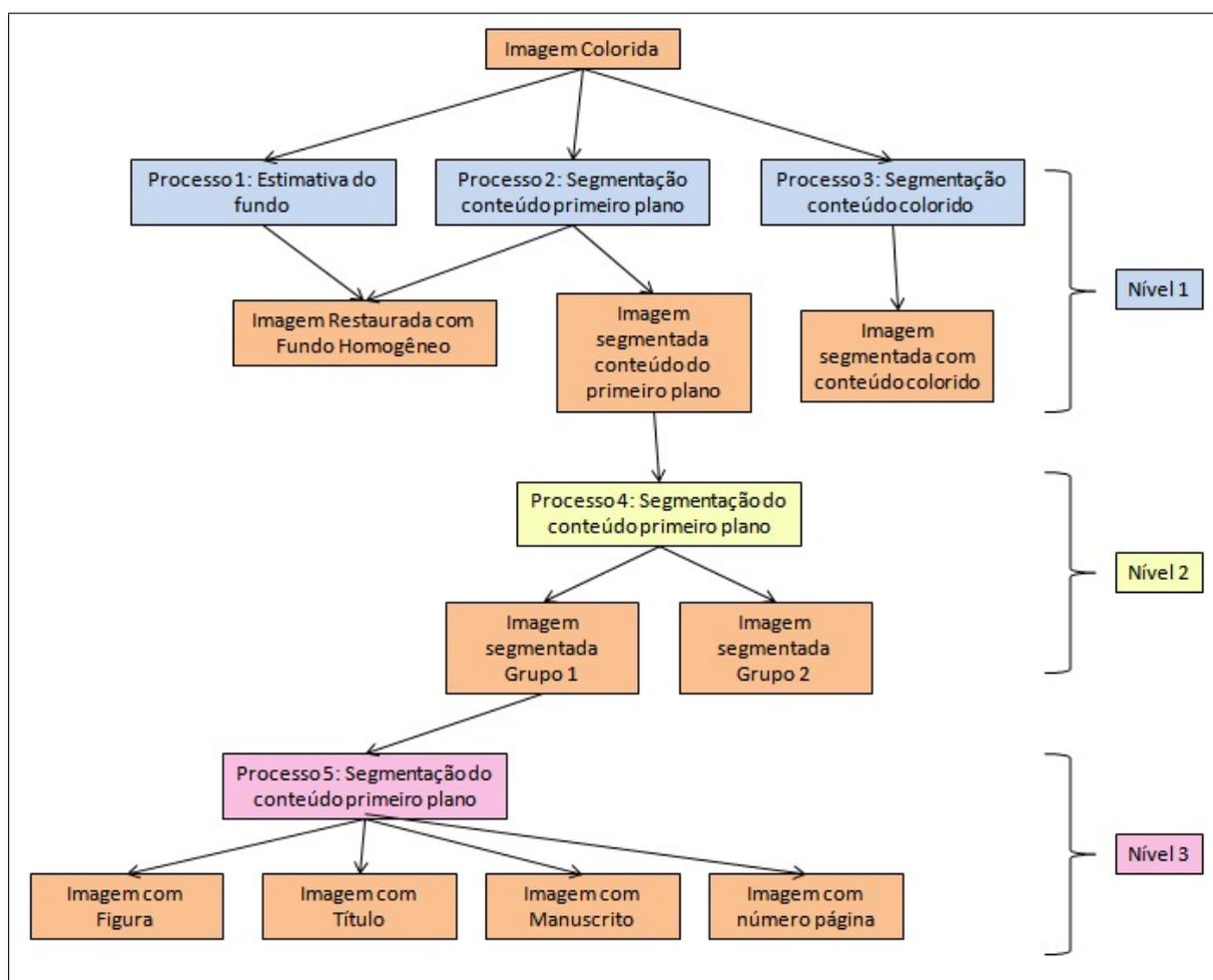


Figura 5.1: Fluxograma do método proposto de segmentação do conteúdo das páginas da Bíblia impressa por Gutenberg.

Para implementação do método proposto, as imagens utilizadas são coloridas, por possuírem muitas informações de suma importância para a relevância do trabalho. A conversão prévia das imagens coloridas em imagens em níveis de cinza acarreta perdas de informações. Portanto, o uso da morfologia em níveis de cinza torna-se inadequado em determinadas aplicações, em que a cor é fator fundamental. Por outro lado, quando o fator cor não for fundamental, técnicas da morfologia em níveis de cinza e/ou binárias foram usadas.

A Figura 5.2 mostra exemplo de uma página da Bíblia colorida e em níveis de cinza. Nota-se que as cores das letras vermelhas presentes na imagem (a), quando transformadas em níveis de cinza em (b), ficam difíceis de serem identificadas, por ser devido o nível de cinza parecido com o nível de cinza das outras letras e do fundo. Também pode-se notar que as linhas do texto em vermelho, da primeira coluna, não podem mais ser diferenciadas do resto do texto, em preto.



Figura 5.2: Exemplo de uma página da Bíblia colorida e em níveis de cinza.

Outro exemplo é mostrado na Figura 5.3, onde a letra capitular azul da imagem (a), quando transformada em níveis de cinza, confunde-se com as letras pretas. As letras vermelhas e com pedaços vermelhos, mostradas na imagem (c), quando transformadas em níveis de cinza, ilustradas na imagem (d), confundem-se com as outras letras.



Figura 5.3: Exemplo de recorte de uma página da Bíblia colorida e em níveis de cinza.

### 5.1.1 Escolha do espaço de cor

Para os testes serão utilizados os espaços de cor descorrelacionados HSI, YCrCb e YIQ, descritos na fundamentação teórica, mostrando a eficiência de cada um para o método proposto.

### 5.1.2 Ordenação para o espaço de cor

Para o método proposto a ordenação lexicográfica será utilizada, sendo, por exemplo, para o espaço de cor HSI a ser utilizado, as suas possibilidades de ordenações: HSI, ISH, SHI, HIS, SIH e IHS. Para os espaços de cor HSL, YCrCb e YIQ, também serão realizados os testes com as suas possibilidades de ordem dos canais. A seguir será descrito detalhadamente cada passo do método proposto.

## 5.2 Descrição do Método Proposto

O método proposto é composto por 05 processos. A seguir são detalhados, passo a passo, os processos, conforme a Figura 5.1, sendo descritas as soluções para cada etapa, com as técnicas adequadas para alcançar os objetivos propostos.

### 5.2.1 Nível 1 - Processo 1: Estimativa do fundo e Processo 2: Segmentação do conteúdo

O objetivo deste nível, com seus processos, é realizar a estimativa do fundo e a segmentação das informações do primeiro plano através da filtragem lexicográfica, pela operação de reconstrução com aproximação do fundo e de limiarização. O fluxograma que representa esses dois processos está ilustrado pela Figura 5.4.

O fluxograma descreve os caminhos para obtenção dos resultados nos dois processos:

- Processo 1: Estimativa do fundo colorido;
- Processo 2: Segmentação das informações do primeiro plano.

Para realização desses 02 processos deve-se proceder da seguinte forma:

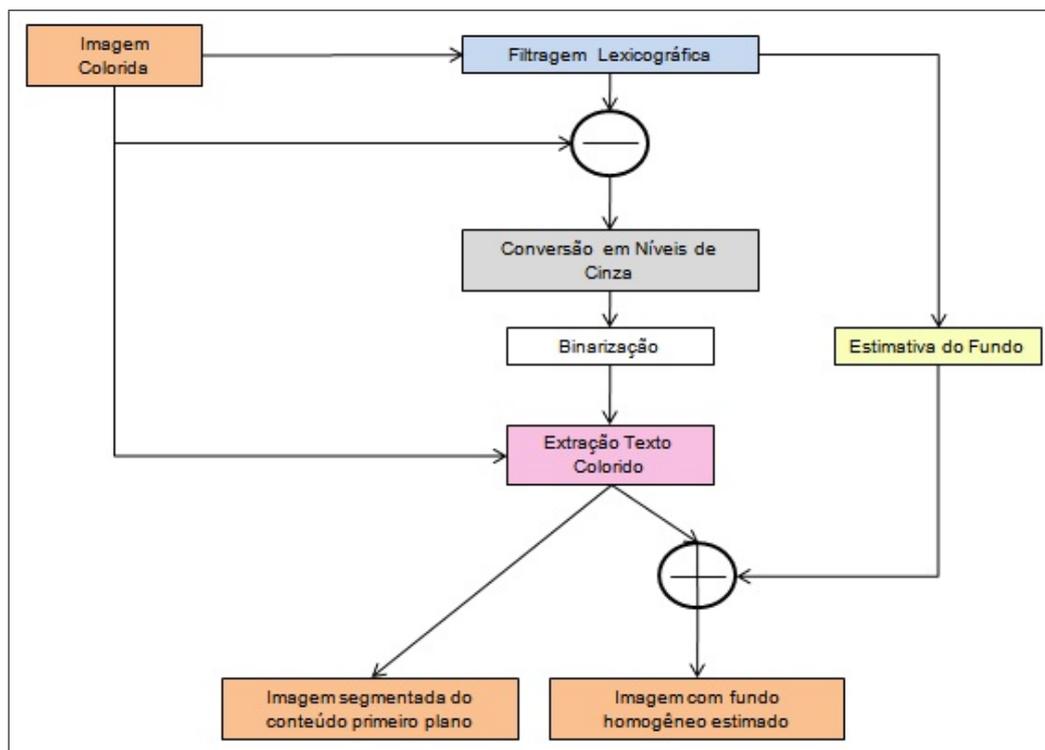


Figura 5.4: Fluxograma da restauração do fundo e extração do conteúdo do primeiro plano de documentos antigos.

### 5.2.1.1 Filtragem Lexicográfica - Operação morfológica reconstrução

As imagens coloridas de documentos antigos possuem fundos deteriorados com o tempo e, para recuperar as informações pertencentes a ele e ao primeiro plano, será utilizada a abordagem descrita no fluxograma, realizando primeiramente as seguintes tarefas:

a) **Operação morfológica de reconstrução em cor:** processo que consiste em recuperar as informações do fundo da imagem antiga, possibilitando a estimativa do fundo. Para isto é preciso realizar um pré-processamento das imagens, como segue.

- Primeiro: a imagem máscara (imagem original)  $g$  é invertida, pois invertendo-a, o fundo fica mais escuro que as letras centrais, que se tornam brancas, e no momento da reconstrução pela borda, será recuperado somente os pixels escuros.
- Segundo: é criada a imagem marcadora  $f$ , que é a imagem borda, composta da seguinte forma: a primeira linha e coluna e a última linha e coluna da imagem marcadora são compostos de pixels da imagem  $g$  invertida, e o resto dos pixels é definido como sendo da cor preta.

A Figura 5.5 ilustra exemplo da imagem máscara (imagem (b)) e da imagem marcadora (imagem(c)).

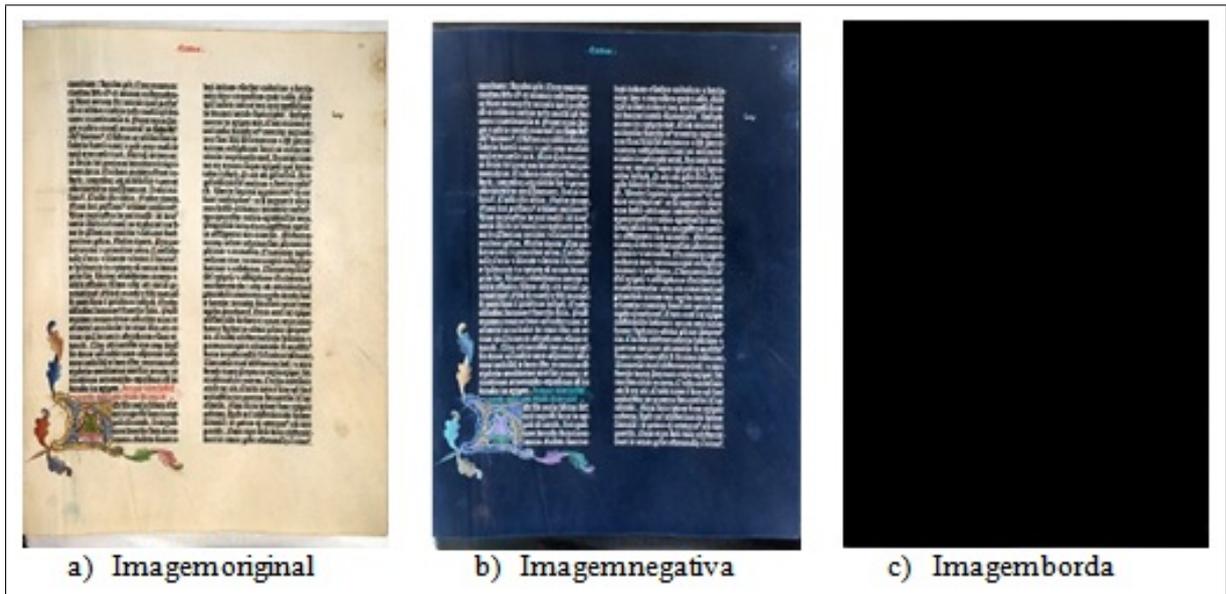


Figura 5.5: Exemplo da imagem da Bíblia, da marcadora e da máscara.

A Figura 5.6 ilustra exemplo da reconstrução do fundo da imagem através da imagem (a). A imagem (b) ilustra o negativo da imagem reconstruída, que servirá para a estimativa do fundo homogêneo da imagem da Bíblia.

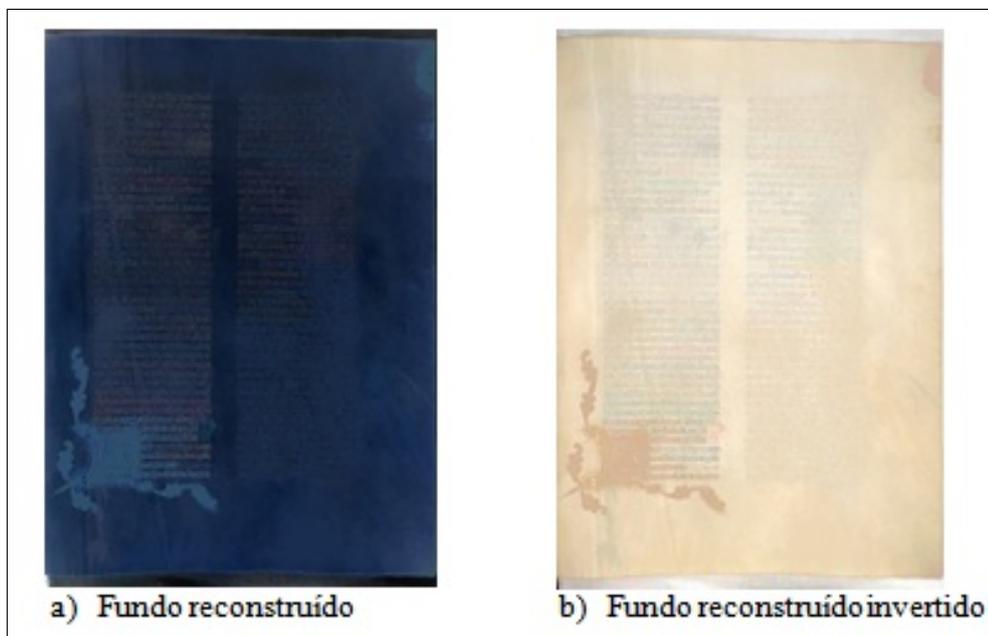


Figura 5.6: Exemplo do fundo reconstruído.

- b) **Operação morfológica Tophat:** subtração da imagem original com a imagem reconstruída, gerando as informações do primeiro plano. Esta subtração tem como objetivo encontrar as informações do primeiro plano da imagem de documentos antigos. A Figura 5.7 representa o resultado do *Tophat* da imagem original com o fundo reconstruído.



Figura 5.7: Exemplo do fundo reconstruído com TopHat.

- c) **Recuperação dos pixels originais:** é realizada a conversão em níveis de cinza e binarização para recuperar os pixels originais do primeiro plano. Com a reconstrução do fundo da imagem, obtém-se uma imagem com os elementos do primeiro plano e do segundo plano, com problema na tonalidade das cores (percebe-se que as cores não são iguais à imagem original). Para resolver este problema, é preciso realizar uma limiarização (binarização) do resultado da imagem reconstruída (imagem (c) da Figura 5.7), para recuperar as tonalidades dos pixels do primeiro plano. Com a imagem binarizada é possível recuperar os pixels originais, pois esses representam as informações do conteúdo do primeiro plano. Para recuperá-los, basta localizar a posição dos pixels pretos da imagem binarizada e recuperar os pixels da mesma posição da imagem original. A Figura 5.8 apresenta o resultado da binarização (imagem(c)) e da recuperação dos pixels originais (imagem (d)).

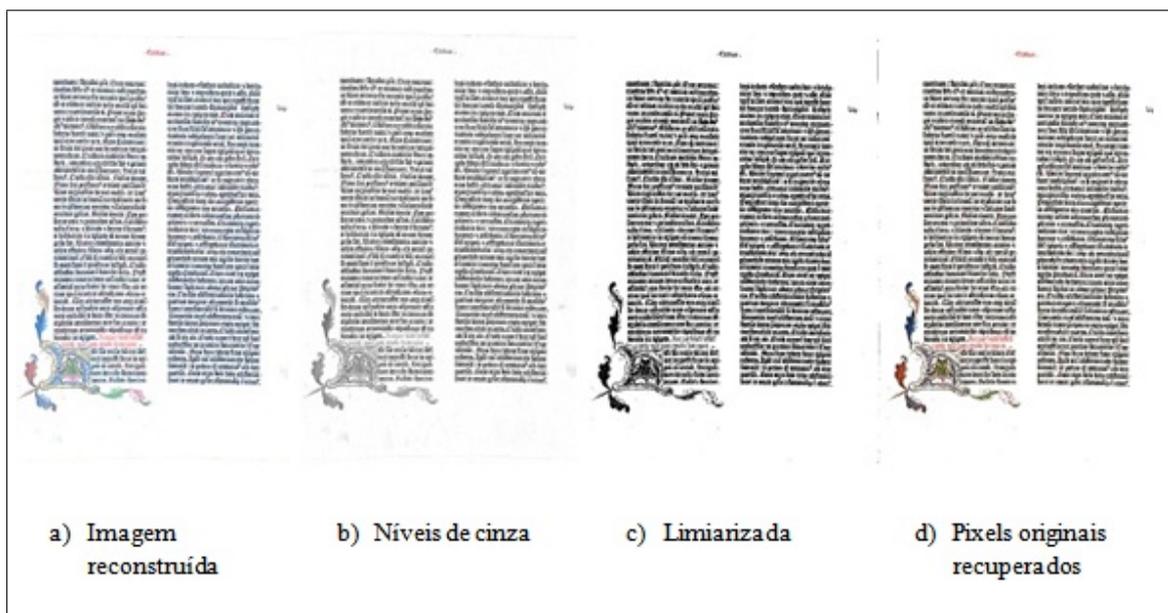


Figura 5.8: Exemplo do fundo reconstruído com a tonalidade original dos elementos do primeiro plano.

**d) Estimativa da cor do fundo e segmentação do conteúdo do primeiro plano:**

A estimativa da cor do fundo é baseada no pico de maior ocorrência dos pixels da imagem reconstruída do fundo. Assim, realiza-se a soma entre duas imagens: a) imagem dos pixels do primeiro plano (com os tons originais) e b) imagem do fundo homogêneo estimado. Para a geração das imagens com o fundo estimado e com o fundo branco, com as informações somente do primeiro plano, é preciso realizar as etapas abaixo relacionadas.

- Primeira: a imagem binarizada é utilizada como máscara da imagem original para encontrar os pixels do primeiro plano da imagem original, gerando uma imagem reconstruída final com fundo branco e os pixels originais do primeiro plano.
- Segunda: é utilizada uma estimativa baseada no valor de maior ocorrência do histograma dos pixels do fundo da imagem reconstruída, para gerar um fundo homogêneo.
- Terceira: a imagem reconstruída final com fundo homogêneo é a soma da imagem do fundo estimado com a imagem original do primeiro plano com o fundo branco.

A Figura 5.9 ilustra os dois resultados alcançados por esta etapa do método proposto, sendo que a imagem (a) ilustra a imagem original, a (b) a segmentação do

conteúdo do primeiro plano com fundo branco e a (c) a imagem com o fundo estimado homogêneo.

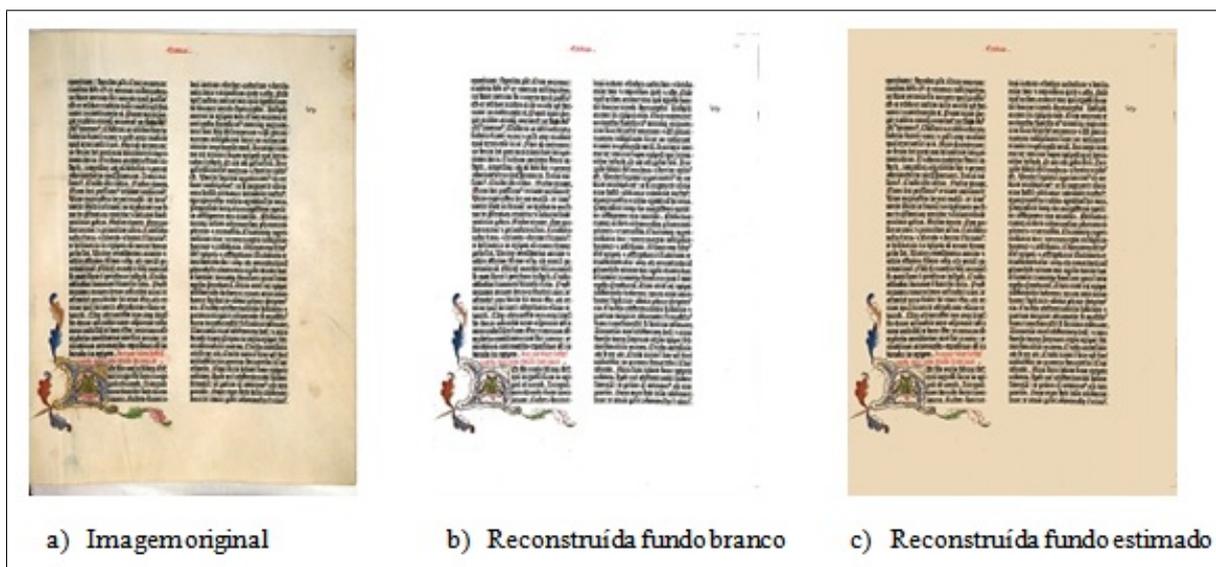


Figura 5.9: Exemplo do fundo reconstruído branco e estimado.

### 5.2.1.2 Nível 1 - Processo 3: Segmentação do conteúdo colorido

Esta etapa consiste na segmentação das letras coloridas, títulos e figuras das páginas da Bíblia através da utilização da morfologia matemática em cor. A Figura 5.10 ilustra exemplo desta segmentação, sendo que as pequenas letras coloridas e os títulos foram identificados por um retângulo vermelho, conforme a imagem (b).

As pequenas letras coloridas na imagem são difíceis de serem identificadas a olho nu, em contrapartida, a figura e o título são mais fáceis de serem identificados. A Figura 5.11 ilustra um pedaço da imagem com as pequenas letras vermelhas, para melhor visualização.

A Figura 5.12 apresenta o fluxograma das etapas da segmentação das letras coloridas.

Nota-se, pelas etapas deste processo, a necessidade de processamentos da filtragem lexicográfica com a dupla operação de reconstrução para atingir os resultados esperados. A primeira operação de reconstrução em cor consegue segmentar as sementes das letras pequenas coloridas vermelhas, da figura e do título, a segunda operação lexicográfica consiste em reconstruir os pixels das letras pequenas, títulos e figuras. Essa é uma maneira de identificar as letras coloridas presentes nas duas colunas, as figuras e os títulos coloridos das páginas da Bíblia.

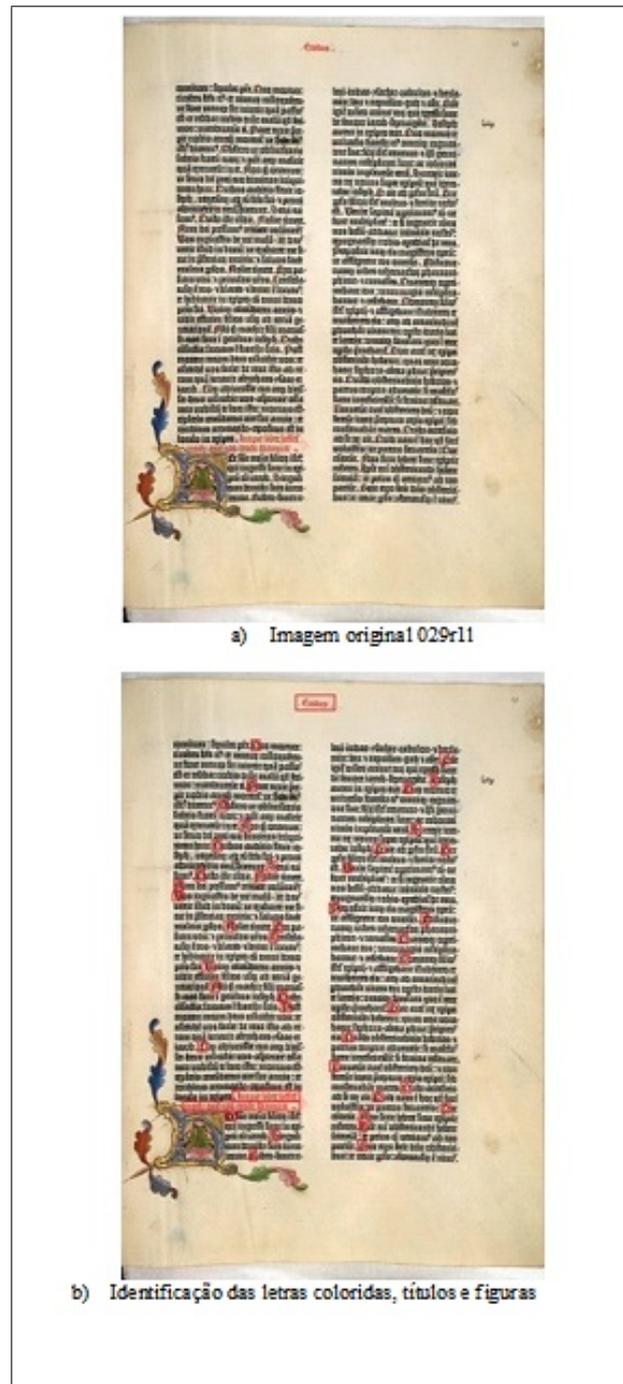


Figura 5.10: Exemplo da identificação do conteúdo colorido.

Funcionamento do fluxograma da Figura 5.12:

- Filtragem Lexicográfica: realiza-se a operação de reconstrução através da ordenação lexicográfica com os espaços de cor HSI ou HSL com as seguintes ordens: SHI ou SHL, para encontrar os pixels que representam os títulos, figuras e letras coloridas. A imagem original invertida é a máscara e a imagem marcadora é a borda da máscara;

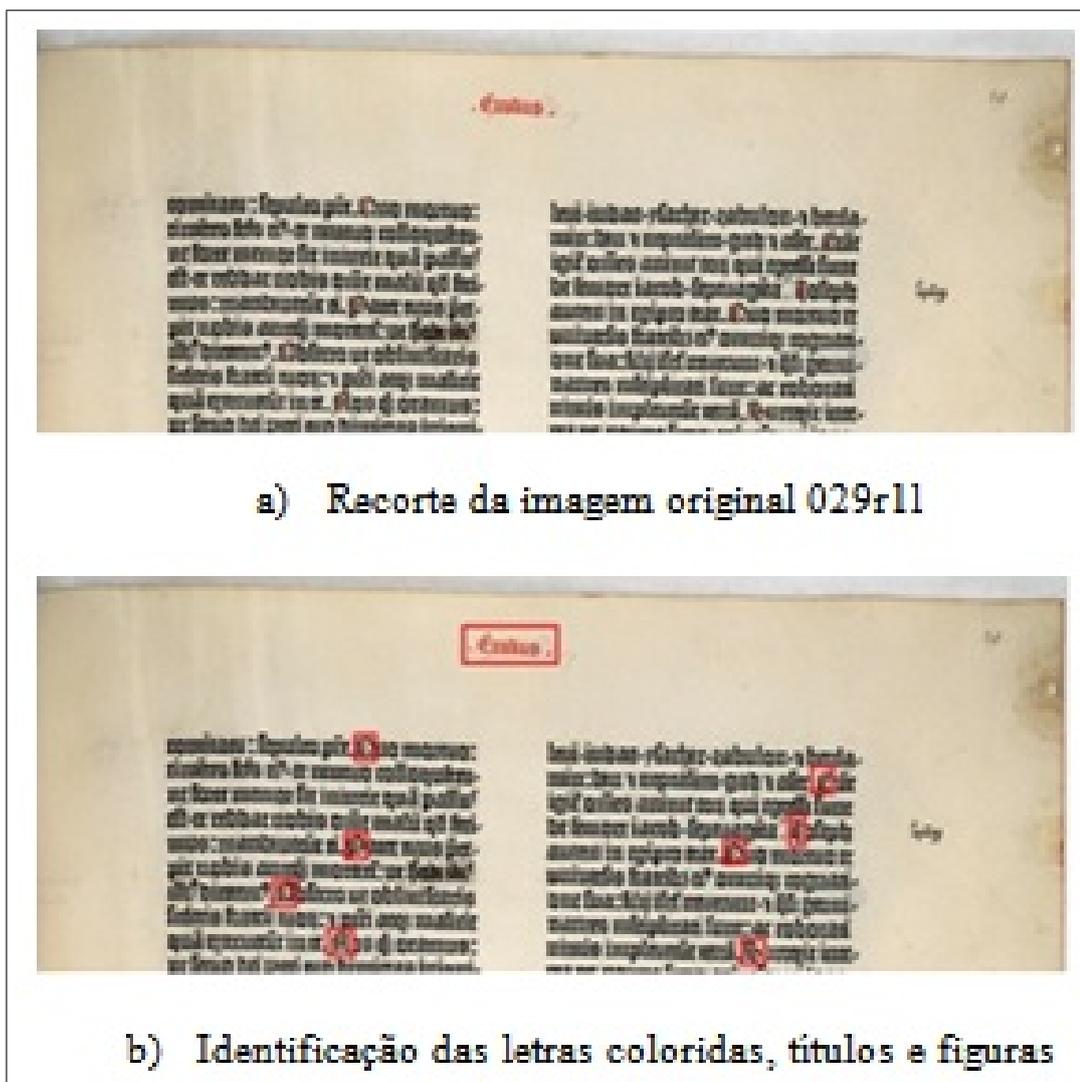


Figura 5.11: Exemplo de identificação das pequenas letras vermelhas.

- Subtração entre a imagem original e a reconstruída SHI ou SHL;
- Conversão em níveis de cinza da imagem subtraída;
- Binarização da imagem em níveis de cinza;
- Filtragem dos ruídos da imagem binarizada por morfologia matemática binária soft (ZMUDA; TAMBURINO; RIZK, 1991);
- Extração do texto colorido: baseada nos pixels pretos filtrados;
- Filtragem lexicográfica: Realiza-se novamente a operação de reconstrução com a ordenação lexicográfica com os espaços de cor HSI, HSL, YCrCb ou YIQ, com as ordens: ISH, LSH, YCrCb ou IYQ, para recuperar todos os pixels pertencentes

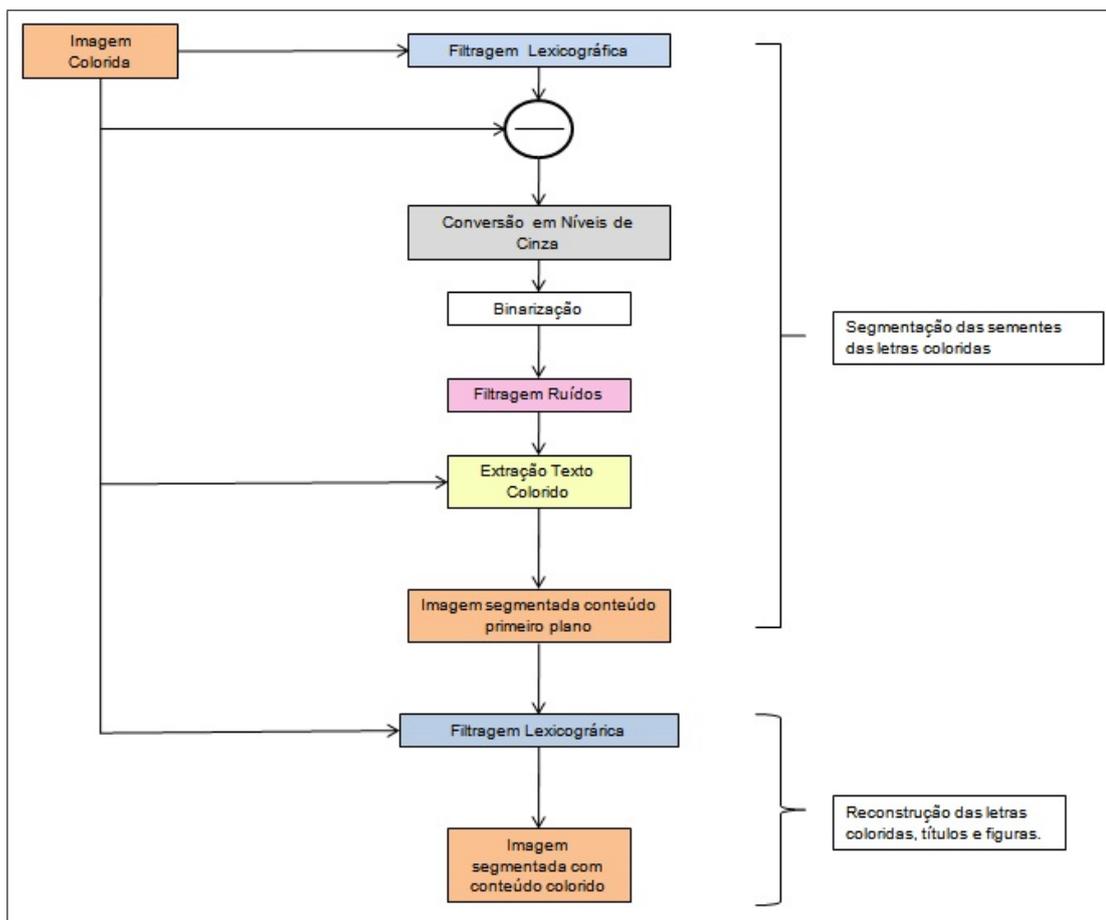


Figura 5.12: Fluxograma da segmentação texto colorido.

aos títulos, figuras e letras coloridas. A máscara é a imagem original invertida e a marcadora é a imagem invertida dos pixels extraídos do texto colorido;

- Resultado final com a imagem segmentada com conteúdo colorido.

A Figura 5.13 ilustra exemplo da operação morfológica de reconstrução lexicográfica com ordem SHI, sendo a imagem (a) original, a (b) o resultado da reconstrução SHI, a (c) o resultado do processo *TopHat* e a imagem (d) a imagem limiarizada.

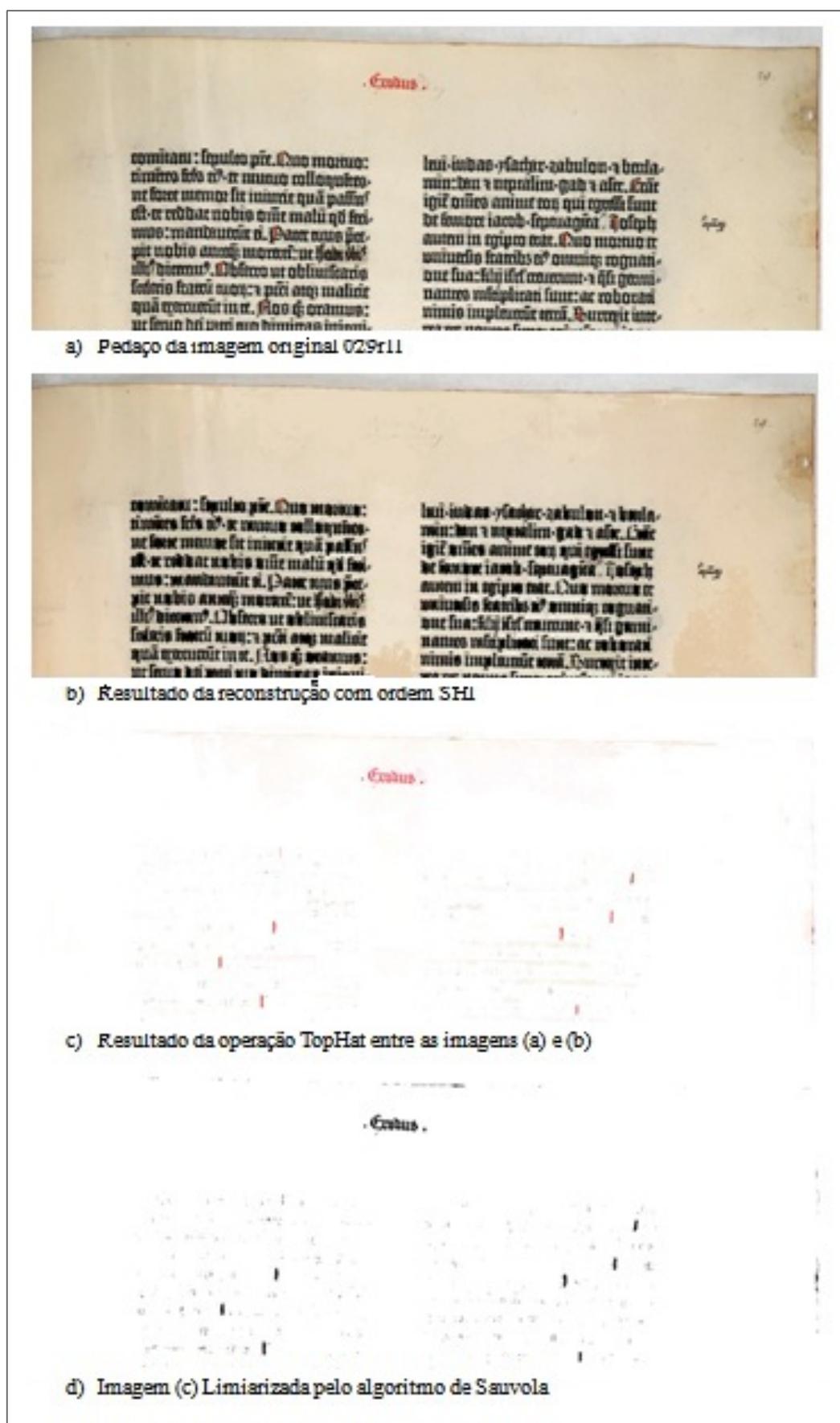


Figura 5.13: Resultado da operação de reconstrução com ordem SHI.

A imagem limiarizada contém muitos ruídos, necessitando de filtragem dos pixels por morfologia matemática binária soft (ZMUDA; TAMBURINO; RIZK, 1991), que analisa um pixel preto da imagem limiarizada e verifica a quantidade de pixels vizinhos pretos, de acordo com o valor de  $m$ , sendo esse valor o de referência para verificação da quantidade de pixels pretos na vizinhança. A Figura 5.14 ilustra a imagem binarizada filtrada com valores de  $m$  variando de 2 a 4.

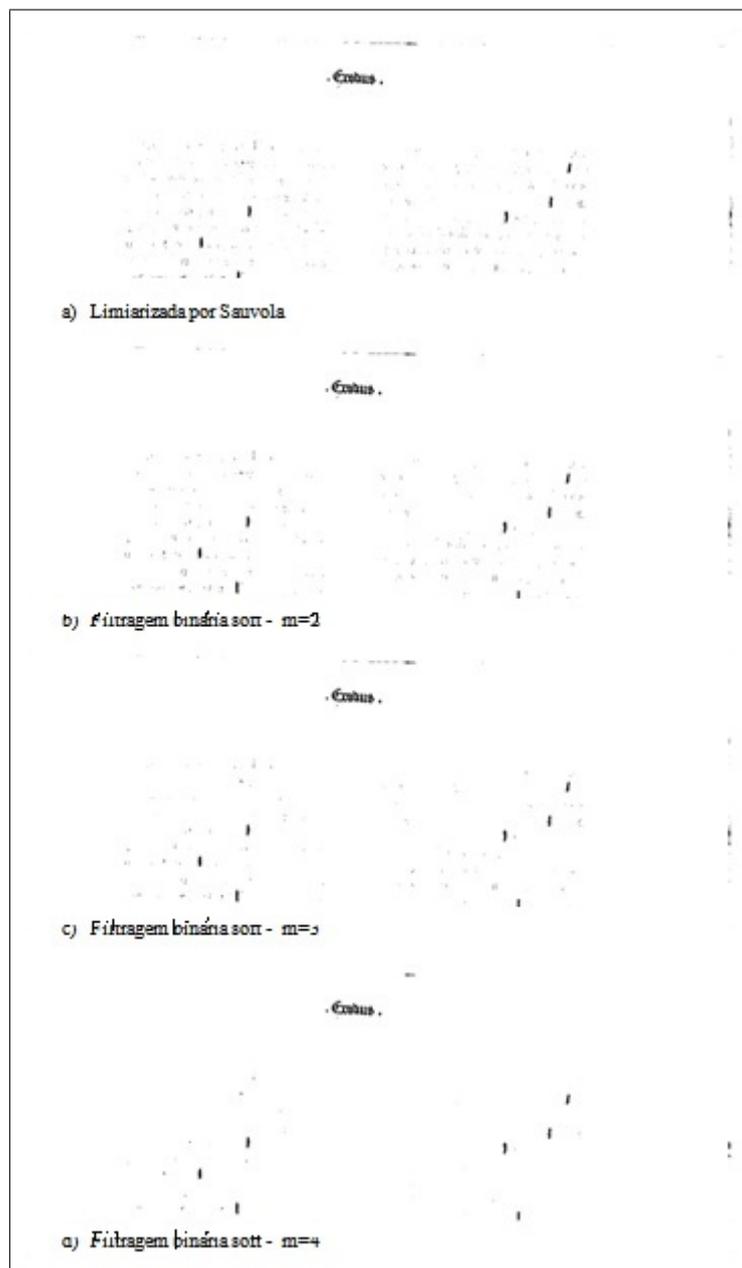


Figura 5.14: Resultados da filtragem matemática binária soft com  $m$  variando de 2 a 4.

A Figura 5.15 ilustra a imagem binarizada filtrada com valores de  $m$  variando de 5 a 8.

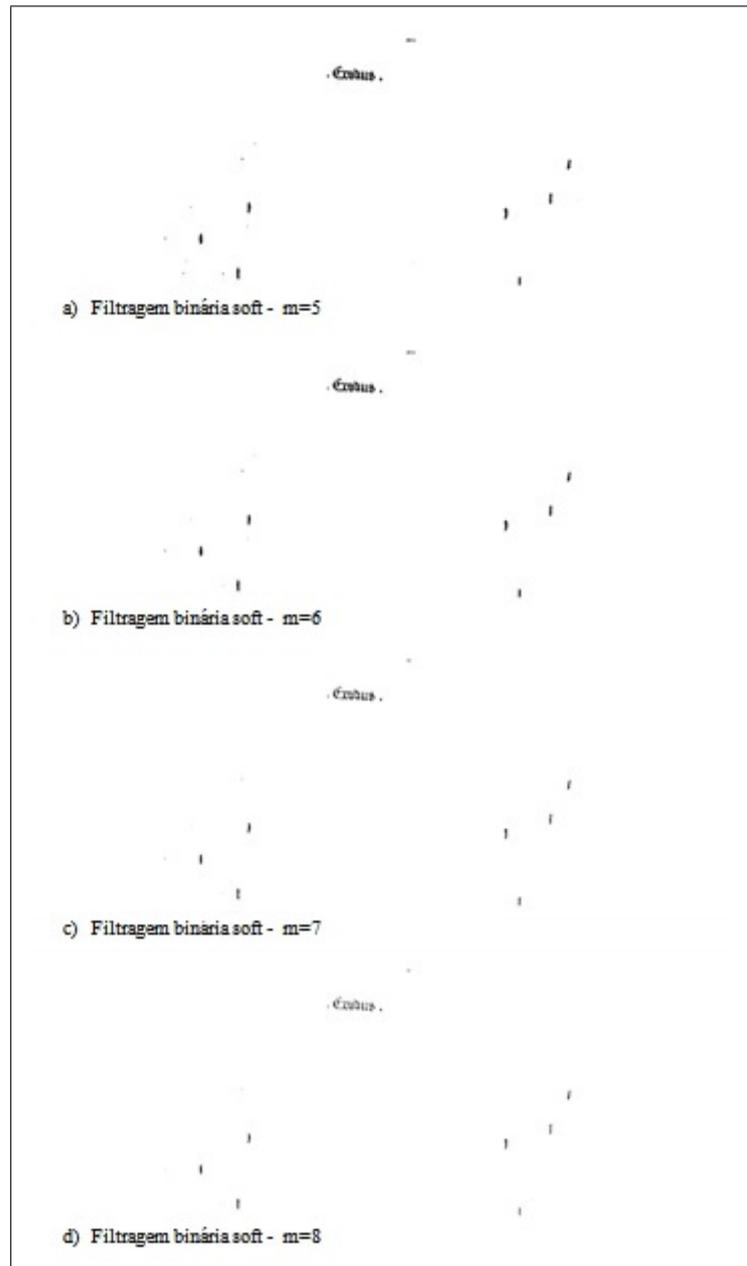


Figura 5.15: Resultados da filtragem matemática binária soft com  $m$  variando de 5 a 8.

Com a imagem filtrada, o próximo passo consiste na reconstrução dos pixels considerados como sendo de letras coloridas, figuras e títulos. A Figura 5.16 ilustra exemplo da reconstrução baseada nas imagens filtradas com  $m$  variando de 2 a 4.

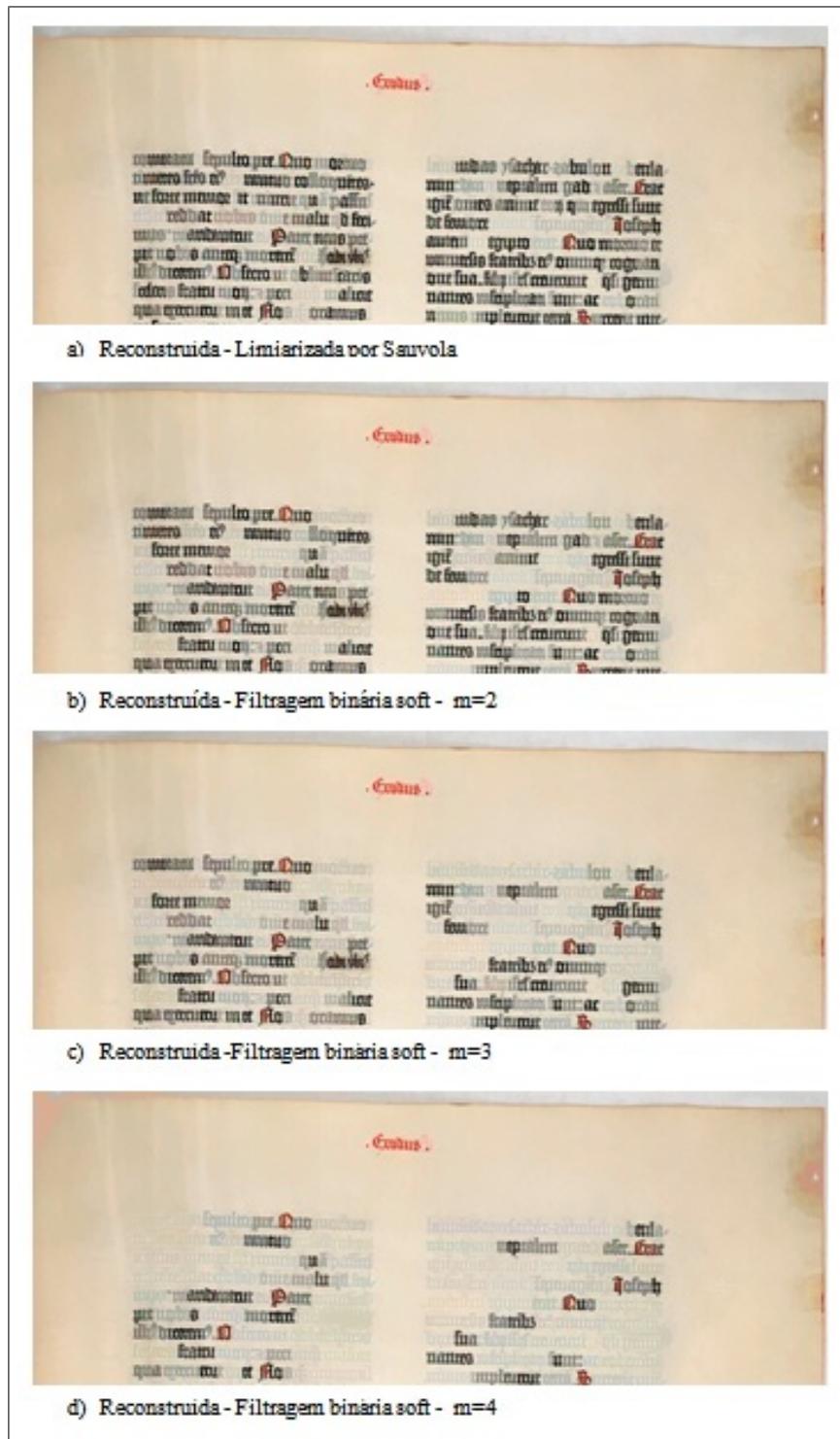


Figura 5.16: Resultados da segmentação com  $m$  variando de 2 a 4.

A Figura 5.17 ilustra exemplo da reconstrução baseada nas imagens filtradas com  $m$  variando de 5 a 8.

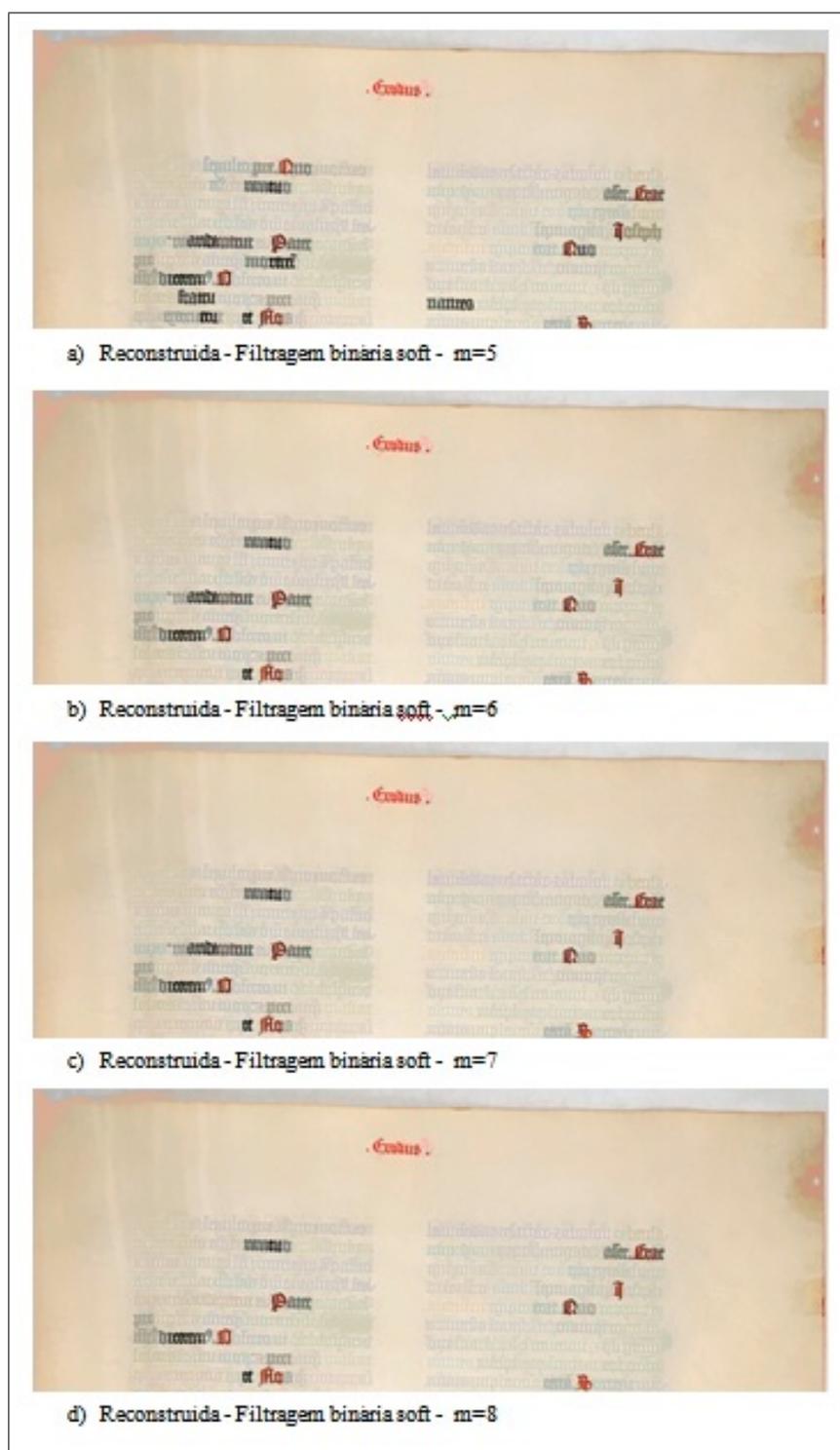


Figura 5.17: Resultados da segmentação com  $m$  variando de 5 a 8.

Os experimentos mostrarão qual é o melhor algoritmo de limiarização e qual é o valor mais adequado para o valor da filtragem  $m$ . Assim, é possível identificar as letras

coloridas maiores e menores presentes nas duas colunas da página da Bíblia e também as figuras e os títulos.

### 5.2.2 Nível 2 - Processo 4: Segmentação do conteúdo do primeiro plano

Essa etapa trabalha com as informações da imagem segmentada do primeiro plano, com o fundo branco, resultado do processo 2 do nível 1, sem a aproximação do fundo colorido. As páginas da Bíblia possuem uma estrutura bem definida, com duas colunas impressas e títulos, algumas imagens possuem figuras e/ou manuscritos. Essa estrutura pode ser dividida em dois grupos:

1. Grupo 1: figuras, títulos, manuscritos e numeração de páginas;
2. Grupo 2: colunas centrais impressas.

Esta etapa visa à separação dos conteúdos do primeiro plano da imagem segmentada pelo nível 1 do método proposto em dois grupos: 1 e 2. A Figura 5.18 mostra a segmentação ideal dos dois grupos.

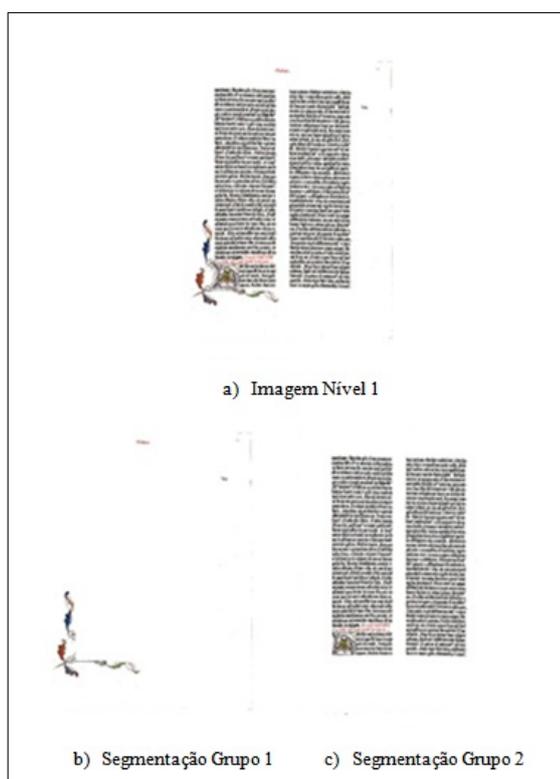


Figura 5.18: Exemplo grupo 1 e 2.

### 5.2.2.1 Diferenciação dos dois grupos

Uma maneira de diferenciar os dois grupos é a identificação/localização das duas colunas centrais impressas. Com a localização destas, realiza-se a segmentação das duas colunas centrais, definindo o conteúdo do grupo 2. O grupo 1 é composto pelas demais informações do primeiro plano. A Figura 5.19 ilustra o fluxograma com as etapas para segmentação do grupo 1.

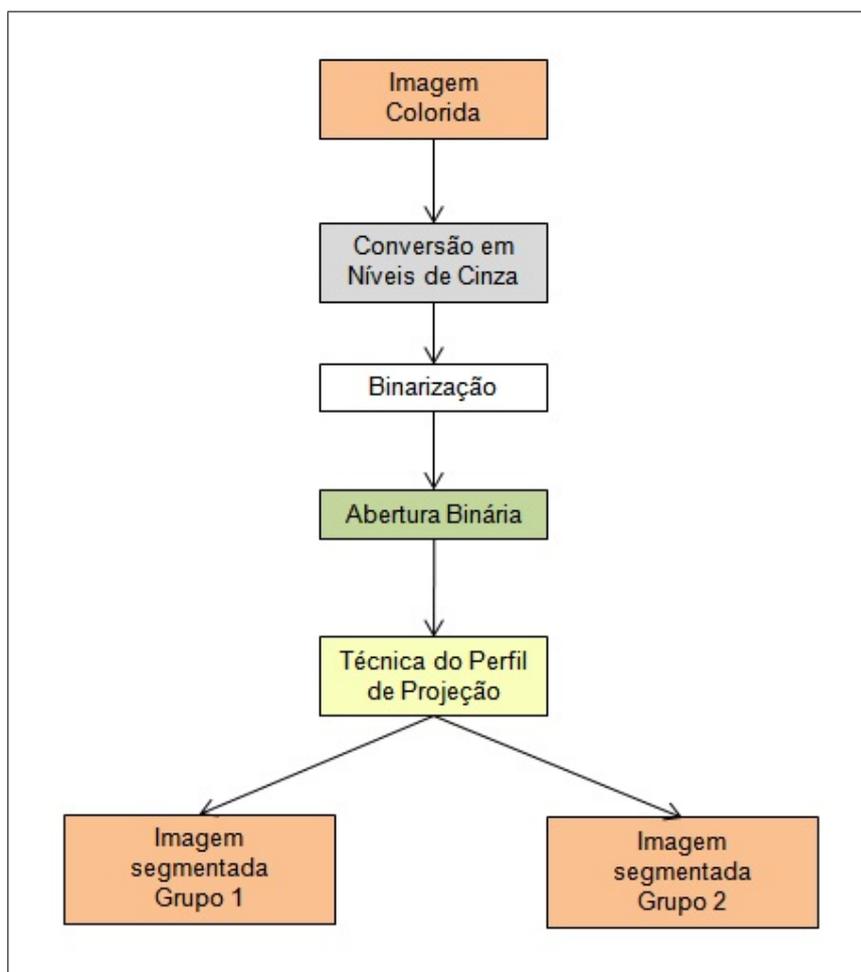


Figura 5.19: Fluxograma da segmentação em grupos 1 e 2.

Para encontrar as duas colunas centrais é preciso realizar as seguintes tarefas:

1. Converter para níveis de cinza e binarizar a imagem resultante do processo 2 do nível 1;
2. Aplicar a operação morfológica abertura binária na imagem binarizada, para assim encontrar os blocos correspondentes dos grupos 1 e 2;
3. Aplicar a técnica do perfil de projeção na imagem resultante da operação de abertura.

### 5.2.2.2 Aplicação da operação de abertura binária

Para diferenciação dos dois grupos, utiliza-se a técnica de perfil de projeção horizontal e vertical. Para um resultado apurado, é preciso realçar as linhas e colunas da imagem, ou seja, realçar o bloco das duas colunas centrais para a sua segmentação. A operação de abertura binária possibilita o realce das duas colunas centrais da imagem da Bíblia, tornando as projeções mais compactadas. Sem o uso desta operação podem ocorrer erros de segmentação das duas colunas centrais das páginas da Bíblia, tornando-se difícil de realizar a segmentação das informações dos grupos 1 e 2. Para esta etapa utilizar-se-á a abertura binária, pois é uma técnica rápida, consolidada, bem conhecida na literatura, sem ambiguidades, sem a necessidade de trabalhar com as informações de cor, pois o objetivo é somente realçar os blocos das duas colunas centrais.

### 5.2.2.3 Abordagem morfológica por abertura binária

A abordagem morfológica por abertura binária consiste na filtragem das imagens com operação morfológica e binarização, sendo que serão testados vários algoritmos de limiarização. Esta operação consiste em duas operações básicas: erosão e dilatação, conforme descritas na fundamentação teórica. Essa abordagem não é automática, pois para realizá-la é preciso escolher o elemento estruturante e o número de iterações. Os elementos estruturantes podem ser: Cruz, Linha Horizontal, Linha Vertical, Quadrado e Rhombus. O elemento estruturante quadrado é o mais utilizado em processamento de imagens, pois analisa as oito direções do pixel para as operações básicas de dilatação e erosão. O número de iterações pode variar, sendo que serão testada várias iterações, descobrindo-se assim qual o melhor valor para processar as páginas da Bíblia. Não é preciso um número exato de iterações, pois esta abordagem precisa realçar os conteúdos do primeiro plano, para assim realizar a segmentação de todas as informações, estando o resultado ideal ilustrado pela Figura 5.20, que apresenta:

- Em (a), a imagem reconstruída resultante da operação do processo 2 do nível 1;
- Em (b), a imagem reconstruída limiarizada;
- Em (c), o resultado da operação morfológica abertura binária com elemento estruturante quadrado e 08 iterações.

Após a operação abertura binária, aplica-se a técnica do perfil de projeção vertical e horizontal para encontrar as duas colunas centrais, segmentando assim os conteúdos relacionados com os grupos 1 e 2.

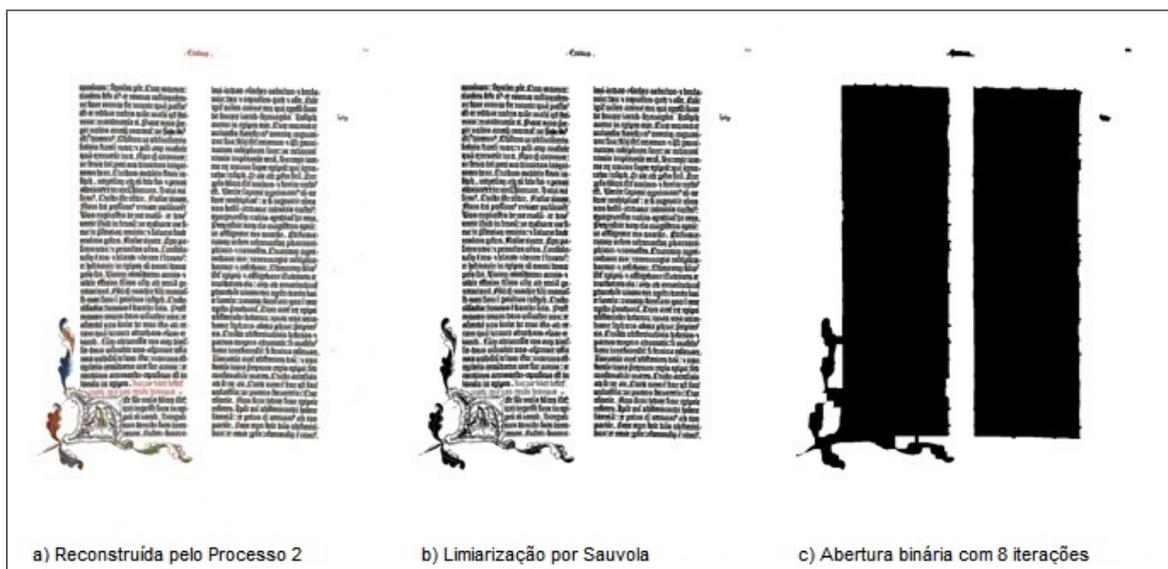


Figura 5.20: Exemplo da segmentação ideal do conteúdo do primeiro plano de uma página da Bíblia.

A Figura 5.21 apresenta a segmentação ideal para os dois grupos, sendo:

- A imagem (a), o resultado da imagem binarizada ideal do conteúdo do primeiro plano.
- A imagem (b), a representação da segmentação ideal do Grupo 1 (figuras, número de páginas, títulos e manuscritos) na cor preta; e do Grupo 2 (colunas centrais impressas), na cor cinza.

Para realizar a segmentação em dois grupos utiliza-se a técnica de perfil de projeção, que será detalhado a seguir.

#### 5.2.2.4 Técnica de perfil de projeção

A técnica de perfil de projeção (MASCARO; CAVALCANTI; MELLO, 2010) preconiza que uma imagem é formada por regiões de texto, organizadas em linhas paralelas, sendo possível determinar as regiões que possuem uma grande concentração de pixels na mesma direção, vertical ou horizontal. São vulneráveis à presença de outros elementos, como figuras, tabelas ou ruído. Evoluções na técnica tentam eliminar componentes gráficos, mas assumem que o documento possui uma região mínima de linhas paralelas de texto.

A Figura 5.22 ilustra o exemplo da aplicação do perfil de projeção com projeção vertical e horizontal.

- Imagem (a), o resultado da binarização do conteúdo do primeiro plano de uma

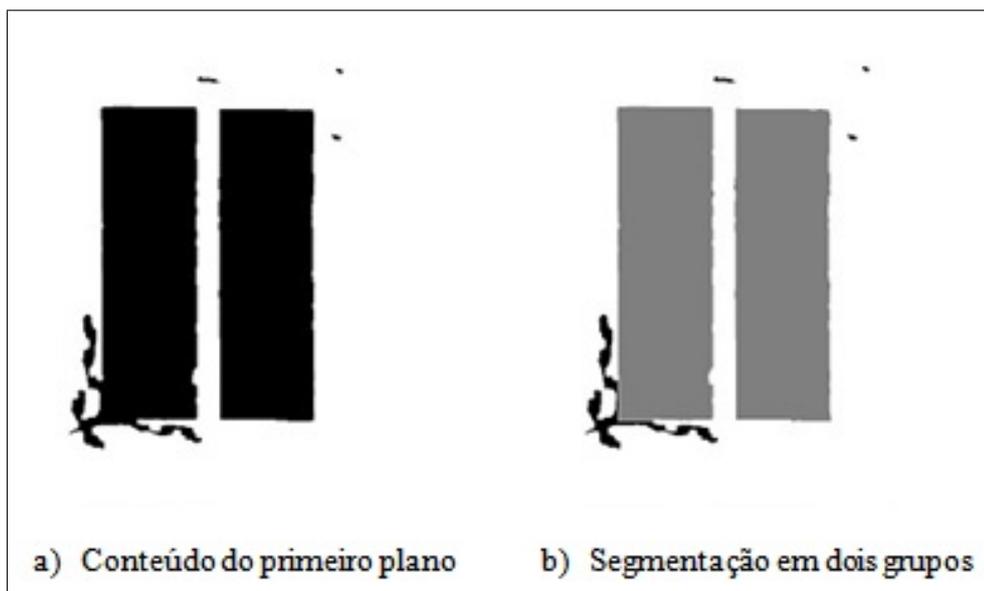


Figura 5.21: Ilustração da segmentação ideal dos dois grupos de conteúdos do primeiro plano.

página da Bíblia.

- Imagem (b), o resultado ideal esperado da aplicação do perfil de projeção horizontal.
- Imagem (c), o resultado ideal esperado da aplicação do perfil de projeção vertical.

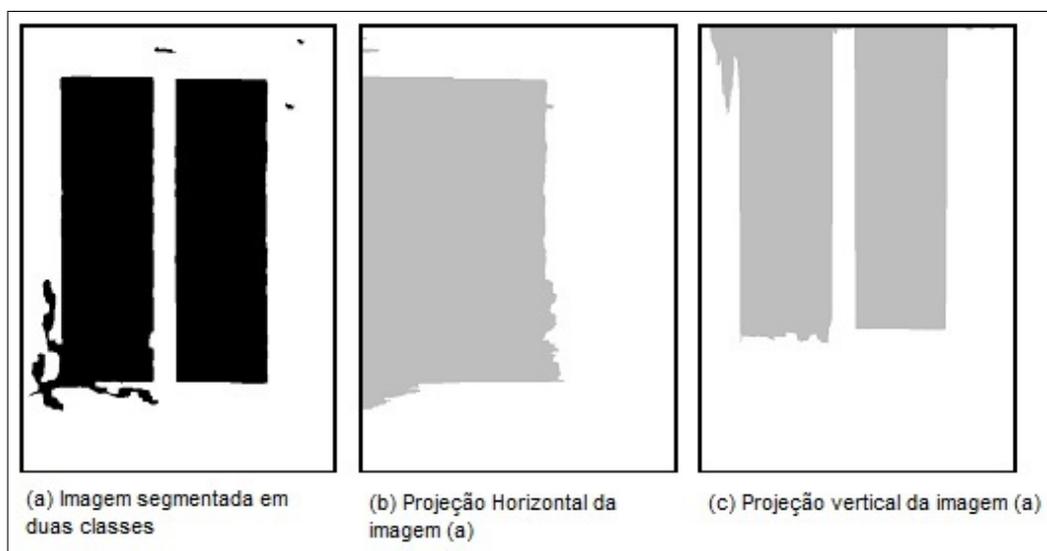


Figura 5.22: Ilustração da aplicação do perfil de projeção horizontal e vertical para segmentação ideal das colunas da imagem.

Com a aplicação do perfil de projeção é possível quantificar os pixels pretos das linhas e das colunas por meio de histogramas, ou seja, encontrar as linhas e colunas

com maior ocorrência de pixels pretos. Para segmentar regiões dos histogramas, pode-se utilizar dos picos dos histogramas ou da média das ocorrências de pixels pretos das linhas e das colunas. Para obtenção das colunas pode-se proceder da seguinte maneira:

- Segmentar as regiões com maior ocorrência de pixels pretos do histograma da projeção horizontal;
- Segmentar as regiões com maior ocorrência de pixels pretos do histograma da projeção vertical;
- Realizar a intersecção do resultado da segmentação do histograma de projeção horizontal com a segmentação do histograma de projeção vertical;

A Figura 5.23 apresenta a imagem ideal da intersecção das duas imagens segmentadas pelo histograma de perfil de projeção horizontal e vertical.

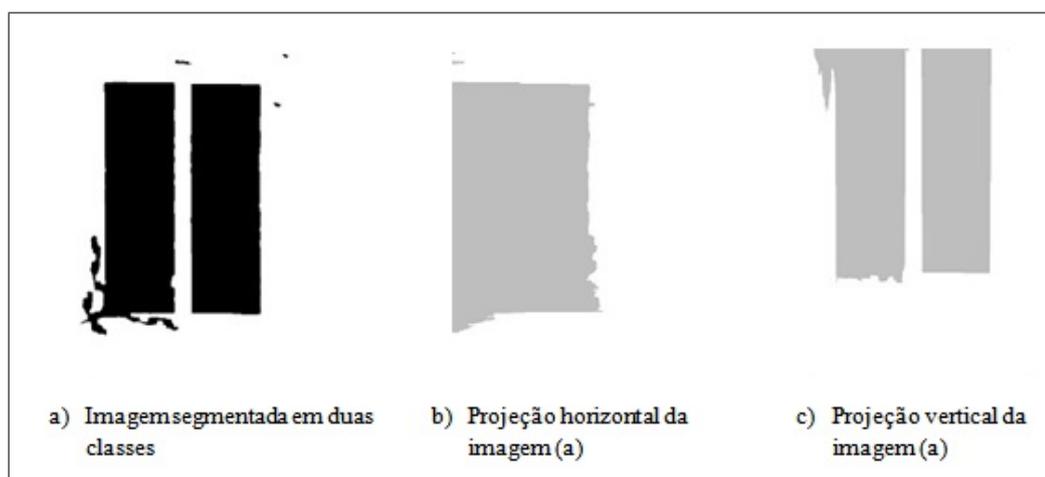


Figura 5.23: Ilustração da intersecção ideal dos histogramas dos perfis de projeções.

Assim, as duas colunas centrais são segmentadas, representando o conteúdo do grupo 2. Para obter o conteúdo do grupo 1 é preciso realizar uma operação de subtração entre as imagens (a) e (b) da Figura 5.24, sendo:

- A imagem (a), apresenta o resultado da binarização do conteúdo do primeiro plano;
- A imagem (b), o resultado da aplicação da técnica de perfil de projeção, com a identificação das duas colunas centrais.
- A imagem (c), representa a subtração, sendo o conteúdo esperado do grupo 1;

Os dois grupos foram segmentados de acordo com o nível 2 do método proposto, com a aplicação de operadores morfológicos e da técnica de perfil de projeção.

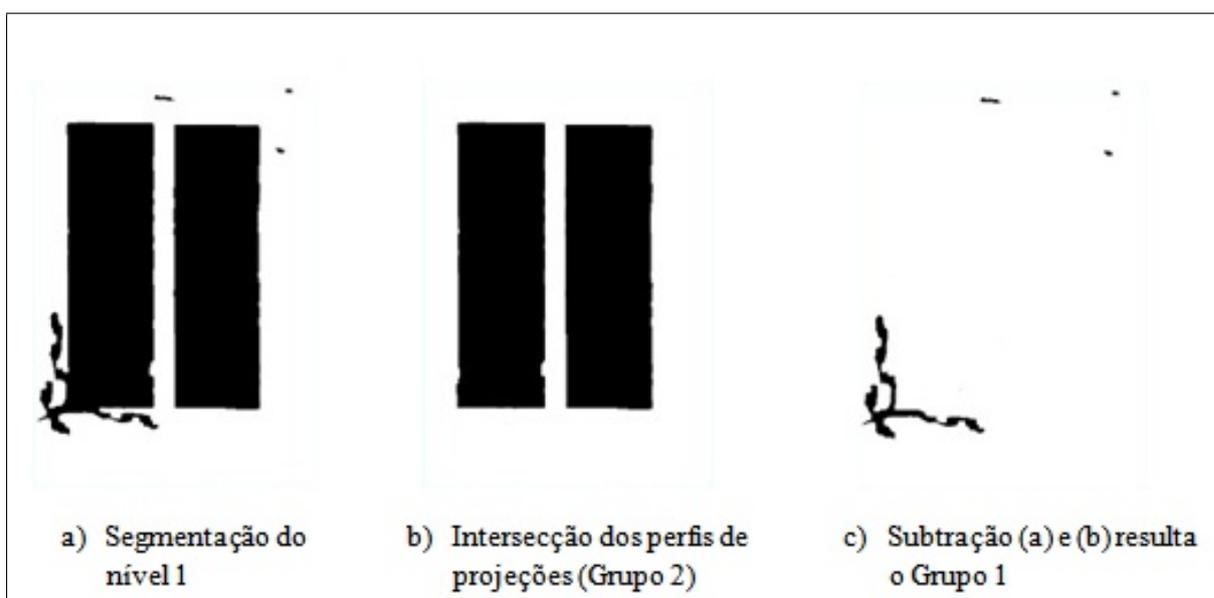


Figura 5.24: Ilustração da segmentação ideal do conteúdo do grupo 1 e do grupo 2.

### 5.2.3 Nível 3 - Segmentação do conteúdo do Grupo 1: figuras, títulos, numeração de páginas e manuscritos

As páginas da Bíblia possuem títulos, manuscritos, numeração de páginas e figuras, mas nem todas possuem as quatro informações. Assim, este nível descreve a segmentação deste grupo, conforme o método proposto.

#### 5.2.3.1 Segmentação de figuras, títulos, manuscritos e numeração de páginas

Para realizar a segmentação do conteúdo do grupo 1, pode-se partir de alguns princípios, que serão enumerados a seguir.

- As figuras podem ser identificadas pela quantidade de pixels conectados em relação a outros conjuntos do grupo, com intersecção com as duas colunas centrais;
- Os títulos podem ser identificados pela localização, posição na imagem, sempre na parte superior da imagem, acima das duas colunas centrais, respeitando o lado esquerdo pelo primeiro pixel da coluna esquerda, e do lado direito pelo último pixel da coluna da direita;
- Os manuscritos podem ser identificados pela quantidade de pixels e pela sua localização na imagem.
- A numeração das páginas está localizada no canto superior direito das páginas,

acima das duas colunas centrais, com início na área após o último pixel da coluna da direita e último pixel da borda da imagem do lado direito.

- Os ruídos estão localizados na borda da imagem, em virtude da aquisição da imagem, e são grupos menores de pixels.

A Figura 5.25 ilustra a segmentação ideal do conteúdo do grupo 1, sendo o título representado pela cor vermelha, o manuscrito pela cor verde, a numeração de páginas pela cor azul e o desenho pela cor preta.

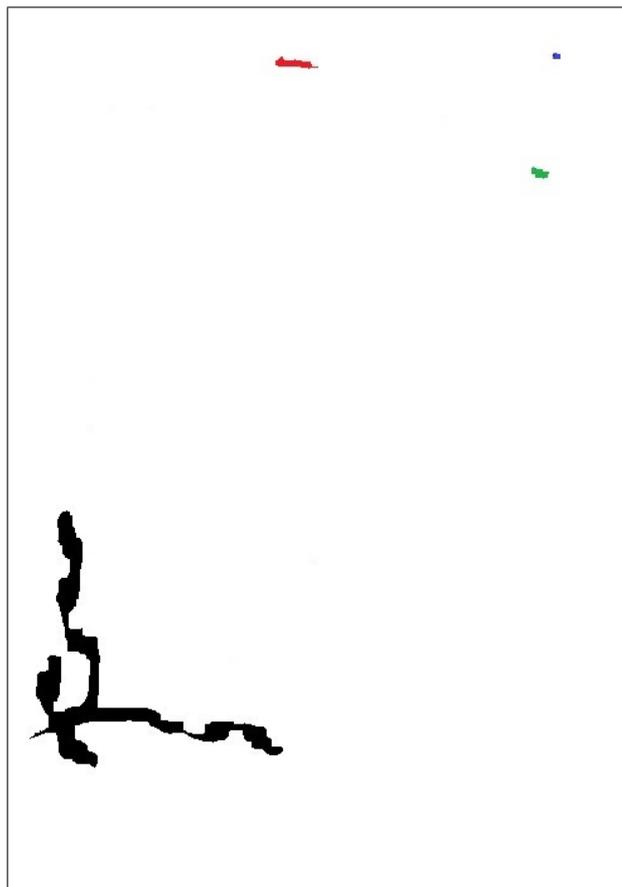


Figura 5.25: Ilustração do conteúdo do grupo 1: título, figura, numeração de páginas e manuscritos.

A Figura 5.26 ilustra o fluxograma do nível 3, da segmentação das figuras, títulos, numeração de páginas e manuscritos.

As figuras podem ter informações entre as duas colunas centrais segmentadas, tornando-se necessária a reconstrução destas para identificá-las corretamente as informações. Uma estratégia para recuperar as informações das figuras, que podem estar presentes nas duas colunas, é através da operação morfológica por reconstrução em cor, que pode recuperar todas as informações pertencentes às figuras. Com essa operação às figuras serão

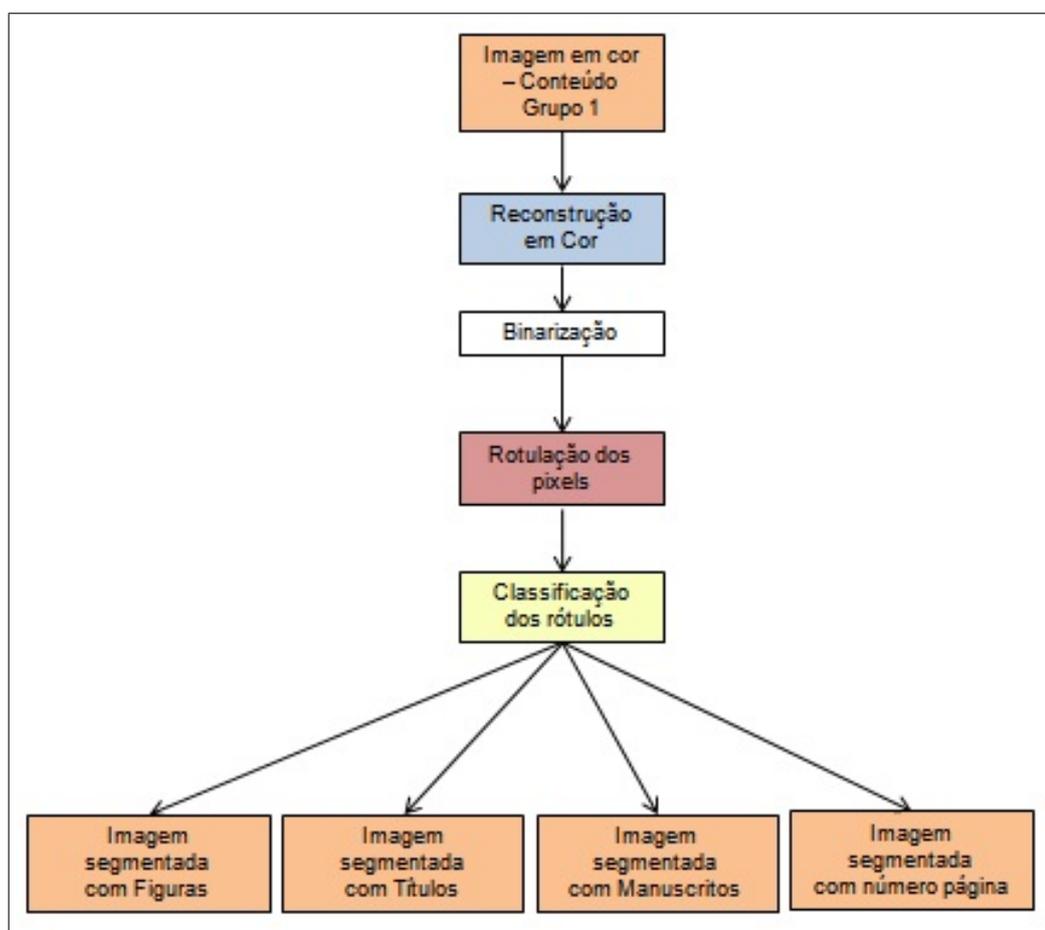


Figura 5.26: Fluxograma da segmentação das figuras, títulos, número de páginas e manuscritos.

reconstruídas. Esta é uma maneira de identificar o que são figuras na imagem e o que são títulos e manuscritos, pois as figuras possuem informações a serem reconstruídas e pertencem às duas colunas centrais.

A Figura 5.27 ilustra a operação de reconstrução em cor, sendo a imagem (a) original, (b) os pixels a serem reconstruídos, a (c) é a máscara (imagem (a) invertida), a (d) a marcadora (imagem(b) invertida), a (e) a imagem reconstruída invertida e (f) os pixels reconstruídos.

Com a operação de reconstrução é possível recuperar as figuras da imagem, os títulos, os manuscritos, as numerações de páginas e o fundo colorido. Assim, o próximo passo consiste na binarização dos pixels para aplicação da técnica de rotulação.

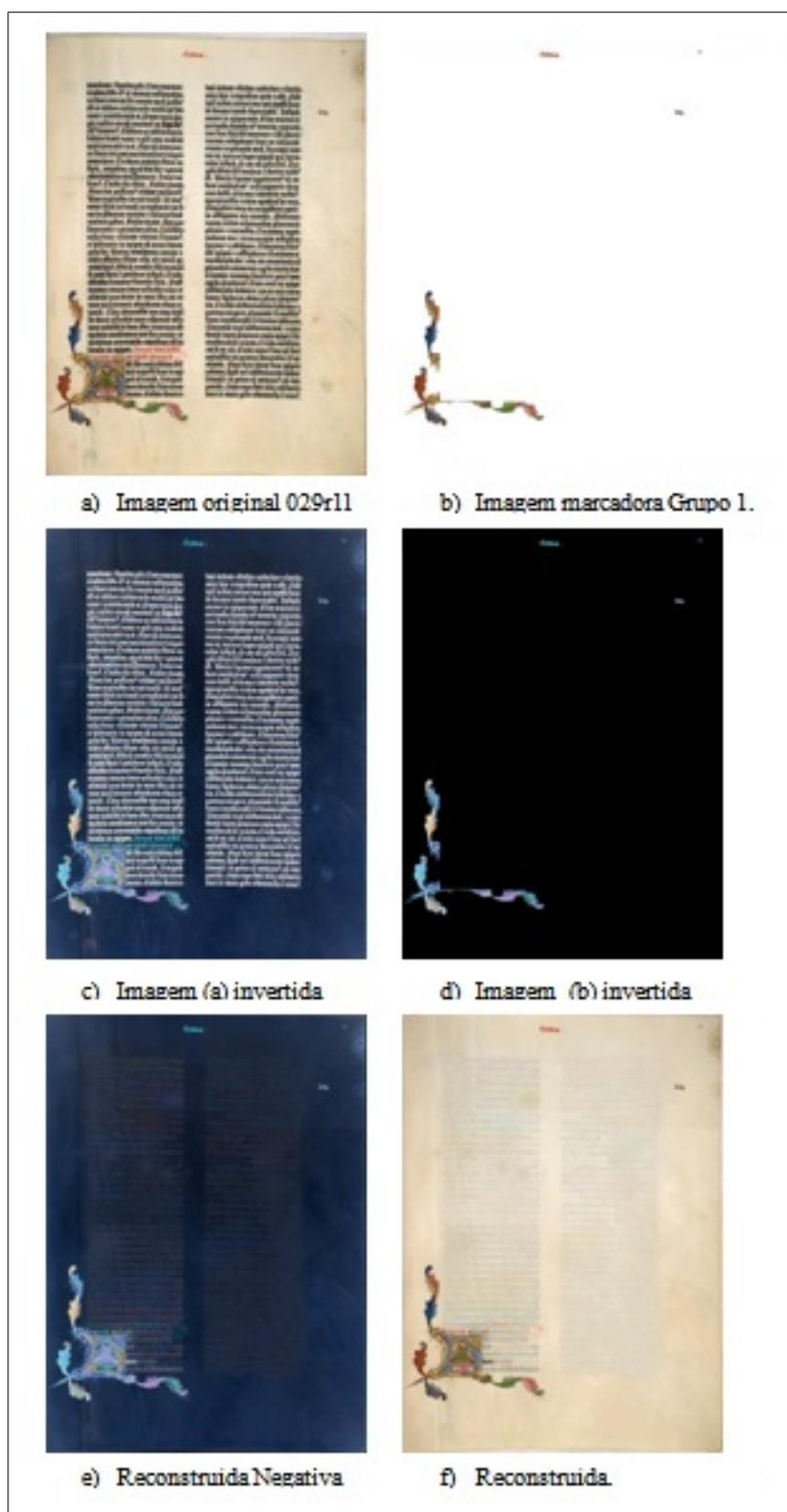


Figura 5.27: Ilustração dos pixels originais da reconstrução do grupo 1.

### 5.2.3.2 Técnica de Rotulação

Esta técnica é utilizada para identificar as regiões da imagem, sendo atribuídos rótulos para os pixels que estejam ligados e que representam conjuntos unidos dentro das imagens. Para aplicar a técnica de rotulação (GOSE; JOHNSONBAUGH; JOST, 1996), os blocos da imagem binária devem estar compactos, pois uma imagem com ruídos acarreta muitos rótulos, dificultando a identificação das caixas *bounding box* e da classificação das informações corretas em títulos, manuscritos, números de páginas e figuras. Para realçar e compactar os blocos, é realizada a operação morfológica binária abertura. A rotulação ideal é ilustrada pela Figura 5.28, que representa a rotulação de uma página da Bíblia em 04 rótulos, sendo o rótulo 1 na cor vermelha, o 2 em azul, o 3 em verde e o 4 na cor preta.

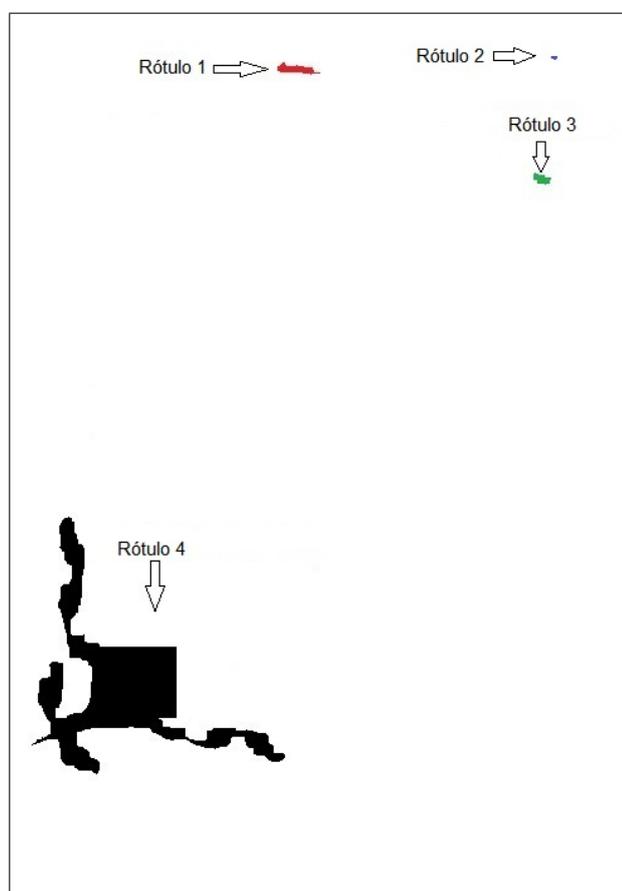


Figura 5.28: Exemplo do grupo 1 rotulada ideal.

Com os pixels da imagem rotulados, o próximo passo é delimitar a caixa de cada rótulo para calcular a quantidade de pixels presentes em cada um, e a localização de cada caixa de rótulos para classificação dos títulos, manuscritos, número de páginas e figuras, conforme as estratégias de classificação.

### 5.2.3.3 Técnica de Bounding Box

A técnica consiste em identificar as caixas de cada rótulo, ou seja, delimitar uma área onde os pixels correspondentes a cada rótulo estão localizados, para assim classificar as caixas pela localização e quantidade de pixels. A Figura 5.29 ilustra as *bounding box* dos rótulos da imagem rotulada.

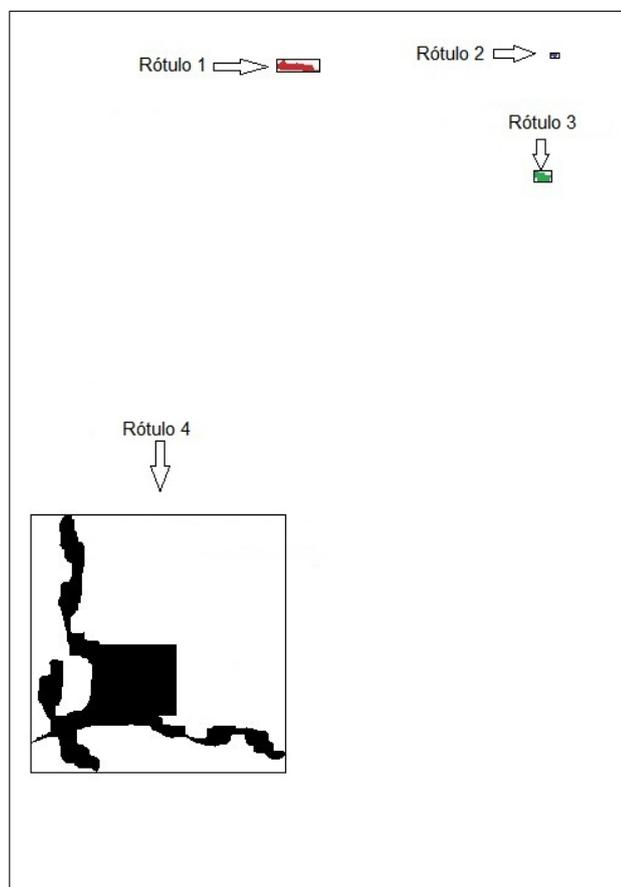


Figura 5.29: Exemplo das bounding box rotuladas para classificação.

### 5.2.3.4 Separação das figuras, títulos, número de páginas e manuscritos

Os títulos e numeração de páginas estão na parte superior das páginas, sendo possível delimitar uma área para essas duas informações (ilustrado pela Figura 5.30), sempre acima das duas colunas centrais. As figuras serão identificadas pela intersecção das informações com pixels pertencentes às duas colunas centrais. Os manuscritos estão presentes nas demais áreas, sem intersecção com as duas colunas centrais. A Figura 5.30 ilustra a posição dos possíveis títulos, número de páginas, figuras e manuscritos presentes nas imagens da Bíblia.

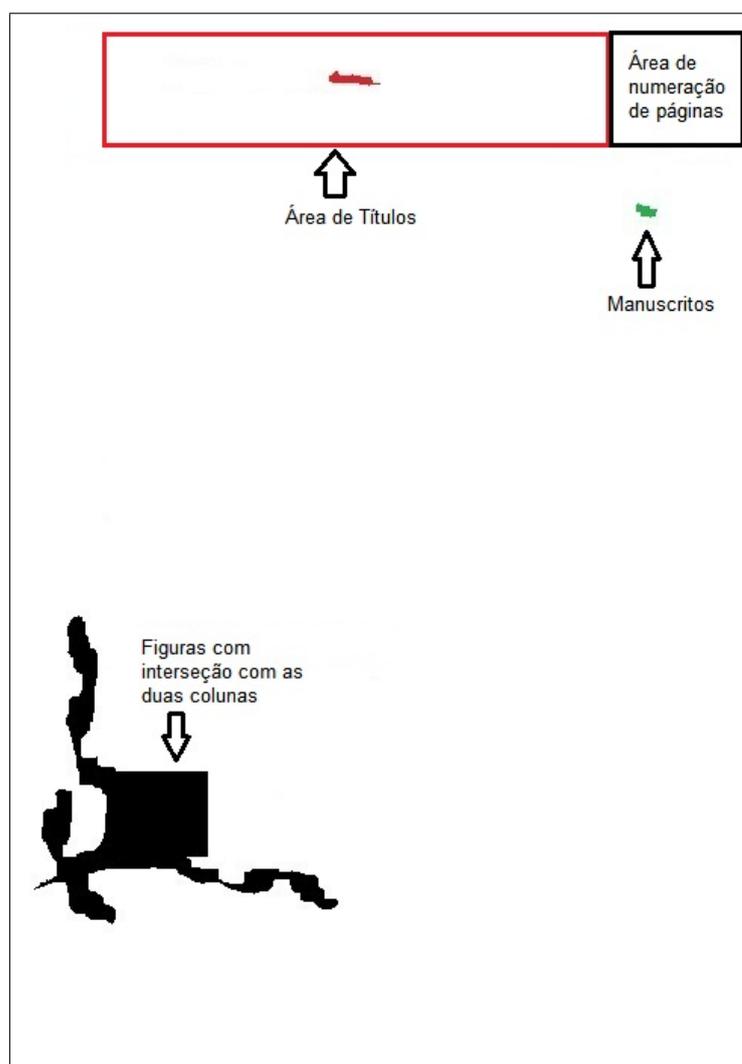


Figura 5.30: Ilustração da localização dos títulos, número de páginas, manuscritos e figuras.

Esta seção apresentou a segmentação das figuras, títulos, número de páginas e manuscritos por meio das operações da morfologia matemática, técnicas de rotulação, *bounding box* e conhecimento *a priori*.

Assim os níveis foram detalhados conforme o método proposto e pode-se medir a eficiência através de avaliações da classificação de cada nível da metodologia por meio das métricas de avaliação, das imagens *ground truth* e imagens sintéticas.

A Figura 5.31 ilustra a identificação ideal do conteúdo do Grupo 1, sendo a imagem (a) o título, (b) número de páginas, (c) manuscritos e (d) figuras.

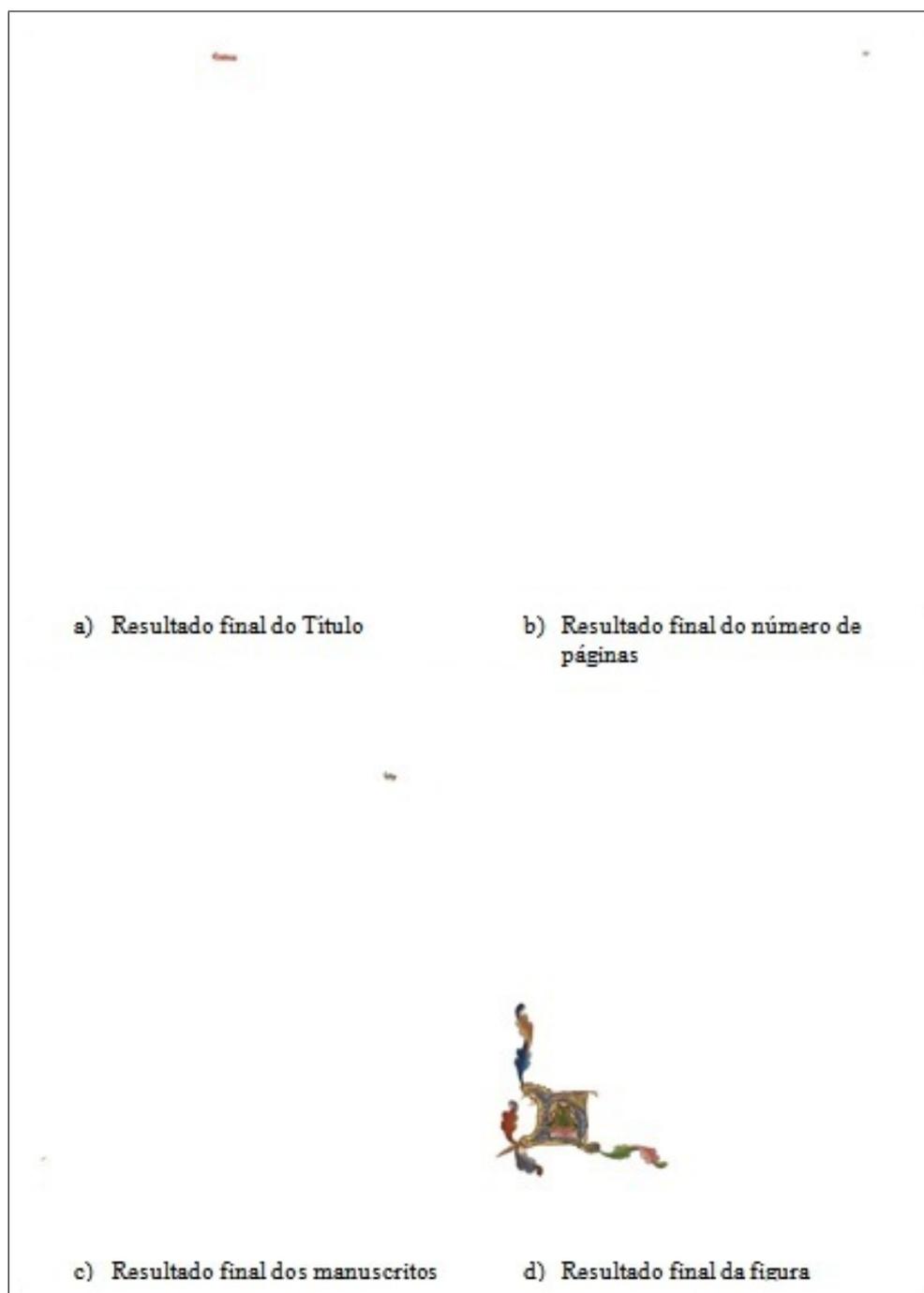


Figura 5.31: Ilustração da segmentação ideal do conteúdo do grupo 1: Título, número de páginas, manuscritos e figura.

## 5.3 Avaliação do método proposto

Os níveis do método proposto devem ser avaliados e medidos pelas métricas já mencionadas na fundamentação teórica. Neste trabalho a forma de avaliação será qualitativa e quantitativa através da geração de imagens *groundtruth* e imagens sintéticas.

### 5.3.1 Geração das imagens *ground truth*

As imagens *ground truth* são utilizadas para medir a eficiência da classificação por meio das métricas, podendo ser geradas de forma automática, manual ou semi-automática e fazem parte da avaliação da eficiência do método proposto. Para os experimentos optou-se optar pela abordagem de Baird e Moll (2008), que gera imagens *ground truth* de forma manual por um "Zoner" (profissional da área), por blocos. Para avaliação do método proposto, serão geradas imagens para os níveis em que é possível realizar a quantificação dos resultados. Algumas observações devem ser consideradas na condução do método.

**Nível 1 - Processo 1: Estimativa do fundo-** não é possível gerar uma imagem *ground truth* para avaliação. Pode-se estimar a quantidade de pixels do primeiro plano, que foram extraídos (recuperados), e desprezar o fundo estimado. Neste caso, o processo é somente baseado na estimativa do fundo reconstruído, sendo, portanto, uma análise qualitativa, observando-se nos testes os resultados da reconstrução dos fundos das imagens de documentos antigos e as imagens geradas com o fundo estimado.

**Nível 1 - Processo 2: Segmentação do conteúdo do primeiro plano -** não existem imagens *ground truth* verdadeiras com um fundo branco e todas as informações do primeiro plano intactas, em virtude de as páginas serem antigas e o processo manual poder levar a erros de geração, pois as letras das duas colunas são escritas em latim, muito complexas, as entrelinhas são pequenas e letras muito conectadas. Assim para esta etapa será possível medir quantitativamente os resultados através da geração de uma imagem sintética (artificial) parecida com uma página da Bíblia, com o fundo de uma página da Bíblia real (ou aproximada) e o conteúdo do primeiro plano sendo gerado por um "Zoner", de forma manual com o uso de uma fonte gótica. A Figura 5.32 ilustra um exemplo da imagem sintética gerada, sendo a (a) o fundo de uma imagem da Bíblia, a (b) a imagem *ground truth* e a (c) a imagem sintética gerada com fundo da Bíblia e conteúdo do primeiro plano gerada artificialmente com fonte gótica. O desenho utilizado nas informações do primeiro plano foi recortado de uma imagem verdadeira da base da Bíblia de Gutenberg.

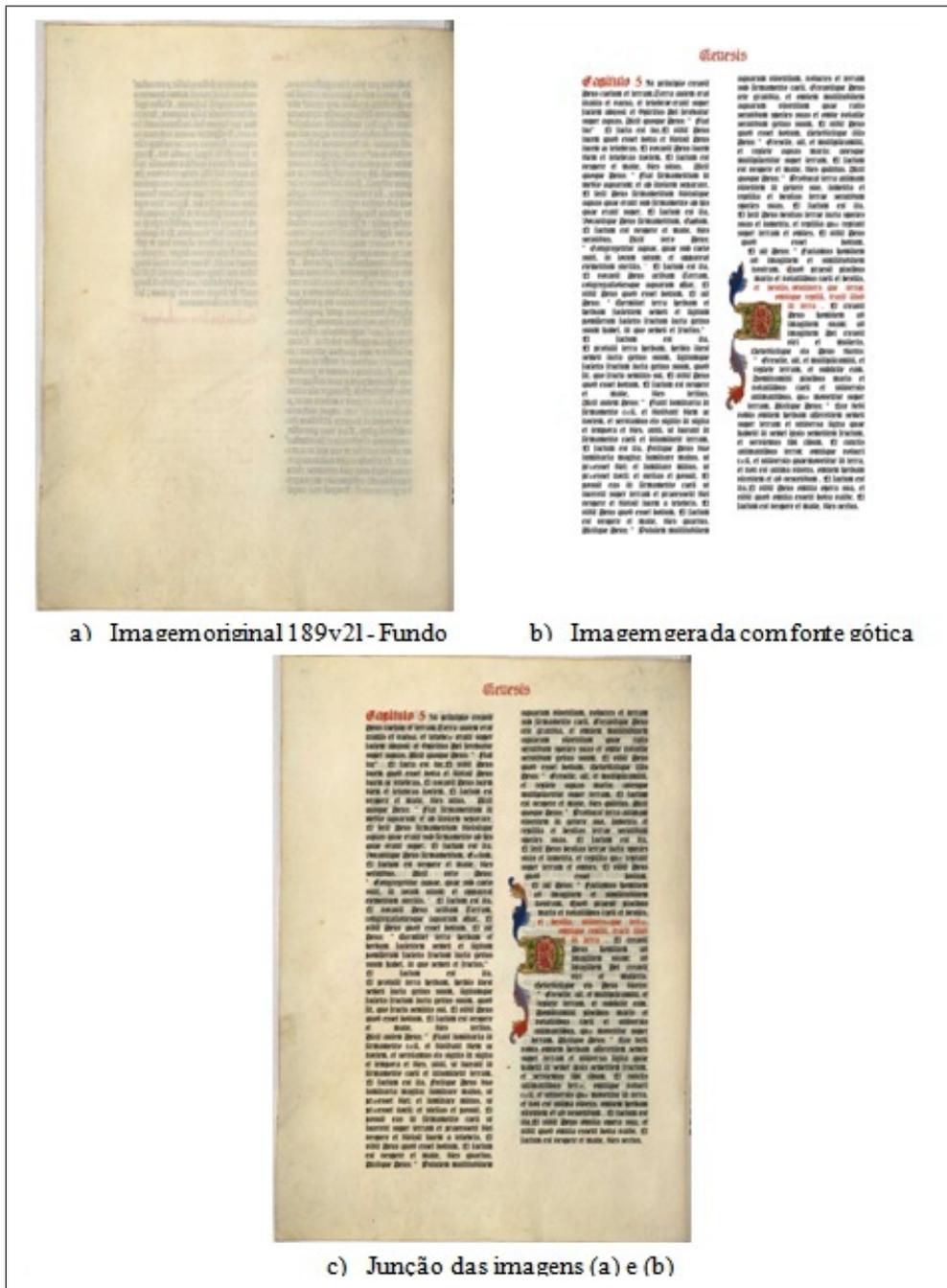


Figura 5.32: Exemplo da geração de uma imagem sintética.

Nível 1 - Processo 3: Segmentação do texto colorido - - a imagem *ground truth* para a análise quantitativa foi gerada com uma imagem sintética de forma manual com auxílio do Photoshop. A Figura 5.33 ilustra em (a) uma imagem sintética gerada e em (b) a imagem *ground truth* desta, com as seguintes informações do primeiro plano: título, figuras, letras coloridas e letras pequenas vermelhas.

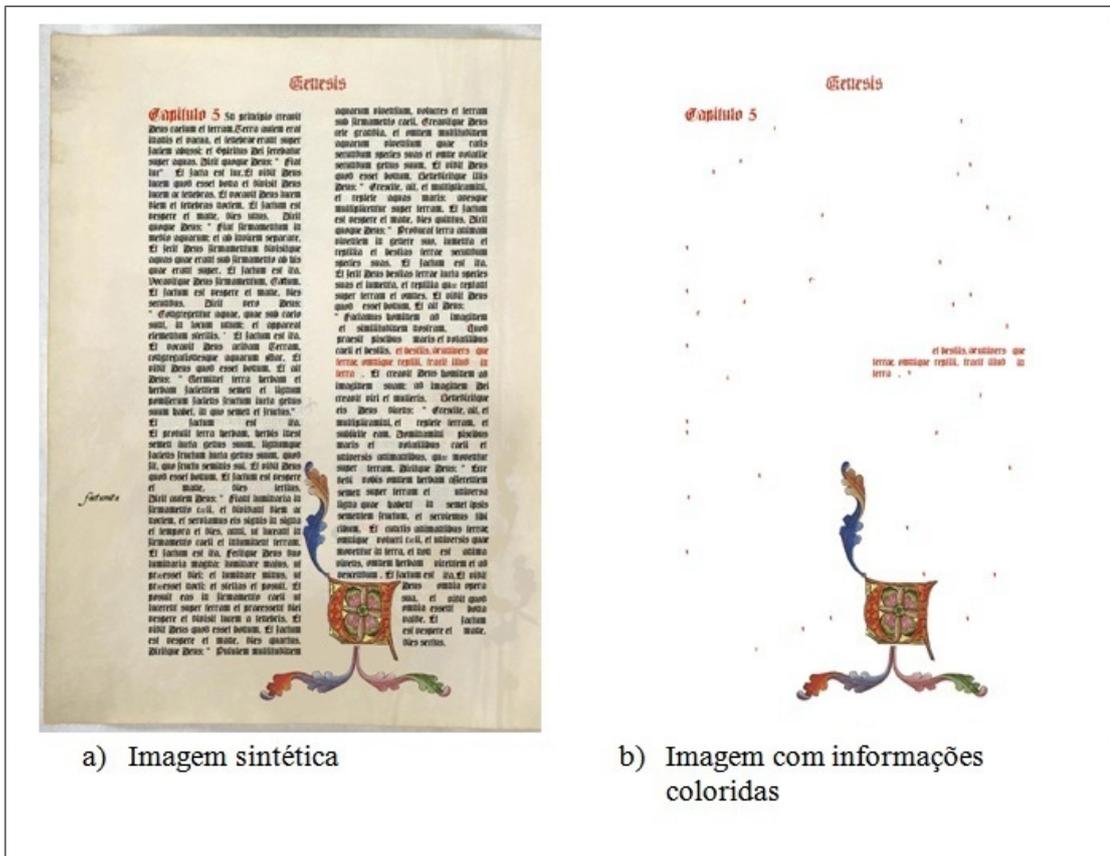


Figura 5.33: Exemplo da geração de uma imagem sintética e sua imagem *ground truth*.

A Figura 5.34 ilustra a imagem sintética gerada em (a), a imagem (b) com as informações somente da parte colorida sem o fundo e sem as letras pretas, e a imagem (c) o resultado da imagem (b) reconstruída, sendo a imagem (c) considerada a imagem *ground truth* que será utilizada nos experimentos.

Portanto, para o Nível 1, Processo 3, têm-se duas imagens *ground truth*.

- Primeira Imagem: representada pela imagem (b) da figura 5.33, somente com as pequenas letras vermelhas, título, letras vermelhas e desenho, com fundo branco.
- Segunda Imagem: representada pela imagem (c) da figura 5.34 com um fundo reconstruído, título, letras vermelhas, desenho e as pequenas letras vermelhas com contornos.

Assim, essas duas estratégias de geração de imagens *ground truth* serão utilizadas nos experimentos.



Figura 5.34: Ilustração da geração da imagem *ground truth* para análise quantitativa.

**Nível 2 - Processo 4: Segmentação em dois grupos** - para este nível será possível gerar a imagem *ground truth* para identificar o conteúdo dos dois grupos da imagem da Bíblia. A Figura 5.35 ilustra um exemplo da imagem *ground truth* em dois grupos.

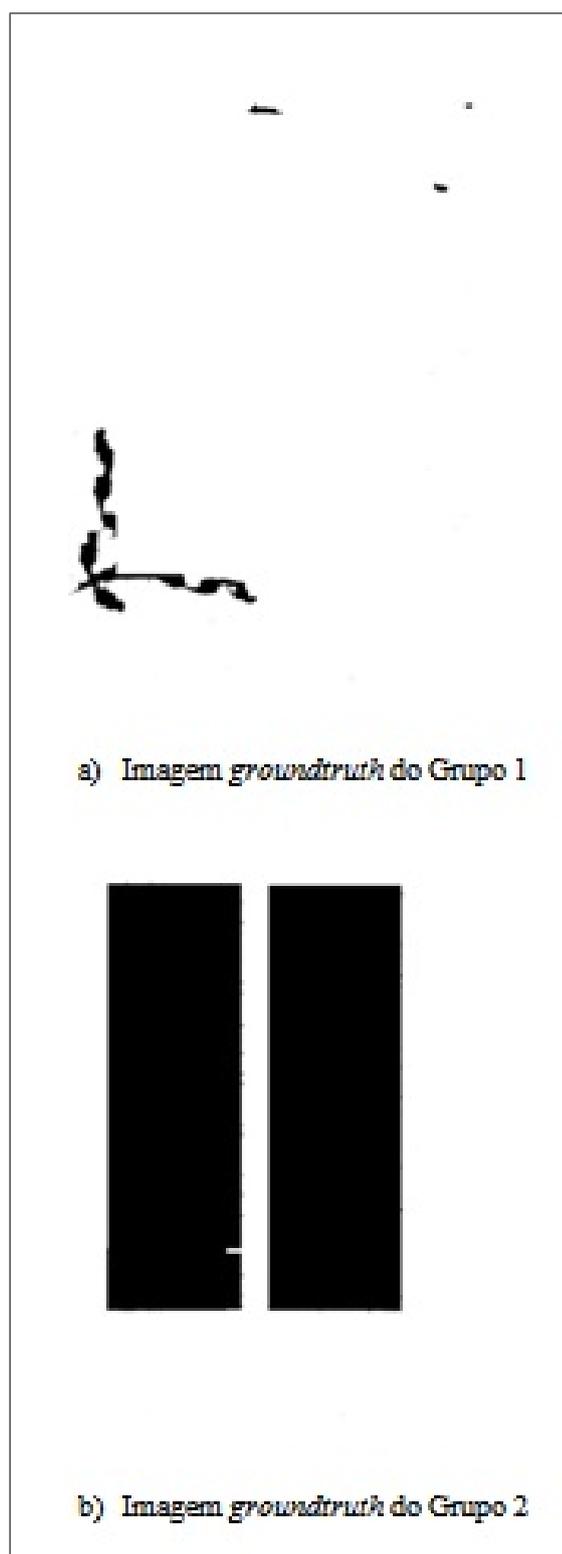


Figura 5.35: Exemplo da geração de uma imagem *ground truth* dos Grupos 1 e 2.

**Nível 3 - Processo 5: Segmentação do Grupo 1** - para este nível será gerada a imagem *ground truth* para a segmentação do grupo 1 (figuras, títulos, número de páginas e manuscritos). A Figura 5.36 ilustra um exemplo da imagem *ground truth* da segmentação do grupo 1.

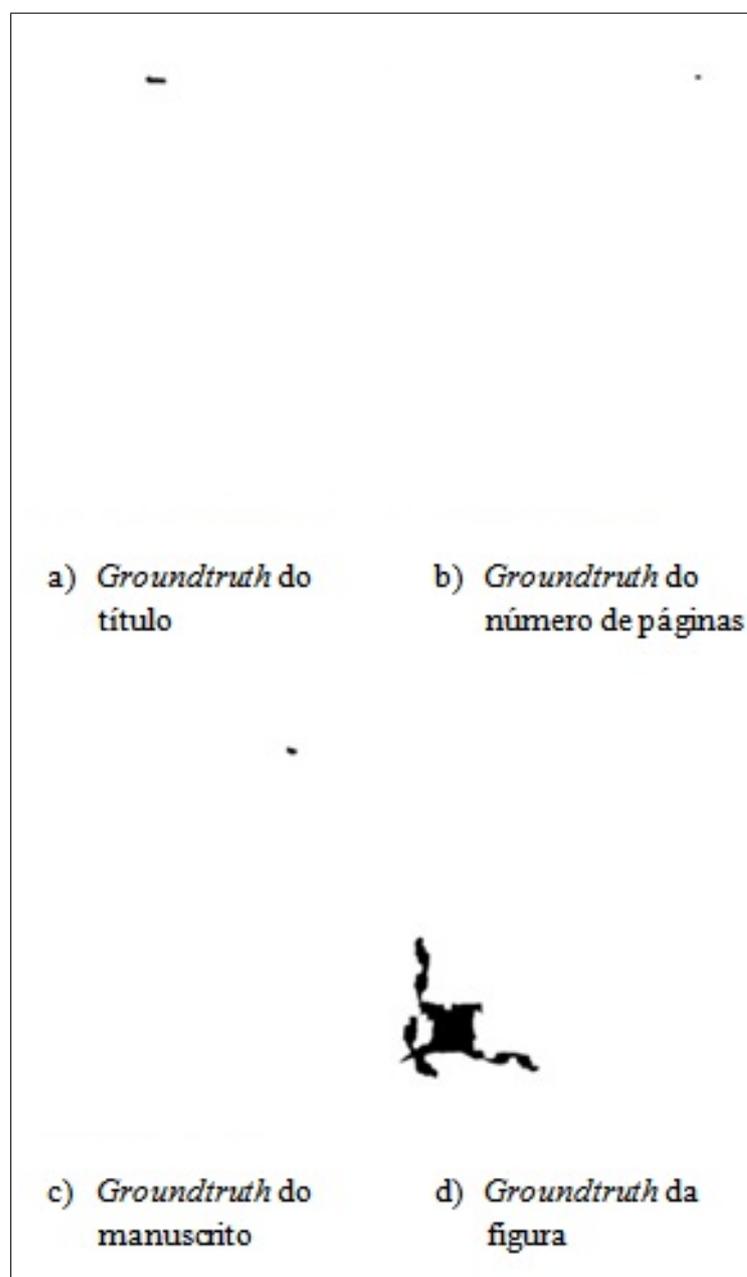


Figura 5.36: Exemplo da geração da imagem *ground truth* para o grupo 1.

Com a geração das imagens *ground truth*, o próximo passo será utilizar as métricas para realizar a avaliação do método proposto em cada nível possível de quantificação.

### **5.3.2 Avaliação da classificação baseado nas métricas**

Para medir a eficiência e avaliar o desempenho dos níveis do método proposto, serão utilizadas as métricas apresentadas no capítulo 3.

## **5.4 Considerações**

Neste capítulo, foram apresentados os níveis do método proposto para a estimativa e segmentação das páginas da Bíblia de Gutenberg, utilizando técnicas da morfologia matemática binária, cinza e em cor, mostrando os resultados para cada nível e as formas de avaliação, qualitativa e quantitativa, utilizadas nos experimentos. O próximo capítulo apresenta os experimentos realizados, com os resultados alcançados, utilizando as imagens da Base de Dados da Bíblia de Gutenberg e as sintéticas geradas.

## Capítulo 6

### Resultados Experimentais

Neste capítulo, são apresentados os testes realizados e os resultados obtidos pelo método proposto. O capítulo está organizado da seguinte forma: na seção 6.1, apresenta-se a base de dados utilizada; na seção 6.2 são apresentados os resultados dos níveis do método proposto: I) nível 1, com a estimativa do fundo, segmentação do conteúdo do primeiro plano, das letras coloridas e das figuras; II) apresentação dos resultados do nível 2, e III) os resultados do nível 3.

#### 6.1 Base de Dados - Bíblia de Gutenberg

Para a realização dos experimentos foi utilizada a base de dados da primeira Bíblia impressa por Gutenberg, composta por 1282 imagens digitalizadas, adquirida do museu Gutenberg em Mainz - Alemanha. As imagens, com dimensões de 965x1390 *pixels* no formato JPEG, resolução de 72dpi e faixa tonal de 24 bits, perderam muitas informações por estar neste formato compactado. A Figura 6.1 ilustra o recorte de um título de uma destas páginas, demonstrando o efeito do formato JPEG, na qual foram percebidas distorções geradas por este formato, estando os pixels ao redor do título, em vermelho, deteriorado, não sendo visível a tonalidade pura dos pixels tanto ao redor do título quanto no próprio título, dificultando muito o processamento das imagens.

As várias características que as imagens apresentavam foram analisadas e podem ser resumidas em 06 blocos de categorias, nos quais são relacionados, respectivamente, sua representatividade (em porcentagem) na base de dados utilizada.

1. Páginas com grafismos (desenhos), duas colunas, título e manuscritos. Essas páginas foram consideradas complexas pela quantidade de informações presentes. Na base de dados representavam aproximadamente 8%



Figura 6.1: Exemplo do Título de imagem da Bíblia no formato JPEG.

2. Páginas com grafismos (desenhos), duas colunas e título; Na base de dados aproximadamente 4 % .
3. Páginas com título, duas colunas, sem desenhos e com manuscritos; Na base de dados aproximadamente 22 % .
4. Páginas com título, duas colunas, sem desenhos e sem manuscritos; Na base de dados aproximadamente 61 % .
5. Páginas sem título, com duas colunas, sem desenhos e sem manuscritos; Na base de dados aproximadamente 5 % .
6. Páginas sem informações, somente 02 do total de 1282 páginas.

A Figura 6.2 ilustra as seis imagens que representaram as categorias de páginas encontradas na base de dados, sendo a imagem (a) a categoria 1, (b) a categoria 2, (c) a categoria 3, (d) a categoria 4, (e) a categoria 5 e (f) a categoria 6. Esta divisão ajudou no processamento e na condução dos testes.

A base de dados da Bíblia apresentou alguns problemas, que foram analisados.

- O fundo não é igual entre as imagens, não possuindo uma uniformidade, nota-se pelos exemplos da Figura 6.2 a diferença na cor do fundo das imagens.
- A posição dos títulos não segue um padrão, eles não ficavam na mesma posição na imagem e existiam páginas com dois, três ou quatro títulos.
- Problemas na aquisição da imagem, no processo de digitalização, em que o fundo do *scanner* aparece digitalizado.

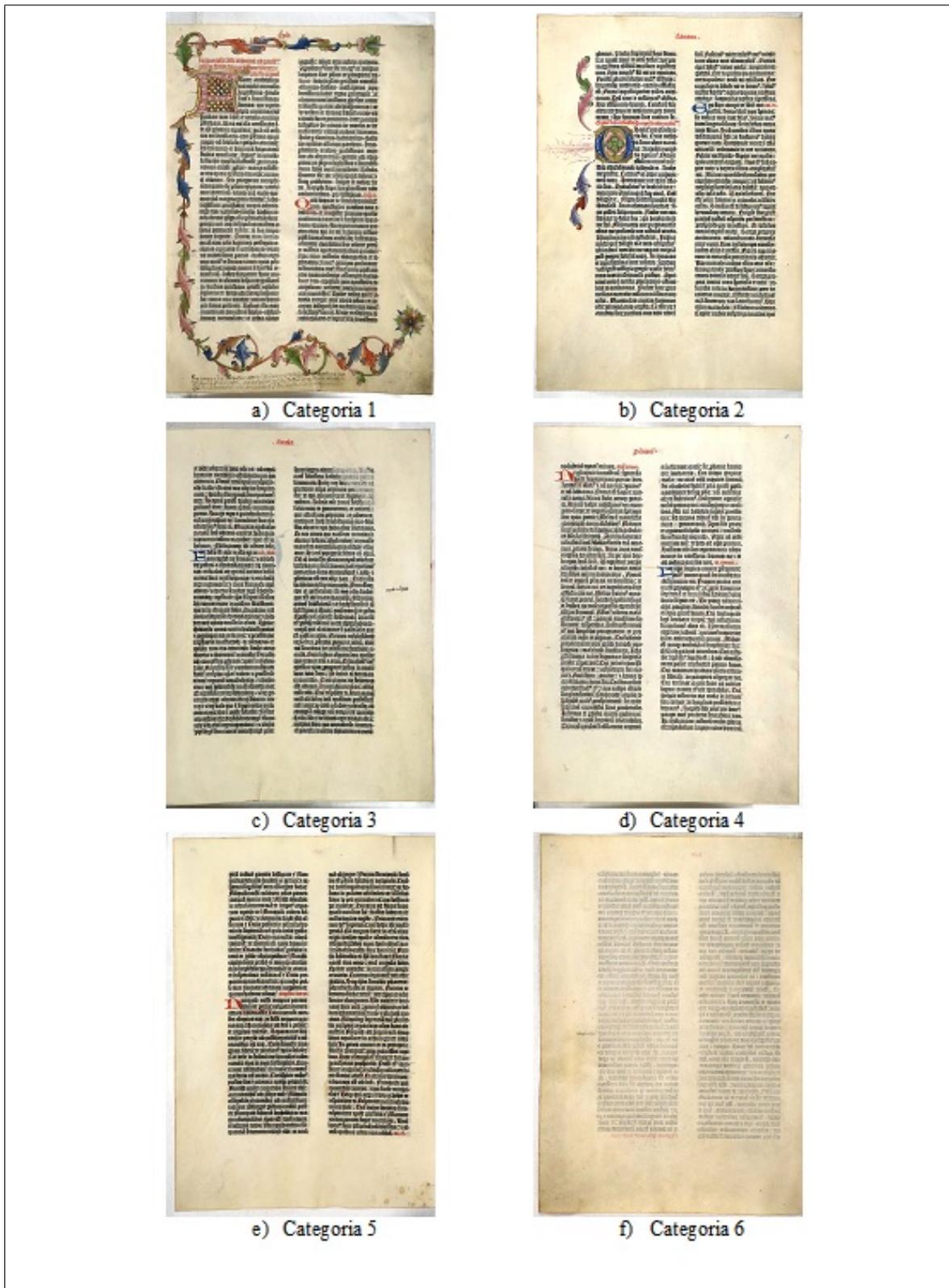


Figura 6.2: Exemplos das categorias das páginas da Bíblia.

- Os desenhos não estão localizados na mesma posição da imagem e apresentam tamanhos diferentes.

- Os manuscritos não estavam presentes em todas as páginas nem localizados na mesma posição da imagem. Algumas páginas apresentavam manuscritos que determinavam a numeração da página.
- As imagens adquiridas estavam no formato JPEG (perdendo informações se comparadas com o formato Bitmap).

Assim, a base de dados da Bíblia era composta por imagens complexas, com problemas e foram utilizadas para os testes do método proposto.

## 6.2 Experimentos

Nesta seção são apresentados os experimentos realizados para os três níveis do método proposto, utilizando as páginas da base de dados e as páginas sintéticas geradas, além da linguagem de programação JAVA para implementação e testes. Como nas imagens reais selecionadas havia a variância na quantidade de títulos, no aparecimento de figuras, números de páginas ou manuscritos, elas foram divididas em 03 conjuntos, de forma manual.

- Conjunto 1: 40 imagens com figuras (letras capitulares), sendo 22 com figuras e títulos e 18 com figuras, títulos e manuscritos;
- Conjunto 2: 40 imagens com manuscritos e com títulos;
- Conjunto 3: 40 imagens com títulos, sem manuscritos e sem figuras;

Para os testes dos processos 1 e 5 foram utilizadas as 120 imagens reais selecionadas da Base de dados da Bíblia de Gutenberg; para os processos 2 e 3, foram utilizadas as imagens sintéticas, geradas por processo manual, com a variação do fundo, das informações do primeiro plano e da opacidade das imagens, totalizando 160 imagens diferentes, para validar os resultados, uma vez que, para comparação dos resultados obtidos, havia necessidade de imagens ideais, com um fundo branco e informações somente do primeiro plano. Para o processo 4 (aplicação da técnica do perfil de projeção), não foi realizada uma avaliação numérica, que foi dada pelo processo 5, uma vez que 4 auxiliou o 5 com as informações somente do conjunto do grupo 1, os dados referentes ao título, possíveis manuscritos, numeração de páginas e figuras.

Nas próximas seções são descritos os experimentos dos níveis do método proposto.

## 6.3 Experimentos do Nível 1

Nessa etapa foram realizados os testes de acordo com os três processos do método proposto para o nível 1.

- Processo 1: estimativa do fundo baseado na imagem reconstruída.
- Processo 2: segmentação do conteúdo do primeiro plano.
- Processo 3: segmentação do conteúdo colorido.

### 6.3.1 Processo 1: Estimativa do fundo

Esta etapa consistiu na estimativa do fundo baseada na imagem reconstruída do fundo. Para realizar a operação da reconstrução em cor é preciso escolher um espaço de cor e a ordenação dos canais. Foram testados os seguintes espaços de cor: HSI, HSL, YCrCb e YIQ, com a ordenação lexicográfica, e o uso de 1, 2 e 3 canais, com as suas ordens para o processamento. A seguir estão descritos os resultados obtidos das imagens reconstruídas da estimativa do fundo para cada espaço de cor e suas ordenações.

#### a) Resultados com Espaço de Cor HSI

Para o processamento da operação de reconstrução com a ordenação lexicográfica, foi utilizado o valor de referência  $H$  majoritário da imagem. Por exemplo, para os testes com a imagem *029r1l*, o valor encontrado para  $H$  foi de  $220^\circ$ ; para o espaço de cor HSI, a ordem dos canais com os resultados estão apresentados a seguir.

- **Ordens com uso de 3 canais: HSI, SHI, ISH, IHS, HIS e SIH**

A Figura 6.3 ilustra os resultados com uso dos 03 canais, apresentando a variabilidade dos resultados.

A Figura 6.4 ilustra os resultados com o uso dos 03 canais, com a escolha do canal H na primeira posição, com ordem HSI, com os seguintes valores de H:  $0^\circ$ ,  $120^\circ$ ,  $240^\circ$  e  $360^\circ$ .

- **Ordenações com uso de 2 canais: HS, SI, HI, IH, IS e SH:** a Figura 6.5 ilustra os resultados com uso de 02 canais, apresentando a variabilidade dos resultados.
- **Ordenações com uso de 1 canal: H, S e I:** a Figura 6.6 ilustra os resultados com uso de somente 01 canal, mostrando a variabilidade dos resultados.



Figura 6.3: Resultado da reconstrução do fundo com uso de 03 canais HSI.

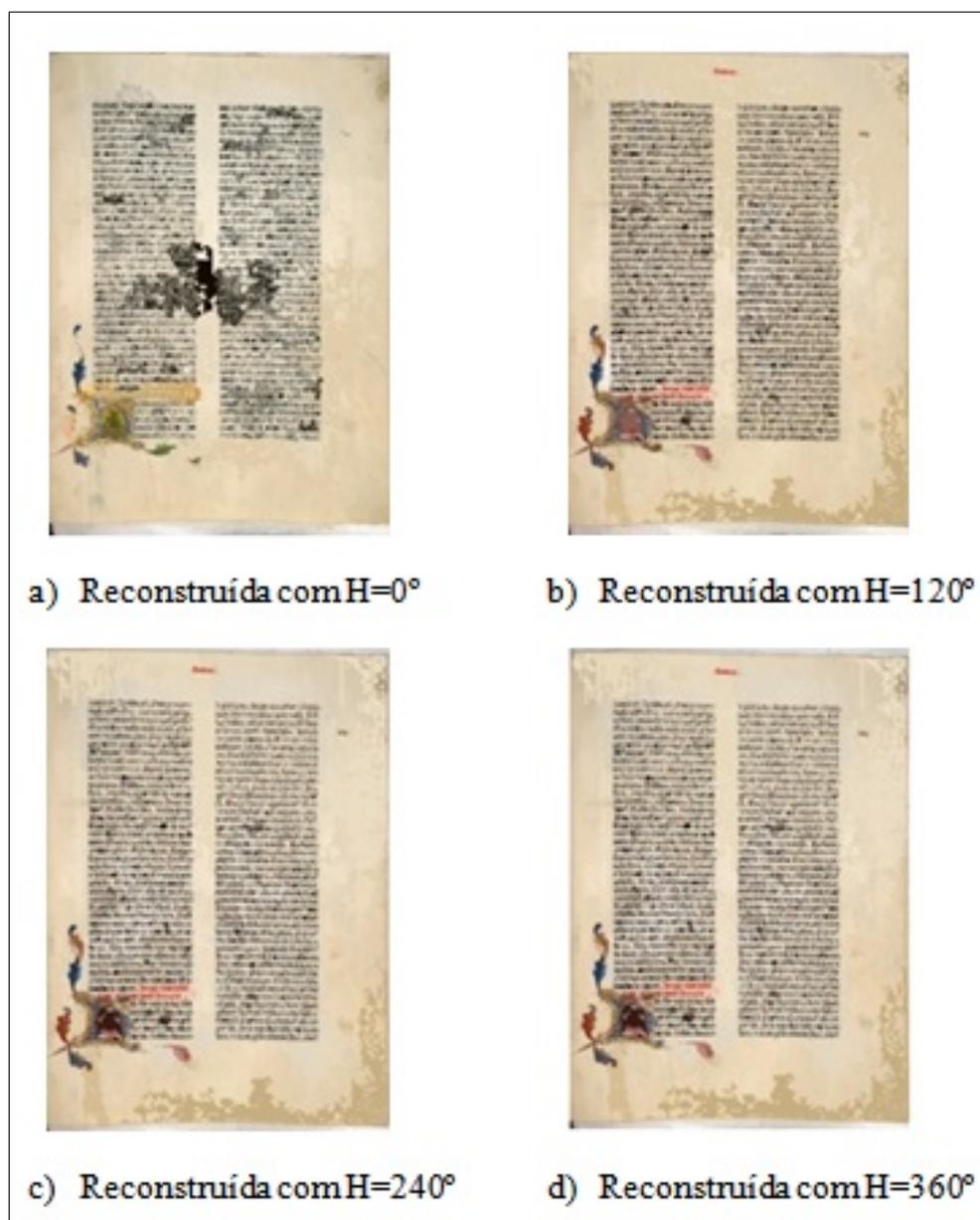


Figura 6.4: Resultado da reconstrução do fundo com uso de 03 canais HSI com H variando.

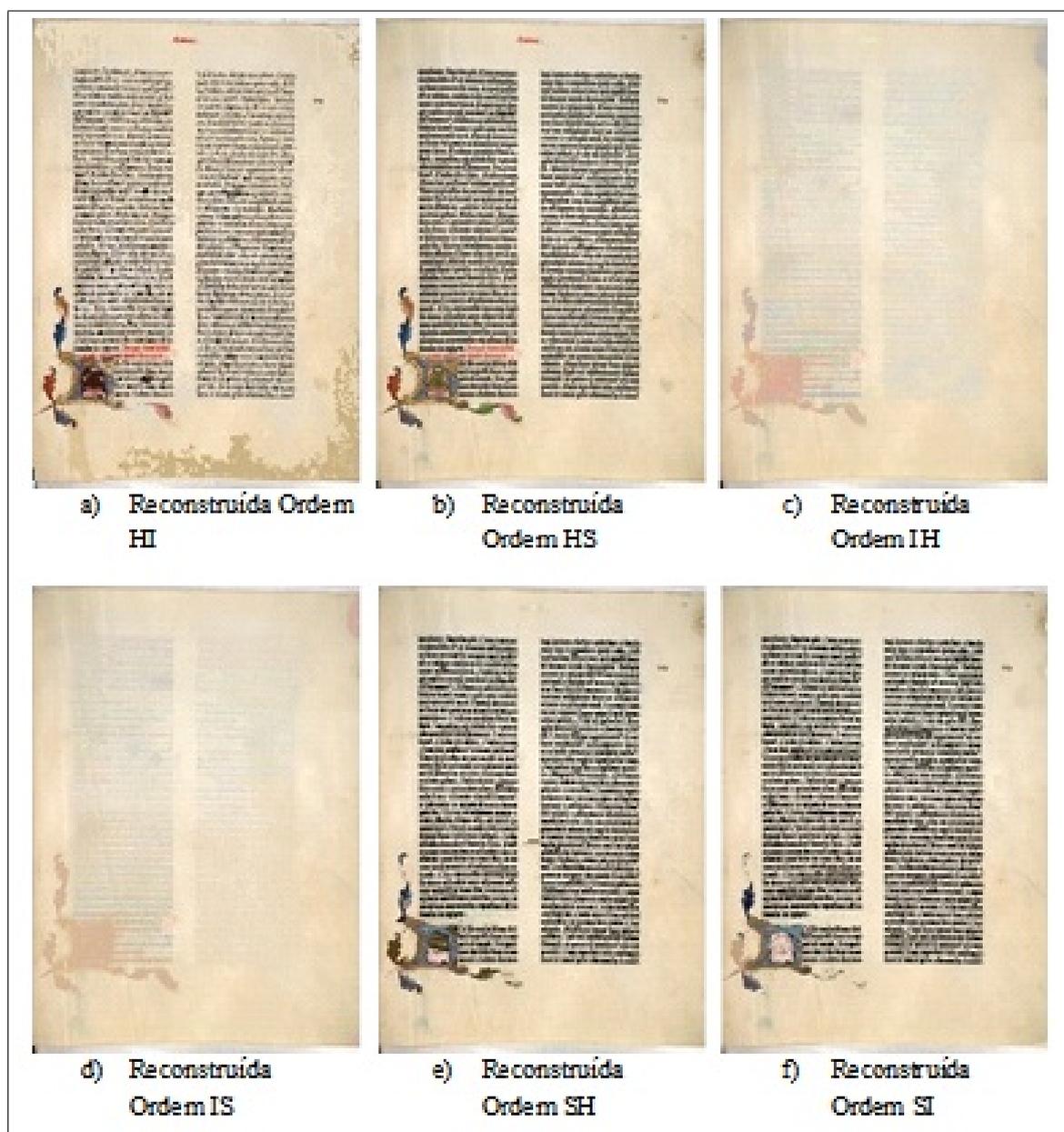


Figura 6.5: Resultado da reconstrução do fundo com uso de 02 canais HSI.

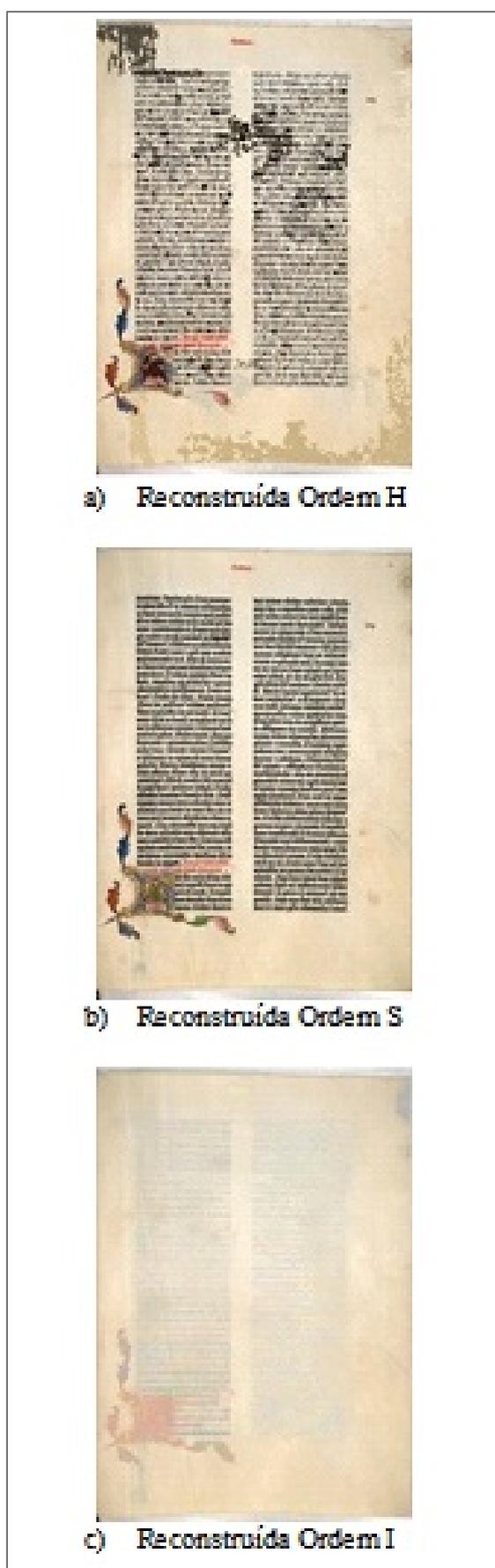


Figura 6.6: Resultado da reconstrução do fundo com uso de somente 01 canal HSI.

## b) Resultados com Espaço de Cor HSL

Para o processamento da operação de reconstrução com a ordenação lexicográfica, utilizou-se o valor de referência  $H$  majoritário da imagem. Para os testes com a imagem 029r11, o valor encontrado para  $H$  foi de  $220^\circ$ ; para o espaço de cor HSL, a ordem dos canais com os resultados estão apresentados a seguir:

- Ordens com uso de 3 canais: HSL, SHL, LSH, LHS, HLS e SLH

A Figura 6.7 ilustra os resultados com uso dos 03 canais, apresentando a variabilidade dos resultados.

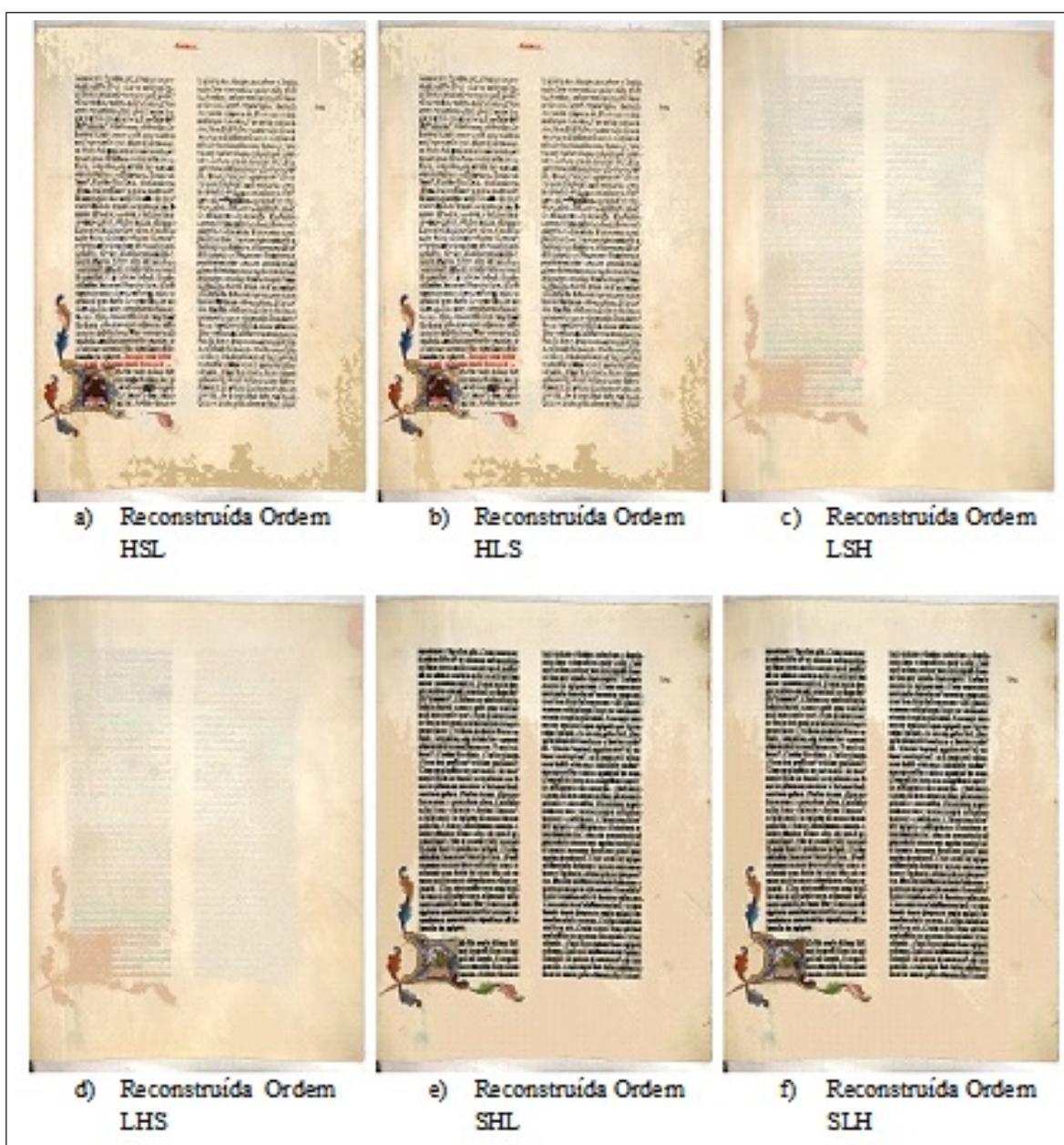


Figura 6.7: Resultado da reconstrução do fundo com uso de 03 canais HSL.

A Figura 6.8 ilustra os resultados com o uso dos 03 canais, com a escolha do canal H na primeira posição, com ordem HSL, com os seguintes valores de H:  $0^\circ$ ,  $120^\circ$ ,  $240^\circ$  e  $360^\circ$ .

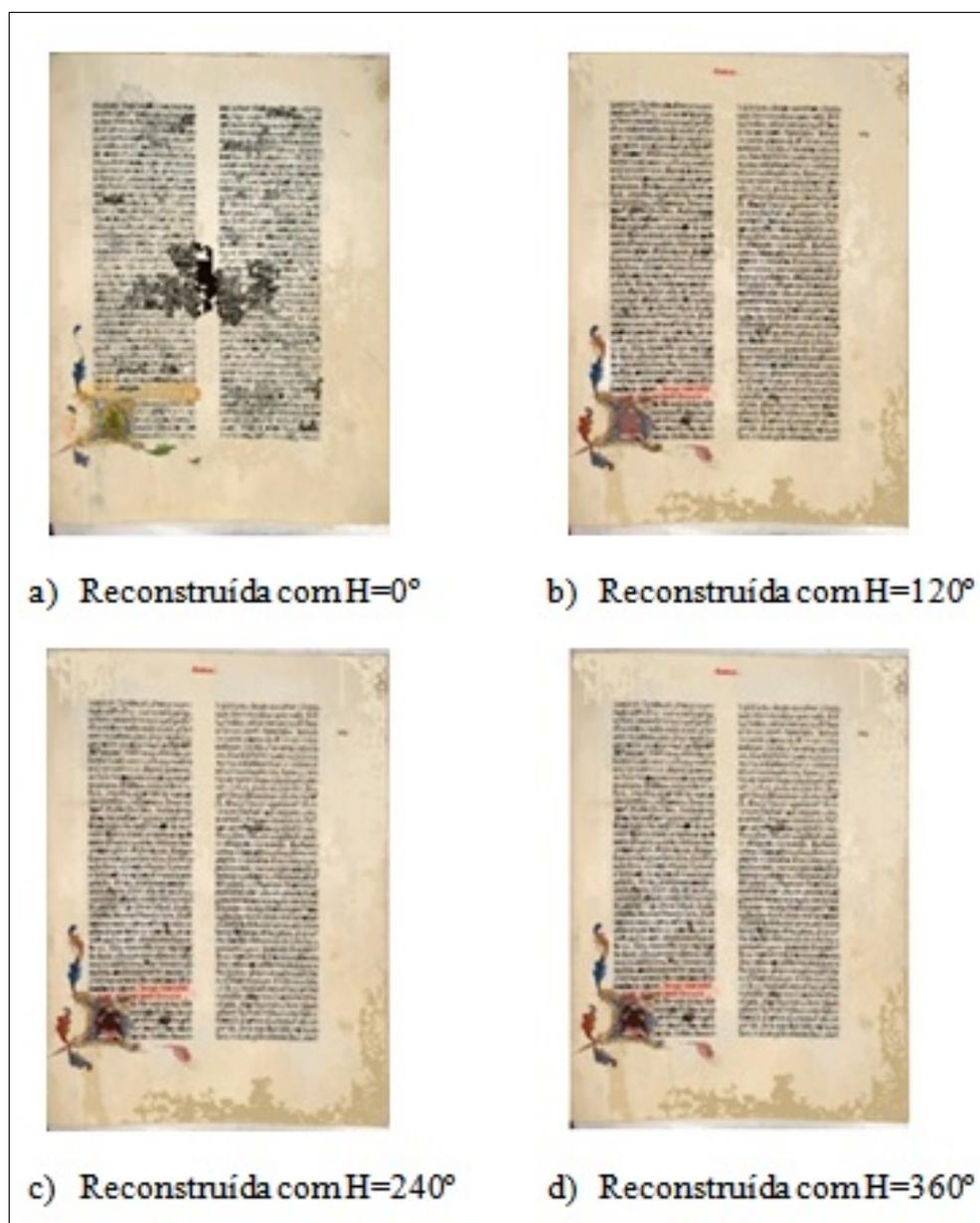


Figura 6.8: Resultado da reconstrução do fundo com uso de 03 canais HSL com H variando.

- **Ordenações com uso 2 canais: HS, SL, HL, LH, LS e SH:** a Figura 6.9 ilustra os resultados com uso de 02 canais, apresentando a variabilidade dos resultados.

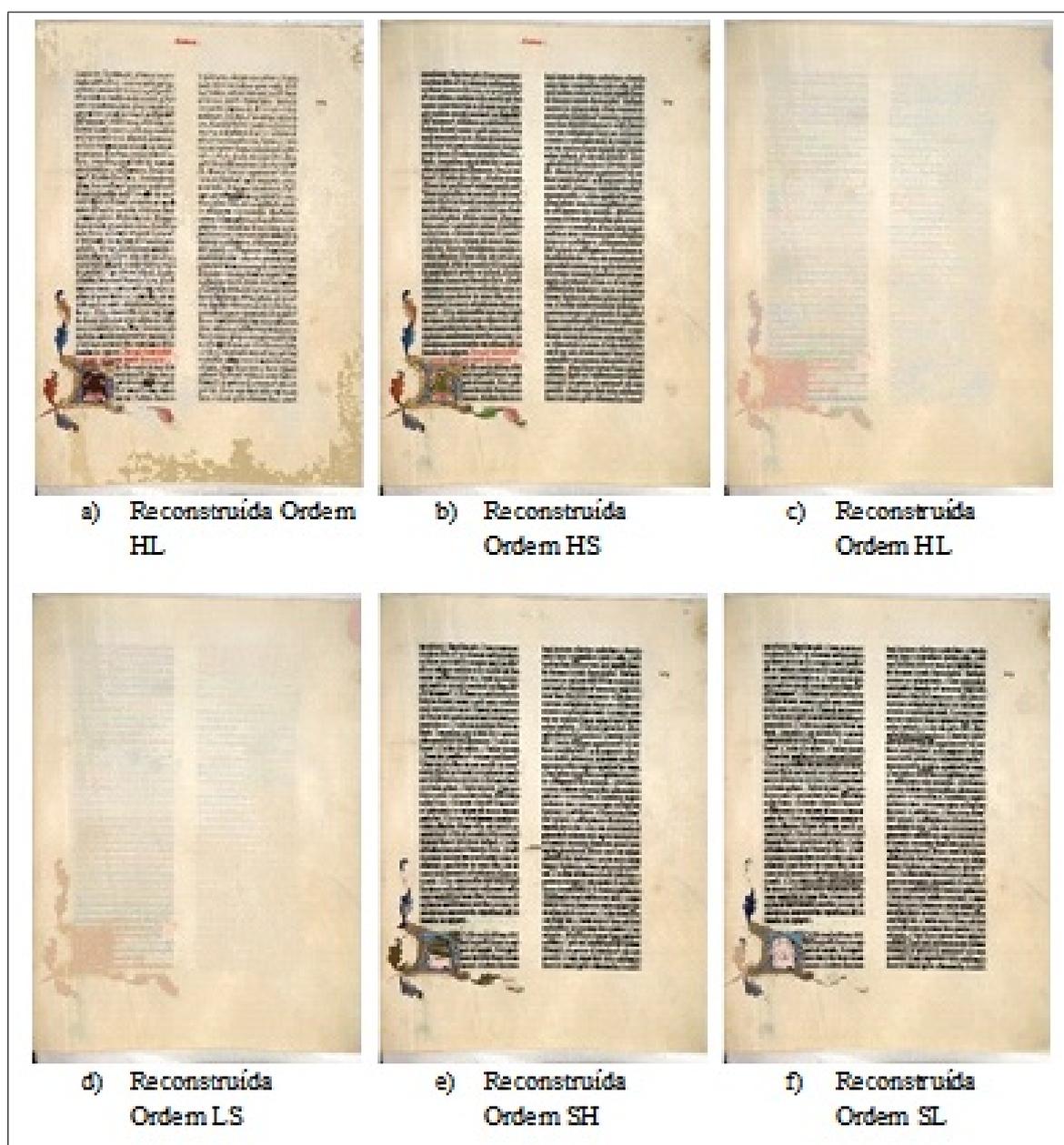


Figura 6.9: Resultado da reconstrução do fundo com uso de 02 canais HSL.

- **Ordenações com uso 1 canal: H, S e L:** a Figura 6.10 ilustra os resultados com uso de somente 01 canal, mostrando a variabilidade dos resultados.

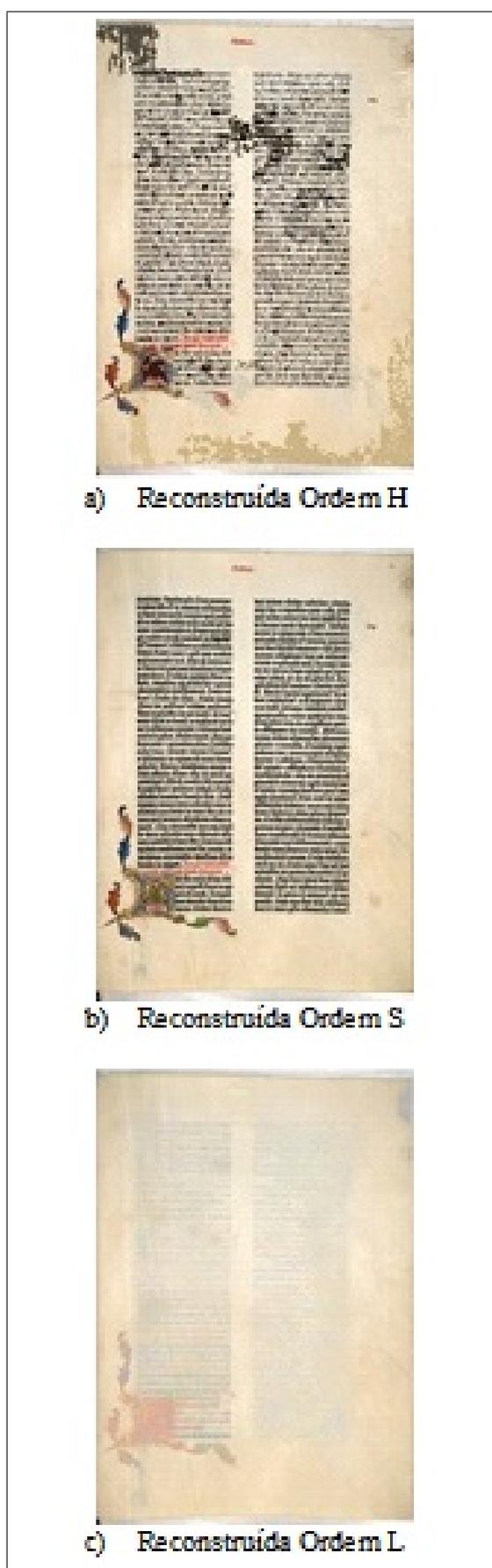


Figura 6.10: Resultado da reconstrução do fundo com uso de somente 01 canal HSL.

### c) Resultados com Espaço de Cor YCrCb

Para o espaço de cor YCrCb, a ordem dos canais com os resultados estão apresentados a seguir.

- **Ordenações com uso dos 03 canais: CrCbY, YCbCr, YCrCb, CrYCb, CbCrY e CbYCr:** a Figura 6.11 ilustra os resultados da reconstrução em todas as possibilidades de ordens, para o espaço de cor YCrCb com uso dos três canais.



Figura 6.11: Resultado da reconstrução do fundo com uso de 03 canais YCrCb.

- Ordenações com uso de 02 canais: YCr, YCb, CrCb, CbCr, CbY e CrY: a Figura 6.12 ilustra os resultados da reconstrução, com todas as possibilidades de ordens, para o espaço de cor YCrCb com uso de dois canais.



Figura 6.12: Resultado da reconstrução do fundo com uso de 02 canais YCrCb.

- Ordenações com uso de 01 canal: Y, Cr e Cb: A Figura 6.13 ilustra os resultados da reconstrução para todas as possibilidades de ordens para o espaço de cor YCrCb com uso de somente 01 canal.

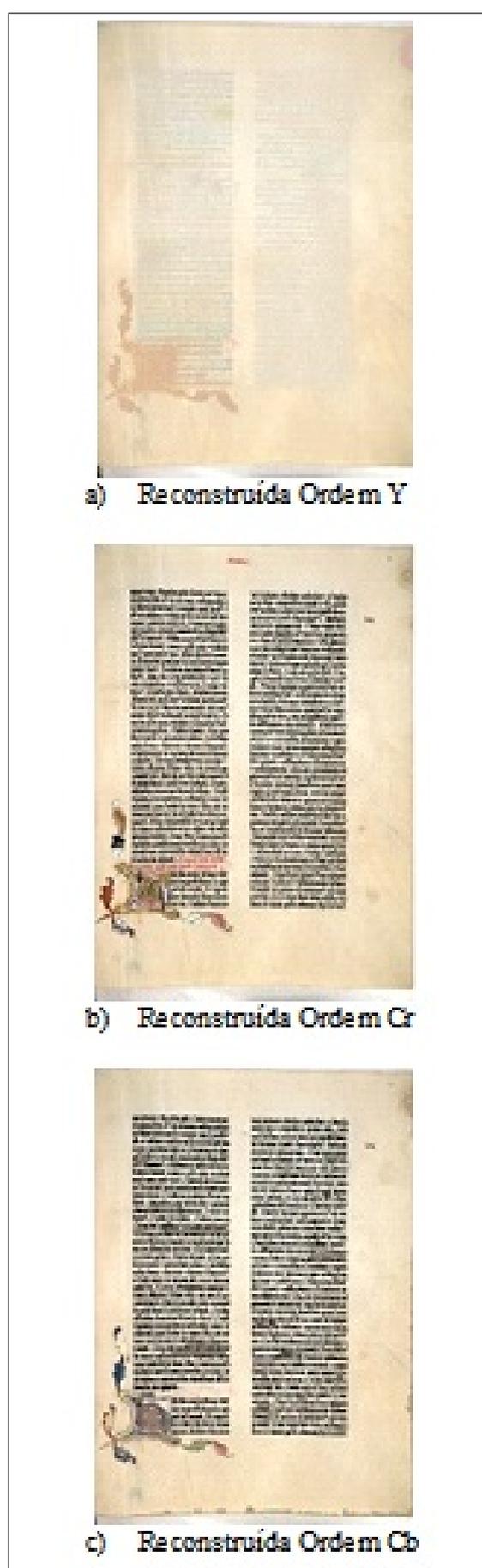


Figura 6.13: Resultado da reconstrução do fundo com uso de 01 canal YCrCb.

#### d) Resultados com Espaço de Cor YIQ

Para o espaço de cor YIQ, a ordem dos canais com os resultados estão apresentados a seguir.

- Ordenações com uso dos 03 canais: YIQ, YQI, QIY, QYI, IQY e IYQ: a Figura 6.14, ilustra os resultados da reconstrução, com todas as possibilidades de ordens, para o espaço de cor YIQ com uso dos 03 canais.



Figura 6.14: Resultado da reconstrução do fundo com uso dos 03 canais YIQ.

- Ordenações com uso de 02 canais: YI, IQ, QI, IY, YQ e QY: a Figura 6.15, ilustra os resultados da reconstrução para todas as possibilidades de ordens para o espaço de cor YIQ com uso de 02 canais.



Figura 6.15: Resultado da reconstrução do fundo com uso de 02 canais YIQ.

- **Ordenações com uso 1 canal: Y, I e Q:** a Figura 6.16 ilustra os resultados da reconstrução, com todas as possibilidades de ordens, para o espaço de cor YIQ com uso de somente 01 canal.

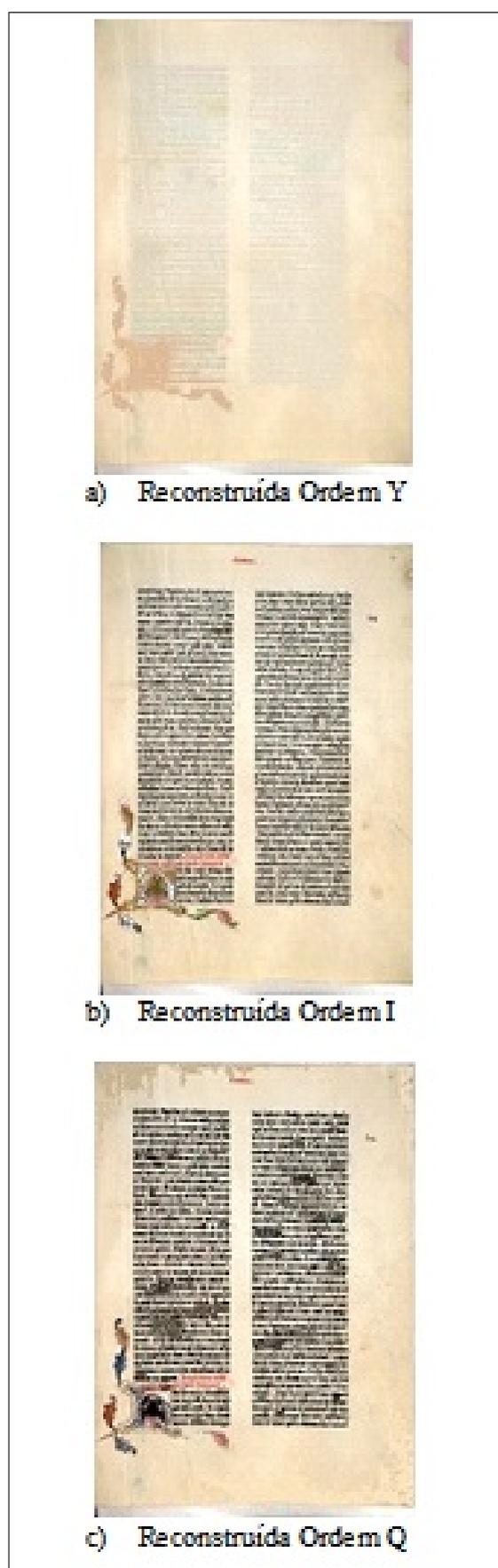


Figura 6.16: Resultado da reconstrução do fundo com uso de 01 canal YIQ.

#### d) Diferenças entre as reconstruções do fundo

Foram realizados vários testes da reconstrução do fundo para os espaços HSI, HSL, YCrCb e YIQ e suas ordenações. Os melhores resultados foram obtidos com as ordenações ISH, LSH, YCrCb e YIQ. Entre esses resultados houve pequenas diferenças. A Figura 6.17 ilustra a diferença entre as reconstruções do fundo com ordem ISH e YIQ, sendo a imagem (a) a reconstrução com ordem ISH, (b) com ordem YIQ, (c) a diferença entre as duas imagens e (d) a diferença limiarizada, mostrando a pequena diferença entre as reconstruções.

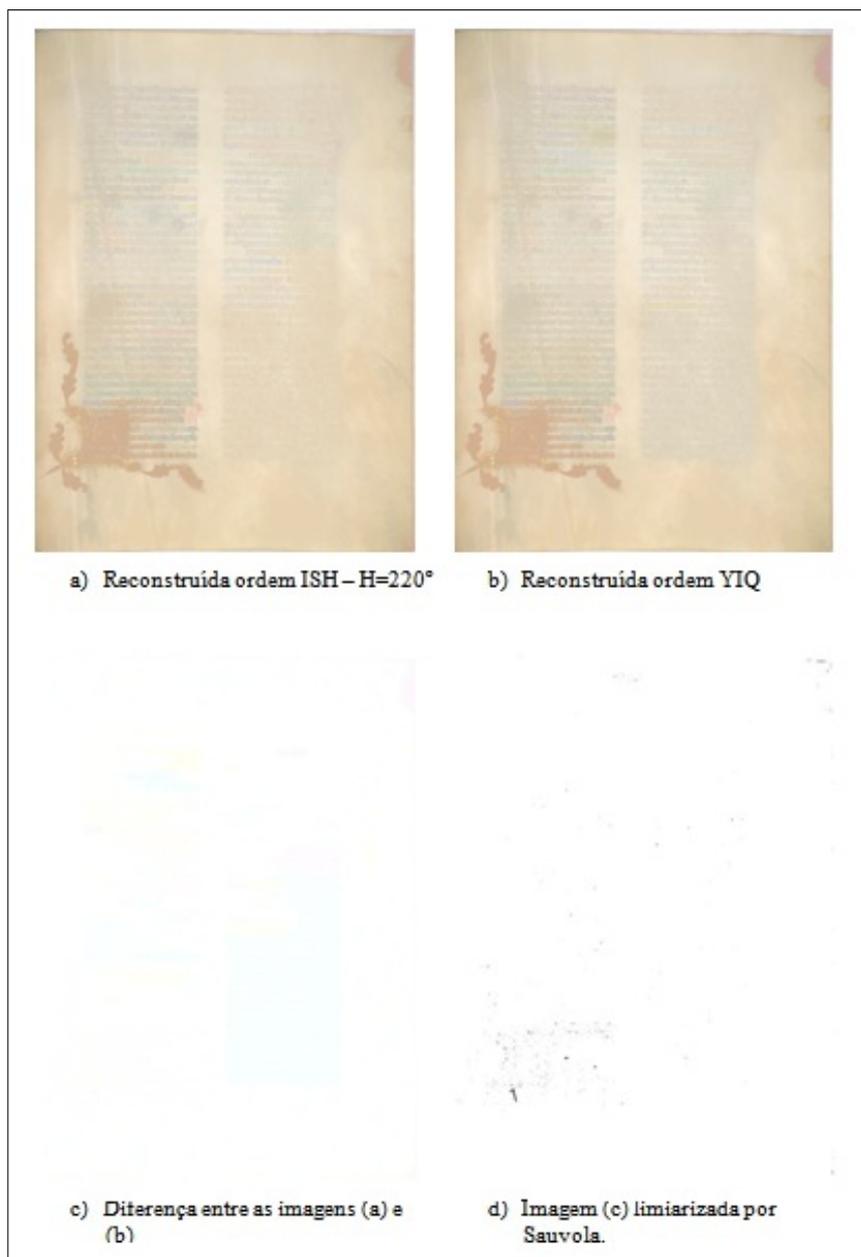


Figura 6.17: Exemplo da diferença entre as reconstruções.

### e) Estimativas do fundo baseados nas imagens reconstruídas

A estimativa dos fundos foi baseada nos histogramas das imagens reconstruídas, conforme descrito no método proposto. A Figura 6.18 ilustra as estimativas dos espaços de cor HSI, HSL, YCrCb e YIQ da imagem original 029r11, sendo a imagem (a) a estimativa da reconstrução com ordem ISH, (b) com ordem LSH, (c) com ordem YCrCb e (d) com ordem YIQ.

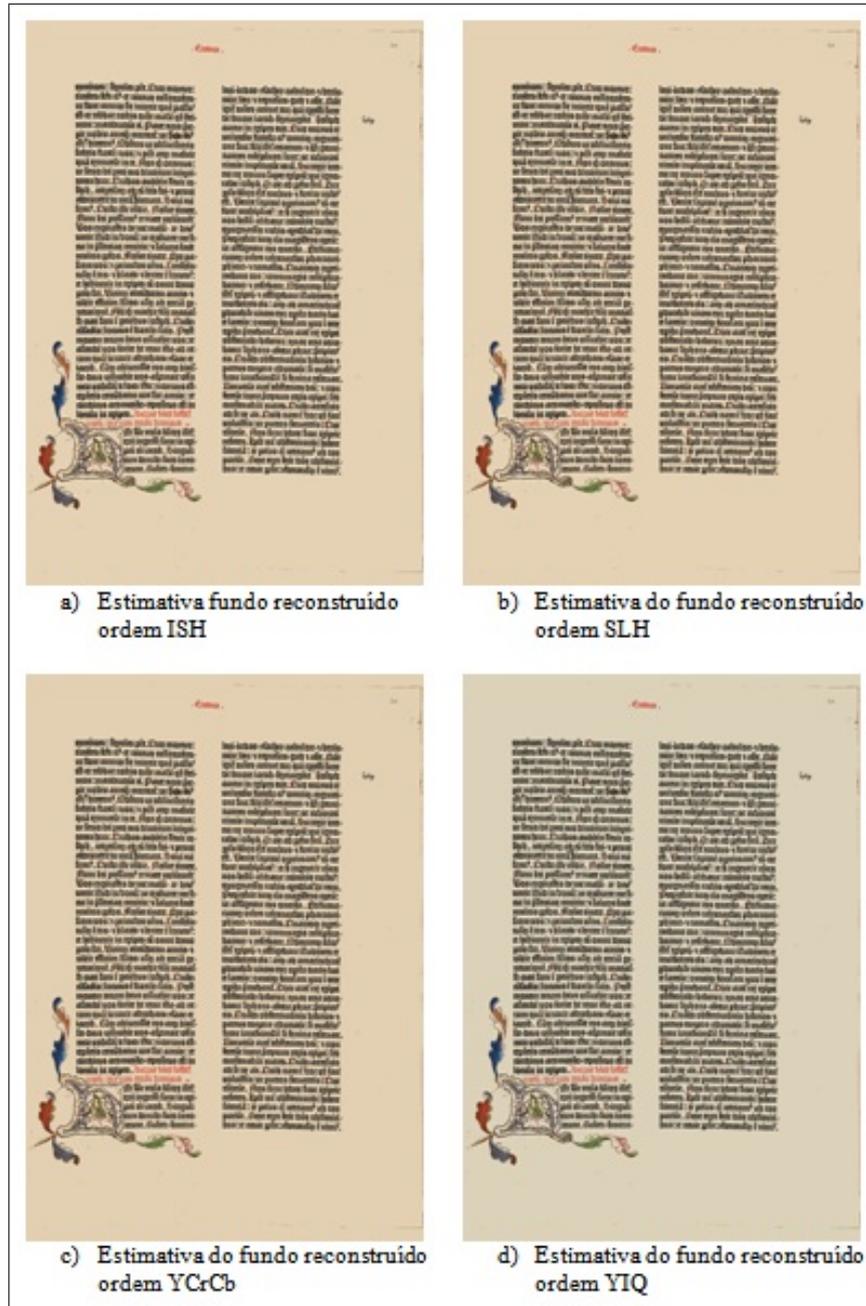


Figura 6.18: Resultado da estimativa do fundo baseada nas imagens reconstruídas do fundo da imagem 029r11.

Os resultados apresentados na Figura 6.18 mostraram-se resultados visualmente similares para as estimativas dos espaços HSI, HSL e YCrCb, no entanto YIQ apresentou uma pequena diferença na tonalidade da cor do fundo, isto devido ao processo de reconstrução em cor, que apresentaram pequenas diferenças.

A Figura 6.19 ilustra outro exemplo da estimativa do fundo da imagem original da Bíblia, identificada como 084r11.

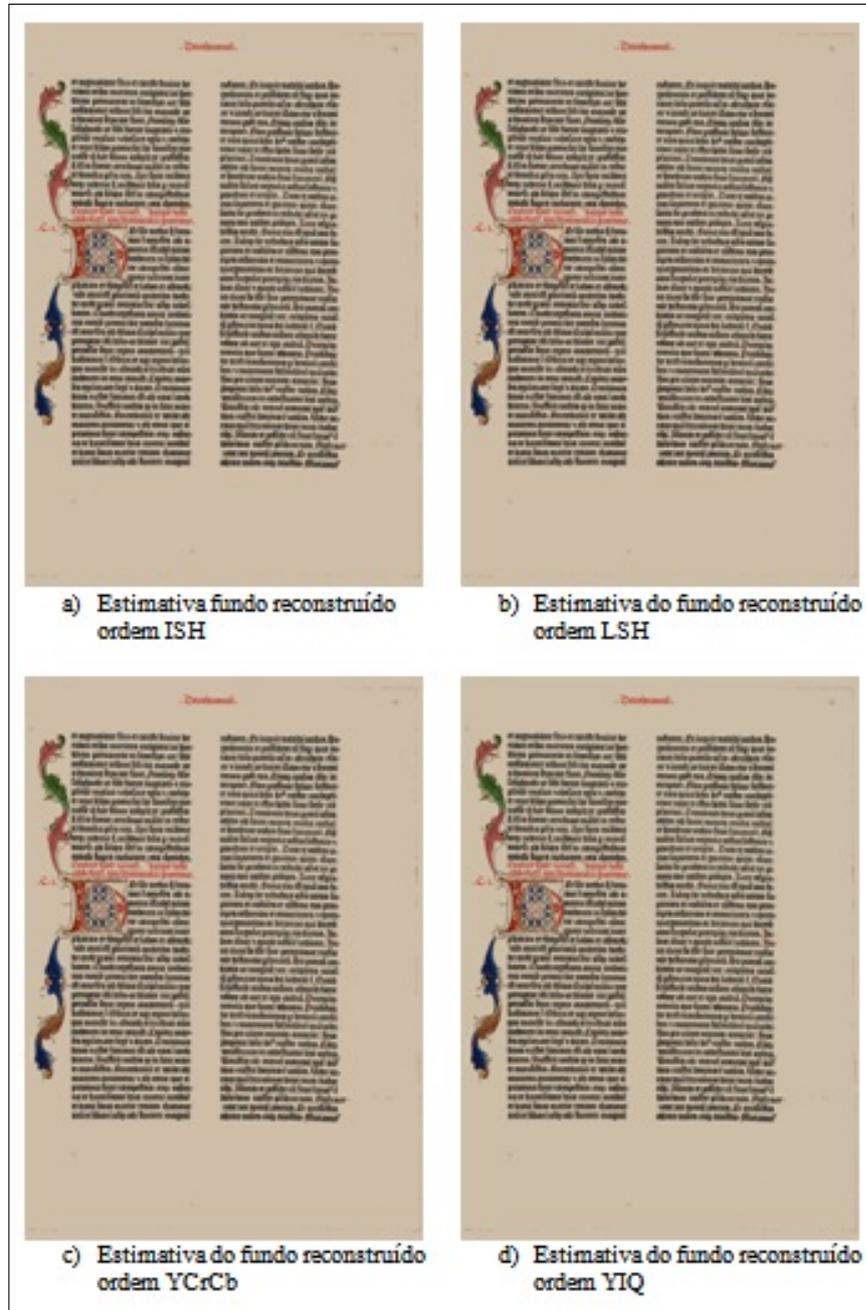


Figura 6.19: Resultado da estimativa do fundo baseada nas imagens reconstruídas do fundo da imagem 084r11.

Nota-se nas imagens da Figura 6.19 que as estimativas para os 04 espaços de cor foram visualmente similares.

#### e) **Considerações Finais do Processo 1: Estimativa do fundo colorido**

Os resultados obtidos na reconstrução do fundo da imagem das páginas da Bíblia confirmaram que todos os espaços de cor retornam bons resultados utilizando o canal intensidade (luminância) na primeira ordem de processamento. Para os espaços de cor HSI/HSL, com o uso dos 03 canais, o melhor resultado foi obtido com as ordens ISH e LSH, para o espaço de cor YCrCb, a ordem adequada foi YCrCb e para o espaço de cor YIQ, a ordem foi YIQ. Para estes quatro espaços, os testes mostraram ser possível utilizar somente o canal intensidade na recuperação do fundo através da operação de reconstrução para estimativa do fundo.

Confirmou-se ainda que foi mais adequado utilizar os espaços de cor YCrCb e YIQ e não HSI/HSL. O grande problema da utilização desses espaços de cor, com os três canais, foi a necessidade de se definir um valor de referência do canal H, o que se apresentaria no processo como uma heurística, pois como indicar o maior valor entre  $H=0$  e  $H=360$ ? Assim, recomenda-se utilizar os outros dois espaços de cor, quando utilizar os três canais para processamento.

A ordenação lexicográfica com o uso dos três canais de cor, juntos, retornam o mesmo resultado quando utilizado somente o canal intensidade, sendo este o mais importante canal para a reconstrução do fundo das páginas da Bíblia de Gutenberg. A Figura 6.20 ilustra o exemplo da operação de reconstrução com espaço de cor YCrCb, com a utilização dos 03 canais, com 02 canais e 01 canal, com as seguintes ordens: YCrCb, YCr e Y, demonstrando os resultados da reconstrução do fundo da imagem para estimativa do fundo homogêneo.

De acordo com os testes e a Figura 6.20, o resultado com o uso do canal Y somente foi igual ao resultado do uso dos canais YCr e YCrCb. Assim, em termos de tempo de execução, é aconselhável utilizar somente um canal. Estes testes não só mostraram ser possível utilizar somente um canal, mas também utilizar os três canais do espaço de cor para o processamento da morfologia em cor. Para chegar a estas conclusões, foram testadas as 120 imagens da base de dados da Bíblia de Gutenberg, verificando uma variação do fundo entre elas.

Os testes foram realizados com a intenção de verificar qual espaço de cor retornaria melhores resultados e qual ou quais ordenações seriam as mais indicadas para esta

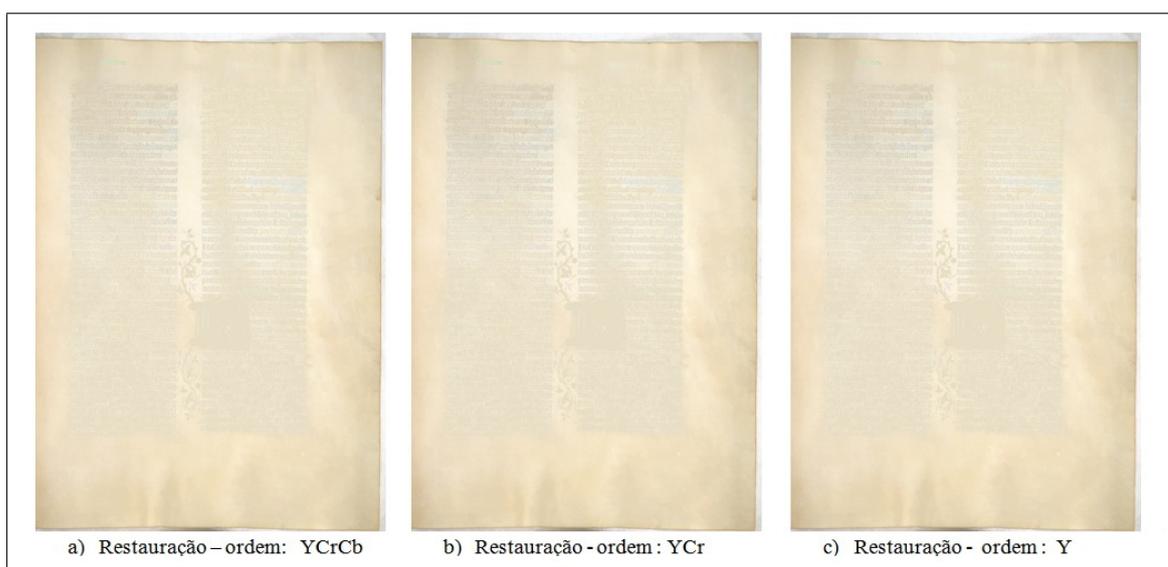


Figura 6.20: Exemplos da restauração do fundo com espaço de cor YCrCb.

etapa.

A estimativa do fundo baseada nas imagens reconstruídas retornaram resultados visualmente muito semelhantes para os espaços HSI, HSL, YCrCb e YIQ.

Portanto o processo 1 apresentou a estimativa do fundo baseada na operação de reconstrução com o uso dos espaços de cor e suas ordenações. O próximo passo consistiu em segmentar o conteúdo do primeiro plano.

### 6.3.2 Processo 2: Segmentação do conteúdo do primeiro plano

A segmentação do conteúdo do primeiro plano consistiu em separar o fundo das informações como títulos, manuscritos, numeração de páginas, figuras e texto colorido ou preto. Com o resultado obtido pelo processo 1, da estimativa do fundo, o próximo passo foi realizar uma operação de subtração entre a imagem original e o fundo reconstruído através da operação morfológica TopHat, a Figura 6.21 ilustra o resultado dessa operação com as informações do primeiro plano.

O resultado ilustrado pela imagem (c) da figura 6.21 possui informações do primeiro plano, mas com uma tonalidade diferente da imagem original, o que levou à necessidade de converter a imagem em níveis de cinza para binarização da imagem. Os pixels binarizados serviram de busca, na imagem original, dos verdadeiros valores dos pixels do primeiro plano, sendo composta a nova imagem com esses valores e com os valores do fundo na cor branca.

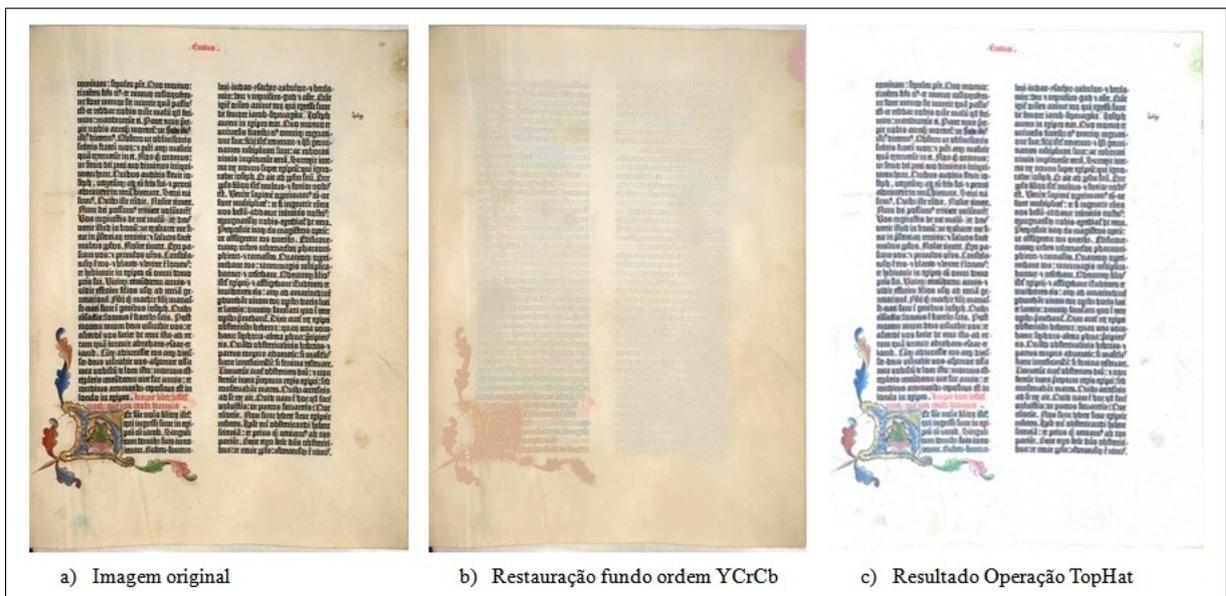


Figura 6.21: Resultados da operação Top Hat.

#### a) Conversão em Níveis de Cinza e Binarização

A imagem de saída do processo de Tophat foi convertida em níveis de cinza e, após isso, foi realizado o processo de binarização através de um algoritmo de limiarização conhecido.

Para os testes foram utilizadas as seguintes técnicas de limiarização:

- Abutaleb (ABUTALEB, 1989);
- Johannsen (JOHANNSEN; BILLE, 1982);
- MCC (YEN; CHANG, 1995) e
- Sauvola (SHAFAIT; KEYSERS; BREUEL, 2008).

Os melhores resultados foram obtidos com a técnica de limiarização de Sauvola. A limiarização MCC retornou muitos ruídos, mesmo utilizando-se a filtragem matemática binária Soft (ZMUDA; TAMBURINO; RIZK, 1991). As outras técnicas de limiarizações citadas segmentaram com falhas os títulos e desenhos. A Figura 6.22 ilustra o resultado das limiarizações para a imagem da Bíblia 001r11, que evidencia a técnica de Sauvola, imagem (d), com o melhor resultado por apresentar poucos ruídos.

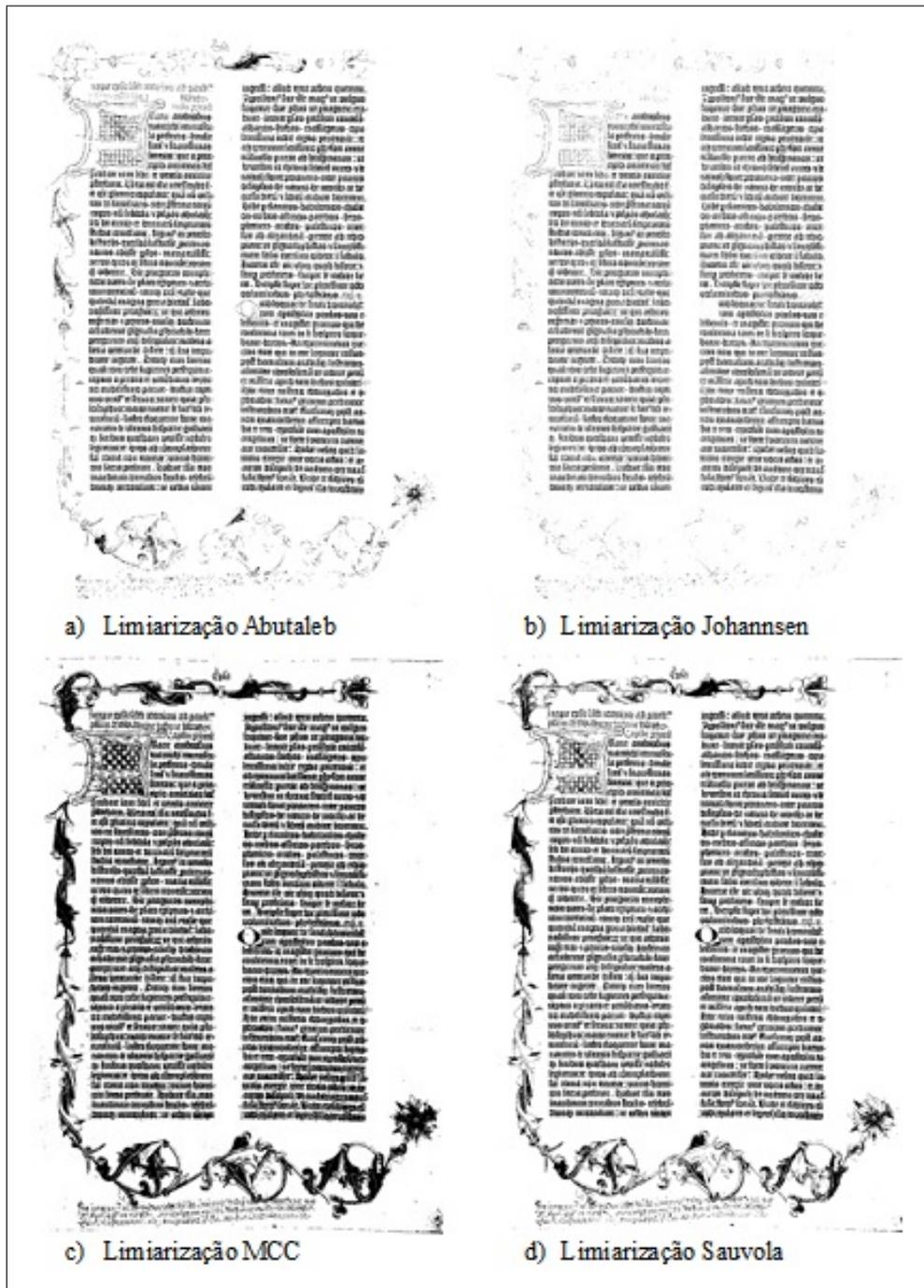


Figura 6.22: Testes com as técnicas de binarização.

A Figura 6.23 ilustra outro exemplo do uso das limiarizações testadas: a imagem (a) não consegue segmentar os desenhos em sua totalidade, a (b) os segmenta muito pouco, a (c) contém muitos ruídos, segmentando um título pertencente ao fundo como sendo informação do primeiro plano. A imagem (d) representou a melhor segmentação para esta imagem da Bíblia, realçando bem os títulos, as figuras e as duas colunas centrais.

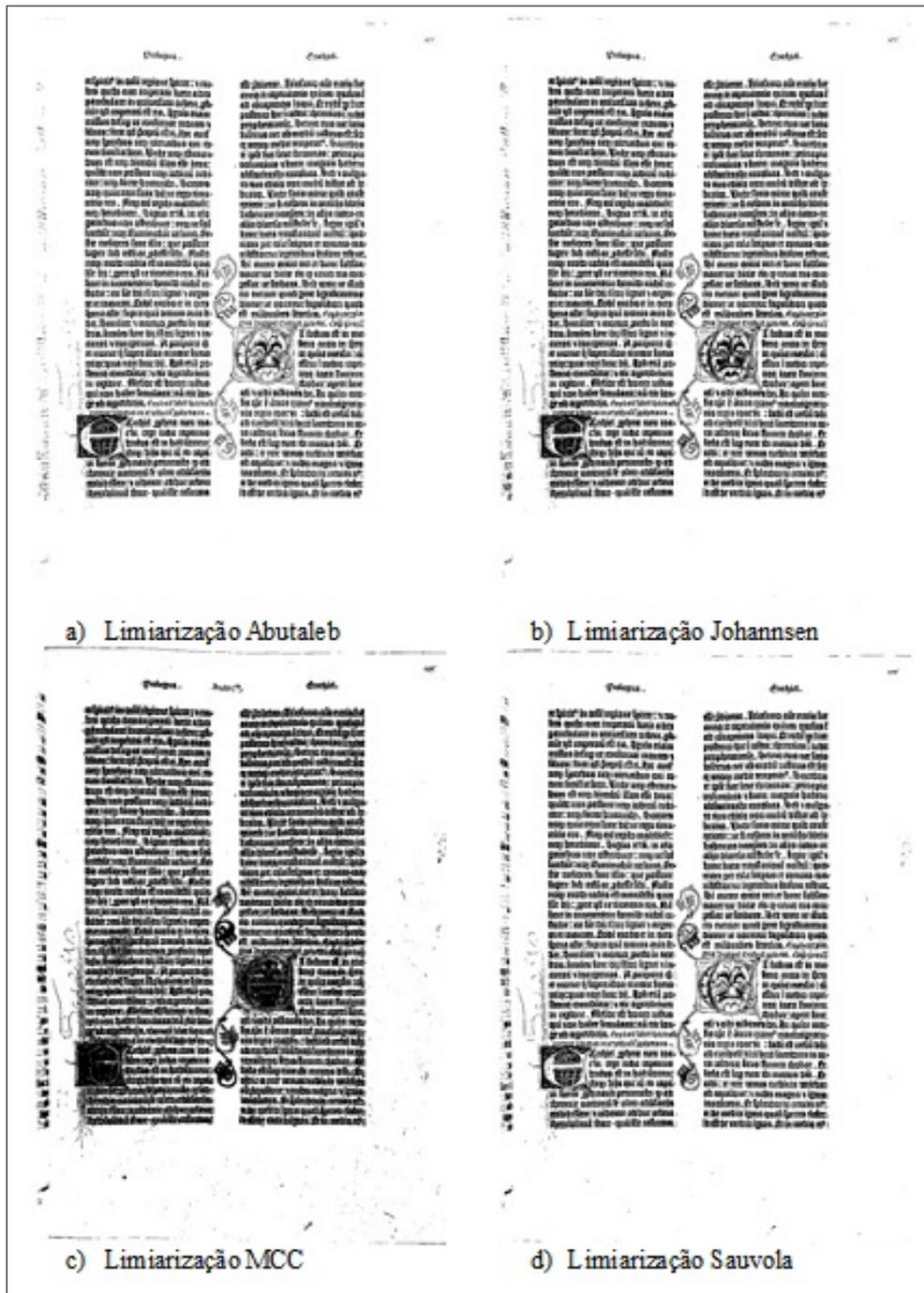


Figura 6.23: Testes com as Limiarizações.

Stathis, Kavallieratou e Papamarkos (2008) apresentaram um ranqueamento dos algoritmos de limiarizações, sendo o melhor o de Sauvola e, em seguida, o algoritmo de Johannsen.

Os melhores resultados obtidos nos testes para a binarização do conteúdo do primeiro plano foram com o algoritmo de Sauvola. Com a imagem limiarizada foi possível extrair da imagem original os tons de cor corretos, baseados nos pixels pre-

tos binarizados, gerando uma imagem com o conteúdo do primeiro plano com as cores originais.

## b) Extração do texto colorido com pixels originais

Com a imagem binarizada, o texto do primeiro plano pôde ser extraído com base nos pixels originais através da verificação dos pixels pretos da imagem binarizada, retornando à posição desses pixels para encontrar o valor do pixel original da mesma posição. A imagem final contém os pixels com o valor original da imagem da Bíblia. Este processo gerou a imagem com o conteúdo do primeiro plano, e está ilustrada pela Figura 6.24, onde aparece a imagem com os pixels verdadeiros.

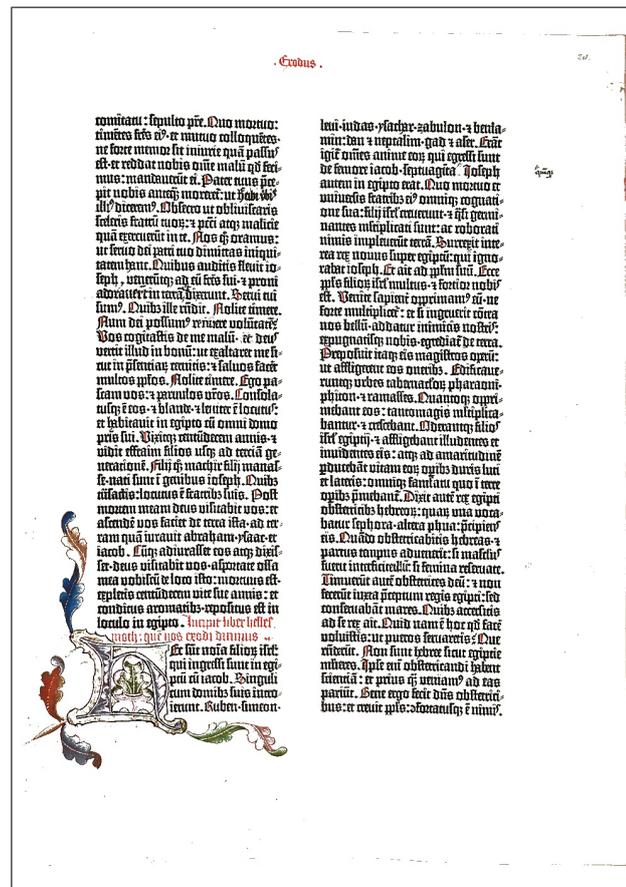


Figura 6.24: Resultado da extração do conteúdo do primeiro plano com pixels originais.

As informações segmentadas do primeiro plano serviram para o processo 4, que consistiu em classificar este conteúdo de acordo com o estabelecido no método proposto.

### 6.3.3 Processo 3: Segmentação do conteúdo colorido

Esta etapa consistiu em encontrar, na imagem antiga da Bíblia, as informações coloridas, como os títulos, as figuras, as letras coloridas e as pequenas letras vermelhas.

Com o processo de segmentação e utilização da morfologia matemática em cor foi possível encontrar estas informações. O processo abrangeu as seguintes etapas, conforme o método proposto:

- Realização do processo de reconstrução com o espaço de cor HSI/HSL, com a ordenação lexicográfica e a ordem dos canais SHI/SHL;
- Conversão em níveis de cinza e binarização;
- Extração do texto colorido;
- Filtragem dos ruídos pelo filtro morfológico binário Soft;
- Realização de novo processo de reconstrução com a ordenação lexicográfica com o espaço de cor HSI, HSL, YCrCb ou YIQ com as seguintes ordens: ISH, LSH, YCrCb e YIQ, respectivamente. Preferencialmente utilizou-se o espaço YCrCb por não precisar indicar valor aos canais *a priori*;
- Imagem segmentada com conteúdo colorido.

Através dos testes realizados, a reconstrução para identificar as sementes das pequenas letras vermelhas, utilizando a ordenação lexicográfica, funcionou somente com o uso dos três canais juntos, com as ordens SHI/SHL. Os resultados mostraram-se iguais para as ordens SHI e SHL. Assim, foram realizados testes com o espaço HSI com ordem SHI.

As imagens reconstruídas pela ordem SHI precisaram ser filtradas pelo filtro matemático binário Soft. Este filtro necessita de um valor para o fator de verificação  $m$ , que pode variar de 1 a 8, para verificar a quantidade de pixels pretos presentes na janela 3x3. Por exemplo, quando-se analisa um pixel preto e quantos pixels pretos sua vizinhança possui, se o fator for igual ou superior a 7, o pixel não será eliminado, ou seja, não é considerado como ruído, mas se a quantidade for menor que 7, o pixel é considerado como ruído e é eliminado. A Figura 6.25 ilustra os resultados dos filtros variando de 2 a 4.

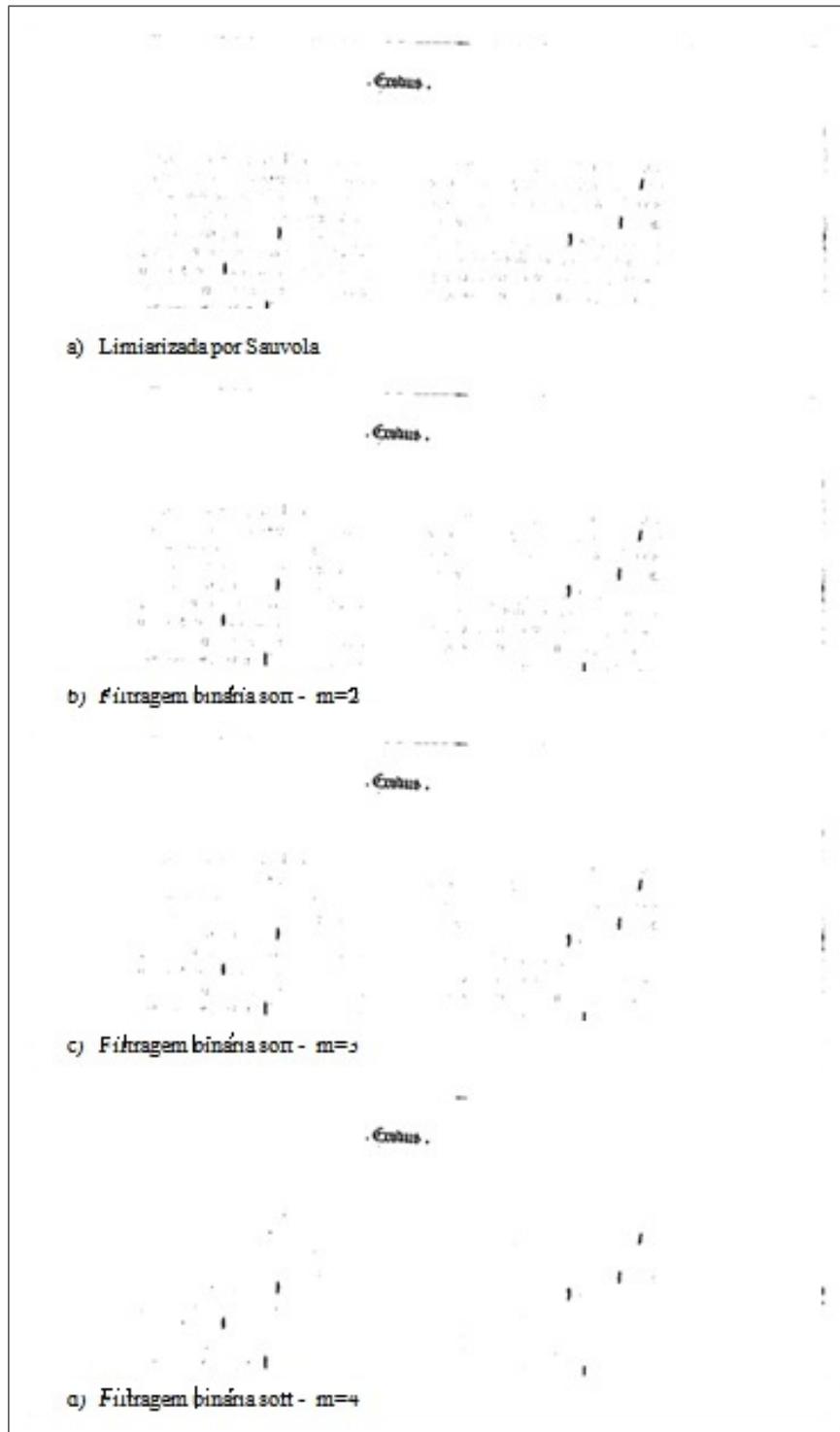


Figura 6.25: Resultados da Filtragem matemática binária Soft com  $m$  variando de 2 a 4.

A Figura 6.26 ilustra a imagem limiarizada filtrada com valores de  $m$  variando de 5 a 8.

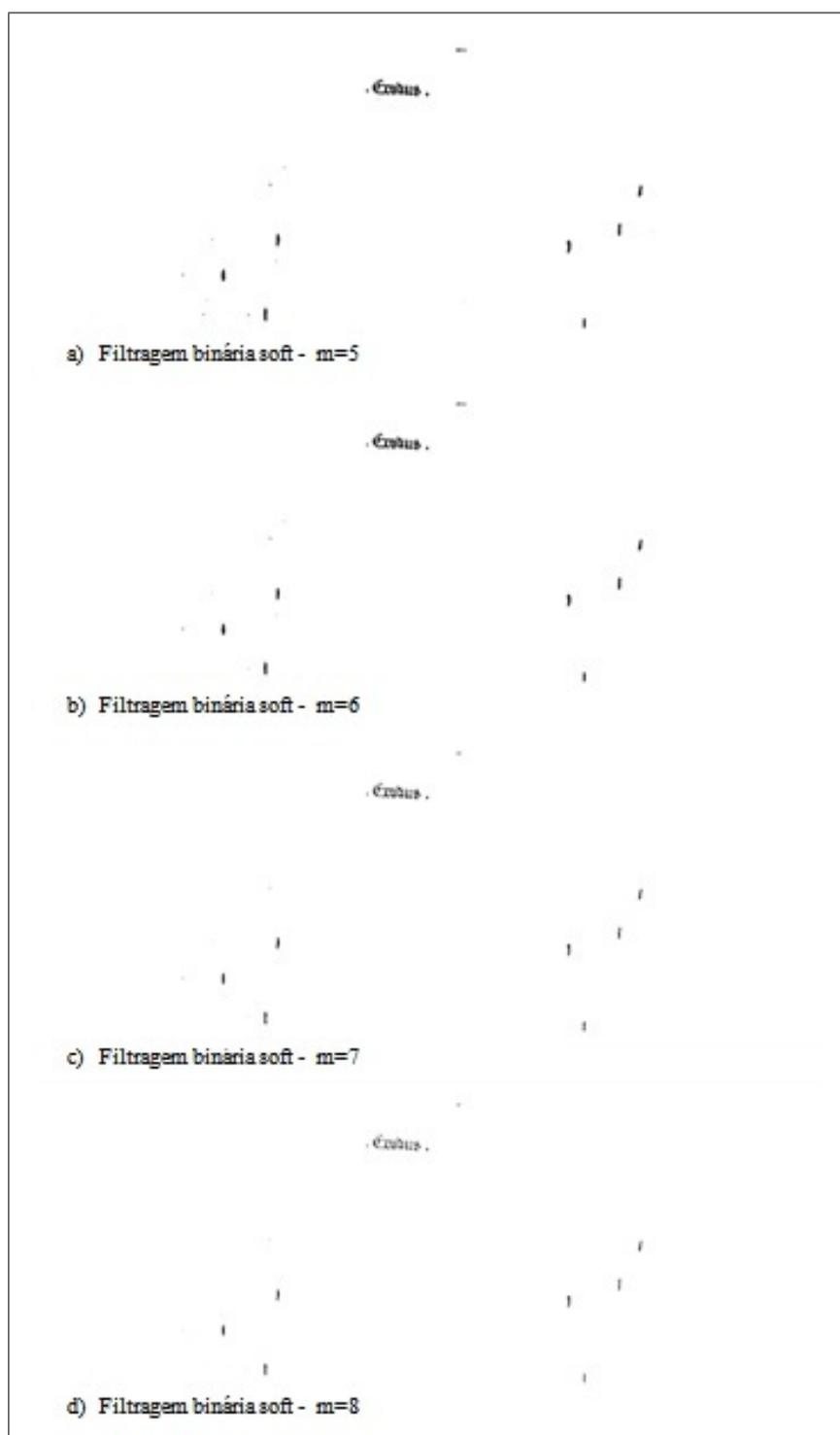


Figura 6.26: Resultados da Filtragem matemática binária Soft com  $m$  variando de 5 a 8.

Pelos resultados obtidos, verificou-se que os melhores resultados, para eliminação dos ruídos, apresentaram valor de  $m \geq 6$ . Com a eliminação dos ruídos, o próximo passo foi a realização do processo de reconstrução dos pixels coloridos presentes na imagem. A Figura 6.27 ilustra exemplo da segmentação do conteúdo colorido, abrangendo o título, a figura e as pequenas partes vermelhas presentes nas letras das duas colunas impressas.



Figura 6.27: Resultado do conteúdo colorido segmentado pelo processo 3.

A próxima seção apresenta os resultados obtidos pelos três processos do nível 1.

### 6.3.4 Resultados finais dos 03 processos do nível 1

#### a) Processo 1: Resultado final com fundo estimado

A imagem final com fundo estimado foi baseada no valor máximo do histograma dos três canais, gerando uma imagem com fundo homogêneo e com as informações reais do primeiro plano. A Figura 6.28 ilustra o resultado final deste processo.

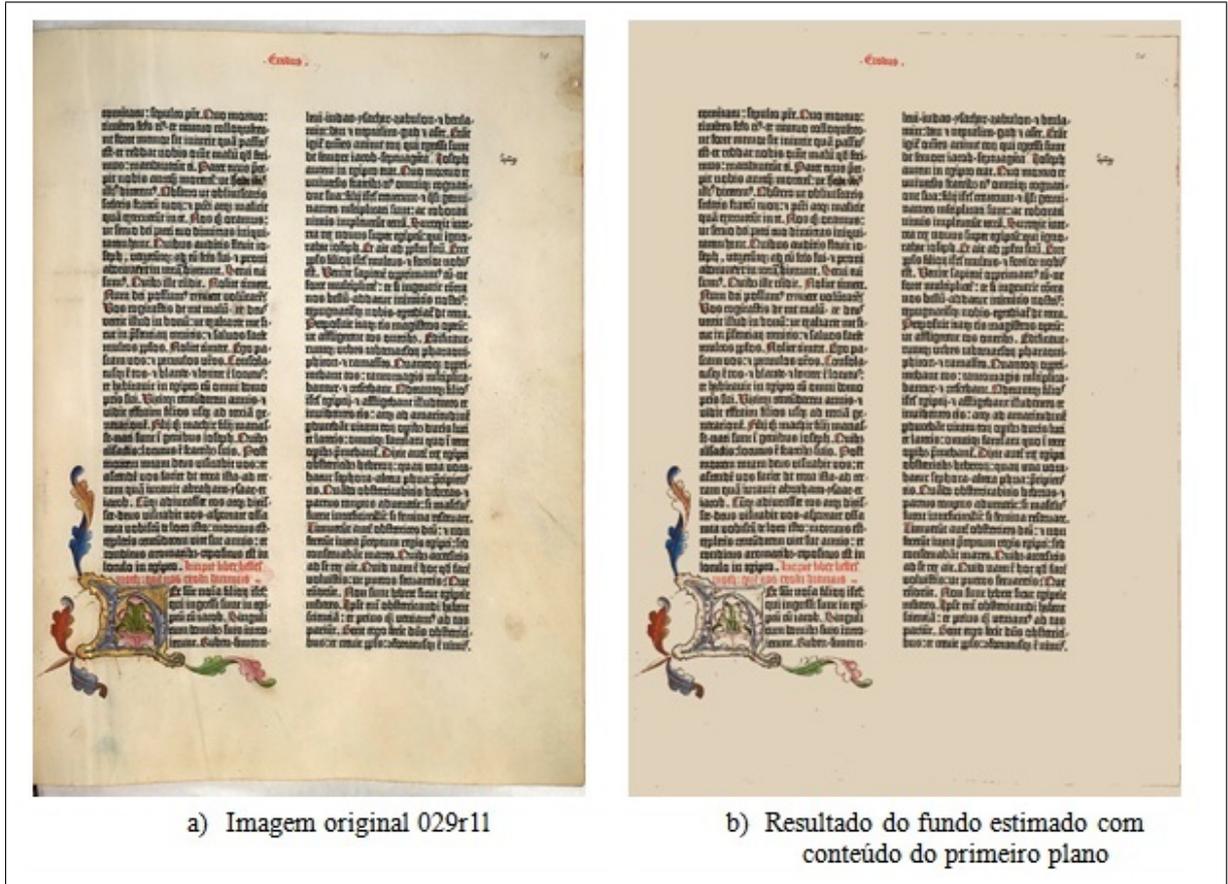


Figura 6.28: Resultado do fundo estimado com conteúdo do primeiro plano.

## b) Processo 2: Resultado final com o conteúdo do primeiro plano

A Figura 6.29 ilustra o resultado final da extração do conteúdo do primeiro plano com base no método proposto.

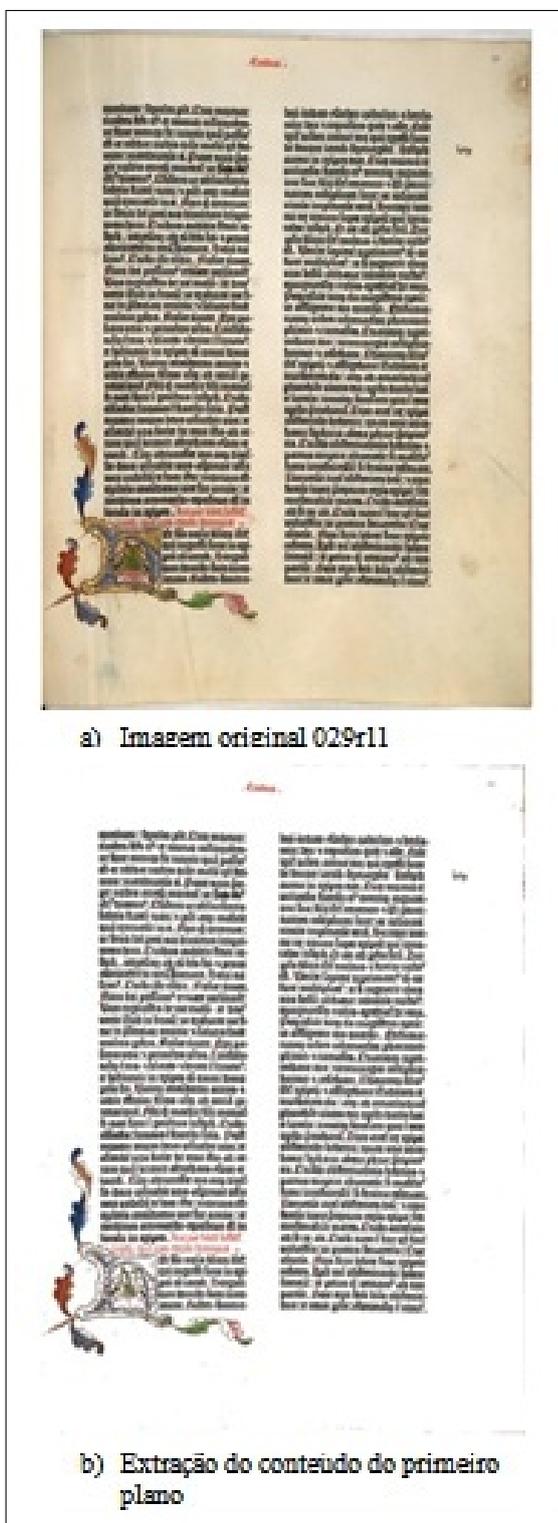


Figura 6.29: Resultado da extração do conteúdo do primeiro plano.

### c) Processo 3: Resultado final com o conteúdo colorido

A Figura 6.30 ilustra o resultado final do conteúdo colorido: títulos, figuras e letras pequenas nas duas colunas.

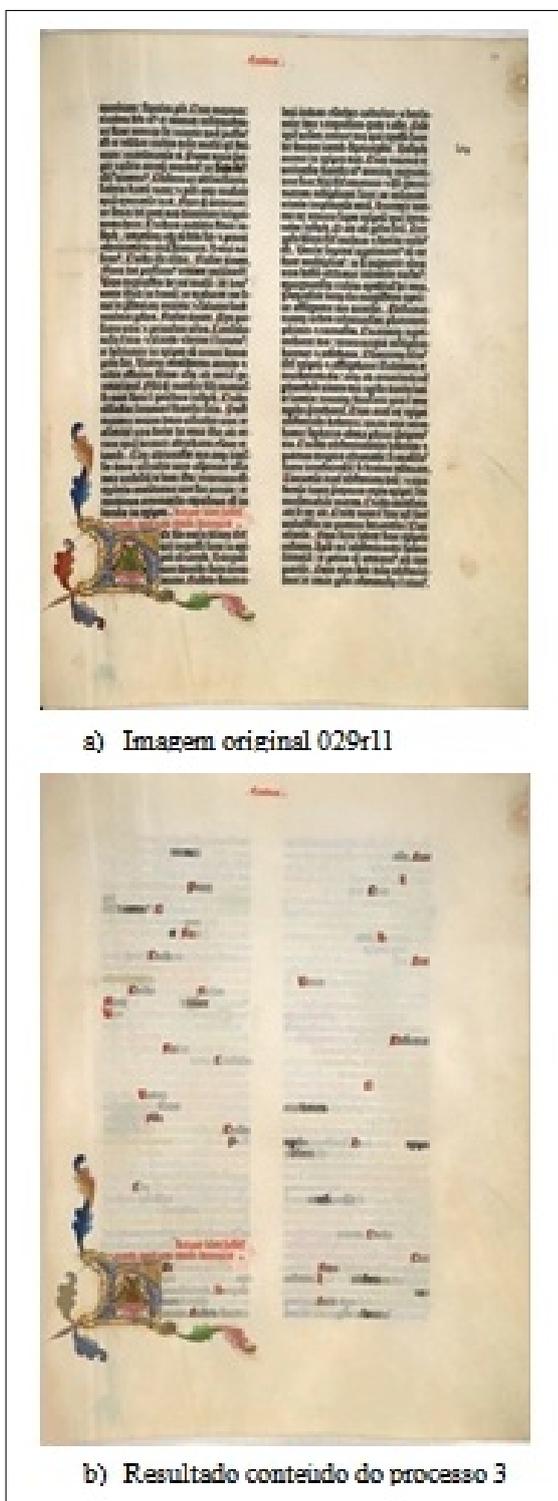


Figura 6.30: Resultado do conteúdo colorido pelo processo 3.

### **6.3.5 Avaliação do Nível 1 e seus 03 processos**

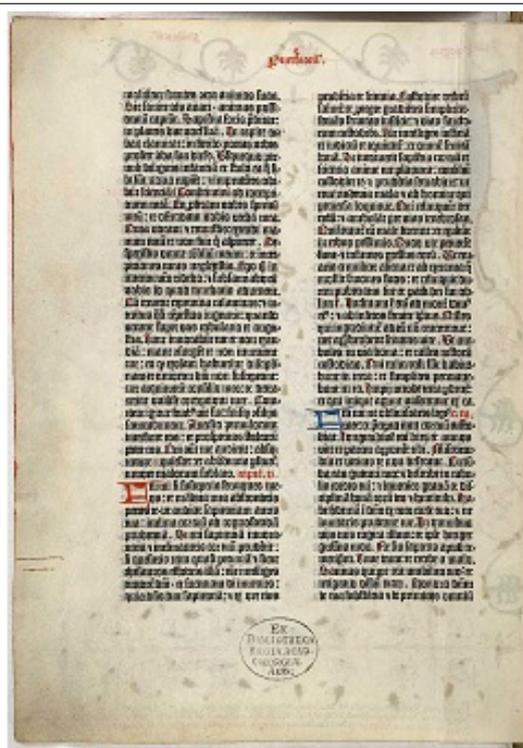
Cada processo do nível 1 foi submetido a uma forma diferente de avaliação. Não foi possível avaliar o processo 1 de forma quantitativa, uma vez que o fundo foi estimado no histograma da imagem. Para a avaliação dos processos 2 e 3 foram geradas imagens sintéticas para validação, uma vez que para realizar as comparações com as métricas era preciso obter as imagens ideais. A imagem sintética era composta de um fundo deteriorado com o tempo e de informações do primeiro plano, reproduzindo uma imagem da Bíblia impressa por Gutenberg. Para os testes foram geradas várias imagens sintéticas variando o fundo, as informações do primeiro plano e a opacidade das imagens, que consiste na variação da união das informações do primeiro plano e do fundo.

#### **a) Geração das imagens do fundo**

Para mostrar a variabilidade nos testes, foi preciso gerar várias imagens com fundos diferentes para reproduzir as imagens da base de dados. Foram selecionadas 08 imagens de forma aleatória, nas quais as informações do verso interferiam nas informações do primeiro plano. A Figura 6.31 ilustra as 04 imagens selecionadas.



a) Imagem 001v11



b) Imagem 001v21



c) Imagem 029v11



d) Imagem 070r21

Figura 6.31: Imagens com diferentes fundos.

A Figura 6.32 ilustra as outras 04 imagens selecionadas.

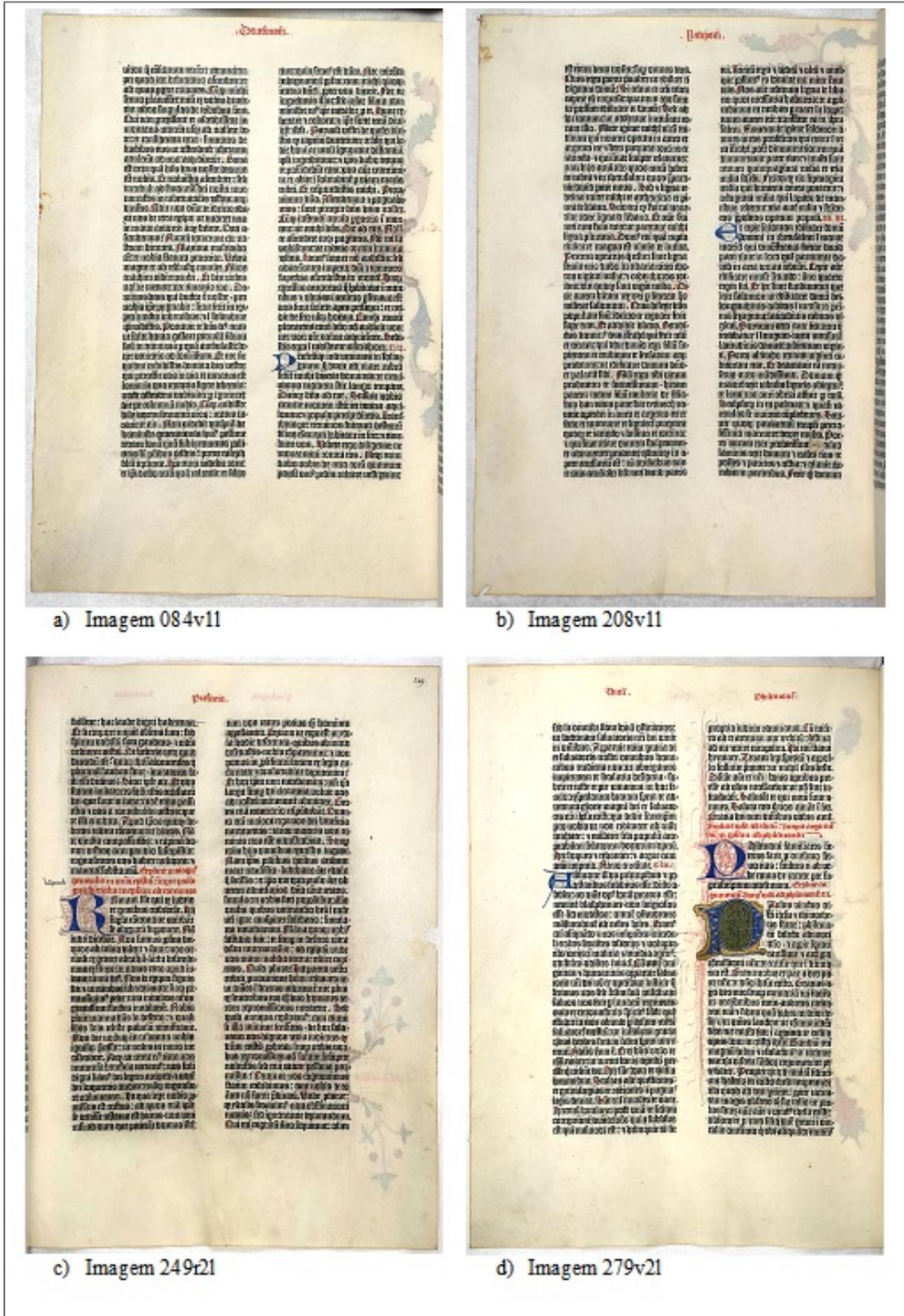


Figura 6.32: Imagens com diferentes fundos.

Com as imagens selecionadas, o próximo passo consistiu na separação somente do fundo para compor as imagens sintéticas. Realizou-se a operação morfológica de reconstrução do fundo para obtenção dos resultados. A Figura 6.33 ilustra os resultados da geração dos fundos das imagens da Figura 6.31.



Figura 6.33: Fundos reconstruídos para geração das imagens sintéticas.

A Figura 6.34 ilustra os resultados da geração dos fundos das imagens da figura 6.32.

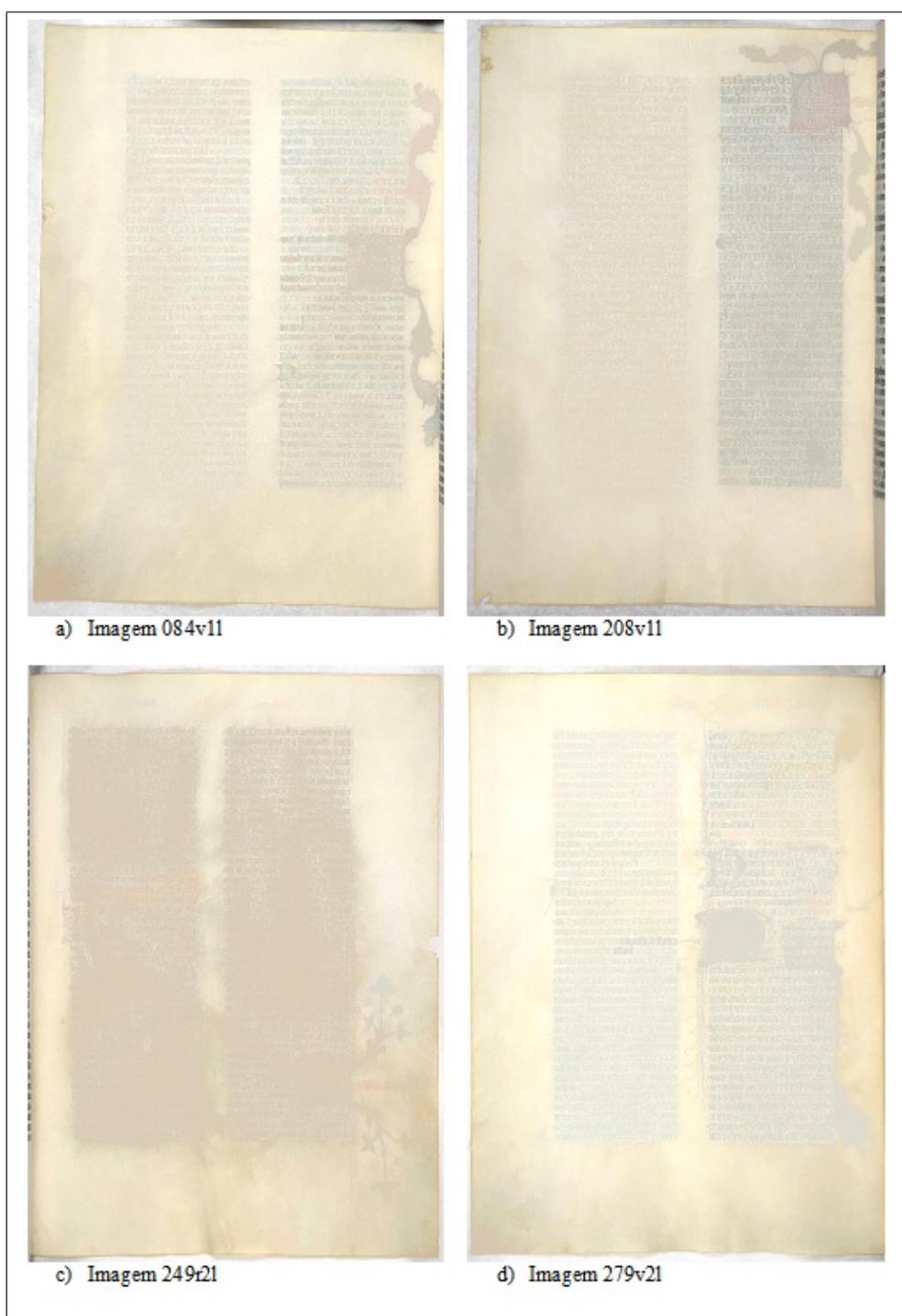


Figura 6.34: Fundos reconstruídos para geração das imagens sintéticas.

Assim, foram geradas várias imagens do fundo com o auxílio do processo 1 do método proposto, demonstrando a variação das interferências do verso nas imagens. Para a formação das imagens sintéticas, foram utilizados os 08 fundos gerados pela operação de reconstrução, precisando gerar ainda as informações do primeiro plano para junção das duas imagens.

#### b) Geração das informações do primeiro plano

Para gerar as informações do primeiro plano com o fundo branco, foram seguidas as estratégias relacionadas a seguir.

- Aquisição de informações de uma página da Bíblia em português (frases possuindo um sentido), traduzida para o latim e formatada para a fonte gótica *Barlos-Random*.
- Aquisição de uma figura, extraída de forma manual, de uma página da Bíblia de Gutenberg, procurando reproduzir uma figura impressa.
- Geração do título com cor vermelha, nos mesmos padrões de uma página da Bíblia.
- Geração de duas colunas com as informações adquiridas, frases e letras coloridas vermelhas.
- Junção das informações em uma única página, reproduzindo uma página da Bíblia com fundo branco e as informações do primeiro plano, como o título, as duas colunas centrais, as frases e letras vermelhas e o desenho.

A Figura 6.35 ilustra exemplo de uma página com as informações do primeiro plano criada de forma sintética.

Para validação e testes foram geradas 05 imagens sintéticas do primeiro plano, com desenhos diferentes, sendo criadas as imagens que constam na Figura 6.36:

## Genesis

**Capitulum 5** In principio creavit Deus caelum et terram. Terra autem erat inanis et vacua, et tenebrae erant super faciem abyssi: et Spiritus Dei ferebatur super aquas. Dixit quoque Deus: " fiat lux" Et facta est lux. Et vidit Deus lucem quod esset bona et divisit Deus lucem ac tenebras. Et vocavit Deus lucem diem et tenebras noctem. Et factum est vespere et mane, dies unus. Dixit quoque Deus: " fiat firmamentum in medio aquarum: et ab invicem separare. Et fecit Deus firmamentum divisitque aquas quae erant sub firmamento ab his quae erant super. Et factum est ita. Vocavitque Deus firmamentum, Caelum. Et factum est vespere et mane, dies secundus. Dixit vero Deus: " Congregentur aquae, quae sub caelo sunt, in locum unum: et appareat elementum sterile. Et factum est ita. Et vocavit Deus aridam Terram, congregationesque aquarum Mar. Et vidit Deus quod esset bonum. Et ait Deus: " Germinet terra herbam et herbam facientem semen et lignum pomiferum faciens fructum iuxta genus suum habet, in quo semen et fructus." Et factum est ita. Et protulit terra herbam, herbis inest semen iuxta genus suum, lignumque faciens fructum iuxta genus suum, quod fit, quo fructu seminis sui. Et vidit Deus quod esset bonum. Et factum est vespere et mane, dies tertius. Dixit autem Deus: " fiant luminaria in firmamento caeli, et dividant diem ac noctem, et serviamus eis signis in signa et tempora et dies, anni, ut luceant in firmamento caeli et inluminent terram. Et factum est ita. fecitque Deus duo luminaria magna: luminare maius, ut praesset diem: et luminare minus, ut praesset noctem: et stellas et posuit. Et posuit eas in firmamento caeli ut lucerent super terram et praesent diem vespere et divisit lucem a tenebris. Et vidit Deus quod esset bonum. Et factum est vespere et mane, dies quartus. Dixitque Deus: " Plentem multitudinem

aquarum viventium, volucres et terram sub firmamento caeli. Creavitque Deus cete grandia, et omnem multitudinem aquarum viventium quae ralis secundum species suas et omne volatile secundum genus suum. Et vidit Deus quod esset bonum. Benedictique illis Deus: " Crescite, ait, et multiplicamini, et replete aquas maris: avesque multiplicentur super terram. Et factum est vespere et mane, dies quintus. Dixit quoque Deus: " Producat terra animam viventem in genere suo, iumenta et reptilia et bestias terrae secundum species suas. Et factum est ita. Et fecit Deus bestias terrae iuxta species suas et iumenta, et reptilia quae reptant super terram et omnes. Et vidit Deus quod esset bonum.

Et ait Deus: " faciamus hominem ad imaginem et similitudinem nostram. Quod praesit piscibus maris et volatilibus caeli et bestis, et bestis, nemivers quae terrae, omnique reptili, traxit illud in terra. Et creavit Deus hominem ad imaginem suam: ad imaginem Dei creavit viri et mulieris.

Benedixitque eis Deus dicens: " Crescite, ait, et multiplicamini, et replete terram, et subicite eam. Dominamini piscibus maris et volatilibus caeli et universis animantibus, quae moventur super terram. Dixitque Deus: " Ecce dedi vobis omnem herbam offerentem semen super terram et universa ligna quae habent in semet ipsis sementem fructum, et serviemus tibi cibum. Et cunctis animantibus terrae, omnique volucri caeli, et universis quae moventur in terra, et non est anima vivens, omnem herbam virentem et ad vescendum. Et factum est ita. Et vidit Deus omnia opera sua, et vidit quod omnia essent bona valde. Et factum est vespere et mane, dies sextus.

Figura 6.35: Imagem sintética do primeiro plano.



Figura 6.36: Imagens sintéticas do primeiro plano.

As imagens geradas do primeiro plano, Figura 6.36 apresentaram variações no posicionamento e número de figuras. As figuras eram diferentes e foram recortadas de páginas originais da base da Bíblia de Gutenberg, para mostrar a variabilidade da formação das imagens sintéticas. Com as figuras do primeiro plano geradas, o próximo passo foi a união das imagens geradas do fundo com as imagens geradas do primeiro plano.

**c) União das imagens geradas do fundo com as imagens geradas do primeiro plano**

Para a união das imagens utilizou-se o Adobe Photoshop CS5, com a utilização de camadas, sendo a imagem com o fundo a camada (1) e a camada (2) somente o texto da imagem gerada artificialmente com fundo branco. Os pixels do primeiro plano e do fundo se uniram pelas bordas (mescladas gradativamente), não sendo simplesmente uma sobreposição. A Figura 6.37 ilustra a união das duas imagens, a camada 1 do fundo, representada pela imagem (a) da Figura 6.33 e a camada 2 sendo representada pela Figura 6.35.

A Figura 6.38 ilustra uma parte da imagem 6.37 ampliada, demonstrando a junção dos pixels pelas bordas.

A Figura 6.37 foi gerada com opacidade igual a 100 %. Para os testes foi preciso variar a opacidade do primeiro plano com o fundo para a validação dos testes realizados. Com a opacidade varia de 0% a 100% para as informações do primeiro plano, não alterando a opacidade do fundo, para os experimentos foram geradas opacidades com valores de 70%, 80%, 90% e 100%. A Figura 6.39 ilustra a geração de uma imagem sintética variando a opacidade.

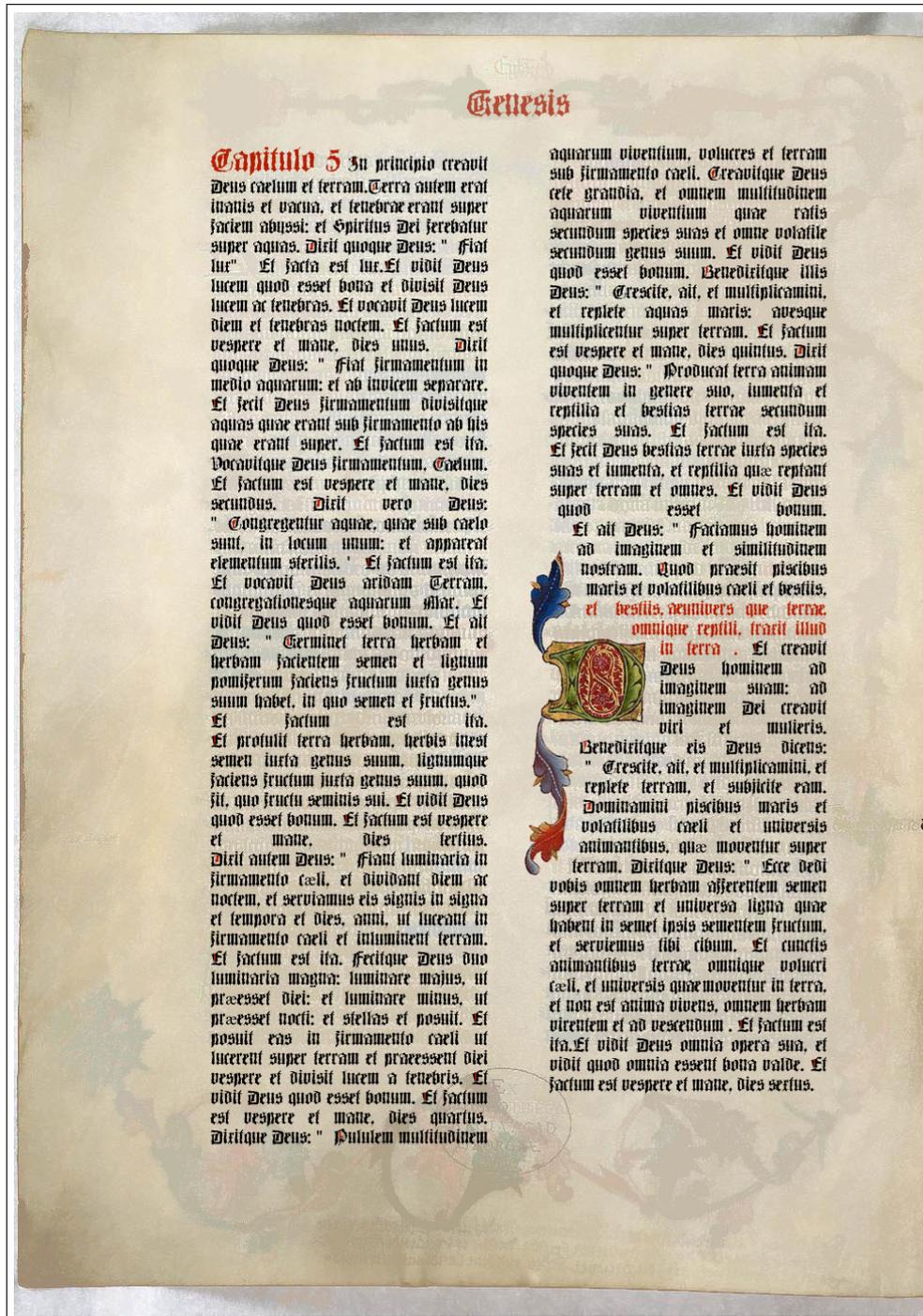


Figura 6.37: Ilustração da junção dos pixels de duas imagens.



Figura 6.38: Ilustração da junção dos pixels de uma imagem sintética com ampliação.



Figura 6.39: Geração das imagens sintéticas com variação da opacidade.

Para cada imagem gerada do primeiro plano foram geradas 08 imagens variando o fundo e a opacidade. A Figura 6.40 ilustra 04 imagens geradas com informações do primeiro plano e com variação de 04 fundos diferentes, com opacidade 100 %, mostrando a variabilidade das imagens para realização dos testes.

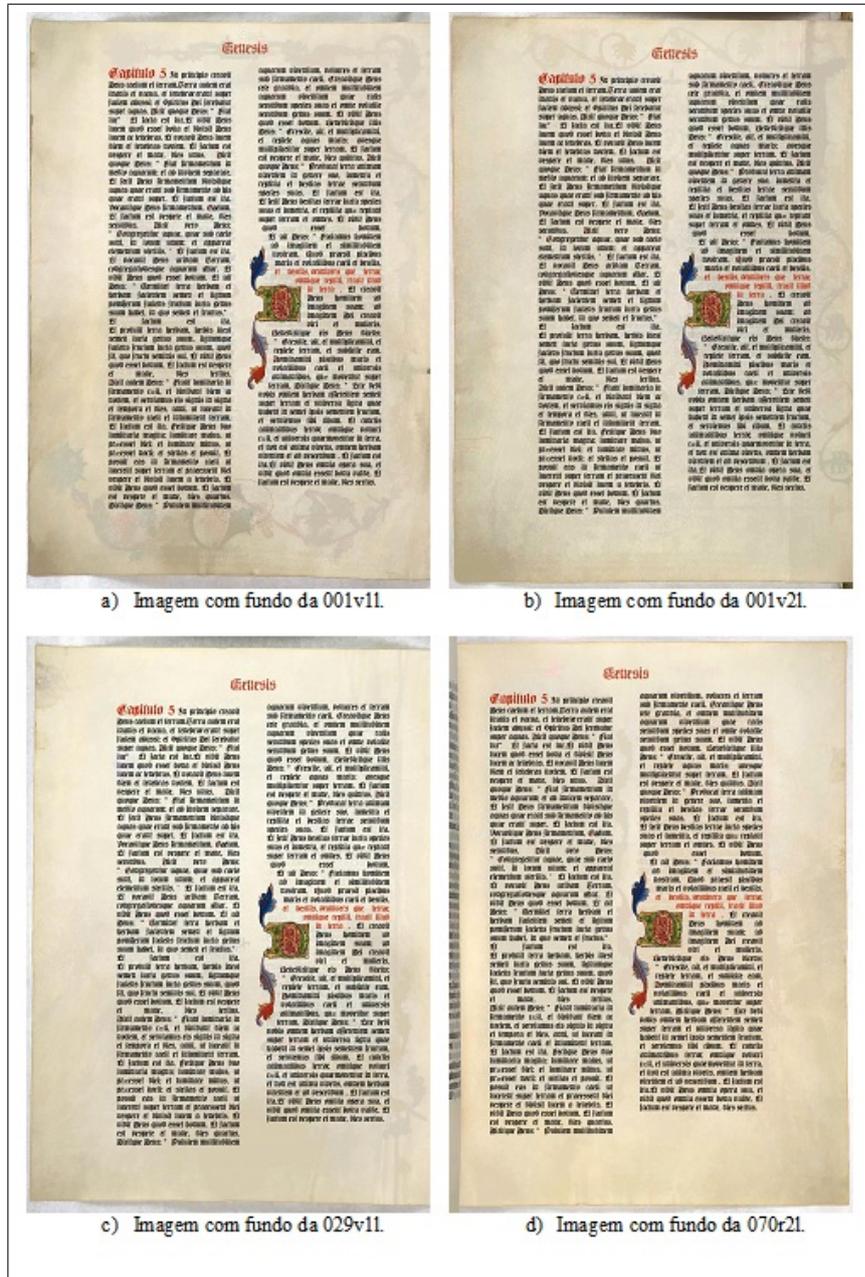


Figura 6.40: Exemplos de uma imagem sintética com 04 fundos diferentes.

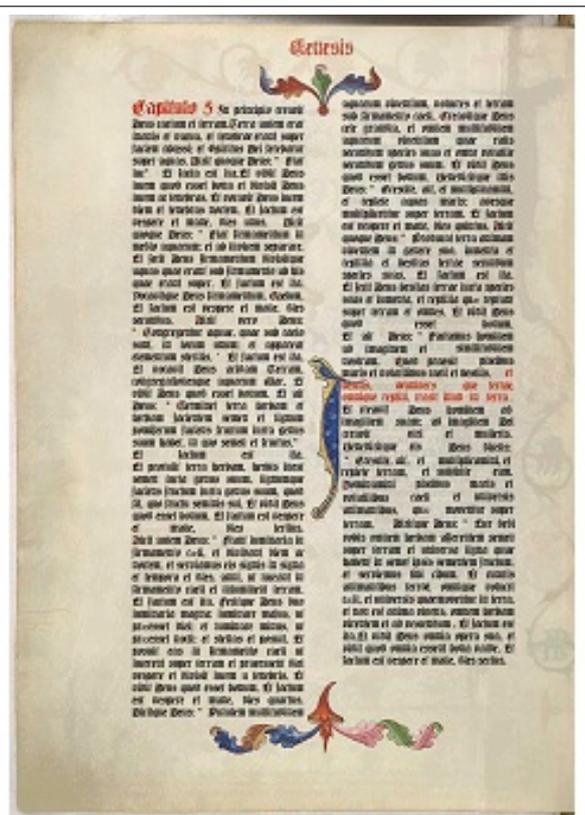
Desta maneira, ficou demonstrada a variação das imagens sintéticas criadas, com vários fundos com intensidades diferentes, com várias figuras recortadas de imagens originais das páginas da Bíblia, diversificando assim os testes e os resultados, mostrando a eficácia do método proposto para diferentes situações de uma mesma base de dados. Foram utilizadas as seguintes imagens para geração das imagens sintéticas:

- 08 imagens de fundo, geradas de forma estimada pelo processo 1 do método proposto, representando o fundo de uma imagem da Bíblia;
- 05 imagens sintéticas com conteúdo do primeiro plano, com figuras recortadas de imagens originais da base de dados da Bíblia de Gutenberg;
- 04 valores de variação da opacidade para cada imagem gerada, com valores de 70%, 80%, 90% e 100%.

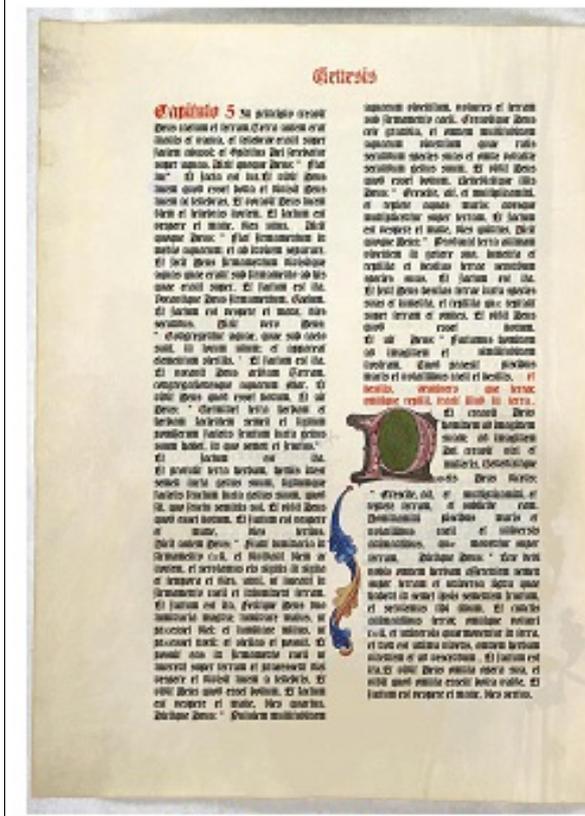
Combinando as imagens de fundo com primeiro plano e com a opacidade foram criadas 160 imagens sintéticas. A Figura 6.41 ilustra 04 imagens sintéticas geradas.



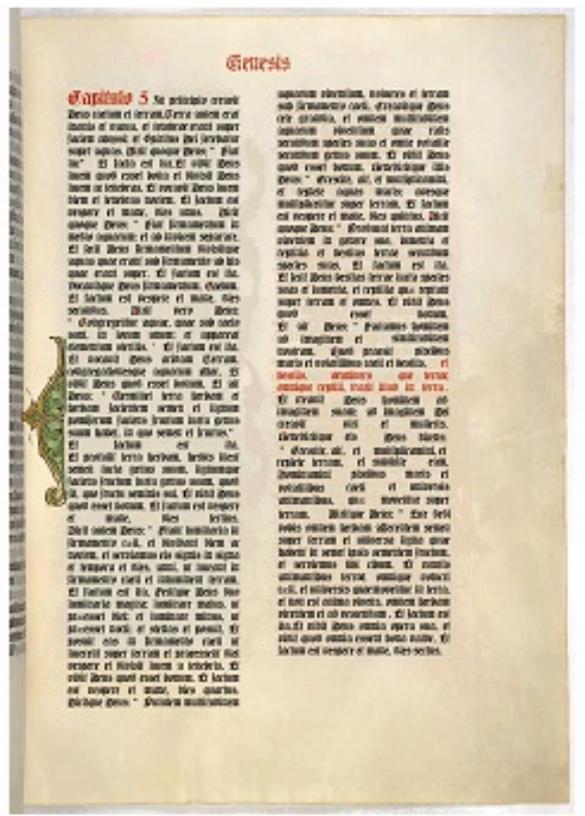
a) Imagem com fundo da 001v11.



b) Imagem com fundo da 001v21.



c) Imagem com fundo da 029v11.



d) Imagem com fundo da 070r21.

Figura 6.41: Exemplos de 04 imagens sintéticas.

Para cada imagem sintética com fundo antigo existiu uma imagem ideal para avaliar os processos 2 e 3. Para o processo 2, a imagem ideal foi representada pelas informações do primeiro plano, geradas sinteticamente com um fundo branco. Para o processo 3, a imagem ideal foi representada com as informações coloridas do primeiro plano com um fundo branco. A Figura 6.42 ilustra exemplo em (a) de uma imagem sintética, a imagem (b) representa uma imagem ideal para o processo 2 e a (c) a imagem ideal para o processo 3.

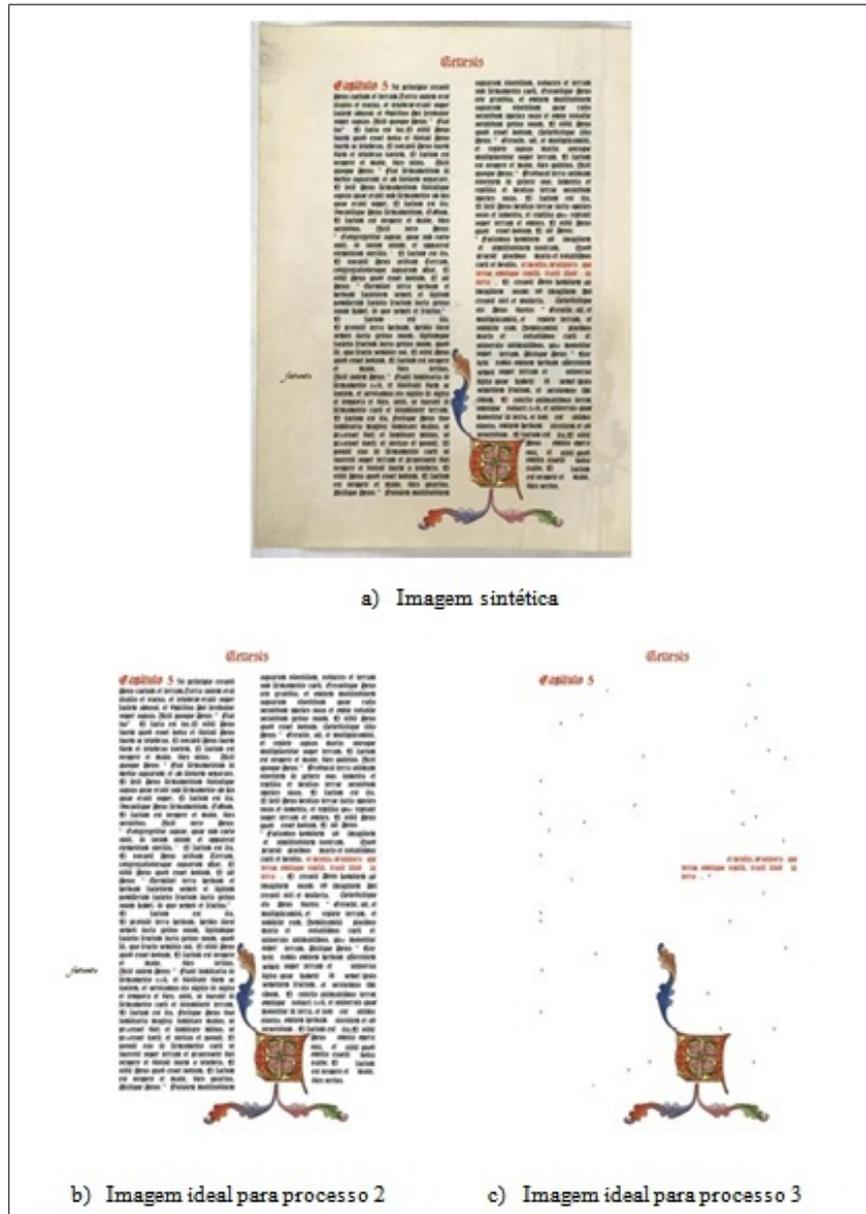


Figura 6.42: Exemplos de imagens sintéticas e ideais.

### 6.3.5.1 Avaliação do Processo 1: Estimativa do fundo homogêneo

A estimativa do fundo homogêneo, realizada pelo método proposto, foi comparada com o trabalho de Ramponi (2005), que restaurou o fundo através da limiarização de Otsu para a segmentação do fundo e do conteúdo do primeiro plano. A diferença entre as duas abordagens consistiu no algoritmo de limiarização adotado por cada um e no processo de reconstrução adotado pelo método proposto antes da limiarização.

Em Ramponi (2005) foi utilizado somente o algoritmo de limiarização de Otsu e calculada a média das crominâncias ( $C_r$  e  $C_b$ ) dos pixels do fundo para geração do fundo homogêneo. O método proposto realizou primeiramente a operação de reconstrução do fundo, em seguida procedeu-se a operação Tophat entre a imagem original e a reconstruída do fundo e finalmente aplicou-se o algoritmo de limiarização de Sauvola para a segmentação do conteúdo do primeiro plano e do fundo. A estimativa da cor do fundo pelo método proposto foi baseada nos valores máximos dos histogramas para os canais de cor da imagem reconstruída do fundo, gerando uma cor homogênea com a a limiarização de Otsu, para as imagens da Bíblia, perderam-se muitas informações das figuras se comparada com a de Sauvola. De acordo com os testes realizados, a melhor limiarização para documentos antigos que possuem imagens é a de Sauvola. Outra diferença entre as duas abordagens consistiu no uso dos espaços de cor, a abordagem de Ramponi (2005) utiliza somente o espaço de cor  $YCrCb$ , já o método proposto fez uso de quatro espaços de cor:  $HSI$ ,  $HSL$ ,  $YCrCb$  e  $YIQ$ .

### 6.3.5.2 Avaliação do Processo 2: Segmentação do conteúdo do primeiro plano

Esta etapa consistiu em quantificar os pixels que foram segmentados corretamente do primeiro plano e do fundo após as etapas de reconstrução do fundo, da operação TopHat e da limiarização de Sauvola. O resultado da aplicação do processo 2 do método proposto é ilustrado pela Figura 6.43, sendo (a) a imagem sintética utilizada como entrada, com opacidade 100 %, (b) o resultado da aplicação da restauração do fundo (processo 1), (c) o resultado Tophat das imagens ((a) e (b)), (d) o resultado da limiarização de Sauvola, (e) os pixels originais segmentados do primeiro plano (resultado do processo 2) e (f) os pixels ideais do primeiro plano.

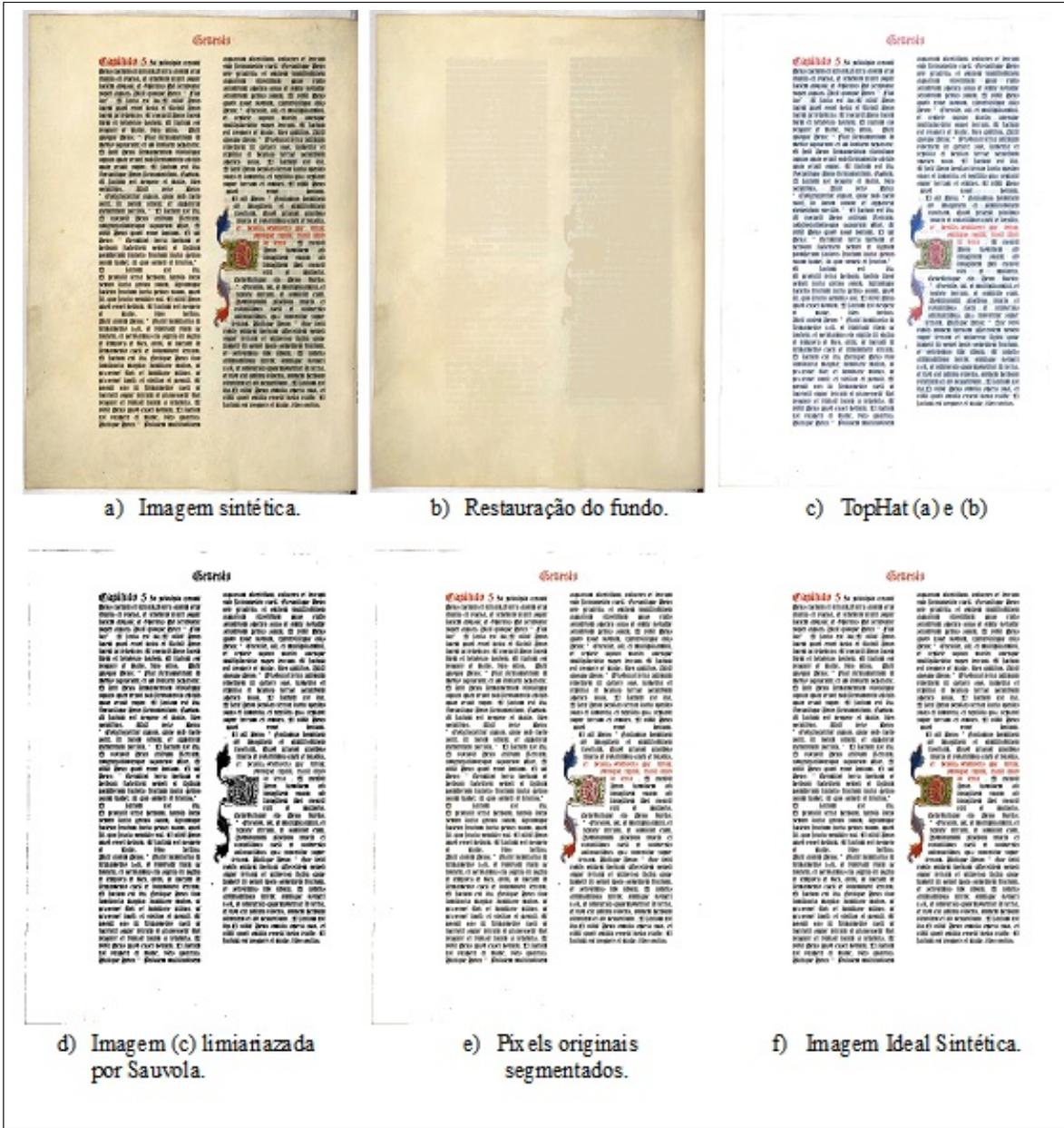


Figura 6.43: Ilustração da segmentação do conteúdo do primeiro plano de uma imagem sintética.

Com o conteúdo segmentado do primeiro plano da imagem sintética, o próximo passo foi realizar a comparação destes pixels com os da imagem ideal sintética, sendo possível saber a quantidade de pixels segmentados corretamente. Para realizar a comparação foram binarizadas as duas imagens para obtenção numérica dos resultados. A Figura 6.44 ilustra as imagens que foram comparadas e suas limiarizações. Utilizou-se o algoritmo de limiarização de Sauvola, que apresentou os melhores resultados pelos experimentos das imagens da Bíblia de Gutenberg, pois preservou mais detalhes das figuras, em comparação com os outros algoritmos de limiarizações testados.

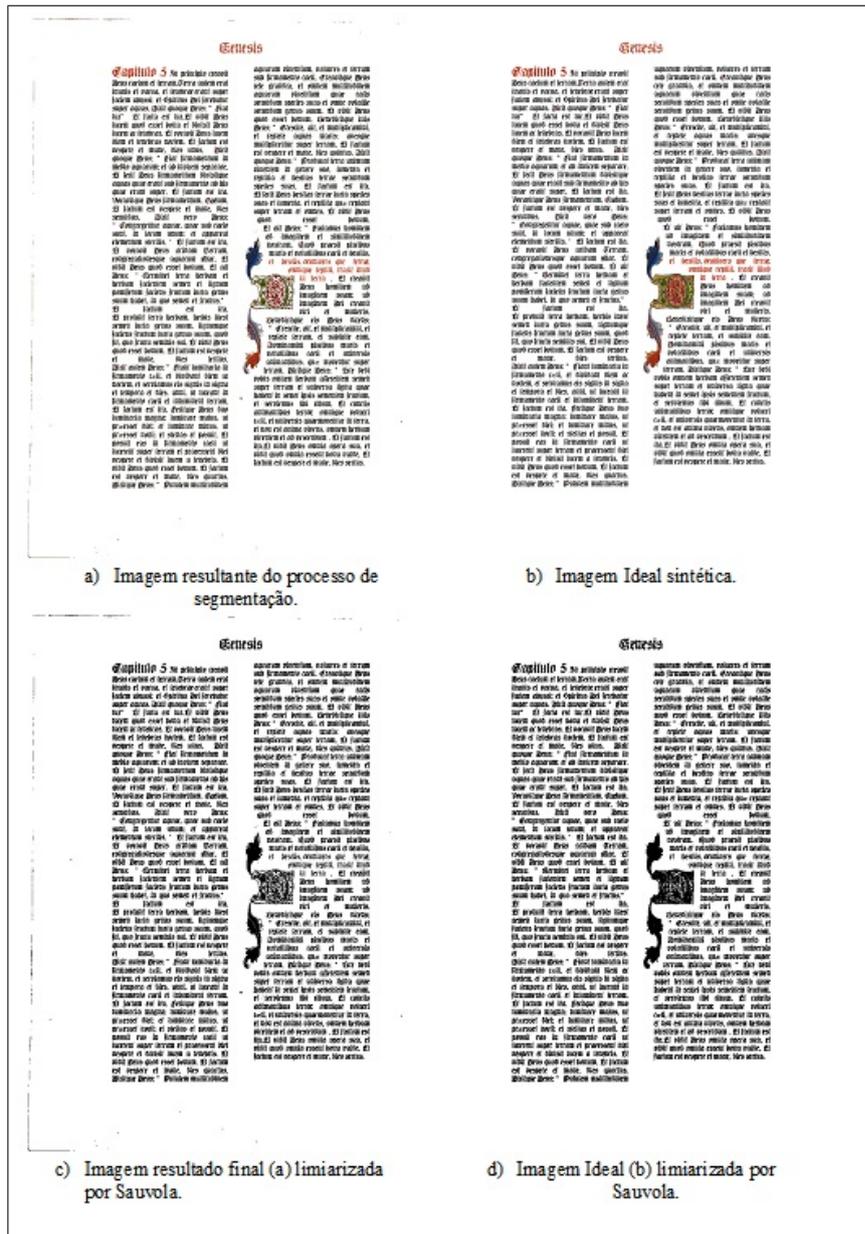


Figura 6.44: Ilustração do resultado da comparação das limiarizações para a imagem sintética.

Com as imagens binarizadas, realizou-se a comparação com a aplicação dos cálculos das métricas de avaliação entre as duas imagens. A Tabela 6.1 ilustra os resultados numéricos da comparação da Figura 6.44.

Tabela 6.1: Resultado da avaliação das limiarizações pelas métricas.

<b>Métrica</b>	<b>Sigla</b>	<b>Valor %</b>
Erro de classificação	ME	01,05
Erro Área Relativa do primeiro plano	RAE	04,34
Verdadeiro Positivo	VP	98,36
Falso Positivo	FP	00,96
Falso Negativo	FN	01,63
Verdadeiro Negativo	VN	99,03
Precisão	P	99,02
Acurácia	A	98,69
Revocação	R	98,36
Erro	E	01,30
Medida F	FM	98,69
Métrica Taxa Negativa	MTN	01,30

- A taxa do Verdadeiro Positivo atingiu 98,36% dos pixels considerados do primeiro plano.
- A taxa do Verdadeiro Negativo representou os pixels do fundo, que foram classificados corretamente como segundo plano, atingindo 99,03%.
- A precisão atingiu 99,02%, pois neste cálculo foram considerados somente os pixels segmentados corretamente e os do fundo, que foram segmentados como sendo do primeiro plano.
- A revocação indicou somente os pixels do primeiro plano que foram segmentados e os que não foram, ou seja, o cálculo foi baseado somente nos pixels do primeiro plano.
- A taxa da Acurácia significou o quanto foi alcançado da meta esperada, conseguindo-se um valor de 98,60%.
- A taxa relativa ao erro da área do primeiro plano atingiu 04,34%. Isto indica quanto da área total do primeiro plano não foi totalmente segmentado. Este problema ocorreu na parte das figuras, pois os pixels não foram totalmente recuperados no momento da limiarização.

- A medida F apresentou um valor numérico de 98,69% e envolveu valores das medidas de precisão e revocação, mostrando que os pixels do primeiro plano foram segmentados quase em sua totalidade.

O resultado apresentado pela Tabela 6.1 foi alcançado com uma imagem com opacidade 100 %, realizando-se assim, foram realizados os testes com a variação da opacidade.

A Tabela 6.2 ilustra os resultados obtidos variando opacidade em 70 %, 80 % , 90 % e 100 %.

Tabela 6.2: Resultado da comparação das técnicas de limiarizações com a variação da opacidade.

<b>Métrica</b>	<b>70%</b>	<b>80%</b>	<b>90%</b>	<b>100%</b>
Erro de classificação	01,40	01,27	01,16	01,06
Erro Área Relativa do primeiro plano	06,15	05,43	04,81	04,35
Verdadeiro Positivo	97,59	97,92	98,16	98,36
Falso Positivo	01,26	01,15	01,05	00,97
Falso Negativo	02,41	02,08	01,84	01,64
Verdadeiro Negativo	98,74	98,85	98,95	99,03
Precisão	98,73	98,84	98,94	99,03
Acurácia	98,17	98,39	98,55	98,70
Revocação	97,59	97,92	98,16	98,36
Erro	01,84	01,61	01,45	01,30
Medida F	98,15	98,38	98,55	98,36
Métrica Taxa Negativa	01,84	01,61	01,45	01,30

- A média da taxa do Verdadeiro Positivo atingiu 98,01% dos pixels considerados do primeiro plano.
- A taxa do Verdadeiro Negativo representou os pixels do fundo que foram classificados corretamente como segundo plano, atingindo 98,89%.
- A taxa da Acurácia significou que 98,45% da meta foram alcançados.
- A taxa relativa ao erro da área do primeiro plano atingiu 5,19%, isto em virtude do desenho, que não foi segmentado em sua totalidade.
- Os resultados apresentados pela Tabela 6.2 reforçaram a idéia de que o fundo interferiu no conteúdo do primeiro plano, quanto menor a opacidade, maior a interferência, menores são os resultados alcançados. Com a opacidade 100%, os resultados são maiores para precisão e revocação. A segmentação do primeiro plano, pelo método

proposto, conseguiu boas taxas, mesmo com a variação da opacidade entre 70% e 100%.

- A variação da Medida F apresentou-se pequena para as quatro variações de opacidade.
- A variação da métrica da taxa negativa pouco variou entre os quatro resultados apresentados.

## a) Resultados dos testes do Processo 2 para a base das imagens sintéticas

Foram testadas as imagens sintéticas da base criada, obtendo-se os resultados pelas métricas de avaliação, com as variações de fundo e da opacidade.

- 1) Primeiro teste: com os resultados numéricos de uma imagem com conteúdo do primeiro plano e as 08 variações de fundo.
- 2) Segundo teste: contempla os resultados numéricos de todas as imagens testadas com a variação da opacidade.

### 1) Primeiro teste - Resultado com a variação dos fundos

Para este teste foram utilizados os oito fundos gerados e o conteúdo do primeiro plano sintético ilustrado pela Figura 6.45.



Figura 6.45: Conteúdo do primeiro plano utilizado para os testes com os 08 fundos.

Serão apresentados neste primeiro teste os resultados numéricos obtidos com a variação da opacidade tem 70 %, 80 %, 90 % e 100 % para uma imagem sintética. A Tabela 6.3 mostra os resultados utilizando-se o conteúdo do primeiro plano da Figura 6.45 com a variação de 08 fundos, com opacidade 70%.

Tabela 6.3: Resultado com variação dos fundos com opacidade 70 %

Sigla	001r1l	001v2l	029v1l	070r2l	084v1l	208v1l	249r2l	279v2l	Média	Desvio Padrão
ME	01,38	01,50	01,21	01,28	01,38	01,62	01,32	01,23	01,36	00,14
RAE	10,53	11,46	09,14	08,82	09,47	11,53	09,09	07,95	09,75	01,30
VP	99,99	99,99	99,97	99,49	99,43	99,53	99,46	99,24	99,64	00,30
FP	01,57	01,70	01,36	01,38	01,49	01,77	01,43	01,29	01,50	00,17
FN	00,02	00,00	00,03	00,51	00,57	00,47	00,54	00,76	00,36	00,30
VN	98,44	98,30	98,64	98,62	98,51	98,23	98,58	98,71	98,50	00,17
P	98,46	98,33	98,65	98,63	98,53	98,25	98,59	98,72	98,52	00,16
A	99,21	99,15	99,30	99,05	98,97	98,88	99,02	98,97	99,07	00,14
R	99,99	99,99	99,97	99,49	99,43	99,53	99,46	99,24	99,64	00,30
E	00,79	00,85	00,70	00,95	01,03	01,12	00,98	01,03	00,93	00,14
MF	99,22	99,15	99,31	99,06	98,75	98,88	99,02	98,98	99,05	00,18
NRM	00,79	00,85	00,70	00,95	01,03	01,12	00,98	01,03	00,93	00,14

As Tabelas 6.4, 6.5 e 6.6 apresentam os resultados utilizando o conteúdo do primeiro plano da Figura 6.45 com os 08 fundos e opacidade 80 % 90 % e 100 % respectivamente.

Tabela 6.4: Resultado com variação dos fundos com Opacidade 80 %

Sigla	001r1l	001v2l	029v1l	070r2l	084v1l	208v1l	249r2l	279v2l	Média	Desvio Padrão
ME	01,26	01,36	01,09	01,15	01,24	01,44	01,19	01,09	01,23	00,13
RAE	09,20	10,04	07,97	07,61	08,20	09,86	07,84	06,80	08,44	01,14
VP	99,90	99,99	99,96	99,53	99,47	99,57	99,50	99,34	99,66	00,25
FP	01,44	01,55	01,24	01,25	01,34	01,58	01,28	01,15	01,35	00,15
FN	00,00	00,00	00,00	00,48	00,53	00,43	00,50	00,66	00,32	00,27
VN	98,56	98,45	98,76	98,76	98,66	98,42	98,72	98,85	98,65	00,15
P	98,59	98,48	98,78	98,77	98,67	98,44	98,73	98,86	98,66	00,15
A	99,28	99,22	99,36	99,14	99,07	99,00	99,11	99,10	99,16	00,12
R	99,93	99,99	99,96	99,53	99,47	99,57	99,50	99,34	99,66	00,26
E	00,72	00,78	00,64	00,86	00,94	01,00	00,89	00,90	00,84	00,12
MF	99,28	99,23	99,37	99,14	99,07	99,00	99,11	99,10	99,16	00,12
NRM	00,72	00,78	00,64	00,86	00,94	01,00	00,89	00,90	00,84	00,12

Pelos resultados numéricos apresentados pelas Tabelas 6.3, 6.4, 6.5 e 6.6 algumas considerações são importantes.

- Quanto menor a opacidade, pior foram os resultados obtidos. A métrica de revocação atingiu 99,69% para opacidade 100% e 99,64% para opacidade 70%, notando-se que a diferença foi pequena para essa métrica.
- A taxa de precisão atingiu em média 98,88% de acerto para imagens com opacidade 100%. Para imagens com opacidade 70%, a taxa caiu para 98,52%.

Tabela 6.5: Resultado com variação dos fundos com Opacidade 90 %

Sigla	001r1l	001v2l	029v1l	070r2l	084v1l	208v1l	249r2l	279v2l	Média	Desvio Padrão
RAE	08,14	08,85	06,91	06,74	07,15	08,77	06,85	06,01	07,43	01,04
ME	01,13	01,23	00,97	01,04	01,12	01,31	01,07	00,98	01,11	00,12
VP	99,99	99,99	99,97	99,61	99,52	99,65	99,53	99,44	99,71	00,23
FP	01,30	01,41	01,11	01,13	01,21	01,45	01,16	01,04	01,23	00,14
FN	00,00	00,01	00,03	00,39	00,48	00,36	00,47	00,56	00,29	00,23
VN	98,70	98,59	98,89	98,87	98,79	08,55	98,84	98,96	98,77	00,14
P	98,72	98,61	98,91	98,88	98,80	98,57	98,85	98,96	98,79	00,14
A	99,35	99,29	99,43	99,24	99,15	99,10	99,19	99,20	99,24	00,11
R	99,99	99,99	99,97	99,61	99,52	99,65	99,53	99,44	99,71	00,23
E	00,65	00,71	00,57	00,76	00,85	00,90	00,82	00,80	00,76	00,11
MF	99,35	99,30	99,43	99,24	99,16	99,10	99,19	99,20	99,25	00,11
NRM	00,65	00,71	00,57	00,76	00,85	00,90	00,82	00,80	00,76	00,11

Tabela 6.6: Resultado com variação dos fundos com Opacidade 100 %

Sigla	001r1l	001v2l	029v1l	070r2l	084v1l	208v1l	249r2l	279v2l	Média	Desvio Padrão
ME	01,03	00,88	00,88	00,94	01,01	01,19	00,99	00,89	00,98	00,10
RAE	07,29	06,17	06,17	05,89	06,38	07,84	06,23	05,40	06,42	00,78
VP	99,99	99,97	99,97	99,58	99,56	99,66	99,58	99,50	99,73	00,21
FP	01,18	01,01	01,01	01,02	01,10	01,32	01,08	00,95	01,08	00,12
FN	00,00	00,03	00,03	00,41	00,43	00,33	00,41	00,49	00,27	00,21
VN	98,81	98,98	98,98	98,97	98,89	98,67	98,91	99,04	98,91	00,12
P	98,82	98,99	98,99	98,98	98,90	98,68	98,92	99,04	98,92	00,12
A	99,40	99,48	99,48	99,28	99,22	99,17	99,25	99,27	99,32	00,12
R	99,99	99,97	99,97	99,28	99,56	99,66	99,58	99,50	99,69	00,26
E	00,59	00,51	00,51	00,71	00,77	00,83	00,74	00,72	00,67	00,12
MF	99,40	99,48	99,48	99,28	99,23	99,17	99,25	99,27	99,32	00,12
NRM	00,59	00,51	00,51	00,71	00,77	00,83	00,74	00,72	00,67	00,12

- A taxa média da acurácia atingiu 99,20 % da meta de segmentação do conteúdo do primeiro plano, demonstrando uma segmentação quase perfeita.
- A taxa da medida F apresentou maior média com 99,32% para opacidade 100% e 99,05% para opacidade 70%, mostrando uma pequena diferença nos resultados numéricos, confirmando que, mesmo variando a opacidade, os resultados se mantiveram próximos.
- A variação da métrica da taxa negativa atingiu uma média de 0,93% para opacidade 70% e 0,67% para opacidade 100%, demonstrando que com opacidade maior a segmentação dos falsos negativos e falsos positivos foi menor. Isto aconteceu em virtude de as imagens com opacidade menores possuírem mais ruídos e porque, no momento da limiarização, muitos falsos negativos não foram segmentados e os falsos positivos apareceram por causa da borda do *scanner*, ocorrida no momento da digitalização.

## 2) Segundo teste - Resultados de toda a base sintética para o Processo 2

Nesta seção são apresentados todos os resultados numéricos obtidos em todas as imagens sintéticas geradas, demonstrando os valores médios e desvio padrão de cada métrica, com variação da opacidade em 70 %, 80 %, 90 % e 100 %, totalizando 160 imagens testadas. A Tabela 6.7 apresenta esses resultados.

Tabela 6.7: Resultado com variação da opacidade nas 160 imagens testadas para o processo 2.

Métrica	Opacidade 70%		Opacidade 80%		Opacidade 90%		Opacidade 100%	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
ME	1,35	9,21	1,23	0,19	1,1	0,18	0,99	0,16
RAE	9,95	1,63	8,73	1,44	7,64	1,32	6,7	1,13
VP	99,89	0,09	99,9	0,08	99,91	0,07	99,92	0,06
FP	1,52	0,24	1,39	0,22	1,26	0,21	1,36	0,77
FN	0,11	0,09	0,1	0,08	0,09	0,07	0,08	0,07
VN	98,48	0,24	95,84	9,03	95,72	8,73	98,87	0,18
P	98,5	0,23	98,63	0,21	98,76	0,2	98,88	0,18
A	99,18	0,13	99,25	0,12	99,33	0,1	99,39	0,11
R	99,87	0,16	99,88	0,12	99,89	0,13	99,89	0,13
E	0,82	0,13	0,75	0,12	0,67	0,11	0,6	0,11
MF	99,18	0,14	99,26	0,12	99,33	0,11	99,4	0,1
NRM	0,82	0,13	0,75	0,12	0,67	0,11	0,6	0,11

Os resultados apresentados pela Tabela 6.7 levam às seguintes conclusões:

- Variando a opacidade, os valores numéricos das métricas variam, os melhores resultados foram apresentados com opacidade 100%, e os piores com 70%, devido à mistura das informações do primeiro plano com o fundo deteriorado.
- Os resultados obtidos pelo cálculo do desvio padrão mostraram que, mesmo com a variação dos fundos, das informações do primeiro plano e da opacidade, houve uma discrepância relativamente pequena.
- A taxa média do Verdadeiro Positivo ( $VP$ ) está acima de 99,90%, com desvio padrão de 0,06%, mostrando que está muito próximo da segmentação perfeita.
- A taxa média da acurácia ( $A$ ) atingiu 99,28%, com desvio padrão de 0,11, mostrando que a meta quase foi alcançada em sua totalidade.
- A taxa média da revocação ( $R$ ) indicou que 99,88% dos pixels do primeiro plano foram segmentados corretamente, demonstrando que os pixels que re-

presentaram o primeiro plano foram segmentados quase na sua totalidade, se comparados somente com os pixels do primeiro plano.

- A taxa média da precisão ( $P$ ) foi menor que a revocação ( $R$ ), indicando que alguns pixels do fundo foram considerados como sendo verdadeiros.
- A medida F ( $MF$ ) indicou que em média, 99,29% dos pixels foram segmentados corretamente, pois foram envolvidos neste cálculo os valores da precisão e revocação, indicando maior fidelidade dos resultados alcançados.
- A variação da métrica da taxa negativa atingiu uma média de 00,82% para opacidade 70% e 00,60% para opacidade 100%.
- A taxa da área do primeiro plano que não foi segmentada atingiu 9,95%, em média, com opacidade 70%, e 6,70% com opacidade 100%, porque as figuras não foram serem segmentadas em sua totalidade em virtude da operação de limiarização, que não conseguiu a totalidade. Outro problema aconteceu onde alguns pixels referentes à imagem eram muito claros, e no momento da limiarização, foram considerados como sendo da classe do fundo, e também havia pixels dos desenhos com tonalidade parecida com o fundo e foram considerados como fundo no momento da limiarização.

### 6.3.5.3 Avaliação do Processo 3: Segmentação do conteúdo colorido

A avaliação deste processo abrangeu a quantificação dos pixels coloridos segmentados através das métricas, pelas imagens obtidas pelo método proposto e imagens ideais *ground truth*.

Para esta avaliação foram utilizadas as imagens sintéticas geradas com as variações do fundo e do primeiro plano.

A Figura 6.46 ilustra a imagem sintética em (a), (b) o resultado da segmentação das informações coloridas pelo processo 3 e (c) a imagem ideal sintética *ground truth*.



Figura 6.46: Imagens para avaliação quantitativa do processo 3.

Para realizar a comparação das imagens (b) e (c) da Figura 6.46, elas foram limiarizadas pelo algoritmo de Sauvola para avaliação pelas métricas, através de duas maneiras.

- Limiarização direta dos resultados do método proposto e da imagem ideal;
- Limiarização após a operação de reconstrução da imagem ideal sintética.

#### a) Limiarização direta dos resultados do método proposto e da imagem ideal

Esta etapa consistiu na limiarização, pelo algoritmo de Sauvola, do resultado da imagem do processo 3 do método proposto e da imagem ideal sintética (*ground truth*). A Figura 6.47 ilustra o resultado do processo 3 limiarizada e a imagem (b) a limiarização da imagem ideal.

A imagem limiarizada continha o contorno das pequenas letras vermelhas e a imagem ideal não, isso influenciou na comparação dos resultados, por isso realizou-se a operação de reconstrução na imagem ideal para posterior limiarização e comparação. A Tabela 6.8 ilustra os resultados obtidos na comparação das duas imagens da Figura 6.47.

Os resultados apresentados pela Tabela 6.8 mostraram:

- A taxa da revocação ( $R$ ) indicou que 90,21 % dos pixels considerados do primeiro plano, que são as letras coloridas, figuras e letras pequenas, foram segmentados corretamente. A perda da segmentação está em: a) algumas letras

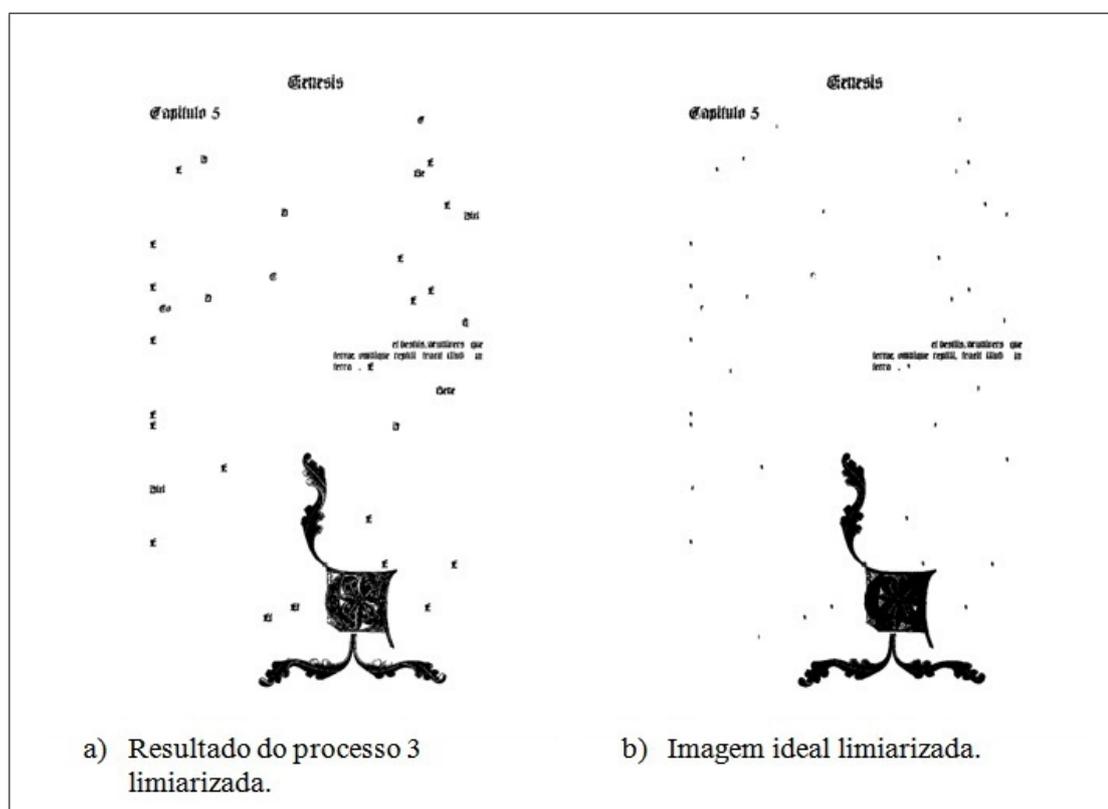


Figura 6.47: Resultado da limiarização das imagens para avaliação.

Tabela 6.8: Resultado do Processo 3 - Limiarização direta da imagem ideal.

Métrica	Sigla	Valor
Erro de classificação	ME	00,60
Erro Área Relativa do primeiro plano	RAE	02,53
Verdadeiro Positivo	VP	90,21
Falso Positivo	FP	00,34
Falso Negativo	FN	09,80
Verdadeiro Negativo	VN	99,66
Precisão	P	99,62
Acurácia	A	94,93
Revocação	R	90,21
Erro	E	05,07
Medida F	MF	94,68
Métrica Taxa Negativa	NRM	05,07

pequenas coloridas não foram identificadas no processo de limiarização e após o processamento não foram reconstruídas; b) as letras pequenas possuíam um contorno que influenciou na segmentação correta dos pixels do primeiro plano, devido ao processo de reconstrução.

- A taxa de verdadeiro negativo ( $VN$ ) indicou que os pixels do fundo foram

classificados corretamente quase na sua totalidade.

- A taxa de acurácia ( $A$ ) indicou que a meta de segmentação dos pixels coloridos não foi atingida na sua totalidade, em virtude de alguns pixels vermelhos não terem sido identificados pelo processo 3.
- A medida  $F$  ( $MF$ ) combinou informações da precisão ( $P$ ) e revocação ( $R$ ), indicou que 94,68% dos pixels do primeiro plano foram segmentados corretamente.
- A Métrica Taxa Negativa ( $MTN$ ) indicou que 5,07%, representando a média dos pixels falsos negativos e falsos positivos, não foram segmentados corretamente.
- A precisão ( $P$ ) atingiu 99,62%, pois foi calculada somente com os valores dos pixels verdadeiros segmentados e dos pixels falsos positivos, por isso mostrou-se elevado o valor numérico, não sendo condizente com a realidade, que deve ser representada pela métrica da revocação.

## b) Limiarização após a operação de reconstrução da imagem ideal

Esta etapa consistiu em realizar a operação de reconstrução em cor da imagem ideal para recuperar o contorno das pequenas letras vermelhas e o fundo da imagem, para posterior limiarização e comparação. A Figura 6.48 ilustra a operação de reconstrução.

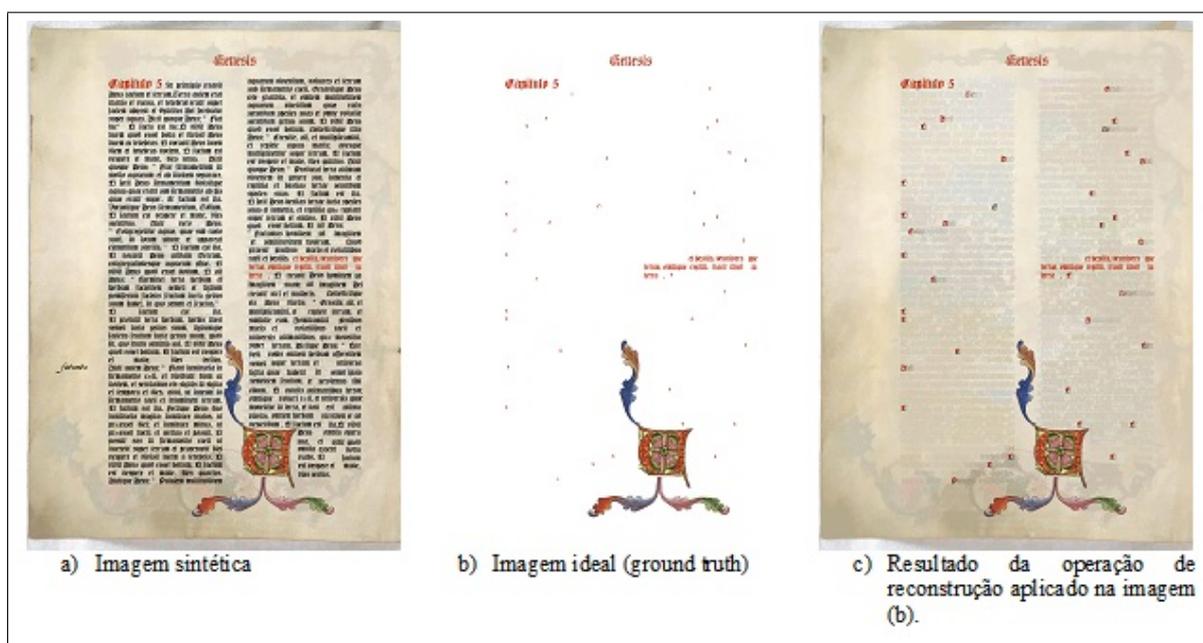


Figura 6.48: Resultado da operação de reconstrução da imagem ideal.

Com o resultado do processo 3 do método proposto e a imagem ideal reconstruída, o próximo passo foi realizar a operação de limiarização das duas imagens para comparação e avaliação das métricas. A Figura 6.49 ilustra as imagens limiarizadas pelo algoritmo de Sauvola e a Tabela 6.9 apresenta os resultados da comparação das duas imagens limiarizadas.

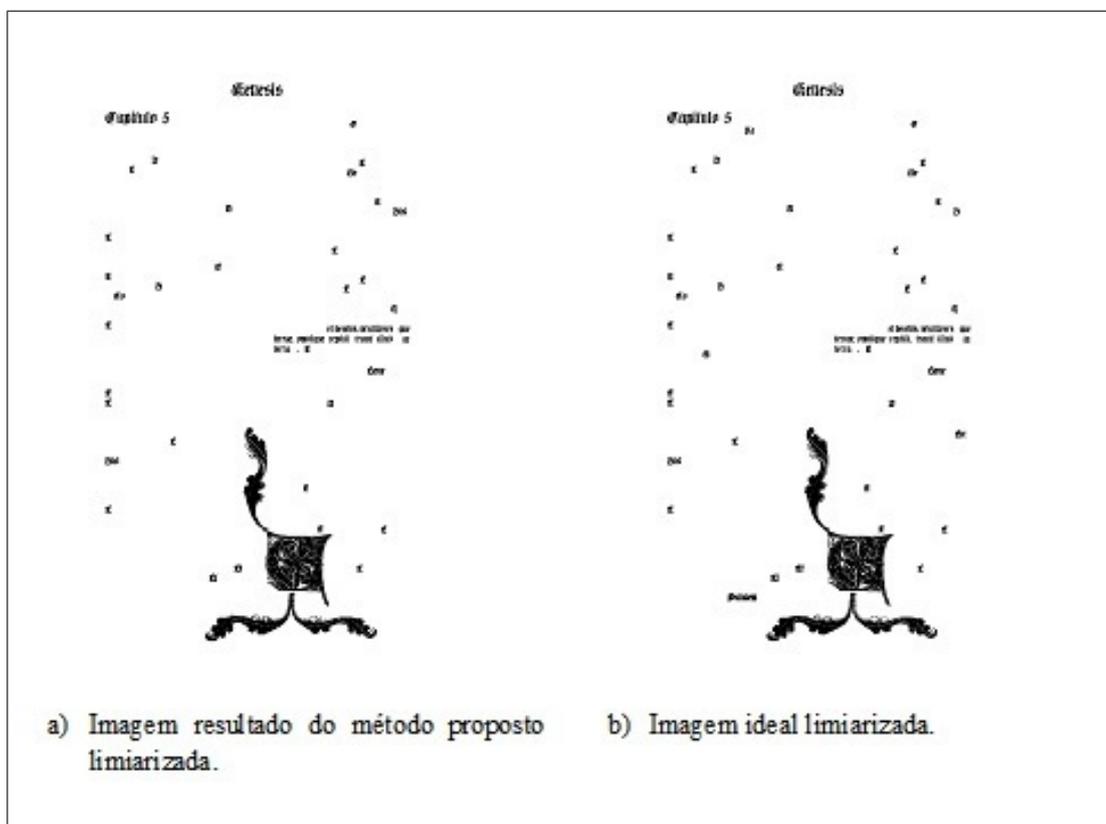


Figura 6.49: Resultado da limiarização das imagens.

- A taxa de revocação ( $R$ ) indicou que 98,38% dos pixels considerados do primeiro plano, letras coloridas, figuras e letras pequenas, foram segmentados corretamente. A perda da segmentação ocorreu em algumas letras pequenas coloridas que não foram identificadas no processo de limiarização e, após, não foram reconstruídas.
- A taxa de verdadeiro negativo ( $VN$ ) indicou que os pixels do fundo foram classificados corretamente quase na sua totalidade.
- A taxa de acurácia ( $A$ ) indicou que 99,19% da meta de segmentação dos pixels coloridos foram alcançados.
- A medida F ( $MF$ ) combinou informações da precisão ( $P$ ) e revocação ( $R$ ),

Tabela 6.9: Resultado do Processo 3 - Limiarização após a reconstrução da imagem ideal.

<b>Métrica</b>	<b>Sigla</b>	<b>Valor</b>
Erro de classificação	ME	00,13
Erro Área Relativa do primeiro plano	RAE	01,85
Verdadeiro Positivo	VP	98,48
Falso Positivo	FP	00,09
Falso Negativo	FN	01,52
Verdadeiro Negativo	VN	99,91
Precisão	P	99,90
Acurácia	A	99,19
Recall	R	98,38
Erro	E	00,81
Medida F	MF	99,19
Métrica Taxa Negativa	MTN	00,81

indicando que 99,19% dos pixels do primeiro plano foram segmentados corretamente.

- A Métrica Taxa Negativa (*MTN*) indicou que 00,81% representou a média dos pixels falsos negativos e falsos positivos que não foram segmentados corretamente.
- A precisão (*P*) atingiu 99,90%, pois foi calculado somente com os valores dos pixels verdadeiros segmentados e dos pixels falsos positivos, por isso foi elevado o valor numérico, não sendo condizente com a realidade, que é representada pela métrica da revocação.

Comparando os resultados da Tabela 6.8 e 6.9, os valores numéricos indicaram que:

- A diferença da taxa de revocação (*R*) entre as duas tabelas foi de 08,17%, indicando que a segunda estratégia apresentou os melhores resultados, em virtude do processo de reconstrução das pequenas letras vermelhas, sendo recuperados os contornos das letras.
- Em geral, as taxas apresentaram-se melhores com a utilização da estratégia da reconstrução da imagem *ground truth* antes do processo de limiarização e comparação.

A seguir serão apresentados os resultados numéricos dos testes realizados na base de dados sintética.

### c) Resultados numéricos dos testes realizados para as 160 imagens sintéticas para o Processo 3

Os testes realizados foram descritos numericamente com os valores médios e desvio padrão considerando os 4 valores de opacidade e as duas estratégias de comparação. A Tabela 6.10 apresenta os resultados das 160 imagens com a estratégia da limiarização direta. A tabela 6.11 apresenta os resultados com a estratégia da limiarização após a operação de reconstrução.

Tabela 6.10: Resultado com a variação da opacidade para as 160 imagens testadas para o processo 3 - Com Limiarização Direta.

Métrica	Opacidade 70%		Opacidade 80%		Opacidade 90%		Opacidade 100%	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
ME	1,16	0,31	0,87	0,45	0,82	0,38	0,81	0,22
RAE	16,70	4,61	9,72	7,76	7,77	6,71	6,62	6,25
VP	82,21	10,64	84,17	9,52	84,85	8,39	85,29	06,40
FP	0,64	0,15	0,42	0,24	0,39	0,22	0,47	0,18
FN	17,79	10,64	15,83	9,52	15,15	8,39	14,15	05,11
VN	99,36	0,15	99,58	0,24	99,61	0,22	99,63	0,18
P	99,23	0,11	99,50	0,29	99,53	0,26	99,56	0,21
A	90,79	5,26	90,52	5,06	92,23	4,23	92,69	02,55
R	82,21	10,64	84,17	9,52	84,85	8,39	85,32	06,42
E	9,21	5,26	8,12	4,77	7,77	4,23	7,31	2,55
MF	89,59	6,67	91,03	5,99	91,65	5,03	92,06	3,08
NRM	9,21	5,26	8,12	4,77	7,77	4,23	7,31	2,55

- O valor numérico da revocação ( $R$ ) indicou que quanto menor a opacidade piores foram os resultados, não sendo segmentados corretamente os pixels em cor. O valor médio para opacidade 100% registrou 85,32 % dos pixels segmentados considerados do primeiro plano das letras coloridas, figuras e letras pequenas. A perda da segmentação deve-se porque algumas letras pequenas coloridas não foram identificadas no processo de limiarização, não sendo reconstruídas.
- A taxa do falso negativo ( $FN$ ) indicou que 14,15% dos pixels, que eram verdadeiros, não foram segmentados, significando que os pixels das pequenas letras vermelhas não foram segmentados.
- A taxa de acurácia ( $A$ ) indicou que 92,69% da meta de segmentação dos pixels coloridos foram atingidos, o problema residiu nos pixels que não foram identificados, localizados e reconstruídos.
- A Medida F ( $MF$ ) indicou que 91,72% dos pixels do primeiro plano foram segmentados.
- No geral, os números mostraram que, quanto menor a opacidade, piores são os resultados.

Tabela 6.11: Resultado com a variação da opacidade para as 160 imagens testadas para o processo 3 - Com Limiarização após Reconstrução.

Métrica	Opacidade 70%		Opacidade 80%		Opacidade 90%		Opacidade 100%	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
ME	0,91	0,45	0,79	0,46	0,53	0,40	0,43	0,32
RAE	13,90	8,67	10,40	8,65	10,18	7,38	9,62	6,70
VP	85,39	11,42	88,57	10,86	94,88	7,46	96,09	5,42
FP	0,45	0,27	0,42	0,16	0,39	0,29	0,34	0,24
FN	14,33	11,83	11,43	10,86	5,12	7,46	3,91	5,42
VN	99,55	0,27	99,58	0,16	99,61	0,29	99,66	0,24
P	99,47	0,28	99,51	0,19	99,58	0,30	99,64	0,26
A	92,47	5,67	94,07	5,45	97,25	3,76	97,88	2,76
R	85,39	11,42	88,57	10,86	94,91	7,48	96,09	5,42
E	7,53	5,67	5,93	5,45	2,76	3,76	2,12	2,76
MF	91,56	6,72	93,44	6,27	97,04	4,19	97,76	2,99
NRM	7,53	5,67	5,93	5,45	2,76	3,76	2,12	2,76

- A taxa de revocação ( $R$ ) indicou que, com opacidade 100%, os resultados obtidos atingiram 96,09% dos pixels considerados do primeiro plano, e que com 70% atingiram 85,39%, devido ao processo de limiarização da imagem com muitas informações das pequenas letras que não foram localizadas e reconstruídas.
- A taxa de acurácia ( $A$ ), com 100% de opacidade, indicou que 97,88% da meta de segmentação dos pixels coloridos foram atingidos, ocorrendo o mesmo problema citado na revocação.
- A Medida F ( $MF$ ), com 100% de opacidade, indicou que 97,76% dos pixels do primeiro plano foram segmentados.

Como os melhores resultados foram alcançados com a opacidade 100 %, a Tabela 6.12 apresenta a comparação das duas estratégias utilizadas, mostrando a variação numérica.

Tabela 6.12: Resultado da avaliação do processo 3 para imagens com opacidade 100 %.

Métrica	Limiarização Direta		Limiarização após a Reconstrução	
	Média	Desvio Padrão	Média	Desvio Padrão
Erro de classificação	0,81	0,22	0,43	0,32
Erro Área Relativa do primeiro plano	6,62	6,25	9,62	6,7
Verdadeiro Positivo	85,29	6,40	96,09	5,42
False Positivo	0,47	0,18	0,34	0,24
False Negativo	14,15	5,11	3,91	5,42
Verdadeiro Negativo	99,63	0,18	99,66	0,24
Precisão	99,56	0,21	99,64	0,26
Acurácia	92,69	2,55	97,88	2,76
Revocação	85,32	6,42	96,09	5,42
Erro	7,31	2,55	2,12	2,76
Medida F	92,06	3,08	97,76	2,99
Métrica Taxa Negativa	7,31	2,55	2,12	2,76

Os resultados numéricos ilustrados pela Tabela 6.12 demonstraram que:

- A imagem *ground truth* reconstruída ofereceu os melhores resultados, pois os

contornos das pequenas letras foram reconstruídos e, no momento da comparação, o resultado obtido pelo método proposto continha também os contornos das pequenas letras.

- O valor numérico da revocação ( $R$ ) indicou uma diferença de 10,77% entre as duas estratégias.
- A medida F ( $MF$ ) indicou que, com a estratégia de limiarização após a operação de reconstrução, o valor numérico foi melhor. O desvio padrão para a Medida F foi menor com a estratégia 1.

#### d) Considerações sobre os resultados do processo 3

Com esses testes foi possível avaliar esta etapa do método proposto, de segmentação do conteúdo colorido com o uso da morfologia matemática em cor. Algumas considerações devem ser feitas:

- A parte de segmentação das pequenas letras coloridas vermelhas não atingiram a totalidade real de identificação devido à reconstrução e filtragem dos ruídos, pois alguns pixels que representavam as pequenas letras coloridas foram eliminados pelo processo de filtragem. Se não fosse realizada a filtragem, muitos pixels falsos positivos seriam reconstruídos, deixando as taxas muito ruins.
- Esta abordagem apresentou uma maneira de identificar as pequenas letras vermelhas que estavam presentes nas imagens das páginas da Bíblia, que passavam despercebidas a olho nu, podendo servir para estudos pelos historiadores.
- Este processo conseguiu demonstrar que foi possível identificar características coloridas das páginas da Bíblia utilizando a morfologia matemática em cor com suas ordenações e ordem.

#### e) Considerações Finais do Nível 1

- Bons resultados foram apresentados para os espaços de cor em que os canais eram descorrelacionados e a ordem dos canais interferiram, ou seja, cada ordem resultava em imagens diferentes, precisando definir qual a melhor ordem para atingir determinado objetivo.
- Para a estimativa do fundo das imagens da Bíblia, foi melhor utilizar os espaços de cor YCrCb ou YIQ, com o uso da ordenação lexicográfica com um ou mais canais, sem a preocupação de definição do valor de qualquer canal *a priori*. No

caso do uso dos espaços de cor HSI/HSL, foi preciso definir *a priori* um valor para o canal H, ou calcular o H majoritário para realização das operações com uso deste canal, consumindo assim mais tempo de processamento e também ficando dependente da indicação de cada usuário o valor de H.

- A segmentação do conteúdo do primeiro plano, pelo processo 2, do método proposto apresentou ótimos resultados avaliados pelas imagens sintéticas geradas. Para esta etapa o melhor algoritmo de limiarização foi o de Sauvola, sendo testados vários algoritmos.
- Para a etapa de segmentação do conteúdo colorido - Processo 3, verificou-se os melhores resultados com o uso do espaço de cor HSI ou HSL, com ordem dos canais SHI ou SHL. O espaço de cor YCrCb com ordem CbCrY não apresentavam bons resultados, quando comparados com o SHI ou SHL. O espaço de cor YIQ não apresentaram resultados satisfatórios para este processo.

## 6.4 Experimentos do Nível 2

Nesta etapa estão descritos os experimentos do nível 2, que consiste na segmentação do conteúdo do primeiro plano em dois grupos através do Processo 4 do método proposto.

### 6.4.1 Processo 4: Segmentação do conteúdo do primeiro plano

O conteúdo do primeiro plano pode ser dividido em dois grupos:

- Grupo 1: giguras, títulos, manuscritos e numeração de páginas;
- Grupo 2: colunas centrais impressas.

Uma maneira de diferenciar e segmentar os dois grupos é localizando as colunas centrais impressas, encontrando o conteúdo do grupo 2. As demais informações do primeiro plano foram consideradas do grupo 1. Para encontrar as duas colunas, utilizou-se a técnica de perfil de projeção na imagem, calculando a quantidade de pixels pretos (que são as informações do primeiro plano) para cada linha e coluna, resultando em dois histogramas, horizontal e vertical:

1. Histograma de perfil de projeção horizontal - representou os pixels pretos encontrados na imagem nas linhas.
2. Histograma de perfil de projeção vertical - representou os pixels pretos encontrados na imagem nas colunas.

A Figura 6.50 ilustra em (a) a imagem da Bíblia segmentada pelo nível 1, limiarizada pelo algoritmo de Sauvola, (b) o resultado da compactação pela operação de abertura binária, (c) o histograma do perfil de projeção horizontal da imagem (b) e (d) o histograma do perfil de projeção vertical da imagem (b).

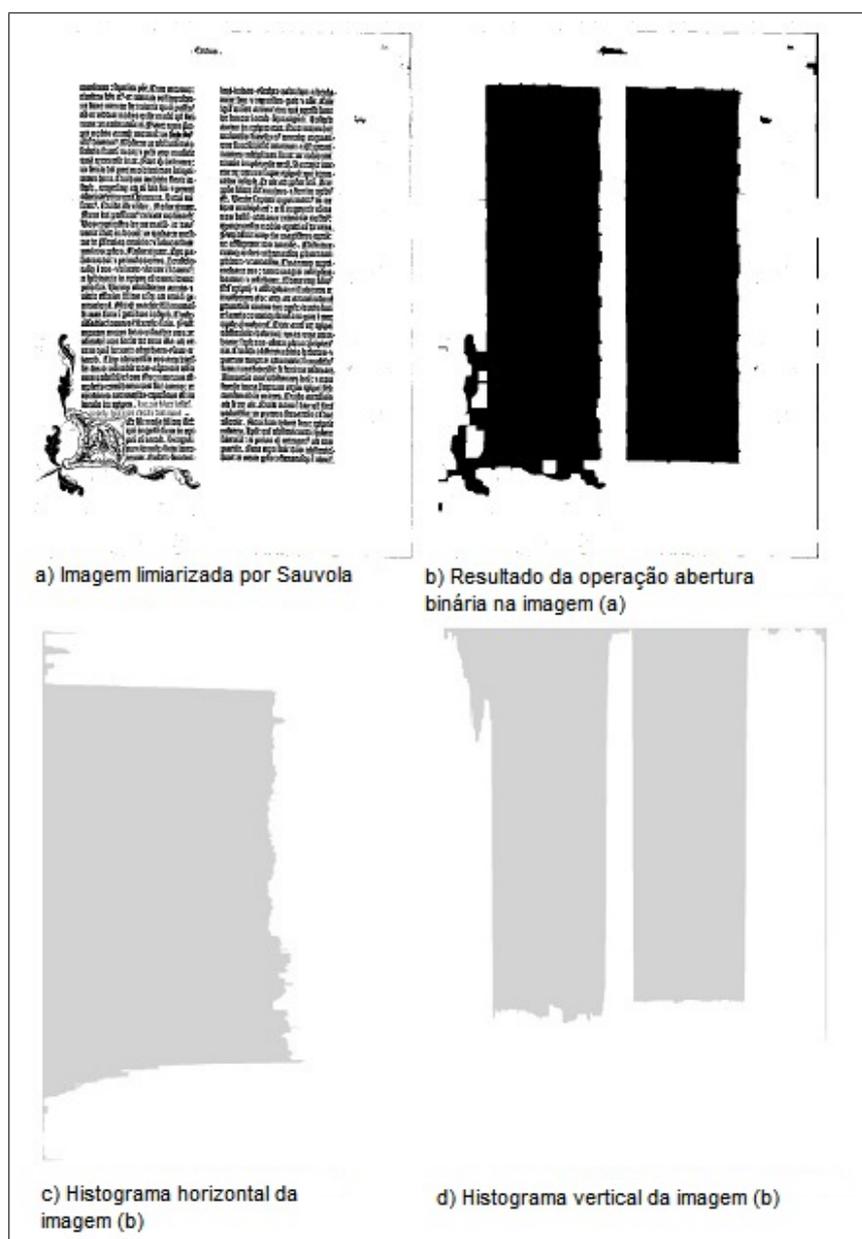


Figura 6.50: Ilustração do resultado da aplicação da técnica de perfil de projeção horizontal e vertical em uma imagem segmentada da Bíblia pelo nível 1.

Com os histogramas foi possível perceber que as linhas do texto preto impresso (duas colunas centrais) apresentam a maior quantidade de pixels pretos, tanto no histograma horizontal quanto no vertical.

### 6.4.1.1 Busca automática das regiões

Essa operação consistiu em segmentar do histograma horizontal as linhas com maior número de pixels encontrados e segmentar do histograma vertical as colunas com maior ocorrência de pixels pretos.

### 6.4.1.2 Limiar de corte para linhas e colunas

Para determinar o limiar, utilizou-se do cálculo automático da média de pixels pretos nas linhas, sendo encontrado o valor de corte para o histograma horizontal. Para o histograma vertical procedeu-se da mesma maneira, com o cálculo da média dos pixels pretos encontrado nas colunas.

A Figura 6.51 ilustra as regiões com maior ocorrência de pixels que foram segmentados de acordo com a média dos pixels para linhas e colunas, sendo a imagem (a) a segmentação baseada no histograma horizontal, a (b) a segmentação do histograma vertical e a (c) o resultado da intersecção da segmentação dos dois histogramas, encontrando-se as duas colunas centrais impressas, que representam as informações do grupo 2 (colunas centrais) do primeiro plano.

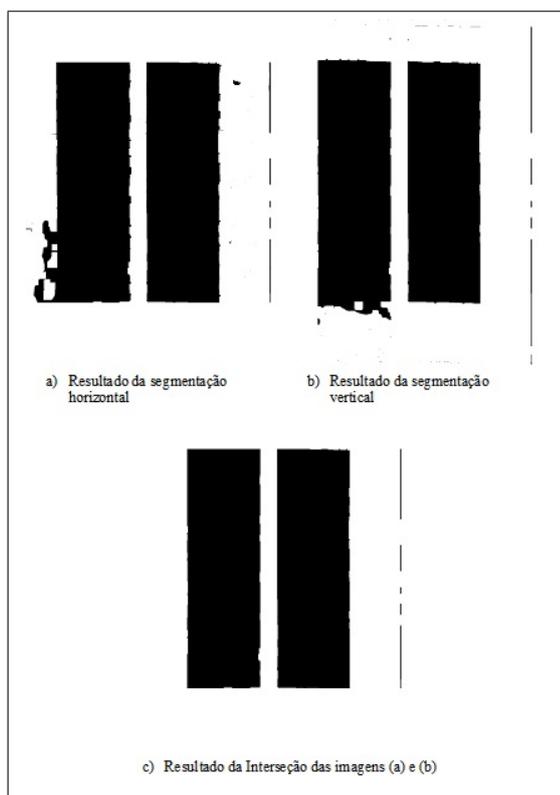


Figura 6.51: Ilustração do resultado da intersecção entre os histogramas horizontal e vertical.

Para encontrar o conteúdo do grupo 1, realizou-se uma operação de subtração entre a imagem limiarizada com o resultado da intersecção do perfil de projeção, que são as informações do grupo 2. A Figura 6.52 ilustra o resultado dessa subtração.

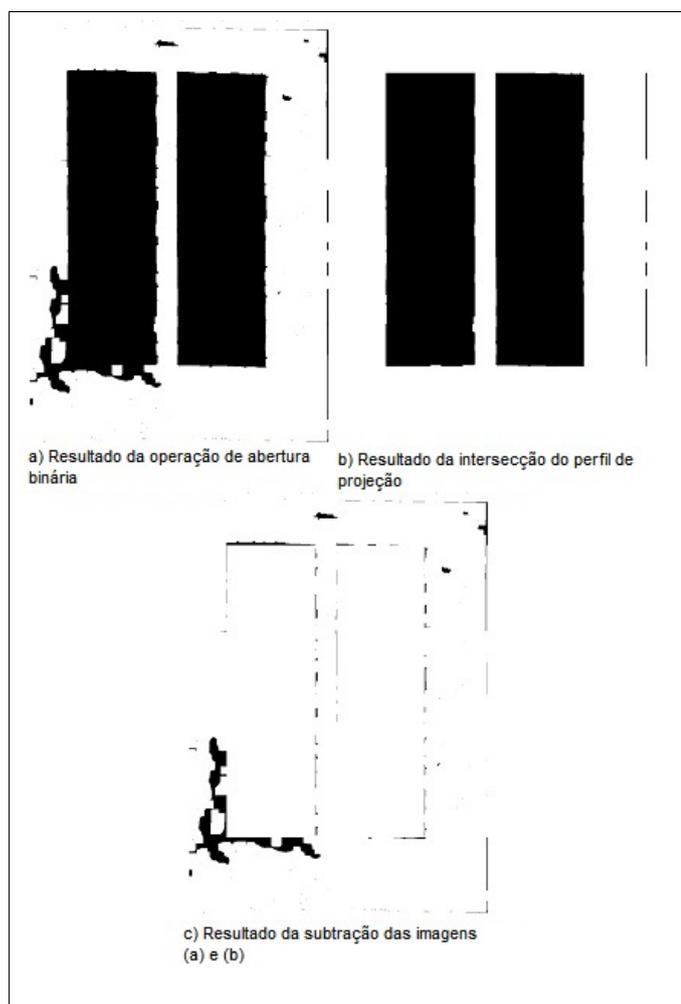


Figura 6.52: Ilustração do resultado da subtração da intersecção e da imagem limiarizada por Sauvola.

Os experimentos mostraram que era preciso realizar uma operação complementar para segmentar todas as letras das duas colunas centrais, uma vez que, pela operação de intersecção dos histogramas horizontal e vertical resultantes do método do perfil de projeção, alguns pixels de algumas linhas ou colunas não foram segmentados em virtude de serem poucos. Por exemplo, a imagem (c) da Figura 6.52 ilustra bem o problema que ocorreu na segmentação das duas colunas centrais, pois a letra "l" na primeira linha superior das duas colunas não foi identificada na intersecção dos histogramas, ficando de fora das duas colunas.

Uma estratégia para resolver este problema, foi realizar a operação morfológica binária dilatação, para aumentar a área das duas colunas segmentadas pelo processo

da técnica do perfil de projeção, a fim de garantir a segmentação de todos os pixels pertencentes às duas colunas centrais. A Figura 6.53 ilustra exemplo da operação de dilatação aplicada na imagem resultante da intersecção do perfil de projeção e a subtração com a imagem resultante do processo da abertura binária.

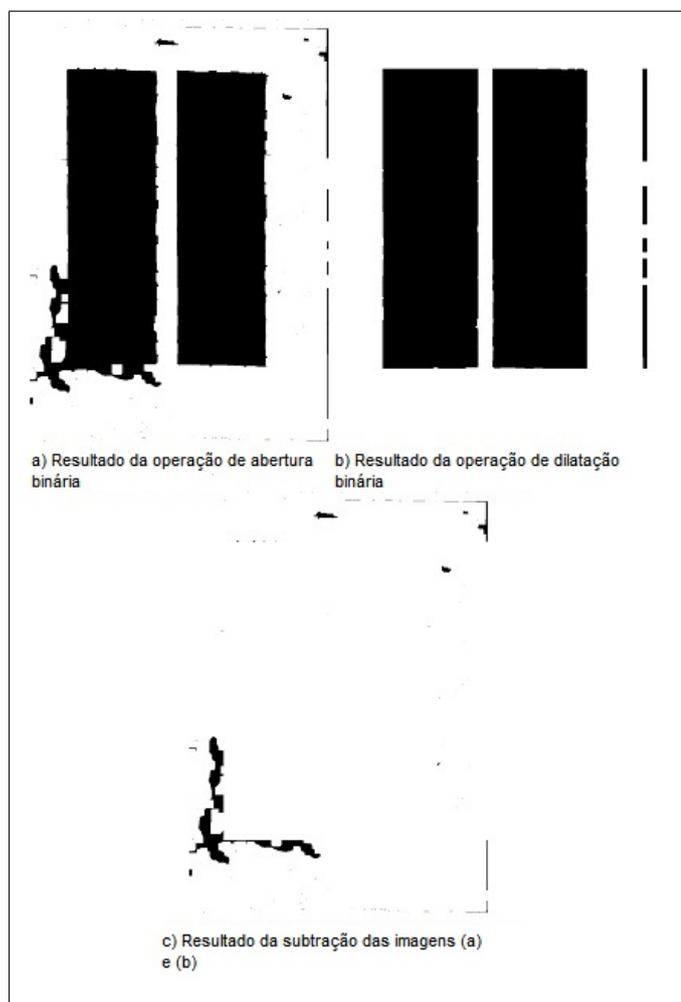


Figura 6.53: Ilustração do resultado da operação de dilatação da imagem resultante da intersecção dos histogramas.

A imagem resultante apresentou vários ruídos, foi feita a aplicação do filtro matemático binário soft duas vezes, considerando as oito direções dos pixels, eliminando os ruídos das imagens, restando os pixels de interesse que representaram os títulos, os manuscritos e as figuras. A Figura 6.54 ilustra o resultado do filtro matemático binário soft.

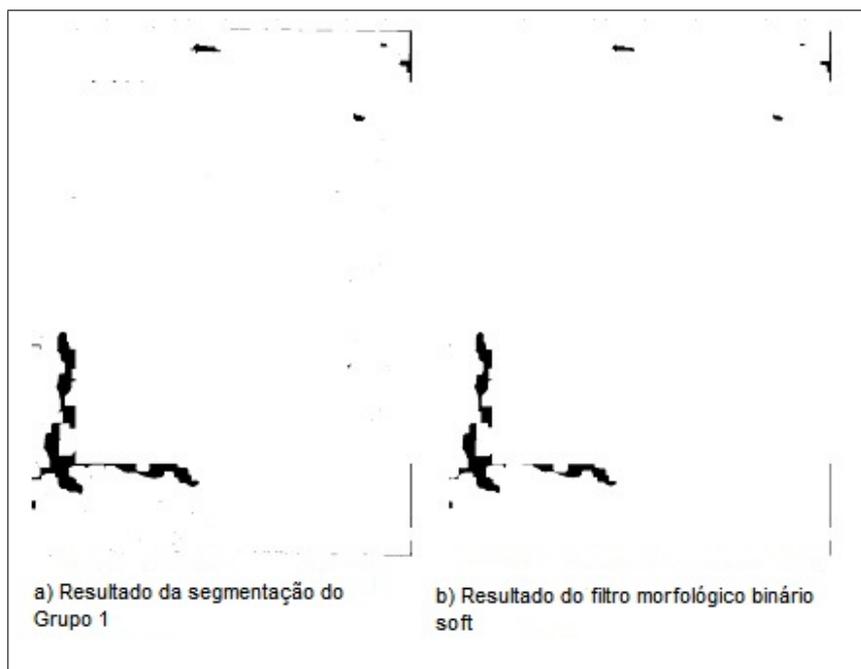


Figura 6.54: Ilustração do resultado da aplicação do filtro para eliminar os ruídos.

#### 6.4.1.3 Considerações do processo 4: Segmentação do grupo 1 e 2

- Os testes mostraram ótimos resultados com o desta estratégia para as 120 imagens dos experimentos, sendo testados vários valores de iterações para a operação de dilatação binária, chegando à conclusão de que com, no mínimo 6 iterações esta operação funciona bem.
- Se no momento da segmentação das duas colunas centrais, com o auxílio da operação de dilatação, alguns pixels de figuras forem segmentados pelas duas colunas, na operação de reconstrução das figuras, esses pixels serão recuperados, não prejudicando assim os resultados.
- As avaliações para este processo não foram necessárias, uma vez que essas imagens serviram de subsídio para o processo 5 do nível 3 do método proposto.

## 6.5 Experimentos do Nível 3

Etapa que consistiu na rotulação e classificação do conteúdo do grupo 1 (figuras, títulos, numeração de páginas e manuscritos) de acordo com o Processo 5 do método proposto.

### 6.5.1 Processo 5: Segmentação do Grupo 1

A classificação deste grupo pôde ser realizada a partir de operações morfológicas, de acordo com os seguintes passos:

- Aplicação na imagem da operação morfológica de reconstrução em cor para reconstruir os títulos e as figuras, caso necessário.
- Aplicação da operação abertura binária para compactar as regiões.
- Aplicação da técnica de rotulação dos pixels da imagem.
- Aplicação da classificação dos rótulos da imagem.

#### 6.5.1.1 Aplicação da operação morfológica reconstrução em cor

A operação morfológica reconstrução em cor se tornou-se necessária para a reconstrução dos títulos e figuras, pois no momento da aplicação da técnica de perfil de projeção parte das figuras (letras capitulares) foram perdidas. A Figura 6.55 ilustra o resultado da operação de reconstrução em cor, sendo a imagem (a) informações da página original, a (b) o conteúdo do grupo 1 (sementes para a reconstrução) e (c) o resultado da reconstrução.

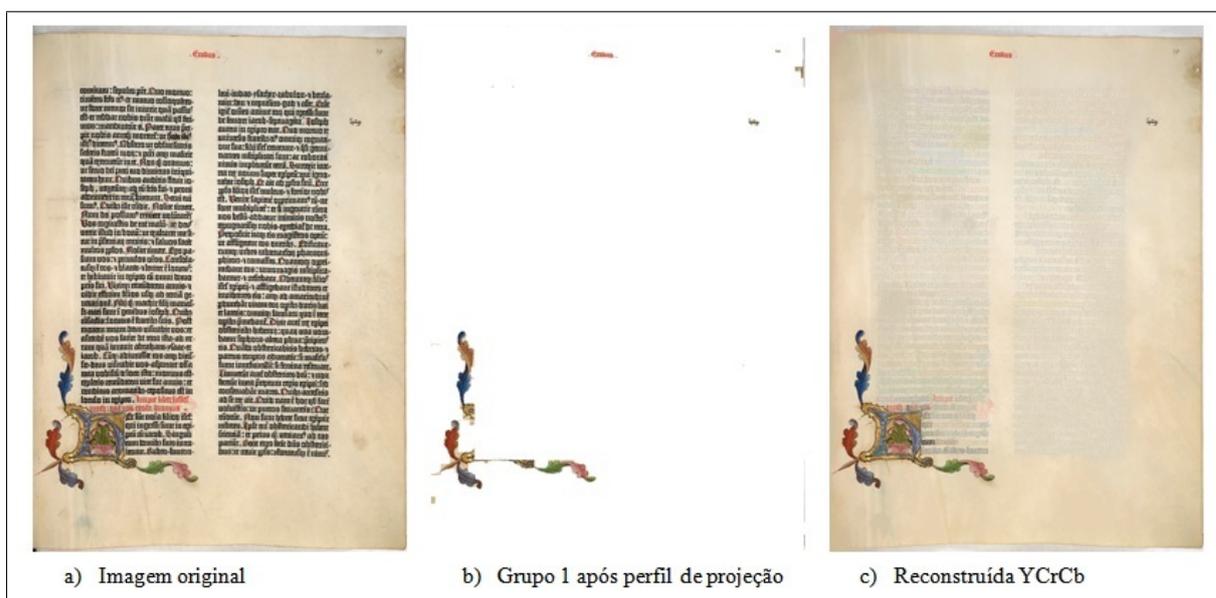


Figura 6.55: Ilustração da reconstrução em cor do Grupo 1.

### 6.5.1.2 Aplicação da operação abertura binária

Antes de realizar a rotulação foi preciso compactar as áreas para melhor rotulação, diminuindo a quantidade de rótulos e facilitando a classificação dos mesmos. A operação abertura binária possibilitou a compactação das áreas, sendo preciso, no mínimo, 6 iterações para uma boa compactação. A Figura 6.56 ilustra em (a) a limiarização da imagem reconstruída e em (b) o resultado da operação de abertura binária com 6 iterações, compactando as regiões de interesse.

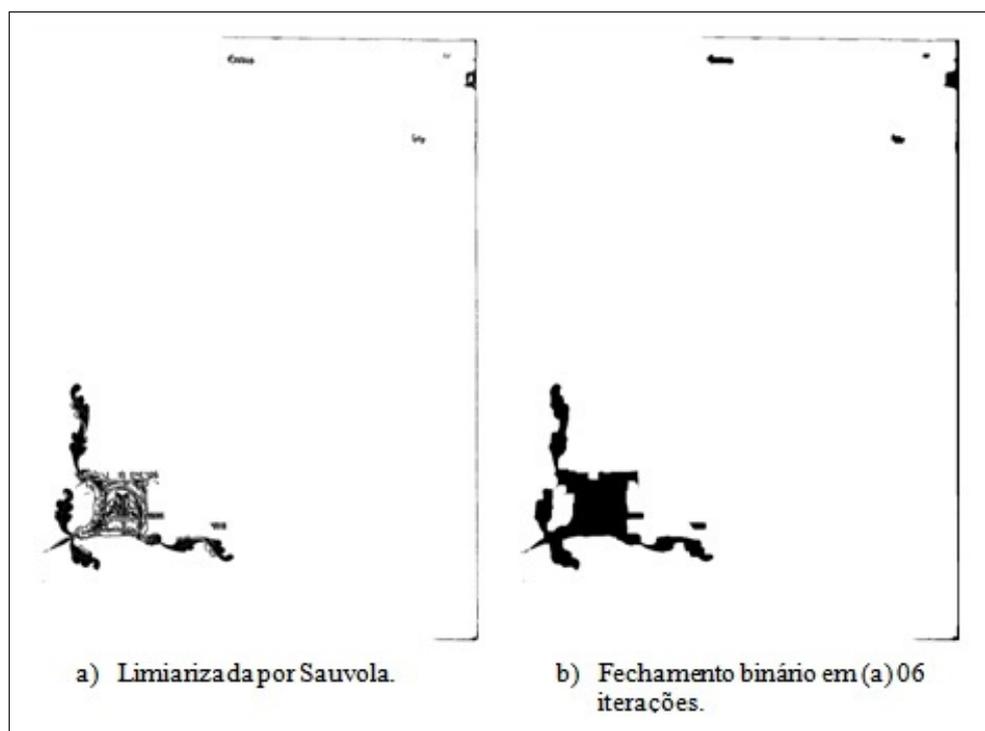


Figura 6.56: Ilustração da operação abertura binário.

### 6.5.1.3 Aplicação da técnica de rotulação dos pixels do Grupo 1

A técnica de rotulação consistiu em etiquetar as áreas conectadas com um valor para posterior classificação dessas áreas. Para este processo não foram rotulados os pixels que se encontravam nas bordas da imagem, sendo considerados como ruídos adquiridos no momento da aquisição da imagem pelo processo de digitalização. Isto visou à diminuição da quantidade de rótulos, possibilitando a classificação somente do conteúdo do primeiro plano que interessava figuras, títulos, números de páginas e manuscritos. A Figura 6.57 ilustra a rotulação realizada em 06 áreas da imagem, após aplicada a compactação, para posterior classificação.

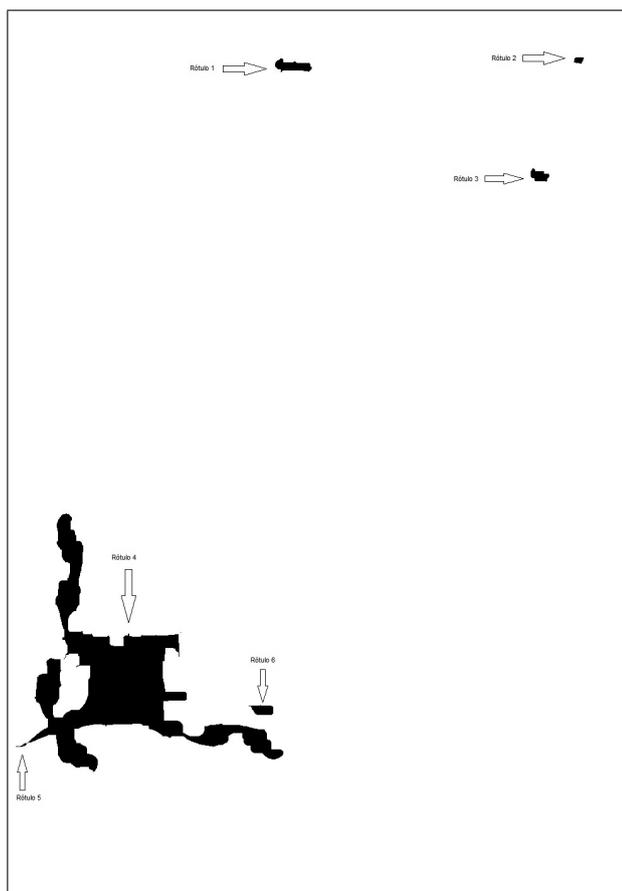


Figura 6.57: Resultado da Imagem Rotulada.

#### 6.5.1.4 Classificação dos pixels rotulados

A classificação consistiu em definir, entre os rótulos, quais eram títulos, figuras, numeração das páginas e manuscritos. Isto pôde ser realizado através das seguintes estratégias:

- **Títulos:** puderam ser localizados e diferenciados das demais informações da imagem pela sua posição, sempre na parte superior da página, acima da localização das duas colunas centrais, respeitando o lado esquerdo pelo primeiro pixel da coluna esquerda e do lado direito pelo último pixel da coluna da direita.
- **Figuras:** puderam ser identificadas pela conexão com as duas colunas centrais, ou seja, possuíam intersecção, informações tanto fora das duas colunas como dentro delas.
- **Manuscritos:** foram diferenciados dos ruídos pela sua localização e pela quantidade de pixels conectados e também por não possuírem intersecção com as duas colunas centrais nem com as bordas.

- **Numeração das páginas:** considerada manuscritos e está localizada no canto superior direito das páginas, acima das duas colunas centrais, com início da área após o último pixel da coluna da direita e último pixel da borda da imagem do lado direito.
- **Ruídos:** eram pequenas regiões, nas quais os rótulos que estiveram dentro das duas colunas centrais somente foram desconsiderados, ou seja, não foram classificados. Os ruídos das bordas não foram rotulados, sendo desconsiderados também. A figura 6.57 ilustra exemplo de um ruído etiquetado como Rótulo 6 , localizado dentro das duas colunas centrais, não sendo classificado.

Portanto, essas foram as estratégias adotadas para a classificação do grupo 1 do processo 5 do método proposto. A Figura 6.58 ilustra a classificação do conteúdo, sendo a imagem (a) classificada como título , a imagem (b) como numeração de páginas, (c) como manuscritos e (d) como figuras.

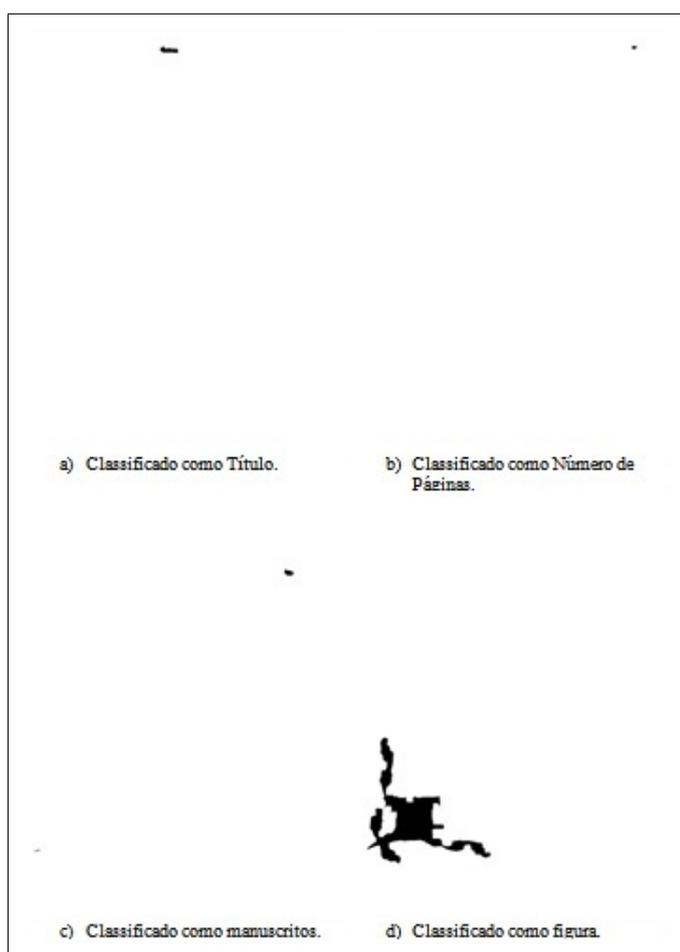


Figura 6.58: Classificação dos rótulos.

A Figura 6.59 ilustra os pixels originais da classificação dos títulos, numeração de páginas, manuscritos e figuras.

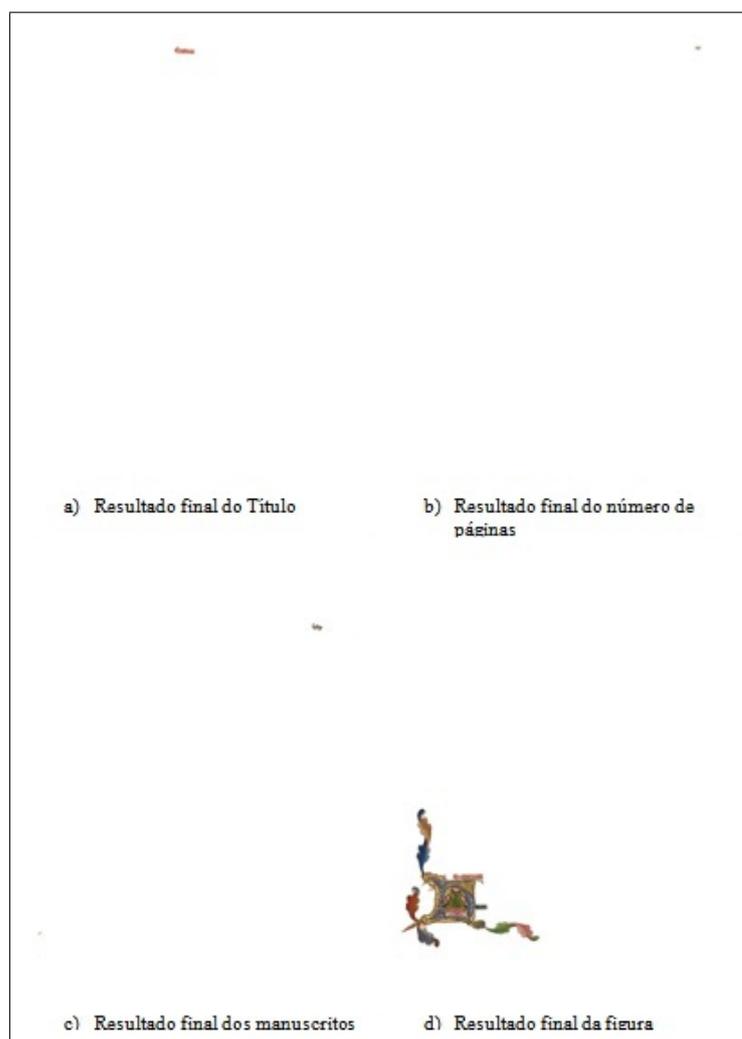


Figura 6.59: Ilustração dos pixels originais classificados.

Assim, foi apresentada uma maneira de segmentar e classificar o conteúdo do Grupo 1. Para esse processo foram testadas as 120 imagens originais da base de dados da Bíblia de Gutenberg para obtenção dos resultados.

#### 6.5.1.5 Avaliação do Nível 3 - Processo 5: Segmentação do Grupo 1

A avaliação deste nível consistiu na verificação das classificações realizadas no conjunto das 120 imagens originais da base de dados da Bíblia de Gutenberg. Nesta seção a avaliação verificou se todos os títulos, figuras, número de páginas e manuscritos foram rotulados e classificados corretamente. Foi uma classificação subjetiva, não levando em consideração a quantidade de pixels de cada conjunto.

- **Títulos:** das 120 imagens testadas 19 possuíam somente um título, 17 com 02 títulos, 02 com 03 títulos e 02 com 04 títulos, totalizando 145 títulos presentes nas imagens. Os resultados mostraram que:
  1. Dos títulos classificados, 02 conjuntos eram ruídos do fundo, em virtude deste estar deteriorado e ser escuro, sendo limiarizado como informações do primeiro plano.
  2. A imagem 070v2l possuía 02 títulos, mas no momento da limiarização e reconstrução os pixels do título do verso foram considerados, sendo classificados 03 títulos para esta imagem, ou seja, um falso positivo.
  3. Do total de títulos, 02 foram classificados como parte de figuras, porque os desenhos estavam acima das duas colunas centrais e no momento da compactação dos pixels, foram unidos aos pixels das figuras. Essas imagens foram: 001r1l e 005r1l.
  4. Do total de 145 títulos, 143 foram classificados corretamente, 01 falso título do verso e 02 falsos títulos do fundo, totalizando 146 títulos classificados. A classificação dos títulos corretos atingiu 98,62 %.

A Figura 6.60 ilustra exemplo de um título de uma imagem da Bíblia de Gutenberg classificado corretamente.

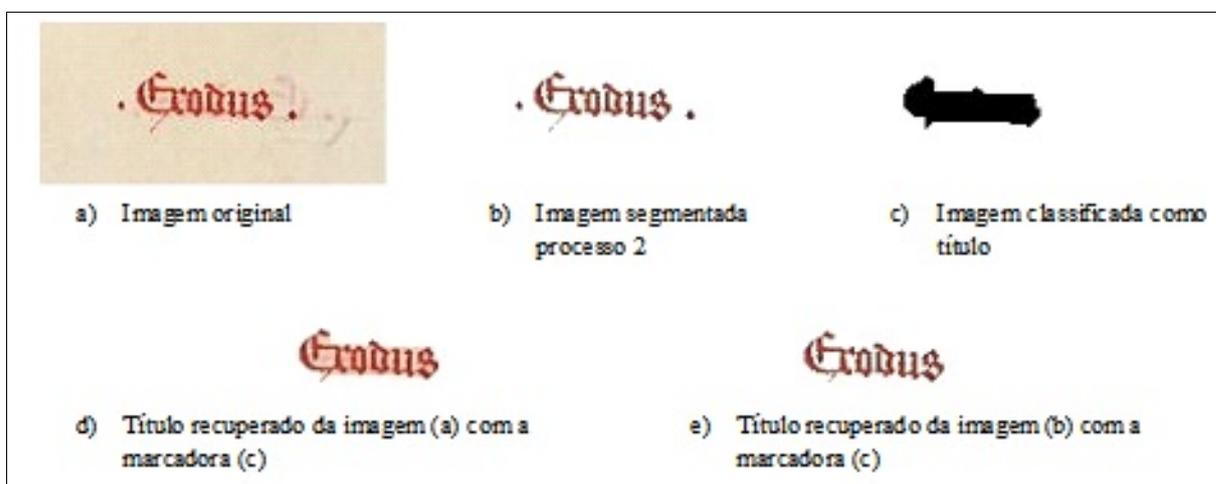


Figura 6.60: Ilustração da classificação do conjunto Título.

- **Manuscritos:** Das 120 imagens testadas, 40 possuíam um ou mais manuscritos, totalizando 70. Os resultados mostraram que todos eles foram classificados corretamente.

A Figura 6.61 ilustra exemplo de manuscrito classificado corretamente.

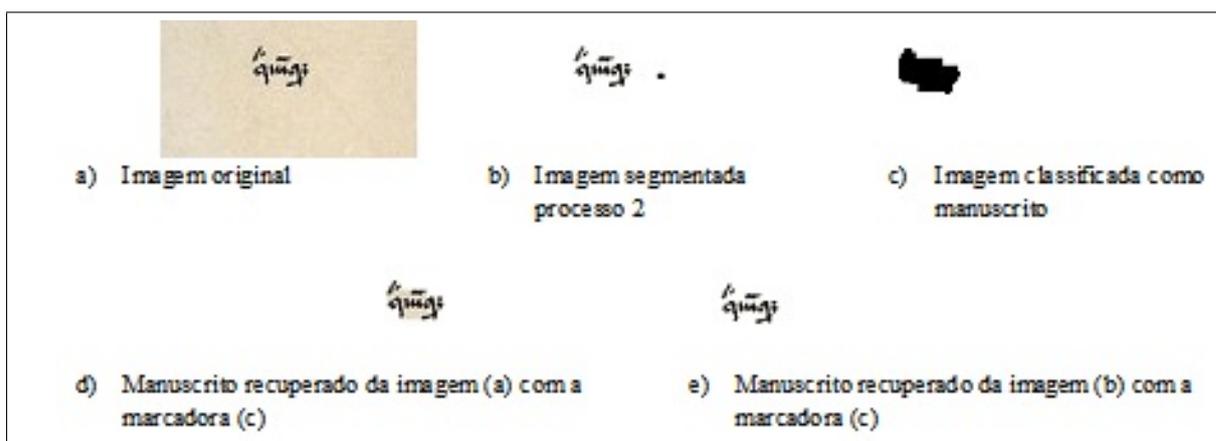


Figura 6.61: Ilustração da classificação do conjunto Manuscrito.

Problemas encontrados: Muitos ruídos foram classificados como manuscritos, devido aos seguintes fatores: I) No momento da limiarização, as bordas das imagens e o fundo do *scanner* influenciaram no processo, sendo classificados como informações do primeiro plano. II) O resultado da intersecção do perfil de projeção, mesmo com a operação de dilatação binária, deixou pequenos ruídos, que não foram eliminados pela filtragem morfológica binária soft, ocasionando o problema de classificação como falsos manuscritos. Do total de 120 imagens analisadas, havia 122 ruídos distribuídos da seguinte maneira: 95 de bordas, 12 de fundo e 15 das duas colunas centrais. Os ruídos foram classificados como manuscritos, pois estavam perdidos na imagem e eram pequenos pedaços.

- **Numeração de Páginas:** Das 120 imagens testadas, 43 possuíam numeração de páginas, destas, 41 foram classificadas corretamente, ou seja, atingindo 95,35 % de acerto. As 02 numerações que não foram classificadas ocorreram em virtude do processo de limiarização e filtragem dos pixels, eliminando as sementes que representavam os números. Como os números das páginas eram pequenos, o processo de identificação e classificação foi dificultado.
- **Figuras (letra capitular grande e pequena):** Das 120 imagens testadas, 40 possuíam um ou mais desenhos que representaram: letras capitulares grandes, pequenas e desenhos manuscritos. Algumas imagens apresentaram mais de um desenho, no total das 40 imagens analisadas, havia 74 desenhos grandes (letras capitulares grandes) e 96 letras capitulares pequenas.

- **Letras Capitulares Grandes:** Do conjunto de 74 imagens com letras capitulares grandes, 52 foram classificadas em quase sua totalidade, 18 parcialmente e 4 foram classificadas somente com os pixels pertencentes fora das duas colunas, em virtude de não terem conexão com as informações das duas colunas centrais. Com esta classificação qualitativa, as taxas de acertos para a quase na totalidade atingiram 70,27% e para parcialmente, 24,32%.

A Figura 6.62 ilustra o resultado na imagem 015v2l, que foi classificada pela metade por não haver conexão dos desenhos entre o conteúdo presente nas duas colunas e o conteúdo fora das duas colunas.

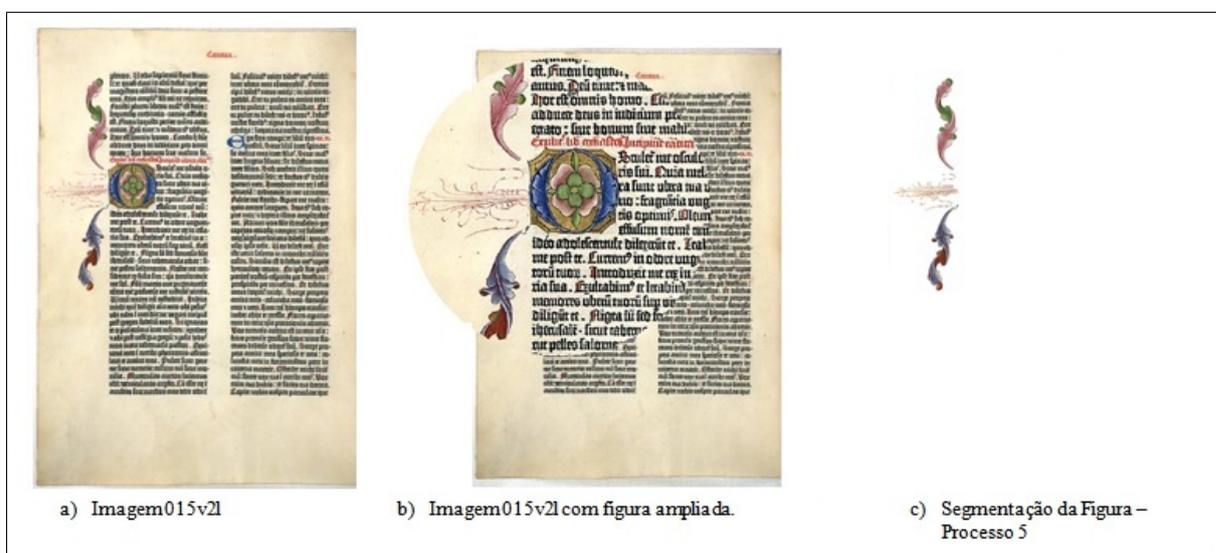


Figura 6.62: Ilustração da imagem sem conexão.

Na Figura 6.63, a imagem 029r2l ilustra exemplo em que a letra capitular possui conexão com a parte que está dentro das duas colunas centrais, ocorrendo, então, a reconstrução. As figuras classificadas não atingiram 100 % da totalidade dos pixels devido ao processo de limiarização, em que alguns deles foram perdidos.



Figura 6.63: Ilustração da imagem com conexão.

A Figura 6.64 ilustra três imagens classificadas, sendo (b) uma segmentação quase na totalidade, (d) uma segmentação parcial e (f) uma segmentação pela metade. Neste processo ocorreram alguns problemas que merecem ser mencionados.

- \* A imagem (b) apresentou algumas letras, pretas ou coloridas, que não pertenciam à figura, segmentadas juntas. Tal fato ocorreu porque essas letras estavam conectadas aos pixels da figura e, no momento da operação de reconstrução em cor, esses pixels também foram recuperados.
- \* A imagem (d) apresentou uma segmentação parcial da letra capitular grande, ocorrida pelo algoritmo de limiarização, em que as partes claras da imagem foram classificadas como sendo do fundo, não sendo possível classificar o interior da letra.
- \* A imagem (f) apresentou a segmentação pela metade, isto ocorreu, como descrito anteriormente, por não haver conexão com os pixels da figura, que estavam dentro das duas colunas centrais.

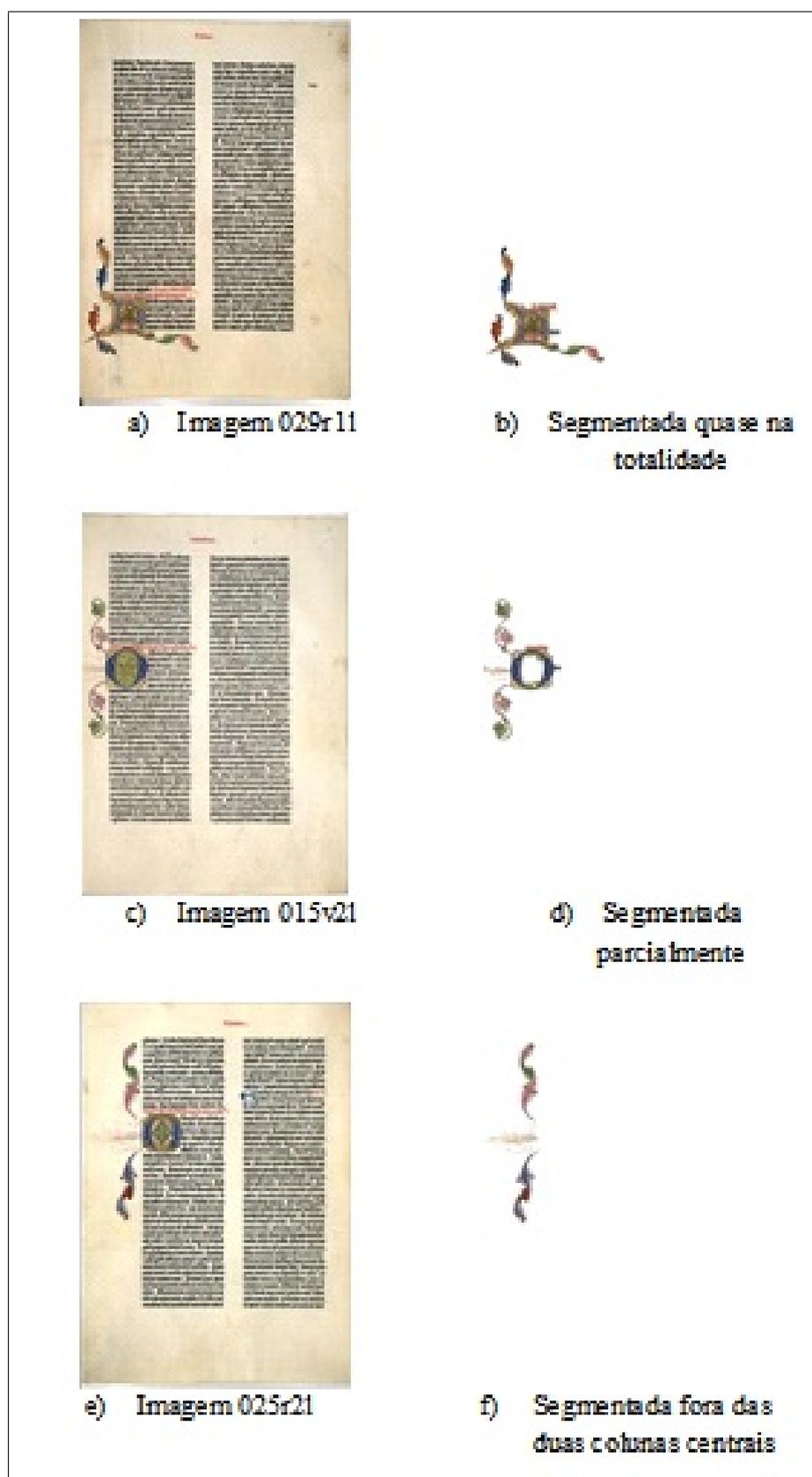


Figura 6.64: Exemplos de 03 figuras segmentadas pelo processo 5.

- **Letras capitulares pequenas:** Do conjunto de 96 letras capitulares pequenas, 75 foram classificadas quase na sua totalidade e 21 não foram localizadas. Isto ocorreu porque elas estavam, em sua totalidade, dentro das duas colunas centrais. Assim, no momento da subtração dessas colunas, com as informações do primeiro plano, os pixels dessas letras também foram subtraídos, não ocorrendo o processo de reconstrução em cor. As letras capitulares pequenas, que foram localizadas e classificadas, possuíam pixels que foram localizados tanto nas duas colunas centrais quanto fora delas.

A Figura 6.65 ilustra na imagem (a) uma letra capitular pequena que não foi segmentada, e em (b) uma que foi.



Figura 6.65: Exemplos de letras capitulares pequenas.

### 6.5.1.6 Considerações dos testes do Nível 3 - Processo 5

- A segmentação dos títulos, figuras, número de páginas e manuscritos conseguiu bons resultados, de forma subjetiva, porém não levou em consideração a quantidade dos pixels de cada região.
- Ocorreu um problema quando existiam figuras na parte superior das duas colunas centrais, não sendo possível classificar os possíveis títulos.
- Alguns ruídos foram classificados como manuscritos devido ao processo de limiarização, pois no momento da compactação das áreas, alguns pixels não foram compactados e eram pertencentes à figura. Ainda, quando havia manchas escuras no fundo, estas foram consideradas pertencentes ao primeiro plano no processo de limiarização.
- O problema de muitos rótulos aconteceu em virtude do processo de digitalização das páginas, sendo as bordas segmentadas como pixels do primeiro plano. Estes ruídos puderam ser eliminados se não foram considerados no momento da rotulação.

## 6.6 Considerações Finais dos Experimentos

Foram apresentados os resultados do método proposto referentes à segmentação do conteúdo e estimativa do fundo de documentos antigos, com auxílio da morfologia matemática binária, cinza e em cor e de algoritmos de limiarização.

Os resultados mostraram-se promissores, em face da complexidade de se trabalhar com imagens coloridas, com documentos antigos possuindo fundos deteriorados com o tempo, com figuras desenhadas, com letras coloridas pequenas, com numerações de páginas pequenas e com manuscritos.

A recuperação do fundo com base na morfologia matemática em cor, através da operação de reconstrução, apresentou bons resultados com o uso dos espaços de cor YCrCb e YIQ, além de se usar a intensidade como primeiro parâmetro para a ordem dos canais na operação de reconstrução. Esta abordagem auxiliou a reconstrução do fundo da imagem, das páginas da Bíblia e também a segmentação do conteúdo colorido do primeiro plano, mostrando esta ótimos resultados, mas esbarrando no algoritmo de limiarização, que acarretou perda de informações nos desenhos.

A segmentação das pequenas letras vermelhas coloridas, presentes nas duas colunas centrais, foi possível através da morfologia matemática em cor com o uso do espaço de cor HSI/HSL e com as ordenações lexicográficas SHI e SHL para encontrar as sementes para a localização dessas letras. A segunda reconstrução para este processo, utilizou-se de qualquer um dos 4 espaços de cor, com o canal intensidade na primeira posição de processamento para a ordenação lexicográfica.

A segmentação do conteúdo do primeiro plano foi realizada através utilizou da técnica de análise do perfil de projeção vertical e horizontal, para separar informações do primeiro plano em dois grupos, mas para as imagens em que alguns pixels não foram segmentados como sendo das duas colunas centrais, houve necessidade de um ajuste. Esse ajuste consistiu em dilatar o conteúdo do primeiro plano para aumentar as áreas das duas colunas centrais, para posterior segmentação de todos os pixels das duas colunas centrais.

A classificação do grupo 1, processo 5, apresentou ótimos resultados, embora baseado em conhecimento *a priori* sobre a localização e estrutura das páginas da Bíblia.

Os testes para os níveis 1 e 5, com as 120 imagens genuínas da Bíblia de Gutenberg, mostraram uma grande variabilidade entre as imagens, havendo fundos, figuras e textos diferentes.

Os testes para os níveis 2 e 3, com as 160 imagens sintéticas geradas, também mostraram uma grande diversidade entre as imagens.

Portanto, com esses testes foi possível mostrar o funcionamento do método pro-

posto, sendo comprovada a viabilidade de se trabalhar com a morfologia matemática em cor para a restauração de documentos antigos, como as páginas da Bíblia de Gutenberg.

## Capítulo 7

### Conclusão

A abordagem proposta objetivou estimar o fundo e segmentar o conteúdo das páginas da Bíblia impressa por Gutenberg a partir de ferramentas baseadas na morfologia matemática em cor, em cinza e binária, e limiarização.

Os conceitos da morfologia matemática em cor foram amplamente estudados e implementados, empregando a ordenação lexicográfica com os espaços de cor HSI, HSL, YCrCb e YIQ.

Os experimentos foram realizados e validados com 120 imagens genuínas da base de imagens da Bíblia de Gutenberg e 160 imagens sintéticas, que foram geradas com textos genuínos da Bíblia, traduzindo em latim e com fonte medieval variando os fundos e os desenhos do primeiro plano, reproduzindo, assim a grande variabilidade de imagens que existentes na base de imagens.

As imagens mais complexas foram utilizadas para reproduzirem o fundo, apresentando desenhos do fundo que influenciaram nas informações do primeiro plano. Essas imagens possuíam fundo muito deteriorado, desgastado com o tempo, e havia desenhos complexos no primeiro plano. Os desenhos utilizados nas imagens sintéticas foram recortados de imagens genuínas da Bíblia, dos quais foram escolhidos vários para formar uma base ampla, com grande variabilidade, tanto nas informações do primeiro plano quanto nos fundos escolhidos.

Os experimentos mostraram que a ordem dos canais, para qualquer um dos 4 espaços usados neste estudo, resultou em estimativas diferentes do fundo e das segmentações do conteúdo.

Os experimentos envolvendo a estimativa do fundo mostraram que a operação de reconstrução em cor foi determinante para gerar um fundo homogêneo, e também que a escolha da intensidade, como primeiro canal a ser processado na ordenação lexicográfica, foi a estratégia ideal para a estimativa eficiente do fundo, qualquer que fosse o espaço de

cor escolhido (HSI, HSL, YCrCb ou ainda YIQ) e as imagens para a estimativa do fundo mostraram-se iguais nos quatro espaços. Assim, ficou evidente que o canal intensidade dominou sobre os demais para a estimativa do fundo de imagens antigas.

A segmentação das letras pequenas coloridas presentes nas páginas da Bíblia, que são dificilmente diferenciáveis a olho nu, foi uma tarefa árdua e necessitou de duas etapas de reconstrução em cor envolvendo a ordenação lexicográfica com diferentes espaços de cor e ordem dos canais. Os experimentos mostraram que, na primeira etapa, os melhores espaços foram HSI e HSL, com as ordens SHI e SHL, e, na segunda, foram HSI, HSL, YCrCb e YIQ com as ordens ISH, LSH, YCrCb e YIQ, respectivamente.

Os resultados dos experimentos foram avaliados por métricas de avaliação de segmentação, utilizando imagens *ground truth* e sintéticas, mostrando a excelência na qualidade dos resultados obtidos. Para o processo 2, os resultados médios obtidos para as 160 imagens sintéticas testadas com opacidades 70%, 80%, 90% e 100%, foram de 99,29% para a taxa da Medida  $F$  e a taxa de erro médio foi de 00,71%.

Para o processo 3, os resultados médios da limiarização direta atingiram 92,06% para a medida  $F$  e com erro de 07,31%. Com a limiarização após a reconstrução, foram atingidos 97,76% para Medida  $F$  e erro de 02,12%. Para obtenção dos resultados do processo 5 foram utilizadas as imagens genuínas. As taxas de acertos para a classificação dos títulos atingiram 98,62%, para as numerações de páginas, a taxa de acerto foi de 95,35% e para as figuras, a taxa foi de 70,27% na classificação das figuras quase na sua totalidade e 24,32% na classificação parcial.

Os resultados dos experimentos permitiram concluir que a metodologia proposta, baseada na ordenação lexicográfica com espaços de cor diferentes mostrou-se eficiente na estimativa do fundo e na segmentação do conteúdo complexo das imagens da Bíblia. Os experimentos mostraram ainda, que definir a ordem adequada dos canais foi de fundamental importância.

Uma conclusão importante que ressaltou deste estudo foi que o uso dos espaços de cor YCrCb e YIQ tiveram uma enorme vantagem sobre o espaço HSI/HSL: não havendo necessidade de escolha do matiz de referência, o que representa um obstáculo sempre presente no uso do espaço HSI/HSL.

Por fim, a metodologia proposta apresentou um novo caminho na estimativa de fundo e na segmentação de documentos antigos, podendo ser aplicada em vários tipos de documentos antigos.

## 7.1 Trabalhos Futuros

Como trabalhos futuros, pretende-se:

- Gerar mais imagens sintéticas de diversas páginas da Bíblia para validação do método proposto e disponibilização da base.
- Efetuar a impressão e digitalização das imagens sintéticas geradas, com o intuito de reproduzir o processo de aquisição de imagens de documentos antigos.
- Elaborar outra estratégia para diferenciação, nos casos de confusão entre áreas com títulos e figuras.
- Testar outras bases de dados de documentos antigos existentes.
- Comparar os resultados com outras técnicas existentes como, por exemplo, a segmentação do conteúdo do primeiro plano com uso de redes neurais e treinamentos.
- Realizar o reconhecimento de caracteres em latim gótico presentes nas páginas da Bíblia.

## Referências Bibliográficas

- ABUTALEB, A. Automatic thresholding of gray-level pictures using two-dimension entropy. *Computer Vision Graph, Image Processing*, v. 1, p. 22–32, 1989.
- ANGULO, J. Unified morphological color processing framework in a lum/sat/hue representation. *Proceedings of the 7th ISMM - Computational Imaging and Vision, Springer-Verlag, Dordrecht, Germany*, v. 30, 2005.
- APTOULA, E.; LEFÉVRE, S. A comparative study on multivariate mathematical morphology. *Preprint submitted to Elsevier Science*, v. 1, n. 1, p. 1–37, 2007.
- ARCIDIACONO, G.; PORTUESE, F.; SCIFO, S. Digital restoration of antique books. *IRT Engineering - University Catania - Italy*, 2005.
- BAIRD, H. S.; MOLL, M. A. Segmentation-based retrieval of document images from diverse collections. *Proc. IST/SPIE Document Recognition and Retrieval XV Conf. (DRR XV), San Jose, CA, January 28 - February 1, 2008*.
- BAIRD, H. S. et al. Document image content inventories. *Proc. IST/SPIE Document Recognition and Retrieval XIV Conf., San Jose, CA, 28 January - 1 February, 2007*.
- BARNET, V. The ordering of multivariate data. *Journal of the Statistical Society*, v. 139, n. 1, p. 318–355, 1976.
- BERTALMIO, M. Image inpainting. *ACM Computer Graphics Proceedings (Siggraph)*, v. 1, p. 417–424, 2000.
- BIMBO, A. D. Visual information retrieval. *Morgan Kaufmann*, 1999.
- BOTTOU, L.; HAFFNER, P.; HOWARD, P. High quality document image compression with djvu. *Journal of Electronic Imaging, SPIE*, v. 7, p. 410–425, 1998.

- BOUSSELLAA, W.; ZAHOUR, A.; ALIMI, A. A methodology for the separation of foreground/background in arabic historical manuscripts using hybrid methods. *Journal of Universal Computer Science*, v. 14, n. 2, p. 284–298, 2008.
- CALDEIRA, C. Do papiro ao papel manufaturado. *Espaço Aberto - USP - SP*, v. 24, 2002.
- CALIXTO, E. Granulometria morfológica em espaços de cores: estudo da ordenação espacial. *Depto de Pós-Graduação, Universidade Federal Fluminense.*, Niterói, RJ, 2005.
- COUSTATY, M. et al. Segmenting and indexing old documents using a letter extraction. *Graphics Recognition, Achievements, Challenges, and Evolution, Jean-Marc Ogier, Wenyin Liu, Josep Lladós (Ed.) (2010)*, DOI : 10.1007/978-3-642-13728-013, p. 142–149, 2010.
- DEBORA, P. Debora. *European Project nº. LB 5608A*, v. 1, p. 1–198, 2000.
- FACON, J. Morfologia matemática: Teoria e exemplos. *CDD 006.42 CDU 681.3 Curitiba, Brasil*, v. 1, 1996.
- FACON, J. *Processamento e Análise de Imagens*. Curitiba-Paraná: [s.n.], 2005. 128 p.
- GARAIN, U.; PAQUET, T.; HEUTTE, L. On foreground-background separation in low quality color document images. *IEEE/Computer Society/ICDAR - Eight International on Document Analysis and Recognition*, 2005.
- GATOS, B.; NITIROGIANNIS, K.; PRATIKAKIS, I. Document image binarization contest. *ICDAR - International Conference on Document Analysis and Recognition*, p. 1375 – 1382, 2009.
- GONZAGA, A. Apostila de processamento de imagens coloridas. p. Acessado em 2008, 2008.
- GONZALEZ, R. *Processamento de Imagens Digitais*. Editora Edgard Blucher Ltda.: [s.n.], 2000.
- GOSE, E.; JOHNSONBAUGH, R.; JOST, S. Pattern recognition and image analysis. *1a Edição. New Jersey Prentice Hall PTR*, v. 1, n. 1, p. 329–335, 1996.
- GRANADO, I.; PINA, P.; MUGE, F. Automatic feature extraction on pages of antique books through a mathematical morphology based methodology. *10 Encontro Portugues Compuatação Gráfica 2001*, 2001.

- HANBURY, A. Lexicographical order in the hls colour space. *Technical Report N-04/01/MM Centre de Morphologie Mathématique. École des Mines de Paris*, 2001.
- HANBURY, A.; SERRA, J. Mathematical morphology in the hls colour space. *BMVC 2001, in Manchester, UK*, p. 451–460, 2001.
- HARDIE, R.; ARCE, G. Ranking in  $r^p$  and its use in multivariate image estimation. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 1, n. 2, p. 197–209, 1991.
- HARRY, R. The university of texas at austin. 2001.
- HASE, H. et al. Color segmentation for text extraction. *IJDAR, International Journal Document Analysis and Recognition*, p. 271–284, 2004.
- HECKBERT, P. Color image quantization for frame. *Comput Graph*, 1982.
- HEITLINGER, P. Tipografia: origens, formas e uso das letras. *ISBN 10 972-576-396-3*, v. 1º Edição, 2006.
- JOHANNSEN, G.; BILLE, J. A threshold selection method using information measures. *Proceedings, 6th Int. Conf. Pattern Recognition, Munich, Germany*, v. 1, n. 1, p. 140–143, 1982.
- JOURNET, N. et al. Dedicated texture based tools for characterisation of old books. *In DIAL '06: Proceedings of the Second International Conference on Document Image Analysis for Libraries*, p. 60–69, 2006.
- JOURNET, N.; RAMEL, J. Y.; MULLOT, R. Document image characterization using a multiresolution analysis of the texture: application to old documents. *Spring-Verlag - IJDAR 2008 International Journal on Analysis and Recognition*, v. 1, p. 9–18, 2008.
- KOSKINEN, L.; ASTOLA, J.; NEUVO, Y. Soft morphological filters. *SPIE Image Algebra and Morphological Image Processing II*, v. 1, n. 1, p. 262–270, 1991.
- KUMAR, S. Lecture notes on rgb to hsi color conversion. 2006.
- LAMBERT, P.; CHANUSSOT, J. Extending mathematical morphology to color image processing. *CNRS - G720 - ISIS - Information Signal Images viSion*, v. 1, p. 272–279, 2000.

- LEDDA, A.; PHILIPS, W. Quantitative image analysis with mathematical morphology. *Department of Telecommunications and Information Processing (TELIN)*, v. 1, n. 1, p. 399–406, 2002.
- LI, C.; LEE, C. Minimum cross entropy thresholding. *Pattern Recognition*, v.26, n. 4, p. 616–626, 1993.
- LINS, R. D.; SILVA, G. F. O. Assessing strategies to remove back-to-front interference in color documents. v. 1, n. 1, p. 5, 2007.
- MAINZ, G. M. Gutenber museum mainz. <http://www.gutenberg-museum.de>, 1900.
- MAINZ, G. M. Gutenberg - museum mainz. 2008.
- MASCARO, A. A.; CAVALCANTI, G. D. C.; MELLO, C. A. B. Detecção de inclinação em imagens de documentos. *SIBGRAPI 2010 - Workshop of Theses and Dissertations (WTD)*, Gramado., 2010.
- MELLO, C.; LINS, R. Generation of images of historical documents by composition. *Proceedings of the 2002 ACM symposium on Document engineering, New York, NY, USA 2002*, v. 1, 2002.
- MILLER, R. Beyond anova, basics of applied statistics. *John Wiley and Sons: New York*, 1986.
- OROZ, M. A. L. et al. Mathematical morphology in polar-logarithmic coordinates application to erythrocyte shape analysis. *IbPRIA 2005, LNCS 3523 Springer-Verlag Berlin Heidelberg, LNCS 3523*, p. 199–206, 2005.
- ORTIZ, F.; TORRES, F.; GIL, P. A comparative study of highlights detection and elimination by color morphology and polar color models. *Pattern Recognition and Image Analysis - LNCS Lecture Notes in Computer Science*, v. 3523/2005, n. DOI:10.1007/1149254237, p. 147–171, 2005.
- ORTIZ, F. et al. Colour mathematical morphology for neural image analysis. *Real-Time Imaging 8 Elsevier Science Ltd*, v. 8, p. 455–465, 2002.
- ORTIZ, F. et al. Morphological processing of colour images using hue/saturation/intensity color spaces. *Proc. of first International Conference on Color in Graphics and Image Processing (CGIP 2000)*, Saint-Etienne, France, Oct, 2000.

- OZAWA, S.; MASSAKI, K. Application of image processing technology to digital archiving and study of rare books. *Project, Keio University. IEEE 2001*, v. 1, p. 272–279, 2001.
- PARETI, R.; VINCENT, N. Ancient initial letters indexing. *In ICPR 06: Proceedings of the 18th Conference on Pattern Recognition Washington DC USA 2006 IEEE Computer Society*, p. 756–759, 2006.
- PETERES, A. Mathematical morphology for angle-valued image. *In: Proceedings of SPIE, Non-Linear Image Processing, VIII*, p. 84 – 94, 1997.
- PING, Z.; LIHUI, C.; ALEX, K. Text document filters using morphological and geometrical features of characters. *In: Proceedings of ICSP - Int. Conf on Signal Processing*. Baltimore, USA: [s.n.], 2000. p. 472–475.
- PITAS, I.; VENETSANOPOULOS, A. N. Order statistics in digital image processing. *Proceedings of the IEEE*, v. 80, n. 12, p. 1893–1923, 1992.
- POPOV, A. T. Fuzzy mathematical morphology and its applications to colour image processing. *15th Internacional Conference on Computer Graphics, Visualization e Computer Vision 2007*, 2007.
- PRESTON, K. Xi filters. *IEEE Transaction Acoustics*, v. 31, n. 4, p. 861–876, 1983.
- PUENTE, M. T. et al. Estudio comparativo de Órdenes en los espacios de color para su aplicación en morfología matemática. *XXVVII Jornadas de Automatica - Almeria*, 2006.
- RAMPONI, G. et al. Digital automated restoration of manuscripts and antique printed books. *EVA 2005 , Electronic Imaging and the Visual Arts.*, 2005.
- SANGWINE, S.; HORNE, R. The colour image processing handbook. *London Chapman Hall*, 1998., 1998.
- SERRA, J. Image analysis and mathematical morphology. *Academic Press, London*, 1982.
- SERRA, J. Espaces couleur et traitement d images. *Tech Report N-34/02/MM*, 2002.
- SEZGIN, M.; SANKUR, B. Selection of thresholding methods for non destructive testing application. *ICIP2001*, v. 3, p. 764–767, 2001.

- SHAFAIT, F.; KEYSERS, D.; BREUEL, T. M. Performance evaluation and benchmarking of six page segmentation algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 30, n. 6, p. 941–954, 2007.
- SHAFAIT, F.; KEYSERS, D.; BREUEL, T. M. Efficient implementation of local adaptive thresholding techniques *Proceedings of the 15th Document Recognition and Retrieval Conference (DRR-2008)*, v. 6815, 2008.
- SHAHBAHRAMI, A.; JUURLINK, B.; VASSILIADI, S. Accelerating color space conversion using extended subwords and the matrix register file. *Eighth IEEE International Symposium on Multimedia (ISM 2006) San Diego, CA, USA*, p. 37–46, 2006.
- SILVA, F.; COSTA, R. Um sistema computacional para extração de características faciais. *Universidade Federal do Rio Grande do Sul, 2006, Rio Grande - RS*, 2006.
- STATHIS, P.; KAVALLIERATOU, E.; PAPAMARKOS, N. An evaluation technique for binarization algorithms. *Journal of Universal Computer Science*, v. 14, n. 18, p. 3011–3030, 2008.
- TRAHANIAS, P.; VENETSANOPOULOS, A. Colour edge detectors based on multivariate ordering. *In: Proceedings of SPIE, Visual Communications and Image Processing'92, Petros Maragos, University of Toronto*, v. 1818, p. 1396–1407, 1992.
- WILSON, S. Vector morphology and iconic neural networks. *IEEE Transaction Systems, Man and Cybernetics*, v. 1, n. 1, p. 636–644, 1989.
- WITTE, V. D. *Colour Morphology with Application to Image Magnification*. Universiteit Gent: Faculteit Wetenschappen, 2006. 220 p.
- WU, B. et al. A novel image segmentation method for complex document images. *16th IPPR Conference on Computer Vision Graphics and Image Processing CVGIP*, p. 646–647, 2003.
- YEN, J. C.; CHANG, F. A new criterion for automatic multilevel thresholding. *IEEE Transactions on Image Processing*, v. 1, p. 370–378, 1995.
- ZMUDA, M.; TAMBURINO, L.; RIZK, M. Efficient software techniques for morphological image processing with desktop computers. *SPIE Data Structures and Target Classification*, v. 1, p. 183–189, 1991.

ZMUDA, M. A.; TAMBURINO, L. A. Efficient algorithms for the soft morphological operators. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 18, n. 11, p. 1142–1147, 1996.