

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Um método para classificação de imagens de satélite
usando Transformada Cosseno Discreta com detecção e
remoção de nuvens e sombras

NOME DO AUTOR

Ana Carolina Quintão Siravenha

DM_18/2011

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2011

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

NOME DO AUTOR

Ana Carolina Quintão Siravenha

Um método para classificação de imagens de satélite
usando Transformada Cosseno Discreta com detecção e
remoção de nuvens e sombras

DM_18/2011

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2011

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

NOME DO AUTOR

Ana Carolina Quintão Siravenha

**Um método para classificação de imagens de satélite
usando Transformada Cosseno Discreta com detecção e
remoção de nuvens e sombras**

Dissertação submetida à Banca Examinadora do Programa de Pós-graduação em Engenharia Elétrica da UFPA para a obtenção do Grau de Mestre em Engenharia Elétrica, ênfase em Telecomunicações.

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2011

S619m Siravenha, Ana Carolina Quintão

Um Método para classificação de imagens de satélite usando Transformada Cosseno Discreta com detecção e remoção de nuvens e sombras / Ana Carolina Quintão Siravenha; orientador, Evaldo Gonçalves Pelaes.-2011.

Dissertação (Mestrado) - Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, Belém, 2011.

1. Processamento de imagens - técnicas digitais. 2. Sensoriamento remoto. 3. Filtros digitais (matemática). I. orientador. II. título.

CDD 22. ed. 621.367

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Um método para classificação de imagens de satélite usando Transformada Cosseno Discreta com detecção e remoção de nuvens e sombras

AUTOR: ANA CAROLINA QUINTÃO SIRAVENHA

DISSERTAÇÃO DE MESTRADO SUBMETIDA À AVALIAÇÃO DA BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA DA UNIVERSIDADE FEDERAL DO PARÁ E JULGADA ADEQUADA PARA OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA ELÉTRICA NA ÁREA DE TELECOMUNICAÇÕES.

APROVADA EM 28/04/2011

BANCA EXAMINADORA:

.....
Prof. Dr. Evaldo Gonçalves Pelaes (ORIENTADOR - UFPA)

.....
Prof. Dr. Aldebaro Barreto da Rocha Klautau Jr (MEMBRO - UFPA)

.....
Prof. Dr. Ronaldo Freitas Zampolo (MEMBRO - UFPA)

.....
Prof. Dr. Valquíria Gusmão Macedo (MEMBRO - UFPA)

VISTO:

.....
Prof. Dr. Marcus Vinícius Alves Nunes

COORDENADOR DO PPGEE/ITEC/UFPA

A Deus,
e à minha família,
com amor... .

Resumo

Apresenta-se nesta dissertação a proposta de um algoritmo supervisionado de classificação de imagens de sensoriamento remoto, composto de três etapas: remoção ou suavização de nuvens, segmentação e classificação. O método de remoção de nuvens usa filtragem homomórfica para tratar as obstruções causadas pela presença de nuvens suaves e o método Inpainting para remover ou suavizar a presença de sombras e nuvens densas. Para as etapas de segmentação e classificação é proposto um método baseado na energia AC dos coeficientes da Transformada Cosseno Discreta (DCT). O modo de classificação adotado é do tipo supervisionado. Para avaliar o algoritmo foi usado um banco de 14 imagens captadas por vários sensores, das quais 12 possuem algum tipo de obstrução. Para avaliar a etapa de remoção ou suavização de nuvens e sombras são usados a razão sinal-ruído de pico (PSNR) e o coeficiente Kappa. Nessa fase, vários filtros passa-altas foram comparados para a escolha do mais eficiente. A segmentação das imagens é avaliada pelo método da coincidência entre bordas (EBC) e a classificação é avaliada pela medida da entropia relativa e do erro médio quadrático (MSE). Tão importante quanto as métricas, as imagens resultantes são apresentadas de forma a permitir a avaliação subjetiva por comparação visual. Os resultados mostram a eficiência do algoritmo proposto, principalmente quando comparado ao *software Spring*, distribuído pelo Instituto Nacional de Pesquisas Espaciais (INPE).

PALAVRAS-CHAVE: Processamento de imagens, imagens de sensoriamento remoto, Classificação de imagens, Transformada cosseno discreta (DCT), Remoção de nuvens e sombras.

Abstract

This work proposes a supervised algorithm for classification of remote sensing images. It is composed by three stages: removal or smoothing of clouds, segmentation and classification. The removing clouds method uses homomorphic filtering to deal with obstructions caused by the presence of clouds and the Inpainting method to remove or soften the presence of dense clouds and shadows. The proposed segmentation and classification approaches are based on AC power coefficients of the Discrete Cosine Transform (DCT). Classification is used in the supervised mode. An image database is used to evaluate the implemented algorithm. This database is composed by 14 images obtained from various sensors, which 12 have some kind of obstruction. The Peak Signal-to-Noise Ratio (PSNR) and Kappa coefficient metrics are used to evaluate the removal or smoothing of clouds and shadows method. In this stage, several high-pass filters were compared to choose the most efficient. The image segmentation task is evaluated by the Edge Border Confidence (EBC) and the classification task is evaluated by the measure of the relative entropy and by the mean squared error (MSE). The resulting images are presented to allow the subjective evaluation by visual comparison. The experimental results show the efficiency of the proposed algorithm, especially when compared to the Spring software, distributed by the Instituto Nacional de Pesquisas Espaciais (INPE).

KEYWORDS: Image processing, remote sensory images, image classification, Discrete Cosine Transform (DCT), cloud and shadow removal.

Sumário

Lista de Figuras	iv
Lista de Tabelas	viii
Abreviações e Acrônimos	ix
1 Introdução	1
1.1 Contextualização	1
1.2 Motivação	3
1.3 Objetivos	4
1.4 Estrutura do documento	4
2 Fundamentos	5
2.1 Imagens digitais	5
2.1.1 Atributos de cor	7
2.2 Sensoriamento remoto	8
2.2.1 Assinatura espectral	8
2.2.2 Sistemas sensores	10
2.2.3 Resolução espacial	12
2.2.4 Resolução espectral	14
2.2.5 Resolução temporal	14
2.2.6 Resolução radiométrica	15
2.2.7 Satélites	15
2.2.7.1 Família Landsat	18
2.2.7.2 <i>Satellite Pour l’Observation de la Terre</i> - SPOT	20
2.2.7.3 <i>QuickBird</i>	21

2.2.7.4	<i>Ikonos</i>	21
2.2.7.5	<i>Chinese-Brazilian Earth Resources</i> - CBERS	21
3	Aspectos sobre processamento digital de imagens de satélite	23
3.1	Pré-processamento	23
3.1.1	Distorção geométrica	24
3.1.2	Distorção radiométrica	26
3.1.3	Correção atmosférica	27
3.1.4	Remoção de nuvens	28
3.2	Segmentação	31
3.3	Classificação	34
4	Algoritmo proposto	37
4.1	Remoção de nuvens	38
4.1.1	Detecção de nuvens e sombras	38
4.1.2	Filtro Homomórfico	39
4.1.3	<i>Inpainting</i>	41
4.1.4	Realce não linear no domínio da frequência	42
4.2	Segmentação	44
4.2.1	Transformada Cosseno Discreta	45
4.2.2	Análise de Histograma	47
4.2.2.1	Algoritmo de localização de picos	48
4.2.3	Filtros morfológicos	49
4.3	Classificação	50
5	Testes e resultados	54
5.1	Remoção de nuvens	54
5.1.1	Avaliação	57
5.2	Segmentação	61
5.2.1	Avaliação	62
5.3	Classificação	63
5.3.1	Avaliação	64

6 Conclusão	77
6.1 Trabalhos futuros	78
6.2 Trabalhos publicados	79
Bibliografia	87

Lista de Figuras

2.1	Cubo representativo do espaço de cores RGB [13].	6
2.2	Diagrama de cromaticidade [13].	6
2.3	Espectro eletromagnético. Adaptado de [14].	7
2.4	Curvas de reflectância generalizada para algumas aplicações de classificação de uso da terra. Adaptada de [18].	9
2.5	Comparação entre as assinaturas espectrais de uma imagem com resolução hiperespectral (superior) e multiespectral de 4 bandas (inferior). Fonte: <i>College of Earth and Mineral Sciences</i> - https://www.e-education.psu.edu/	13
2.6	Variação do diâmetro de imageamento do IFOV de um sensor em linha [15].	14
2.7	Técnicas de imageamento: (a)Varredura linear ou <i>whiskbroom</i> , (b) varredura eletrônica ou <i>pushbroom</i> , (c) varredura em quadros [26].	18
2.8	Linha do tempo dos satélites Landsat. Fonte: http://www.landsat.usgs.gov/	20
3.1	Exemplificação dos efeitos inerentes a plataforma de sensoriamento [31].	24
3.2	Impacto da distorção devido a forma da superfície terrestre [31] (a) Formato da cena ideal. (b) Formato da cena distorcido pela rotação da Terra.	25
3.3	Efeito da curvatura da Terra no imageamento de sua superfície [31].	25
3.4	Estimação da radiância de uma imagem pelo seu histograma [15].	27
3.5	Esquema de um filtro homomórfico. Adaptado de [38].	30
3.6	Transições entre pixels de borda: (a)Ideal, (b)Rampa e (c) <i>Roof edge</i> [6]	32
3.7	Histograma separado por um único limite [6].	33
3.8	Histograma separado por múltiplos limites. Adaptado de [6].	33
4.1	Diagrama de fluxo de dados do algoritmo proposto.	37

4.2	Função de transferência $H(u, v)$ de um filtro homomórfico esfericamente simétrico, onde $D(u, v)$ é a distância da origem [66].	40
4.3	Exemplo de RPM com $N = 4$, onde cada nível é subamostrado por um fator de 2.[74]	43
4.4	Diagrama do algoritmo de remoção de nuvens	44
4.5	Destaque das ações de segmentação do algoritmo proposto.	45
4.6	Histogramas (a) bimodal (b) multimodal. Adaptado de [6].	48
4.7	Exemplo de imagem usada nos testes: (a) Imagem Original; (b) DCT da imagem original com regiões em destaque.	52
4.8	Valores de (a) Energia e (b) Média dos blocos da DCT (coeficientes DC) . . .	53
4.9	Detalhe de uma região de confusão na imagem.	53
5.1	Detecção de nuvens e sombras: (a) Figura original retirada do repositório disponível. (b) Figura com as regiões destacadas.	55
5.2	Primeira fase: (a) Figura interpolada pelo algoritmo <i>Inpainting</i> . (b) Figura tratada pelo filtro Homomórfico.	56
5.3	Segunda fase: (a) Melhoramento não-linear da Figura 5.2(a). (b) Melhoramento não-linear da Figura 5.2(b). (c) Figura resultante do processo de filtragem de nuvens.	56
5.4	Valores PSNR para imagens filtradas por filtros homomórficos.	59
5.5	Valores Kappa para imagens filtradas por filtros homomórficos.	59
5.6	Valores PSNR para imagens tratadas por filtragem homomórfica e <i>Inpainting</i>	60
5.7	Valores Kappa para imagens tratadas por filtragem homomórfica e <i>Inpainting</i>	60
5.8	Efeito visual: (a) Imagem de Roma-ITA (Imagem 8) limpa de presença de nuvens. (b) Figura tratada por <i>Inpainting</i> e filtro homomórfico <i>High-boost</i> . (c)Figura tratada por <i>Inpainting</i> e filtro homomórfico Ideal.	61
5.9	Relação de valores de entropia relativa entre as imagens classificadas pelo método proposto e classificadas pelo <i>software Spring</i>	64
5.10	Relação de valores de MSE entre as imagens classificadas pelo método proposto e classificadas pelo <i>software Spring</i>	65

5.11	Imagem 1 - Cena capturada pelo satélite Ikonos 4: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	66
5.12	Imagem 2 - Cena capturada pelo satélite Ikonos 5: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	67
5.13	Imagem 3 - Cena da região do Hawaii-EUA: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	67
5.14	Imagem 4 - Cena da região de Tucuruí-Pa: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	68
5.15	Imagem 5 - Cena da região de Sedona, Arizona-EUA: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	69
5.16	Imagem 6 - Cena da região do lago <i>Leech</i> em Minnesota-EUA: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	69
5.17	Imagem 7 - Cena da região Andina: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	70
5.18	Imagem 8 - Imagem de Roma-ITA: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	71
5.19	Imagem 9 - Imagem da região metropolitana de Belém: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	72
5.20	Imagem 10 - Cena da região da Baía do rio Capim-Pa: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	73

5.21	Imagem 11 - Cena da cidade de Macapá-Ap: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	74
5.22	Imagem 12 - Cena da região de Tucuruí-Pa: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo <i>software Spring</i>	75
5.23	Imagem 13 - Cena da cidade de Tucuruí-Pa: (a) Imagem original (b) Imagem classificada pelo método proposto (c) Imagem classificada pelo <i>software Spring</i>	76
5.24	Imagem 14 - Cena da região do rio Tocantins-Pa: (a) Imagem original (b) Imagem classificada pelo método proposto (c) Imagem classificada pelo <i>software Spring</i>	76

Lista de Tabelas

2.1	Aplicações para as bandas imageadas pelo sensor TM [19].	10
2.2	Relação de sensores hiperespectrais, quantidade de canais imageados e a largura do espectro em que estão compreendidos. Adaptada de [23].	12
4.1	Valores de energia AC para diferentes regiões de uma imagem de satélite.	51
5.1	Matriz de confusão para a imagem de Roma-ITA original (com nuvens) e limpa manualmente.	57
5.2	Matriz de confusão para a imagem de Roma-ITA filtrada de nuvens e limpa manualmente.	58
5.3	Valores EBC para as figuras segmentadas pelo método de segmentação proposto e pelo <i>software Spring</i>	63

Abreviações e Acrônimos

ASTER	-	<i>Advanced Spaceborne Thermal Emission and Reflection Radiometer</i>
AVHRR	-	<i>Advanced Very High Resolution Radiometer</i>
CCD	-	<i>Charge-coupled device</i>
CIE(L*a*b)	-	<i>Espaço de cores tridimensional onde a axial L representa a luminosidade e as coordenadas a* e b* representam respectivamente avermelhar-esverdear e amarelar-azular</i>
DCS	-	<i>Data Collection System</i>
DCT	-	<i>Discrete Cosine Transform</i>
EBC	-	<i>Edge-border coincidence - coincidência entre bordas</i>
ETM	-	<i>Enhanced Thematic Mapper</i>
ETM+	-	<i>Enhanced Thematic Mapper Plus</i>
FFT	-	<i>Fast Fourier Transform - Transformada Rápida de Fourier</i>
FPA	-	<i>Filtro passa-alta</i>
GCP	-	<i>Ground Control Points - pontos de controle terrenos</i>
HIRS	-	<i>High-Resolution Infrared Radiation</i>
HRG	-	<i>High Resolution Geometric</i>
HRS	-	<i>High-Resolution Stereoscopic</i>
HRV	-	<i>High Resolution Vertical</i>
HRVIR	-	<i>High Resolution Vertical Infrared</i>
HSB	-	<i>Variação do espaço de cores HSV, Hue (matiz), Saturation (saturação) e brightness (brilho)</i>
HSL	-	<i>Variação do espaço de cores Hue (matiz), Saturation (saturação) e lightness (claridade)</i>
HSV	-	<i>Espaço de cores Hue (matiz), Saturation (saturação) e Value (valor)</i>

IFOV	-	<i>Instantaneous Field of View (Campo de Visada Instantâneo)</i>
INPE	-	<i>Instituto Nacional de Pesquisas Espaciais</i>
MODIS	-	<i>Moderate-resolution Imaging Spectroradiometer</i>
MSE	-	<i>Mean Square Error - erro médio quadrático</i>
MSS	-	<i>Multispectral Scanner System</i>
ND	-	<i>Digital number</i>
NVDI	-	<i>Normalized Vegetation Difference Index</i>
PDI	-	<i>Processamento Digital de Imagens</i>
PSNR	-	<i>Peak Signal-to-noise ratio - razão sinal-ruído de pico</i>
RBV	-	<i>Return Beam Vidicon</i>
REM	-	<i>Radiação Eletromagnética</i>
RGB	-	<i>Espaço de cores Red (vermelho), Green (verde) e Blue (azul)</i>
RMSE	-	<i>Root Mean Square Error - erro médio quadrático</i>
SEM	-	<i>Solar Environment Monitor</i>
SNR	-	<i>Signal-to-noise ratio, razão sinal-ruído</i>
SPOT	-	<i>Satellite Pour l'Observation de la Terre</i>
TM	-	<i>Thematic Mapper</i>
TOVS	-	<i>Tiros Operational Vertical Sounder</i>
YCbCr	-	<i>Espaço de cores usado em sistemas de vídeo. Y é a componente de luminância, Cb e Cr são componentes cromáticas azul e vermelha, respectivamente</i>

Capítulo 1

Introdução

Este capítulo tem por objetivo apresentar uma introdução abrangente sobre o conteúdo apresentado neste trabalho, as motivações para a pesquisa e os objetivos a serem alcançados. A próxima seção destacará os aspectos que validam a proposta apresentada através da contextualização dos assuntos abordados, como pré-processamento, segmentação e classificação. Em seguida, são apresentados os motivos que levaram ao desenvolvimento deste trabalho e os objetivos pretendidos pela pesquisa desenvolvida. Finalmente, a estrutura do trabalho é exposta na última seção.

1.1 Contextualização

O desenvolvimento de tecnologias direcionadas ao processamento de imagens teve um crescimento notório nas últimas 3 décadas [1]. A expressiva quantidade de trabalhos científicos na área de processamento de imagens deve-se ao fato de que, pela grande heterogeneidade dos dados a serem analisados, os algoritmos desenvolvidos para essas tarefas devem atender a inúmeras variáveis particulares a cada caso ou grupo de imagens.

O avanço da tecnologia, com a possibilidade de comunicação confiável e em grande escala entre observatórios em terra e sensores remotos, proporcionou aos centros de pesquisa a capacidade de monitoramento de grandes superfícies, como, por exemplo, a Floresta Amazônica com seus 5,5 milhões de km².¹ Imagens de satélite são fotos da superfície terrestre retiradas por sensores remotos ou satélites artificiais imageadores presentes na órbita da terra. Sensoriamento remoto, como define [2], é a ciência da obtenção de informação sobre um objeto, área ou fenômeno através da análise de dados adquiridos por um dispositivo que não

¹<http://www.ambientebrasil.com.br/>

está em contato com o objeto. Comumente, a aquisição dessas imagens é feita por sensores eletromagnéticos operados remotamente.

Operações de pré-processamento são necessárias para correção de certas distorções, degradações e incidência de ruídos indesejáveis. Essas incidências podem ser introduzidas durante o processo de imageamento e podem ser causadas por variações na iluminação, no ângulo solar, no ângulo de visada do sensor usado, dentre outros fatores [3]. Dentro do universo de imagens de sensores remotos podem-se destacar algumas operações de pré-processamento bastante comuns como: a correção de distorções geométricas e atmosféricas [4], e filtragem de remoção de nuvens. Trabalhos dedicados somente à remoção de objetos são comuns e utilizam combinações de técnicas como síntese de textura para preenchimento de grandes regiões e a técnica *Inpainting* para o preenchimento de pequenas regiões, que juntas podem suprimir grandes objetos de uma imagem digital [5].

Para a tarefa de particionar uma imagem em regiões homogêneas é dado o nome de segmentação. Por sua importância dentro das rotinas de processamento digital de imagens (PDI), esta é a tarefa que exige mais atenção durante o desenvolvimento, isso porque a separação da imagem em regiões de homogeneidade irá determinar a qualidade da interpretação [6]. Uma segmentação com alta granularidade pode representar uma sobre-segmentação (*over segmentation*) onde as regiões de homogeneidade encontradas separam até as mais sutis regiões de heterogeneidade; no sentido contrário, uma segmentação pobre pode produzir uma sub-segmentação (*under segmentation*), onde feições não homogêneas são agrupadas indiscriminadamente.

Algoritmos de segmentação tradicionais geralmente são baseados em regiões ou a partir de identificação de bordas. Esses algoritmos podem incorporar informação *a priori* adquiridas, em sua maioria, pela intervenção de especialistas e de forma manual [7]. O uso de informação de textura nas tarefas de segmentação também é bastante empregado pois essas informações estão intimamente ligadas à correlação entre os *pixels* formadores da imagem [6, 8]. Além disso, a DCT, comumente usada para propósitos de compressão, pode ser usada para representar características de textura de forma mais eficiente quando comparado a sistemas baseados em Transformada *Wavelet* Discreta [9]

Apesar de todos os estudos publicados sobre a tarefa de segmentação de imagens, não há uma técnica padrão a ser escolhida para uma aplicação em particular. Não há também, um acordo sobre como proceder a comparação de um método sobre outro, com relação ao seu desempenho. Segundo [1], isso é decorrente da ambiguidade inerente ao propósito e ao escopo do processo de segmentação nas diversas aplicações. Portanto, a avaliação dos métodos é feita de modo genérico, com o objetivo de medir atributos que sejam pertinentes ao maior número

de aplicações.

A tarefa de classificação, dentro do contexto de imagens, consiste na adequação dos *pixels* em um determinado número de classes pré-definidas [10]. Diversas abordagens podem ser implementadas para essa tarefa e podem ser divididas em abordagens supervisionadas e não supervisionadas.

Tão complexa quanto a tarefa de segmentação, a tarefa de classificação torna-se desafiadora pelo fato de que o fundo da imagem possui uma grande variação quanto aos desvios espectrais, além da existência de pequenas variações entre classes próximas. Em [11] é apresentada uma abordagem não supervisionada para tentar contornar os desafios citados. Para isso, investiga a remoção do fundo da imagem, o processo de geração de amostras de treinamento - crucial para um classificador não supervisionado - e a determinação de classes.

Os esforços atuais sobre as tarefas abordadas nos parágrafos anteriores estão concentrados principalmente na otimização, modificação e generalização de técnicas que já se mostraram eficientes para alguns casos, sendo justificados através da comparação dos resultados dessas modificações com o algoritmo original.

1.2 Motivação

O desenvolvimento de técnicas de processamento de imagens busca contribuir para a tarefa de interpretação realizada por analistas humanos mediante um melhoramento visual aplicado à imagem. Também visa promover a percepção automática através de máquinas.

O crescimento do uso de imagens remotamente adquiridas observado nas últimas décadas é reflexo do avanço tecnológico dos sensores, que proporciona a obtenção de imagens com resolução de poucos metros. O domínio do uso dessas imagens vem crescendo, passando a contribuir para as mais diversas análises sobre recursos naturais, por exemplo, e de forma mais abrangente, o processamento de imagens agrega, cada vez mais, conhecimento para áreas como biomedicina, análise metalográfica e automação visual.

Os algoritmos dedicados ao processamento de imagens de satélite devem ser robustos na medida em que esse tipo de imagem geralmente possui grandes dimensões e grandes variações espectrais (bastante heterogêneas). Nesse sentido, o encorajamento para o desenvolvimento de novas alternativas para a interpretação de imagens e o aprimoramento de técnicas reconhecidamente eficientes é latente.

1.3 Objetivos

Este trabalho tem como objetivo, em primeiro lugar, apresentar uma revisão sobre o estado-da-arte nos temas desta proposta: pré-processamento, segmentação e classificação de imagens de satélite. Em segundo, apresentação dos algoritmos de remoção de nuvens densas e suaves, uma das abordagens da tarefa de pré-processamento.² Finalmente, propõe-se um algoritmo classificador de imagens de satélite baseado na Transformada Cosseno Discreta (DCT), que utiliza limiarização de histogramas para a segmentação, filtragem homomórfica e interpolação por vizinho próximo para remoção de nuvens.

1.4 Estrutura do documento

A dissertação está estruturada da seguinte forma:

- Capítulo 2 - Apresentação dos conceitos básicos sobre a aquisição e processamento de imagens de forma geral, imagens de satélite, sistemas de imageamento e alguns dos satélites mais produtivos do ponto de vista de sensoriamento remoto.
- Capítulo 3 - Apresentação dos conceitos de pré-processamento, com ênfase na remoção de nuvens, segmentação, classificação e DCT.
- Capítulo 4 - Descrição dos algoritmos implementados, apresentados nesta dissertação.
- capítulo 5 - Descrição dos testes realizados e discussões sobre os resultados.
- Capítulo 6 - Conclusão do trabalho e apresentação das propostas de trabalhos futuros.

²Neste trabalho não serão executadas ações de correção de distorções geométricas e geográficas.

Capítulo 2

Fundamentos

Este capítulo apresenta os conceitos básicos ligados à formação de imagens. O estudo estará focado em imagens bidimensionais. Será destacado o aspecto que tange a formação de imagens por sensores remotos e os sistemas de imageamento responsáveis por essa captação.

2.1 Imagens digitais

Uma imagem é definida como a representação de um objeto através de um desenho ou a representação visual de um objeto dada por um instrumento de ótica. Uma imagem digital pode ser definida por meio de uma função bidimensional $f(x, y)$, onde x e y representam as coordenadas espaciais no plano [12]. A amplitude de f é chamada de intensidade naquele ponto (x, y) , e a esse ponto, que em conjunto com outros pontos formam a imagem digital, dá-se o nome de *pixel*.¹

Sendo uma imagem digital formada por vários pixels que se distribuem em linhas e colunas, pode-se entender uma imagem como uma matriz bidimensional de *pixels*. O termo nível de cinza ou brilho ocorre quando se faz referência a imagens monocromáticas, formadas por apenas uma matriz de *pixels* (níveis de cinza). Imagens coloridas são formadas pela combinação de matrizes individuais de níveis de cinza. Para o caso de imagens no espaço de cor conhecido como RGB, é feita a combinação de três matrizes bidimensionais individuais representando as componentes vermelho (do inglês, **Red**), verde (do inglês, **Green**) e azul (do inglês, **Blue**).

Comumente o espaço de cores RGB é representado por um cubo de cores (Figura 2.1). Nesse cubo, os valores estão normalizados, ou seja, variam de 0 a 1, e os vértices representam

¹O termo *pixels* ou *pels* são abreviações do termo em inglês *picture elements*, ou seja, elementos da figura.

as cores aditivas primárias e secundárias. Os tons de cinza estão localizados na diagonal que liga as cores preto e branco.

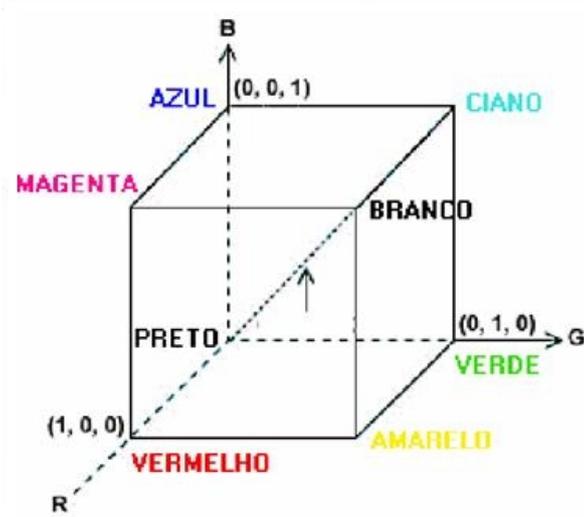


Figura 2.1: Cubo representativo do espaço de cores RGB [13].

As cores saturadas, representadas sobre as arestas do cubo, são cores que só podem ser representadas no espaço de cores em questão. As cores totalmente puras não podem ser representadas pelo espaço de cores RGB. Estas cores estão localizadas no contorno do diagrama da Figura 2.2. Nesta, o triângulo delimita a região que abriga as cores primárias aditivas. Os vértices do triângulo, chamado *gamut*, estão sobre as cores primárias vermelho, azul e verde [13].

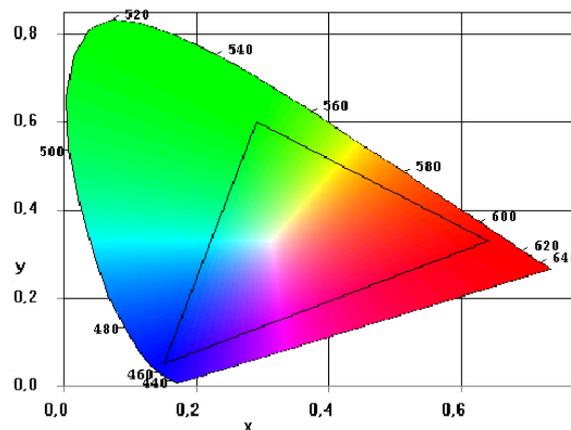


Figura 2.2: Diagrama de cromaticidade [13].

As cores representadas nos mais diversos espaços de cores existente, tais quais HSV, YCbCr, CIE(L*a*b*), por exemplo, são descritas em termos de comprimento de onda, pois este é um parâmetro de fácil medida [14]. A Figura 2.3 mostra o diagrama do espectro eletromagnético, onde o intervalo de comprimento de onda está compreendido em 13 potências

de 10. Esse diagrama, adaptado de [14], não representa todo o espectro, mas destaca o espectro visível. O olho humano usa a luz visível para coletar a informação visual e essa banda visível está compreendida no intervalo de 400nm a 750nm , onde as regiões do espectro não estão delimitadas em intervalos estritamente definidos, por isso há sobreposição de bandas e subclassificação destas.

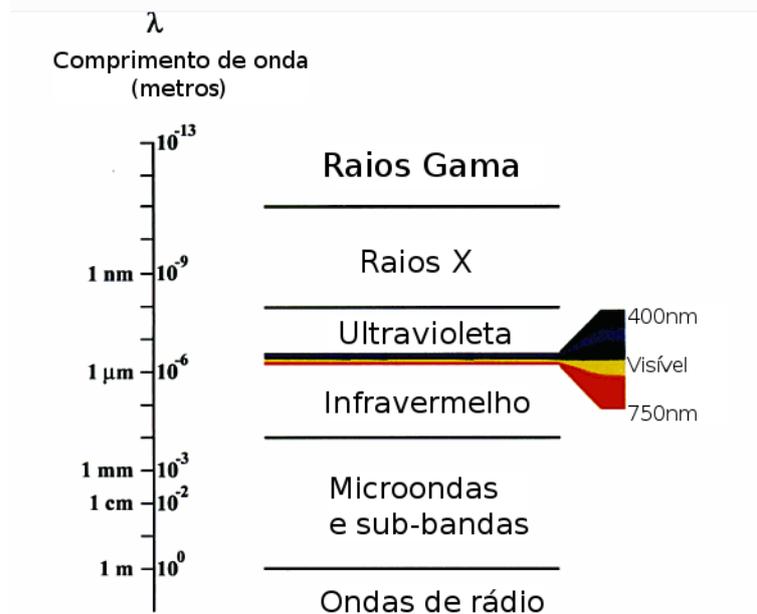


Figura 2.3: Espectro eletromagnético. Adaptado de [14].

2.1.1 Atributos de cor

A representação tridimensional do espaço de cores descreve sobre seus eixos uma quantização da energia referente aos comprimentos de onda das cores primárias e os pontos no interior do espaço são resultado da combinação única dessas três componentes [13].

A associação dessas componentes produz uma intensidade (ou brilho) particular, adicionalmente, a quantidade de luz branca determina a saturação e a cor predominante chamada matiz. As medidas dos valores de brilho e saturação estão compreendidos entre os valores de 0 e 1, enquanto a matiz é medida em valores de graus, entre 0 e 360. Para o grau 0 a matiz será o vermelho, as matizes verde e azul são representados pelos graus 120 e 240, respectivamente.

Os atributos de matiz, saturação e brilho são parâmetros de cores dos espaços HSB e HSL, enquanto que no espaço de cores RGB o brilho é entendido como a média aritmética das três coordenadas.

2.2 Sensoriamento remoto

Sensoriamento remoto é definido pelo uso de sensores instalados em aeronaves ou satélites para detectar a energia eletromagnética emitida ou dispersada pela superfície terrestre [15].

Instrumentos de sensoriamento remoto podem ser projetados para operar em mais de uma banda espectral, e essa escolha depende do objetivo do projeto e principalmente da tecnologia disponível para uso.

Outro fator determinante na escolha das bandas espectrais a serem operadas é a interação dessas bandas com a atmosfera. A absorção e o espalhamento são os principais mecanismos que alteram a intensidade e a direção da radiação eletromagnética (REM), sendo a absorção responsável pela inviabilidade de um projeto de sensoriamento remoto, restringindo o estudo às chamadas janelas atmosféricas, que são regiões de baixa absorção. Todas as bandas do espectro sofrem com as interações atmosféricas, em maior ou menor escala, de forma variada no espaço e tempo.

2.2.1 Assinatura espectral

A assinatura espectral está relacionada à intensidade com que cada corpo absorve, reflete e emite radiação eletromagnética (REM) ao longo do comprimento de onda. A REM proveniente da reflexão ocorre por um espalhamento de fótons que ocorre de forma especular ou difusa (também chamada de Lambertiana), provenientes de superfícies lisas e rugosas, respectivamente. A reflexão por uma superfície perfeitamente lisa possui um ângulo de reflexão igual ao ângulo de incidência, enquanto que uma superfície perfeitamente Lambertiana reflete a energia de forma igualitária em todas as direções, fato bastante raro nas aplicações práticas [15]. A REM por emissão envolve uma absorção seguida da emissão em outros comprimentos de onda de fontes radiantes no ambiente onde o objeto está situado e é basicamente dependente da temperatura desse objeto [16]. A REM por absorção está ligada às características das partículas em seu espectro de absorção, ou seja regiões do espectro onde ela absorve energia. Nessas regiões, ocorrem as janelas atmosféricas, ou seja, regiões do espectro nas quais a atmosfera é transparente [17].

A quantificação da reflectância de objetos é feita através da medição da porção de energia refletida por ele, feita em função do comprimento de onda. Essa medida é chamada de reflectância espectral e pode ser observada na Figura 2.4. Esta imagem mostra as curvas de reflectância para as componentes de solo, vegetação estressada e sadia, além das respostas

refletivas de neve e água, cuja interpretação é extensível à componentes como telhas de barro, alumínio, concreto e asfalto, por exemplo.

A absorção de radiação por um objeto resulta na perda real de energia da radiação, ocasionada principalmente pela ocorrência de vapor de água, dióxido de carbono e ozônio, gases amplamente dispersos na atmosfera terrestre. A absorção de radiação solar por parte desses gases ocorre em bandas espectrais específicas, e por isso a escolha da banda para a observação de um fenômeno deve ser cuidadosamente considerada durante a elaboração do projeto de um sensor.

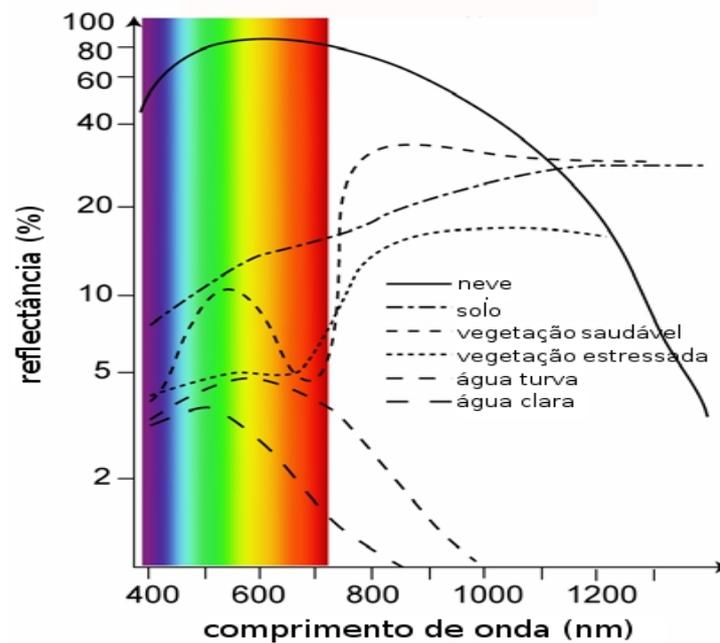


Figura 2.4: Curvas de reflectância generalizada para algumas aplicações de classificação de uso da terra. Adaptada de [18].

A Figura 2.4 mostra que a região de menor reflectância em medidas percentuais dá-se na faixa correspondente às bandas espectrais violeta e azul. O aumento da reflectância é gradual e atinge os valores mais altos na banda do infravermelho próximo e médio (região posterior aos 700nm).

Uma medida comum nas análises espectrais é o índice de vegetação por diferença normalizada, do inglês *Normalized Difference Vegetation Index* - NDVI. Essa medida se vale da característica de reflectância da vegetação que é alto na faixa do infravermelho próximo e baixo na faixa do vermelho. Esse índice mostra baixa sensibilidade à mudanças atmosféricas e pode ser definido de várias formas, sendo o mais conhecido definido pela equação:

$$NDVI = \frac{(IFP) - (V)}{(IFP) + (V)} \quad (2.1)$$

onde IFP representa o valor de reflexão na região do infravermelho próximo e V a região do vermelho.

O conhecimento acerca do comportamento espectral dos alvos é base para a determinação das bandas espectrais a serem escolhidas para uma determinada pesquisa, como monitoramento da vegetação ou a identificação de minérios no solo. Um exemplo dessa diferenciação é mostrada na Tabela 2.1, onde é feita a relação entre as bandas imageadas pelo sensor *Thematic Mapper* (TM) (a ser descrito na Seção 2.2.7) e suas aplicações.

Tabela 2.1: Aplicações para as bandas imageadas pelo sensor TM [19].

Canal	Faixa espectral (μm)	Principais aplicações
1	0,45 - 0,52	Mapeamento de águas costeiras Diferenciação entre solo e vegetação Diferenciação entre vegetação coníferas e decídua
2	0,52 - 0,60	Reflectância de vegetação verde sadia
3	0,63 - 0,69	Absorção de clorofila Diferenciação de espécies vegetais
4	0,76 - 0,90	Levantamento de biomassa Delineamento de corpos d'água
5	1,55 - 1,75	Medidas de umidade da vegetação Diferenciação entre nuvens e neve
6	10,4 - 12,5	Mapeamento de estresse térmico em plantas Outros mapeamentos térmicos
7	2,08 - 2,35	Mapeamento hidrotermal

2.2.2 Sistemas sensores

A aquisição de dados remotos pode ser feita usando métodos analógicos ou digitais, acoplados à satélites artificiais colocados na órbita da Terra por meio de foguetes. O tempo de vida desses veículos permite o monitoramento de grandes áreas durante períodos prolongados de tempo.

Por sua natureza discreta, o valor do nível de cinza² representa a intensidade da energia

²Também representada pela sigla DN, do inglês *Digital Number*, Número Digital.

eletromagnética refletida ou emitida, captada pelo sensor.

Os sensores podem ser divididos em duas categorias [20]: sensores ativos e sensores passivos. Se o sensor emitir energia que irá interagir com o objeto alvo é chamado sensor ativo, dos quais se destacam os radares e os emissores de raios *laser*. Sensores passivos, que representam a maioria dos satélites em órbita, captam a energia originada de uma fonte externa, principalmente o Sol, cuja energia captada pelos sensores é aquela resultante da interação com a superfície terrestre.

Em geral, os sistemas sensores podem ser divididos em quatro sistemas básicos [21, 22]: sistema óptico de coleta, sistema de dispersão, detector e subsistema de processamento. O sistema óptico é responsável pela determinação da resolução espacial (comentada na sub-seção 2.2.3) e pela área de coleta da radiação. O sistema de dispersão possui um arranjo óptico para produzir um espectro a partir da radiação incidente. O sistema de detecção tem a função de transformar a REM em sinal de energia elétrica. Detectores térmicos tem como saída a alteração de temperatura da energia incidente, os fotoemissivos tem seu sinal de saída em função do efeito fotoelétrico, ou seja, o deslocamento de elétrons proporcional à energia incidente, e os quânticos tem seu sinal de saída dependente da incidência de fótons sobre o detector, estes que interagem de forma direta com os níveis de energia dos elétrons, produzindo transportadores de carga livre. O processamento do sinal recebido é a última fase do sistema sensor, e através de técnicas específicas (amplificação e filtragem, por exemplo) produz o dado desejado sobre o alvo.

A maioria dos sensores remotos, bem como algumas máquinas fotográficas, são capazes de adquirir imagens pela captação da interação entre um objeto e as várias bandas do espectro eletromagnético. Sensores hiperespectrais mapeiam as cenas através de várias bandas espectrais contínuas (dezenas ou centenas de bandas) e estreitas (centésimos de μm) de forma simultânea. Essas imagens geralmente possuem resolução espacial de até poucas dezenas de metros, são limitadas pela performance dos instrumentos e pela forte correlação existente entre as faixas de frequência adjacentes. Ainda sim, geram imagens de maior resolução espectral do que imagens multiespectrais, que usam largura de banda maiores. A Tabela 2.2 mostra uma relação de sensores hiperespectrais e suas respectivas bandas de imageamento.

A Figura 2.5 mostra a diferença entre as assinaturas espectrais das componentes de uma imagem obtida por um sensor hiperespectral (imagem superior) e a mesma imagem obtida por um sensor multiespectral (imagem inferior). O bloco espectral formado na imagem superior é composto por 64 bandas espectrais compreendidas entre as faixas de $415nm$ e $900nm$. Por sua vez, o bloco inferior é composto por 4 bandas compreendidas na faixa de $475nm$ e $750nm$.

Outra modalidade de captação de imagens é a pancromática, onde um radiômetro

Tabela 2.2: Relação de sensores hiperespectrais, quantidade de canais imageados e a largura do espectro em que estão compreendidos. Adaptada de [23].

Sensor	Canais	Largura (μm)
AAHIS <i>Advanced Airborne Hyperespectral Imaging System</i>	288	0,39 - 0,84
AISA Airborne Imaging Spectrometer for Applications	286	0,45 - 0,9
ASAS Advanced Solid State Array Spectroradiometer	62	0,4 - 1,06
CASI Compact Airborne Spectrographic	288	0,43 - 0,86
HIRIS High Resolution Imaging Spectrometer	192	0,4 - 2,5
IMSS Image Multispectral Sensing	320	2 - 5
MIDIS Multiband Identification and Discrimination Imagens Spectroradiometer	256	0,4 - 30

registra a energia eletromagnética em uma única banda de frequência, englobando boa parte da luz visível e do infravermelho próximo, formando imagens de uma única camada (escala de cinza).

2.2.3 Resolução espacial

Esse parâmetro se refere à capacidade do sensor de imagear objetos. Nesse sentido, quanto menor o objeto passível de captação por meio do sensor, maior a resolução espacial.

A resolução espacial de um sensor pode ser medida pelo seu campo de visada instantâneo (IFOV - *Instantaneous Field of View*), que é determinado pelas características

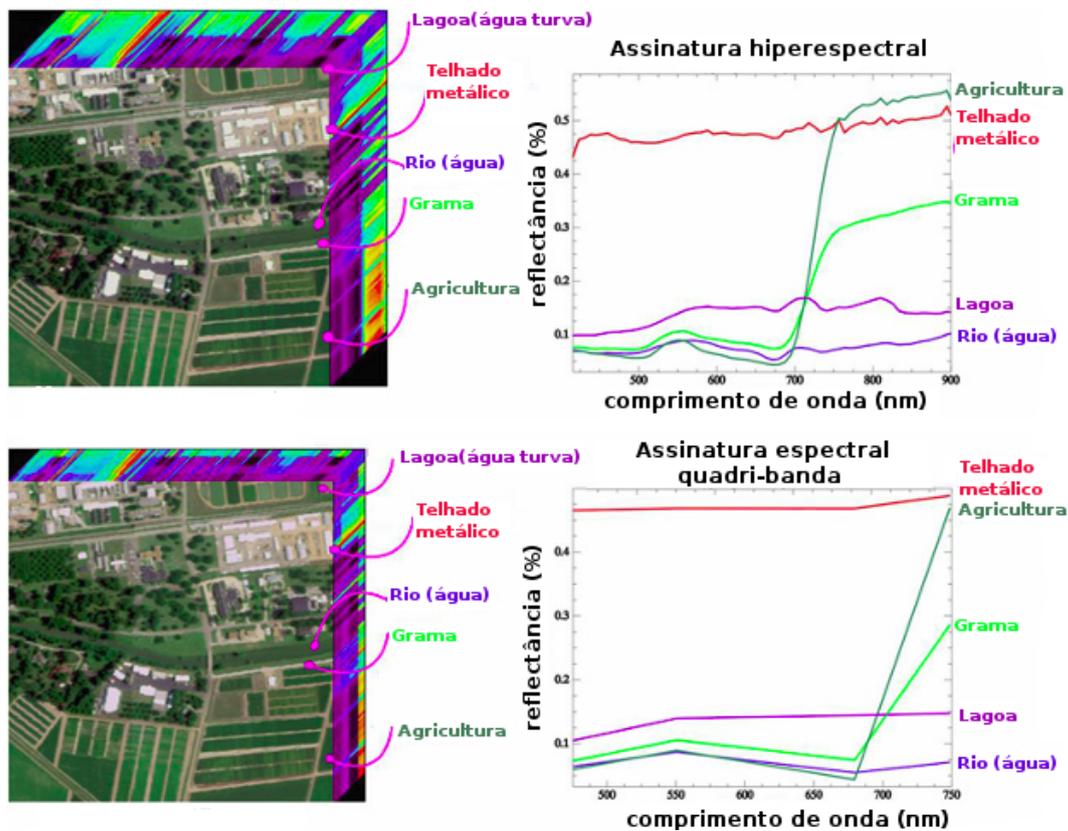


Figura 2.5: Comparação entre as assinaturas espectrais de uma imagem com resolução hiperespectral (superior) e multiespectral de 4 bandas (inferior). Fonte: *College of Earth and Mineral Sciences* - <https://www.e-education.psu.edu/>

geométricas que à uma dada altitude e em um dado momento, determina a área imageada pelo sensor. O IFOV, ilustrado na Figura 2.6, representa o tamanho do *pixel* e sua conceituação é teórica, pois varia de acordo com a altitude do sensor em relação a superfície, valor mutável de acordo com a órbita do satélite. Vale destacar que o conceito de *pixel* acontece somente no produto final à qual um valor de radiância lhe é atribuído.

O termo IFOV *Nadir* ou IFOV Nadiral refere-se à região apontada ortogonalmente pelo sensor à superfície, na direção de escaneamento. Os sensores também podem imagear áreas fora da região nadiral e essa área, apontada como IFOV *Off Nadir* ou IFOV Pós-nadiral na Figura ??, possui um IFOV maior em relação ao primeiro, porém com uma imagem resultante distorcida devido a diminuição da escala espacial em relação ao centro. A correção dessas distorções é feita por correção geométrica.

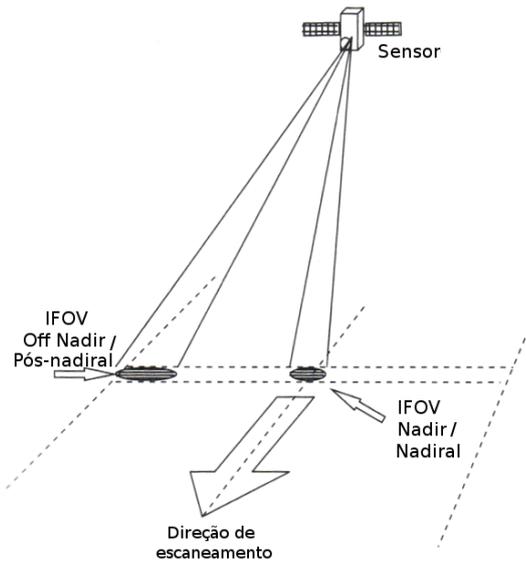


Figura 2.6: Variação do diâmetro de imageamento do IFOV de um sensor em linha [15].

2.2.4 Resolução espectral

Aplica-se aos sensores que produzem imagens multiespectrais e é definida pelo número de bandas espectrais do sistema, associada à largura do intervalo do comprimento de onda coberto por essas bandas. Quanto maior o número de bandas e menor a largura do intervalo entre elas, maior é a resolução espectral do sensor.

A resolução espectral é afetada diretamente pela alteração do IFOV de um sensor. A redução de IFOV causa piora na razão sinal-ruído do sistema de imageamento³, fato que pode ser compensado pelo aumento da largura de banda de escaneamento.

2.2.5 Resolução temporal

Também conhecida como tempo de revisita, é dependente da órbita do satélite e das características do sensor. Por princípio, em órbitas geoestacionárias é possível obter imagens em qualquer instante, porém a captação dessas imagens depende do tempo de aquisição de dados pelo sensor.

³A razão sinal ruído (SNR, sigla em inglês para *signal-to-noise ratio*) estabelece o quanto um sinal foi corrompido por ruído e é descrito pela equação: $SNR = Potência_{sinal}/Potência_{ruído}$.

2.2.6 Resolução radiométrica

Esse parâmetro é dado pelo número de níveis de cinza usados na representação das imagens captadas pelo sensor. A resolução radiométrica será maior quanto maior forem os níveis digitais de cinza, que hoje variam entre 6 e 8 bits (64 e 256 níveis de cinza).

A resolução espacial é influenciada pelo IFOV do instrumento. O uso de um IFOV estreito implica diretamente na quantidade de energia recebida pelo detector, que é menor. Isso acontece pois a área de energia que é coletada é menor, e o tempo disponível para detectar essa energia também é mais curto. Dessa forma, a sensibilidade do instrumento para mudanças nos níveis de energia irá diminuir, podendo o sensor não ser capaz de distinguir pequenas diferenças de energia ao longo da linha imageada. Assim, a resolução radiométrica é degradada [15].

2.2.7 Satélites

Satélites artificiais são plataformas que carregam dispositivos destinados à observação de objetos e fenômenos ao redor da Terra, de outros planetas ou de satélites naturais. Seu desenvolvimento começou no final da década de 50, precisamente em novembro de 1957, quando a extinta União Soviética pôs em órbita o satélite *Sputnik I*. Em menos de um ano, em fevereiro de 1958, os Estados Unidos lançaram o *Explorer I*. Após esses eventos, ano a ano, o sistema solar passou a ser explorado e até 1975, o Sol, a Lua, Marte e Vênus já possuíam em sua órbita um satélite artificial de observação. Até 1977 planetas como Júpiter, Saturno, Urano e Netuno já haviam recebido sondas de observação.

Assim como grande parte da tecnologia desenvolvida, o uso de satélites para monitoramento de grandes regiões foi inicialmente motivado por ações militares. Segundo [24], cerca de 75% dos satélites lançados a partir de 1957 tiveram essa finalidade. Exemplo disso são os satélites da família *Key Hole* ou KH, que possuem uma resolução de aproximadamente 0,15 metros, o que permite, por exemplo, avistar uma bicicleta em um parque. Essa família foi precedida pelos satélites Corona, que os Estados Unidos lançaram entre os anos de 1960 e 1972.

As plataformas orbitais dedicadas ao estudo de fenômenos terrestres, geralmente estão localizados em órbita baixa - poucas centenas de quilômetros de altitude -, e podem servir para múltiplos propósitos. Por exemplo, o satélite NOAA (*National Oceanic and Atmospheric Administration*) foi projetado para atender à fins meteorológicos, mas possui um dispositivo acoplado que permite receber e retransmitir para estações em terra informações sobre bóias marítimas.

Os sensores instalados nas plataformas orbitais atendem aos objetivos da missão estabelecida. Os mais usados estão listados na sequência, atentando para algumas de suas aplicações:

- *Thematic Mapper (TM)*: é usado para diferenciação de solo e vegetação, diferenciação entre espécies de plantas, estudos sobre biomassa ou eventos termais e diferenciação de nuvens e neve (Tabela 2.1).
- *Enhanced Thematic Mapper (ETM)* e *Enhanced Thematic Mapper Plus (ETM+)*: seguem as mesmas aplicações do sensor TM, com melhorias internas e adição de banda Pan-cromática.
- *Multispectral Scanner System (MSS)*: eficiente para estudos sobre zonas costeiras, áreas urbanas, mapeamento de fronteiras água/terra e estudos sobre sedimentação marinha.
- *Moderate-resolution Imaging Spectroradiometer (MODIS)*: monitora aspectos de temperatura e umidade da atmosfera e da superfície terrestre, ocorrência de incêndios, monitora a temperatura do oceano, a cobertura de neve e gelo do mar.
- *Return Beam Vidicon (RBV)*: câmera pancromática usada para obtenção de luz visível e infravermelho próximo. As imagens de alta resolução captadas por esta câmera são usadas para mapeamento de regiões.
- *Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER)*: observa a dinâmica de gelo na terra e no mar, a dinâmica da vegetação, efeitos vulcânicos, ocorrência de incêndios, além das propriedades das nuvens.

Outros instrumentos importantes no processo de sensoriamento remoto descritos em [25] são elencados na sequência.

- *Advanced Very High Resolution Radiometer (AVHRR)*: é imageador radiométrico com seis detectores e pode ser usado para determinação de cobertura de nuvens, gelo e neve e é útil para monitoramento da temperatura de superfícies (terra, topo de nuvens ou água).
- *Visible High-Resolution (VHR)*: instrumento que proporciona observação detalhada de desastres, mapeamento de vegetação e ocupação urbana. Por sua alta resolução de imageamento é útil para estudos sobre os componentes de regiões urbanas (telhados, estradas) e de culturas vegetais.

- *Tiros Operational Vertical Sounder* (TOVS): é uma sonda meteorológica dotada de três instrumentos: o HIRS que é uma sonda de radiação infravermelha, o MSU que é uma sonda de micro-ondas e a SSU, uma sonda estratosférica.
- *Data Collection System* (DCS): instrumento usado para monitoramento dos oceanos através do recebimento de informações enviadas por bóias espalhadas ao redor do planeta. É usado também para estudar os caminhos de migração de animais selvagens.
- *Space Environment Monitor* (SEM): é um espectômetro de partículas carregadas que mede a atividade da radiação solar sobre a Terra. É composto por uma unidade de processamento de dados, um sensor de pacotes de partículas de energia, dois sensores magnetômetros e um sensor de raios-X solar.
- *Charge-coupled device* (CCD): é um dispositivo de movimentação de cargas elétricas para áreas onde estas possam ser manipuladas para, por exemplo, sua conversão em ND.

Os produtos gerados pelos instrumentos citados podem assumir duas classificações [26]:

- Produtos imageados: são aqueles produtos que se apresentam como uma imagem da superfície imageada. O resultado gera informações sobre a variação espacial da superfície com relação à sua resposta espectral.
- Produtos não-imageados: intuitivamente, o produto final não é apresentado em forma de uma imagem da superfície. Esses produtos fornecem informação sobre a assinatura espectral do alvo (resposta espectro-radiométrica) ou em termos radiométricos, ou seja, dígitos ou gráficos.

Os sensores imageadores são responsáveis pela formação de imagens em diferentes sistemas [26]:

- sistema de quadros (*frame system*): uma cena é completamente adquirida em um único instante. Um exemplo de um sensor com varredura em quadros é o RBV.
- sistema de varredura mecânica ou *whiskbroom*: sistema de varredura cujos detectores estão arrançados de forma linear. A imagem é formada pelo movimento do satélite ao longo da sua faixa orbital e digitalizado pela trilha cruzada da sua lente óptica.
- varredura eletrônica ou *pushbroom*: a cena formada por esse sensor é formada por detectores em arranjo linear, e a radiância que passa através da lente óptica em linha, ao contrário da varredura mecânica onde a radiância captada passa em secções de cena pela lente.

Uma idealização dos sistemas de varredura pode ser vista na Figura 2.7. Nela observa-se a formação das cenas mecanicamente (Figura 2.7 (a)), eletronicamente (Figura 2.7 (b)) e em quadros (Figura 2.7 (c)). A varredura eletrônica é geralmente mais leve, menos custosa que a varredura mecânica e capta mais luminosidade pois submete a câmera à uma longa exposição em cada cena. Os detectores dos sensores *pushbroom* são muito sensíveis à variação individual, o que se torna uma desvantagem em relação aos outros tipos.

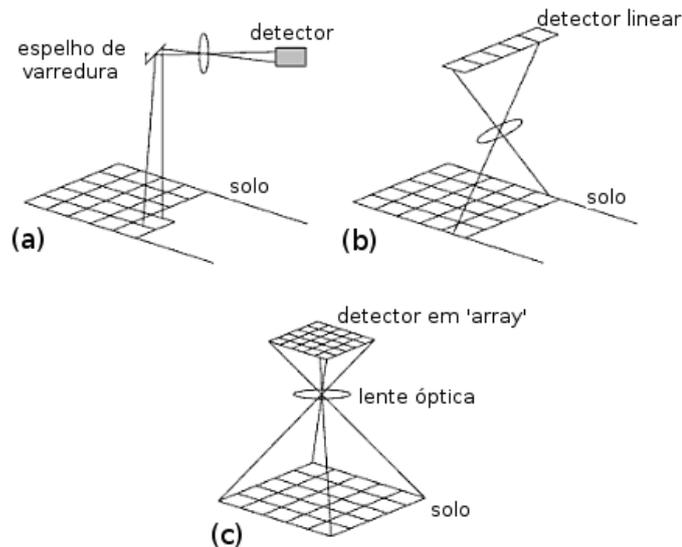


Figura 2.7: Técnicas de imageamento: (a) Varredura linear ou *whiskbroom*, (b) varredura eletrônica ou *pushbroom*, (c) varredura em quadros [26].

Como dito anteriormente, os satélites são plataformas de observação remota e servem a vários propósitos. Nos próximos parágrafos serão elencados alguns dos satélites mais importantes dentro do interesse de sensoriamento remoto, como os satélites da família Landsat e o satélite de alta resolução espacial SPOT.

2.2.7.1 Família Landsat

O programa de observação da terra Landsat é administrado pela Agência Espacial Norte Americana, NASA (sigla em inglês), e pelo Departamento de pesquisa geológica do mesmo país (*U.S. Geological Survey* - USGS) desde 1972. As imagens captadas ao longo desses anos promove estudos abordando vários aspectos e colaborando para a avaliação da dinâmica de mudanças causadas por processos naturais e humanos⁴.

O primeiro produto Landsat foi lançado em Julho de 1972 e se comunicava com estações terrenas à 15 Mbps, quantizado à 6 *bits* (64 níveis), enviando imagens de cenas de 170×185

⁴Fonte: <http://www.landsat.usgs.gov>

quilômetros. O Landsat 1 levou à bordo um sensor RBV, que coletou imagens por apenas 2 semanas, e um sensor MSS. O sensor RBV possuía 3 câmeras que operavam nas bandas 1, do azul ao verde ($475 - 575nm$), 2, do laranja ao vermelho ($580 - 680nm$) e 3, do vermelho até o infravermelho próximo ($690 - 830nm$). Por sua vez, o sensor MSS imageava em quatro bandas espectrais: banda 4 - verde ($0,5 - 0,6\mu m$), banda 5 - vermelho ($0,6 - 0,7\mu m$), banda 6 - infravermelho próximo ($0,7 - 0,8\mu m$) e banda 7 - infravermelho próximo ($0,8 - 1,1\mu m$).

O Landsat 2 foi lançado em 1975, mantendo as mesmas características de comunicação. Transportou os mesmos instrumentos que o seu antecessor, com as mesmas configurações de imageamento.

O terceiro produto da série alterou as configurações dos dois sensores carregados anteriormente. O sensor RBV usou duas câmeras de resposta pancromática de alta resolução espacial de 40 metros para complementar a cobertura do sensor MSS. Ao sensor MSS foi adicionado o imageamento da banda 8, no termal ($10,4 - 12,6\mu m$). Assim como seus antecessores, cada cena imageava regiões de 170×185 quilômetros.

Lançado em 1982, o Landsat 4 aumentou a taxa de transmissão de dados para 85 Mbps e os níveis de quantização para 256 níveis. O sensor RBV foi retirado e inserido um sensor TM que imageava 7 bandas espectrais, sendo 3 dentro do visível ($0,45 - 0,69\mu m$), 2 no infravermelho próximo ($0,76 - 1,75\mu m$), 1 na banda termal ($10,40 - 12,50\mu m$) e 1 no infravermelho médio ($2,08 - 2,35\mu m$). A resolução espacial da banda termal era de 120 metros, enquanto as outras bandas tinham resolução de 30 metros. As bandas imageadas pelo sensor MSS voltaram a ser como os dois primeiros satélites.

Com a colaboração da Companhia de Satélites de Observação da Terra (Earth Observation Satellite Company (EOSAT)) em março de 1984 foi lançado o Landsat 5, ainda em funcionamento, levando à bordo os mesmos instrumentos da sua versão anterior. O sensor MSS foi desligado em 1995, mas seu sensor TM continua mandando imagens para as estações em terra.

A sexta versão do satélite lançada em 93 não atingiu a órbita da Terra e um novo satélite foi enviado em 1999, o Landsat 7. Ao contrário das cinco primeiras versões, as duas últimas levaram à bordo somente um sensor ETM+ que capta imagens em 8 bandas, sendo 7 bandas esquematizadas como sua primeira versão (sensor TM) e uma banda pancromática ($0,52 - 0,90\mu m$). A resolução espacial se manteve para as bandas de 1 à 5 e 7 em 30 metros, a banda 6 termal passou a ter 60 metros de resolução e a banda pancromática foi projetada para obter imagens com resolução de 15 metros.

A linha do tempo da família Landsat pode ser vista na Figura 2.8.

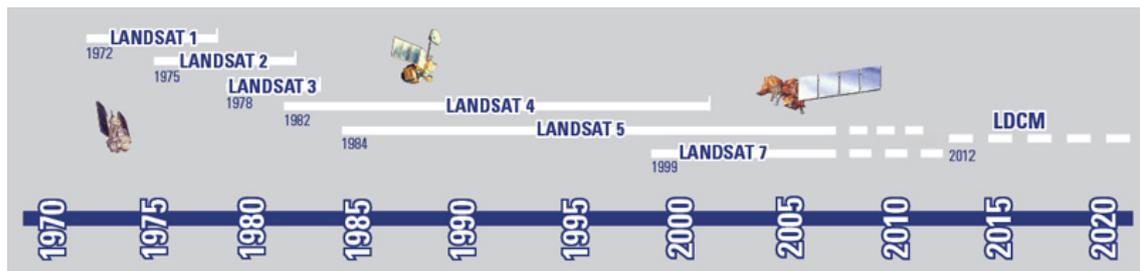


Figura 2.8: Linha do tempo dos satélites Landsat. Fonte: <http://www.landsat.usgs.gov/>

2.2.7.2 *Satellite Pour l'Observation de la Terre - SPOT*

A família de satélites SPOT começou a ser desenvolvida na década de 70 pelo Centro Nacional de Estudos Espaciais da França (Centre national d'études spatiales - CNES) em parceria com o Serviços Científicos, Técnicos e Culturais Belga (Belgian scientific, technical and cultural services - SSTC) e o Conselho Nacional Espacial Sueco (Swedish National Space Board - SNSB).

A fim de investigar os recursos terrestres, em 1986 foi posto em órbita o SPOT 1 que ficou ativo até o ano de 2003. Possuía dois instrumentos tipo *High Resolution Vertical* - HRV idênticos que imageavam nas bandas 1 - verde ($0,50 - 0,59\mu m$), 2 - vermelho ($0,61 - 0,68\mu m$) e 3 - infravermelho próximo ($0,78 - 0,89\mu m$) e na pancromática $0,50 - 0,73\mu m$.

A configuração do satélite SPOT 1 foi a mesma usada nas versões SPOT 2, lançada em 1990, e SPOT 3, lançada em 1993. As resoluções espaciais das bandas multiespectrais eram de 20 metros e a banda pancromática tinha resolução espacial de 10 metros.

O SPOT 4, lançado em 1998 e ainda em operação, está equipado com dois sensores do tipo *High Resolution Vertical Infrared* (HRVIR) idênticos além do sensor *VEGETATION*. Os sensores HRVIR são similares ao HRV, mas passaram a imagear na banda do infravermelho médio. O sensor *VEGETATION* foi projetado para prover medidas precisas sobre as características de cobertura de plantas na Terra. Quase diariamente esse sensor gera imagens com resolução de 1 quilômetro, sendo ideal para observação à longo prazo de mudanças ambientais. Ele imageia em 4 bandas (azul - $0,43 - 0,47\mu m$, vermelho - $0,61 - 0,68\mu m$, infravermelho próximo - $0,79 - 0,89\mu m$ e infravermelho médio - $1,58 - 1,75\mu m$).

O último satélite da família SPOT foi lançado em 2002 e ainda opera na órbita da Terra. O SPOT 5 transporta 3 tipos de sensor: 2 *High Resolution Geometric* - HRG, High-Resolution Stereoscopic - HRS e *VEGETATION 2*. Os dois instrumentos HRG são capazes de gerar dados em 5 bandas espectrais (4 multiespectrais e 1 pancromática) com resoluções que variam de 2,5 à 20 metros. O sensor HRS produz imagens estéreo quase simultaneamente em uma banda pancromática ($0,51 - 0,73\mu m$) com resolução de 10 metros. O sensor *VEGETATION 2* segue

as mesmas configurações do sensor original, em funcionamento no SPOT 4.

2.2.7.3 *QuickBird*

O satélite *QuickBird* [27] foi lançado em 2001 pela empresa *DigitalGlobe* e desde então capta imagens de alta resolução para fins comerciais.

Seus sensores geram imagens em banda pancromática (450 – 900nm) com resolução espacial de 60 centímetros nadiral e imagens multiespectrais de 2,4 metros nos canais azul (450 – 520nm), vermelho (630 – 690nm), verde (520 – 600) e infravermelho próximo (760 – 900nm). As imagens representam áreas de 16 × 16 quilômetros. No ângulo de 25° pós-nadiral a resolução da imagem pancromática é de 72 centímetros e a imagem multiespectral passa a ter resolução de 2,88 metros.

2.2.7.4 *Ikonos*

O satélite *Ikonos* [27] foi lançado em 1999 e foi o primeiro satélite a disponibilizar imagens com resolução de 4 à 1 metros.

Suas imagens pancromáticas tem resolução de 82 centímetros, enquanto suas imagens multiespectrais possuem resolução de 4 metros. No modo *pan-sharpened*, que usa os sensores pancromático e multiespectral associados, consegue imagens multiespectrais com resolução de 1 metro. O sensor pancromático utiliza a faixa de 526 – 929nm para formação de suas imagens e o sensor multiespectral imageia na faixa do azul (445–516nm), verde (506–595nm), vermelho (632 – 698nm) e infravermelho próximo (757 – 853nm).

2.2.7.5 *Chinese-Brazilian Earth Resources - CBERS*

O projeto CBERS⁵ é uma parceria do governo brasileiro, por meio do Instituto Nacional de Pesquisas Espaciais - INPE, e o governo chinês para o lançamento de um satélite de cobertura global de monitoramento ambiental com grande emprego no controle do desmatamento e queimadas na Amazônia Legal (projetos PRODES⁶, DETER⁷ e CANASAT⁸).

Os satélites CBERS-1 e 2 possuem um sensor Wide Field Imager - WFI imageando as bandas 0,63 – 0,69 μ m do vermelho e 0,77 – 0,89 μ m do infravermelho, com resolução

⁵<http://www.cbears.inpe.br/>

⁶<http://www.obt.inpe.br/prodes>

⁷<http://www.obt.inpe.br/deter>

⁸<http://www.dsr.inpe.br/canasat>

espacial de 260 metros; uma câmera CCD de alta resolução, imageando em $0,51 - 0,73\mu m$ (pancromático), $0,45 - 0,52\mu m$ do azul, $0,52 - 0,59\mu m$ do verde, $0,63 - 0,69\mu m$ do vermelho e $0,77 - 0,89\mu m$ no infravermelho próximo, com resolução de 20 metros; e um *scanner* multiespectral de infravermelho (IR-MSS), operando nas faixas $0,50 - 1,10\mu m$ (pancromático), $1,55 - 1,75\mu m$ e $2,08 - 2,35\mu m$ do infravermelho e $10,40 - 12,50\mu m$ do termal, com resolução de 80 metros. Além disso, a série CBERS possui um sistema de controle de altitude de alta precisão que mantém os satélites na altura ideal para o funcionamento pleno dos sensores.

O terceiro satélite da série é chamado de CBERS-2B que é bastante semelhante aos anteriores, onde o sensor IR-MSS foi substituído por uma câmera pancromática de alta resolução ($0,50 - 0,80\mu m$) e foi inserido um sistema de localização global (GPS). Um sistema sensor de estrelas foi adicionado para a assistência aos mecanismos de controle de altitude, que inclui propulsores que também auxiliam na correção orbital.

Os satélites CBERS-3 e 4, previstos para serem lançados em 2011 e 2014, respectivamente, utilizarão 4 câmeras: Câmera Imageadora de Amplo Campo de Visada - WFICAM operando em 4 bandas espectrais do azul ($0,45 - 0,52\mu m$ ao infravermelho próximo ($0,77 - 0,89\mu m$); uma câmera IRMSS como nas versões prévias; duas câmeras PanMux (*Panchromatic and Multispectral Camera*) imageando no pancromático como a câmera CCD, mas com resolução espacial de 5 metros, e no multiespectral como a WFICAM de resolução espacial de 10 metros.

Capítulo 3

Aspectos sobre processamento digital de imagens de satélite

Neste capítulo serão abordados aspectos que concernem o processamento de imagens oriundas de sensores remotos no que tange algumas operações ligadas à tarefa de interpretação de imagens: pré-processamento, segmentação e classificação. Não serão abordadas particularidades de outras áreas da interpretação de imagens, como por exemplo biometria ou biomedicina.

3.1 Pré-processamento

Após a seleção das imagens utilizadas para um estudo, por vezes é preciso manipulá-las para que correspondam à realidade geográfica, ou mesmo para que tenham retirados ruídos provenientes da interação com a atmosfera como a presença indesejável de nuvens e sombras. Nesta seção serão destacados os procedimentos mais comuns no universo das imagens remotas da superfície. Deve-se destacar que as correções nas imagens remotas, citadas nas subseções seguintes, podem ser feitas pela empresa distribuidora que pode incluir o serviço de correção no ato da compra.

Para alguns autores, as distorções ocasionadas por interferência atmosférica são tratadas como um tipo de desvio radiométrico [28, 29]. Porém, dada sua importância na formação de imagens remotas, alguns autores dedicam seções exclusivas a esses efeitos [15, 30].

Como esses desvios radiométricos (causados por efeitos atmosféricos) tratam do espalhamento causado principalmente por presença de partículas suspensas, geralmente em

nuvens e neblina, e isso está diretamente atrelado à tarefa de remoção de nuvens optou-se, neste trabalho, seguir essa última abordagem, dedicando uma subseção (Subseção 3.1.3) somente às distorções atmosféricas.

3.1.1 Distorção geométrica

A correção geométrica visa tratar distorções ocorridas durante a formação da imagem, tais quais aquelas inerentes à plataforma, ao instrumento e ao modelo da Terra [31]. As distorções inerentes à plataforma se referem às variações de altitude (variação de escala) e velocidade do satélite, o que causa sobreposição e lacunas em varreduras sucessivas, nos três eixos (x e y de deslocamento de órbita e z de deslocamento de altitude). Uma ilustração desses efeitos pode ser vista a Figura 3.1, adaptada de [31]. Os diferentes tipos de varredura também causam distorções corrigíveis. A varredura mecânica pode ser acometida, por exemplo, de cisalhamento, quando há movimento simultâneo entre o satélite e o espelho e a distorção panorâmica. Da mesma forma a varredura eletrônica pode ser afetada pelo deslocamento de banda muito próximas ou o desalinhamento das barras detectoras.

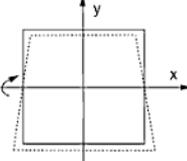
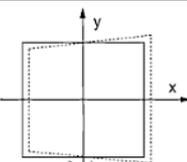
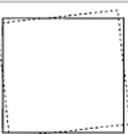
Alteração de varredura	Efeito sobre a imagem	Ilustração
Alteração de altitude	Alteração na escala da imagem	
Alteração de localização	Alteração na área imageada	
Rotação ao longo da direção da órbita no eixo x	Alteração de escala ao longo do eixo y	
Rotação ao longo da direção da órbita no eixo y	alteração de escala ao longo do eixo x	
Rotação ao longo do eixo da altitude	Diferença na área imageada	

Figura 3.1: Exemplificação dos efeitos inerentes a plataforma de sensoriamento [31].

Distorções relacionadas à superfície alvo, aqui entendida como a superfície da Terra, são causadas, principalmente, por dois fatores: rotação e curvatura [31].

A Figura 3.2 mostra o efeito distorcivo que a imagem sofre devido à rotação da Terra. Essa distorção é causada pelo acúmulo de linhas imageadas, o que causa um deslocamento à esquerda cumulativo, da primeira à última linha de cenas adjacentes. Então, uma cena como na Figura 3.2 (a) é formada em quadrado, enquanto que, devido a rotação, a imagem gravada apresenta um deslocamento proporcional (Figura 3.2 (b)).

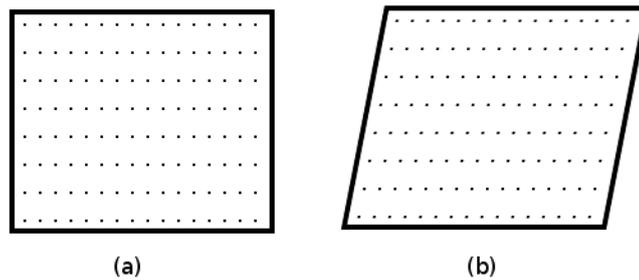


Figura 3.2: Impacto da distorção devido a forma da superfície terrestre [31] (a) Formato da cena ideal. (b) Formato da cena distorcido pela rotação da Terra.

A curvatura da Terra necessita de um modelo tridimensional para sua representação, o imageamento, como é feito atualmente, utiliza de um modelo bi-dimensional de ilustração. Considerando esse fator, a transformação do modelo 3-D em uma imagem 2-D inevitavelmente causa distorções à representação. Essas distorções são dependentes do ângulo de escaneamento (θ) e do alcance da visada (D), ambos relacionados à altitude do sensor H (ilustrados na Figura 3.3).

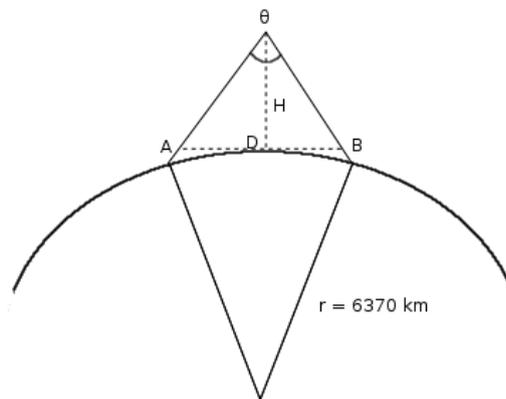


Figura 3.3: Efeito da curvatura da Terra no imageamento de sua superfície [31].

As distorções geométricas são em sua maioria previsíveis, por exemplo a distorção por rotação, e por esse motivo as correções acabam sendo uma operação sistemática, feita pela agência distribuidora ou captadora da imagens. Correções sobre distorções ditas não sistemáticas, como, por exemplo, a alteração na velocidade e/ou altitude da plataforma, devem ser analisadas caso a caso, por meio de algoritmos especialistas.

Durante a captação da imagem são armazenadas, numa espécie de cabeçalho (*header*), todas as informações do sensor no momento da captação. Com base nessas informações, é possível efetuar as correções descritas até o momento.

Outra operação comum de correção de distorção não sistemática é a transformada de coordenadas com base em pontos de controle de superfície (GCP, sigla em inglês) [28, 31]. Um mapa geo-referenciado ou um dispositivo de posicionamento global - GPS - geralmente é usado para guiar essa transformada. A coordenada relativa à posição da cena imageada no globo é um exemplo de informação contida no cabeçalho da imagem.

3.1.2 Distorção radiométrica

Distorções radiométricas são causadas principalmente pela variação de iluminação causada nos diferentes ângulos de iluminação solar¹, por variação indesejada na resposta dos detectores individuais (sensores de bandas específicas) ou mesmo pela perda de linhas de imageamento ou *pixel*.

O ideal é que a medida da intensidade seja proporcional à reflectância do objeto e à irradiância espectral do Sol, assumindo que a reflectância de uma superfície seja uniforme na área de um *pixel* [28].

As distorções não lineares são causadas por comportamento não esperado dos sensores e se refere à resposta não linear da função de entrada de radiância. Relaciona-se também aos ruídos inseridos pelo próprio sensor ou às diferenças em câmeras de multidetecção, onde podem ocorrer pequenas diferenças na função de transferência ou de *offset* [28].²

A correção desses efeitos comumente é feita pelo ajuste de níveis de cinza da imagem de acordo com informações de média e variância tomadas como referência.

¹Considere que uma cena é feita pela concatenação de “sub-cenas”, sendo que uma cena completa é um mosaico de “sub-cenas”.

²*Offset* refere-se à saída de cada detector na ausência de radiação.

3.1.3 Correção atmosférica

As distorções induzidas pela interferência atmosférica são causadas pelo retorno da radiação solar ocasionado pelo encontro dos feixes com partículas suspensas (feixe que deveria atingir o alvo imageado) ou o espalhamento causado por radiação refletida por vizinhos do *pixel* de interesse, ambos causando perdas de detalhes da imagem e diminuindo a precisão da representação dos níveis radiométricos ao longo das bandas. Essas distorções podem ser modeladas em forma de atenuação atmosférica através de informações detalhadas do momento da formação da imagem. Na falta dessas informações detalhadas, são mais comuns as correções em massa, que visam compensar a radiância do sensor ou do caminho (interferência de um *pixel* espalhado sobre o *pixel* alvo).

Um método de correção bastante conhecido é o de remoção de objetos escuros (DOS, sigla em inglês) [15, 28, 29]. O histograma de uma imagem, como mostrado na Figura 3.4, ilustra a representação de *offset*, valor constante, que nesse método é usado para a subtração dos *pixels* da imagem em uma determinada banda espectral. Esse método é baseado na suposição de que os *pixels* com valor de reflectância zero são oriundos de espalhamento, portanto, representam distorção. No espectro visível, áreas de sombra podem representar objetos escuros e pelo seu histograma oferecer o valor de *offset*.

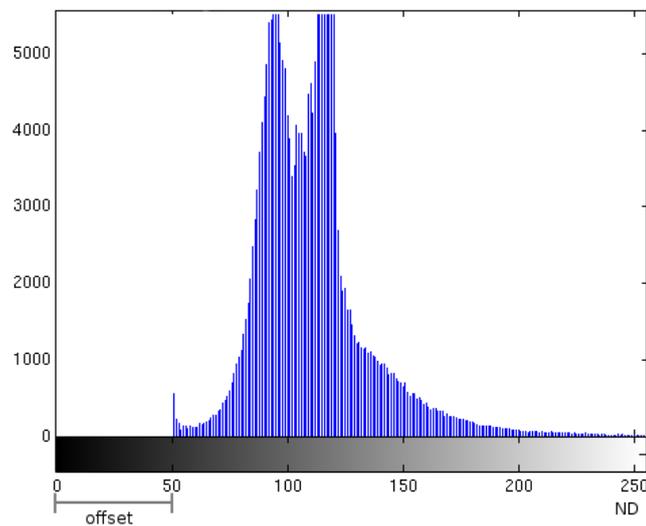


Figura 3.4: Estimação da radiância de uma imagem pelo seu histograma [15].

A modelagem de efeitos atmosféricos supre a deficiência da DOS, uma vez que o método aplica a subtração de uma constante por todos os *pixels* da imagem, promovendo uma correção aproximada e não precisa. Porém, em [32] foram estudados seis modelos de calibração, onde um modelo adaptado baseado em DOS, conhecido como *DOS3* [33] foi superior aos modelos *DOS - 1, 2 e 4*, *Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes* (FLAASH) e

Segunda Simulação de Sinal de Satélite no Espectro Solar (6S).

3.1.4 Remoção de nuvens

Como citado anteriormente, o espalhamento espectral pode ser decorrente de moléculas suspensas no ar, também chamado de espalhamento *Rayleigh*, que é uma função de quarta potência inversa ao comprimento de onda e é responsável pela cor azul no céu [29]. Isso se dá pois a luz azul e ultravioleta espalham-se de quatro a dezesseis vezes mais que a luz vermelha, respectivamente.

Quando o espalhamento é causado por partículas maiores, geralmente associadas com fumaças, neblina e vapores em geral, é chamado de aerosol ou *Mie* e são da ordem de um décimo na potência de dez do comprimento de onda. O aerosol é responsável pela cor branca das nuvens.

Uma vez que as partículas suspensas tornam-se muito maiores que o comprimento de onda (causado por nevoeiros, nuvens e poeira), a dependência que existia até então entre a quantidade de espalhamento e o comprimento de onda desaparece.³ Esse tipo é chamado de espalhamento não-seletivo e acontece nas camadas mais baixas da atmosfera.

As nuvens são descritas como um conjunto de partículas de água líquida ou gelo visíveis a olho nu em suspensão na atmosfera.⁴ Seu aspecto depende da natureza, dimensão, número e distribuição no espaço das partículas que a constituem, além da intensidade e cor da luz incidente, da posição do observador e do elemento irradiante em relação à nuvem.

Imagens de sensoriamento remoto são, constantemente, corrompidas por nuvens, o que, dependendo da porcentagem de cobertura da cena, acarretará na exclusão de uma imagem de uma pesquisa. De forma a evitar a exclusão de cenas, o que pode diminuir a abrangência de um estudo de monitoramento, por exemplo, pensou-se na possibilidade de filtragem dessas áreas corrompidas, a fim de “expor” porções de superfícies encobertas.

Uma imagem com resolução espectral limitada dificulta a remoção de nuvens por métodos multiespectrais, e quanto mais baixa é a resolução temporal, mais importante é o processo de remoção de nuvens [34]. Como os *pixels* que compõem uma nuvem possuem alto brilho e os que compõem sombra, o contrário, métodos de identificação de bordas baseados nessas características tornam-se importantes durante a remoção. No trabalho apresentado por [34], após a identificação de áreas de nuvem e sombra, uma combinação de duas imagens da mesma cena (uma com e outra sem nuvens) foi feita, criando uma imagem com os melhores

³Quanto menor o comprimento de onda maior o espalhamento causado.

⁴<http://www.inmet.gov.br>

pixels entre as duas cenas.

Em [35] foi usado um método conhecido como *Inpainting* que faz o preenchimento das nuvens através de informações coletadas de fora da nuvem. O método *Inpainting* é usado também para outros fins a parte do sensoriamento remoto, como em [5] que usou o método para o preenchimento de grandes áreas de imagens comuns com eficiência.

A técnica apresentada em [35] faz o preenchimento da região contaminada por nuvem com dados geométricos extraídos de regiões de fora da nuvem, depois faz o agrupamento dos *pixels* através de uma transformada. Esse método mostrou-se mais eficiente que outras técnicas de reconstrução, mas os autores atentam para aspectos que possam aprimorar os resultados: a geração de mapas de confiança para basear a reconstrução em dados estatísticos (e não só geométricos) e promover uma reconstrução em imagens monocanal para posteriormente fazer a reconstrução multicanal.

Imagens acometidas por nuvens acontecem quando o feixe de luz não consegue penetrar nas partículas suspensas na atmosfera, cujo tamanho é maior que o comprimento de onda da luz, fato que não ocorre com imagens adquiridas por micro-ondas. Combinar imagens obtidas por micro-ondas e por instrumentos ópticos foi a proposta de [36], onde a detecção de nuvens foi feita pela medição do Índice Total de Radiância Refletida (TRRI, sigla em inglês) e do Índice Nuvem-Sol (CSI, sigla em inglês), estendendo a região do entorno próximo da nuvem para a região reconhecidamente de nuvem, pois assume-se que a informação dessa região sofre muita mistura e não representa a realidade. A partir desse ponto é feita uma combinação entre as imagens ópticas e de micro-ondas, resultando em imagens com aspecto visual convincente.

A combinação de imagens de diferentes resoluções espaciais é comum [37], assim como a combinação de imagens de cenas colhidas em datas diferentes pelo mesmo instrumento [34]. A solução apontada por [37] mostrou que simples detectores e removedores de nuvens podem ter resultados satisfatórios, com alta taxa de precisão.

Muitos métodos utilizam transformadas para agrupamento de *pixels* ou para tornar os dados mais resistentes a ruídos, realçando as componentes da imagem. Trabalhos como o de [38] utilizam transformada rápida de *Fourier* (FFT) no desenvolvimento de seu algoritmo para remoção de nuvens. O uso de FFT geralmente está atrelado aos chamados filtros homomórficos [39, 40], e é usada para converter a imagem do domínio espacial para o domínio da frequência, onde esta passa a ser uma série de ondas senoidais bidimensionais de frequências variadas. Nesse estágio, é possível reduzir ruídos e remover características periódicas que podem estar associadas a estas, além de permitir a extração de informação de realce [38]. A aplicabilidade dos filtros homomórficos se baseia no fato de que imagens são constituídas principalmente da quantidade de luz vinda da fonte incidente na cena e da quantidade de luz

refletida pelos objetos da cena, respectivamente, as componentes de iluminação e reflexão, que se combinam multiplicativamente.

Após a aplicação da FFT, as componentes de iluminação e reflexão passam a ser identificadas por componentes de baixa e alta frequências, sendo que ao passarem por um filtro passa-altas (FPA), as componentes de iluminação serão suprimidas, pois geralmente estão ligadas à incidência de nuvens, como observou [38].



Figura 3.5: Esquema de um filtro homomórfico. Adaptado de [38].

A Figura 3.5 ilustra os passos para a filtragem homomórfica aplicada a tarefa de remoção de nuvens. A transformação logarítmica natural executada na imagem original ($f(x, y)$) é feita para que as componentes multiplicativas de reflectância e iluminação possam ser aproximadamente separáveis. Uma vez que a transformada de *Fourier* do produto de duas funções é não separável, ou seja,

$$\mathcal{F}(f(x, y)) \neq \mathcal{F}(l(x, y))\mathcal{F}(r(x, y)),$$

onde r e l são as componentes de reflexão e iluminação, a operação logarítmica faz com que a imagem original tome a forma:

$$\begin{aligned} f(x, y) &= l(x, y) * r(x, y) \\ \log(f(x, y)) &= \log(l(x, y) * r(x, y)) \\ \log(f(x, y)) &= \log(l(x, y)) + \log(r(x, y)). \end{aligned} \quad (3.1)$$

onde ao final, após o retorno da imagem para o domínio espacial, é feita a operação inversa ao logaritmo e, por fim, a imagem resultante ($g(x, y)$) é encontrada.

A estrutura e sequência da filtragem homomórfica apresentada é bem estabelecida, ficando a cargo da filtragem passa-altas a variabilidade do método. Em [38] foram testados cinco tipos de FPA: *Ideal*, *Bartlett*, *Butterworth*, *Gaussiano* e *Hanning*. Apesar dos resultados obtidos com FPA *Ideal* e FPA *Bartlett* terem sido melhores, a incidência de ruído e confusão nessas imagens foi grande. Diante deste fato, os autores observaram que as imagens filtradas pelo FPA do tipo *Butterworth* obtiveram as melhores relações entre a retirada de nuvens e a inserção de ruído na imagem.

3.2 Segmentação

A tarefa de segmentar uma imagem consiste na divisão da cena em regiões ou objetos que a constituem, unindo regiões adjacentes como características próximas de homogeneidade [6, 41]. Por representar parte crucial de um sistema de classificação de imagens, essa tarefa exige o máximo de precisão, o que a torna digna de atenção redobrada durante seu desenvolvimento.

As implementações de segmentadores geralmente baseiam-se em informação de homogeneidade, ou seja, procuram através de medidas definidas encontrar características próximas e, por isso, aglutináveis em regiões da imagem. Essas informações de homogeneidade (ou similaridade), associadas a dados sobre a descontinuidade dos *pixels*, vêm servindo de base para muitos algoritmos [42].

Como a heterogeneidade de aplicações desses algoritmos é grande, podendo ser aplicado em identificação biométrica, propósitos biomédicos, militares e ambientais, não há uma técnica ou abordagem que sirva a todas as aplicações com bons resultados. Assim como não há consenso sobre a técnica, não há também, consenso sobre como quantificar o quão boa é uma técnica em comparação à outra [1]. Ainda segundo esse trabalho, os autores geralmente apresentam evidências da eficácia de seus algoritmos, expondo algumas imagens e suas respectivas segmentações, o que torna difícil a comparação entre métodos, pois há fatores ligados ao desenvolvimento que não são expostos ou comentados (complexidade do algoritmo e custo de execução, por exemplo). Existem técnicas de avaliação de segmentação [43, 44, 45], mas muitas vezes essas propostas não estão bem desenvolvidas ou compreendidas, ou mesmo atendem somente a propósitos específicos e por isso não são muito aceitas por outros autores.

Imagens de satélite são conhecidas pela heterogeneidade de seus dados, que em uma mesma cena podem apresentar regiões de floresta densa, região devastada por queimada ou desflorestamento, regiões de ocupação urbana recente e antiga, por exemplo. Essa heterogeneidade permite ao desenvolvedor trabalhar com as informações presentes na textura de cada padrão apresentado. Os algoritmos que se baseiam nessas informações de textura e os que utilizam a técnica de crescimento de região são conhecidos como segmentadores baseados em região [46, 47]. Em suma, métodos baseados em regiões assumem que *pixels* vizinhos possuem similaridades e que por isso, representam uma área homogênea e limitada.

Algoritmos baseados em bordas (*edge-based*) [48], assumem que as características de intensidade, cor e textura podem mudar entre regiões diferentes. É uma das aplicações mais comuns [6] e podem abordar três tipos de transição entre *pixels*: ideal, rampa e *roof edge*. As transições estão exemplificadas na Figura 3.6.

A transição tida como ideal ocorre com distância de 1 *pixel*, sem ocorrência de suavização e são usados em modelos de detecção de bordas (Canny e Sobel, por exemplo). Transições em rampa são mais reais a medida que se percebeu que as bordas são borradas e ruidosas devido a limitações de foco em lentes de imagens ópticas e em componentes eletrônicos de sistemas imageadores. Transições tipo *roof edge* são uma descontinuidade apresentada na derivada de primeira ordem do encontro de superfícies com orientações distintas, que delimitam diferentes faces de um objeto ou de objetos distintos na mesma intensidade [49].

Esses dois métodos tidos como tradicionais, baseados em textura e em crescimento de região, tem seus desempenhos melhorados com a inserção de conhecimento *a priori* fornecido por um especialista, de forma não automática, normalmente. Essas informações podem vir também de outras técnicas, ditas não morfológicas que fornecem informações espectrais como em [50], cujo segmentador baseado em região incorporou propriedades texturais à informações espectrais como NDVI e desvio padrão para segmentar, e posteriormente classificar, imagens dos satélites Landsat e Aster. O algoritmo obteve boa desambiguação de classes.

A Matriz de Co-ocorrência de Níveis de Cinza (GLCM, em inglês), usada em [51], é uma técnica de extração de características de textura baseada nos níveis de cinza de uma imagem. Seu conceito assume que a informação de textura é definida pela relação entre *pixels* adjacentes, ou seja, o método denota uma matriz formada pela frequência de tons separados por uma distância determinada e em direções específicas. Matrizes GLCM geralmente são esparsas e grandes, e fornecem subsídio para classificação de forma satisfatória.

Outro método de segmentação bastante difundido e bem mais simples e geral que os apresentados até o momento, é o método baseado em limites (*thresholding*). Uma forma simples de interpretação desse método é utilizando histogramas [6]. Supondo que a Figura 3.7 represente o histograma de uma imagem onde há dois grupos predominantes de níveis de cinza. Uma forma simples de extração de objetos (grupos) seria estabelecer um valor para T , gerando “pontos de fundo” para todos os níveis de cinza menores que T e um “ponto objeto” para outros valores.

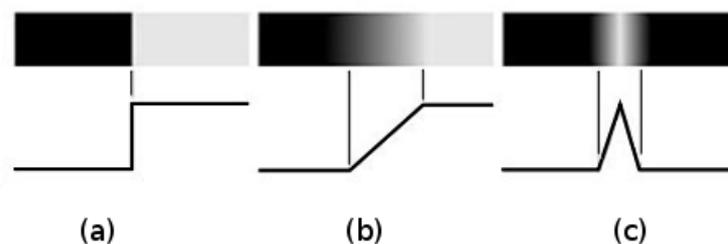


Figura 3.6: Transições entre pixels de borda: (a)Ideal, (b)Rampa e (c)*Roof edge* [6]

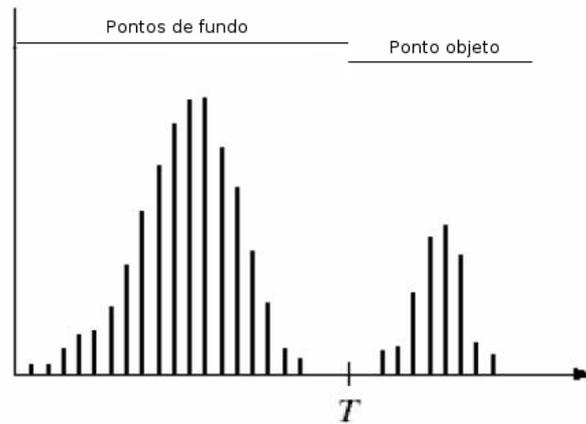


Figura 3.7: Histograma separado por um único limite [6].

Já na Figura 3.8 a separação dos grupos de *pixels* formaria três conjuntos, dos *pixels* menores que o valor $T1$ (fundo da imagem ou objeto escuro), dos *pixels* cujos valores estejam entre $T1$ e $T2$ (primeiro objeto claro) e dos valores maiores que $T2$ (segundo objeto claro). Dessa forma, é fácil perceber que os valores de T que separam grupos de *pixels* dependem diretamente das características dessas concentrações e que, portanto, caracterizam um processo dinâmico e adaptativo.

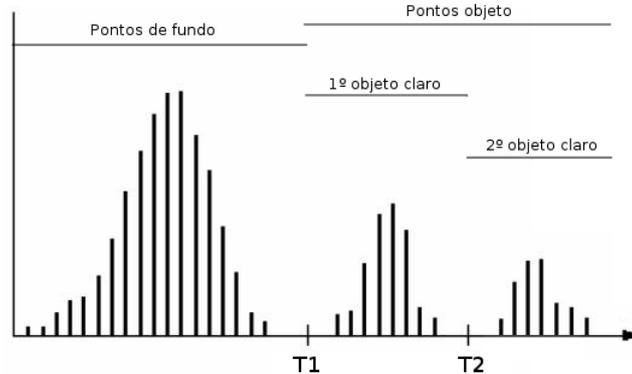


Figura 3.8: Histograma separado por múltiplos limites. Adaptado de [6].

O exemplo mostrado nas Figuras 3.7 e 3.8 são modelos cuja identificação dos limites é algo instintivo e muito simples, porém a dificuldade em encontrar esses limites cresce à medida em que o número de separações cresce, tornando impraticável a utilização de um método tão simples em grande parte dos casos.

Em [52], foi desenvolvido um segmentador hierárquico baseado em histogramas que utiliza o conceito de limites em várias escalas. Trabalhando em dois espaços de cores distintos (RGB e CIE(L*a*b)) e usando conceitos de homogeneidade, conseguiu, com base em picos, vales e limites, estabelecer bons resultados para segmentação de imagens artificiais.

Uma visão geral de 40 métodos de limiarização, divididos em categorias de acordo com a informação explorada é apresentada em [53]. Entre eles estão: o desenho do histograma, medidas de agrupamento, entropia, atributos de objetos, correlação espacial e superfície de níveis de cinza. A comparação dos métodos levou em consideração o desempenho não destrutivo das aplicações e a combinação de medidas de performance sobre imagens de ultrassom, documentos, oriundas de microscópios, entre outras.

A proposta de usar limiarização de histogramas proposta em [54], usou imagens multibandas do satélite Landsat. Primeiramente, foram selecionadas as classes de uso da terra para treinamento. Na sequência, procedeu-se uma seleção automática de limites e por fim, a segmentação foi feita com base nos limites encontrados. Os resultados foram satisfatórios para as classes urbana, solo exposto, floresta e água.

3.3 Classificação

O reconhecimento de padrões tem o propósito de classificar padrões em categorias referenciados como classes ou etiquetas [55]. Pode ser baseado em modelos estatísticos, induzidos por informações sobre as classes buscadas para a classificação de todos os dados, informações comumente obtidas por interferência de especialistas (dados de treino). A performance desses modelos está diretamente ligada à precisão dos parâmetros de estimação usados durante a apresentação dos dados de treino do modelo.

No contexto de imagens de sensores remotos, o objetivo do reconhecimento de padrões é unir cada objeto ou *pixel* a um ou mais elementos de um conjunto de classes definidas pelo usuário, convertendo a informação radiométrica em informação temática (floresta, água, solo) [15].

Um objeto ou *pixel* pode ser associado a uma ou a várias classes. Para o segundo caso, é estabelecida uma relação baseada em probabilidades entre o objeto em questão e as classes em que ele se encaixa. Além disso, é possível proceder uma associação do tipo muitos para uma, ou seja vários *pixels* são classificados em uma classe, o que é comum quando se quer aglutinar vários *pixels* para representar uma região de floresta, por exemplo.

Os métodos costumam ser divididos em duas categorias, métodos supervisionados e não supervisionados.

A classificação supervisionada pressupõe o uso de exemplos sobre as classes que irão ser identificadas, são os dados de treino. Desse modo, a caracterização estatística dos objetos a serem classificados será feita com base nesses dados de treino, de forma a estabelecer a

assinatura dessas classes buscadas. Em [56], foi usada a teoria da incerteza em dois modelos (baseado na teoria das probabilidades ou *bayesiano* e baseado na teoria das evidências) para compor um classificador supervisionado do tipo Maxver (máxima verosimilhança). Esse tipo de classificador foi comparado à abordagem *fuzzy* em [57] para a classificação de imagens do satélite Landsat (ETM+). Os métodos *bayesiano* e os baseados na abordagem *fuzzy* são mais indicadas para regiões que apresentam comportamento similar (comportamento inerentemente *fuzzy*), difíceis de separar.

Em [58] é utilizada a técnica de vizinhos próximos (NN, da sigla em inglês) para classificação, esta que é bastante usada devido a reconhecida relação entre os *pixels*. É proposto um procedimento para encontrar uma métrica adaptativa que combinasse a variância local e a incorporação linear do espaço de observação em um espaço euclidiano. Foram usadas imagens multi e monotemporal do satélite Landsat para comparação entre essa métrica adaptativa e algoritmos Maxver e k-NN, obtendo maior precisão que estes.

De forma a evitar a separação de conjuntos de treino, os algoritmos classificadores não supervisionados, fazem uma busca individual pela formação de grupos ou *clusters* com características similares. É exigido do programador certo conhecimento sobre as características dos materiais componentes da cena, mas não como na classificação supervisionada, com a inserção de dados por especialista, e sim para a determinação de parâmetros de separação, inclusive o número de *clusters* a serem identificados.

As propriedades espectrais de uma classe podem variar com a mudança de estações climáticas, o que torna a relação entre a classe de um objeto e a característica espectral do *cluster* inconstante. Muitas vezes a relação entre um objeto e sua classe também está restrita à uma imagem captada em um determinado momento, não podendo ser estendido à outra imagem colhida em momento diverso. Pensando nisso, [59] usou algoritmo genético para determinar automaticamente o número de *clusters* durante a classificação não supervisionada de imagens do satélite IKONOS, que mostraram bons resultados quando comparados à resultados convencionais de ISODATA,⁵.

Agrupar *pixels* no domínio espacial não é a única alternativa para os classificadores, esta também pode ser feita no domínio espectral. A proposta apresentada em [60] usa um esquema *fuzzy* de agrupamento de *pixels* multi-objetiva combinada com redes neurais artificiais com base probabilística. A abordagem multi-objetiva produz um conjunto de soluções, onde as mais confiáveis são usadas para treinar a rede neural. As soluções pouco confiáveis são então classificadas pela rede neural, que apresentou boa precisão em comparação com outros

⁵Isodata é o acrônimo para *Iterative Self-Organizing Data Analysis Techniques Algorithm* o mais sofisticado algoritmo para determinação e ajuste automático de *clusters*.

classificadores conhecidos.

Explorando o esquema difuso de agrupamento multiobjetivo, [61] propôs o uso de *Support Vector Machine* (SVM) para aumentar a precisão do classificador. Assim como em [60], a agrupamento difusa fornece subsídio (pontos de maior confiabilidade) para o classificador, nesse caso baseado em SVM. O treinamento é feito com os pontos mais confiáveis e a classificação é feita com os pontos de menor confiabilidade.

Capítulo 4

Algoritmo proposto

Este capítulo será dedicado à descrição detalhada dos métodos que compõem o algoritmo de classificação proposto neste trabalho. O algoritmo está dividido em três grandes tarefas: remoção de nuvens, segmentação e classificação. O fluxograma apresentado na Figura 4.1 mostra os passos do algoritmo que serão descritos no decorrer desse capítulo.

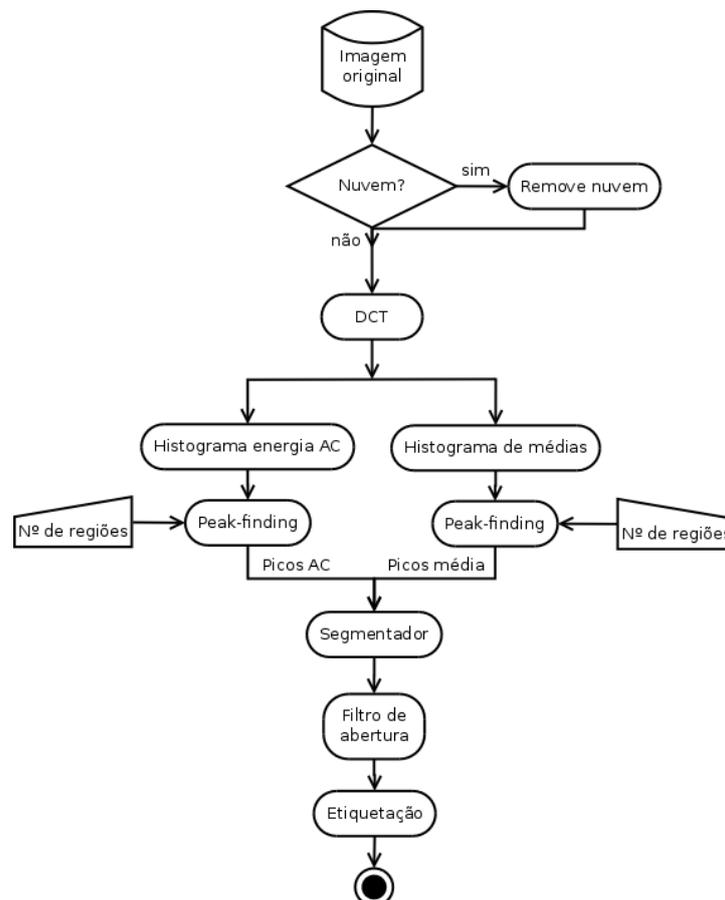


Figura 4.1: Diagrama de fluxo de dados do algoritmo proposto.

A tarefa de remoção de nuvens, que como mostrado no capítulo anterior, torna-se mais importante à medida que a disponibilidade de imagens sem cobertura de nuvens de áreas tropicais diminui, será apresentada na Seção 4.1. A segmentação baseada nas componentes obtidas pela transformada DCT será apresentada na Seção 4.2 e a tarefa de classificação será mostrada na Seção 4.3.

4.1 Remoção de nuvens

A remoção de nuvens torna-se vital quando não há disponibilidade de imagens de uma mesma cena em determinadas épocas cuja cobertura de nuvens seja pequena, além disso, ainda que haja disponibilidade de acervo, em regiões tropicais, a incidência de nuvens ao longo do ano pode impedir o monitoramento de certas áreas, cuja precipitação seja intensa.

De forma a contornar o problema de cobertura de nuvens, alguns trabalhos usam imagens de radar ou micro-ondas em seus estudos [62, 63, 64, 40]. Sistemas de radar podem gerar imagens de alta resolução, independente das condições de iluminação solar ou meteorológicas. Essas imagens exigem a implementação de métodos complexos, uma vez que empregam radiação coerente e as imagens são corrompidas por uma forma multiplicativa de ruído salpicado ou *speckle* [62]. Além disso, [64] expõe que o fato de que áreas com feições de queimada possuem comportamento espectral semelhante ao da floresta e a existência do efeito *double bounce*, que ocorre com o acúmulo hídrico no solo dias antes do imageamento e com a incidência de vestígios de biomassa no solo, limitam o uso de imagens de radar.

Aliando as limitações do uso de imagens de radar ao propósito deste trabalho, de usar somente imagens de satélite de média e alta resolução, optou-se por desenvolver um método de remoção de nuvens para tratar as contaminações que ocorrerem. A primeira etapa do fluxograma da Figura 4.1 será descrita nos próximos parágrafos, sendo dividido em duas fases: (1) identificação de regiões de nuvens e sombras e (2) filtragem de remoção das áreas afetadas.

4.1.1 Detecção de nuvens e sombras

A existência de nuvens em imagens de sensoriamento remoto, está geralmente ligada à presença de sombras, causadas por estas. O ângulo solar no momento da captura da cena e a captura da cena com angulação do tipo *off-nadir* são duas das mais comuns causas de formação de sombras nesses produtos. Com base em [38], que apresenta um método para separação de regiões, baseando-se em estatísticas para identificar áreas de nuvens densas e suaves, propõe-se a adição da capacidade de discernir regiões de sombra em uma imagem.

A separação de regiões é feita segundo a expressão:

$$f(x, y) = \begin{cases} i(x, y) < i_{m-dp}, & i(x, y) \in 0; \\ i_{m-dp} \leq i(x, y) < i_m, & i(x, y) \in 1; \\ i_m \leq i(x, y) < i_{m+dp}, & i(x, y) \in 2; \\ i(x, y) \geq i_{m+dp}, & i(x, y) \in 3. \end{cases} \quad (4.1)$$

onde $i(x, y)$ é igual ao valor do pixel da imagem com ruído (nuvem ou sombra), i_m representa o valor médio da imagem com ruído, $i_{m+dp} = i_m + \sigma_i$ e $i_{m-dp} = i_m - \sigma_i$. Dessa forma, as regiões definidas como 0 são as regiões de sombra da imagem, as definidas como 1 são as livres de nuvens e sombras, já as definidas como 2 são rotuladas como sendo nuvens esparsas, nevoeiros ou neblina, enquanto que as definidas como 3 são detectadas como nuvens densas.

4.1.2 Filtro Homomórfico

A natureza básica de uma imagem, representada pela função bidimensional $f(x, y)$, pode ser caracterizada por duas componentes: (1) a quantidade de luz na cena que está sendo vista, e (2) a quantidade de luz refletida pelos objetos da cena. Estas componentes são chamadas de *iluminação* e *reflectância*, sendo indicadas por $l(x, y)$ e $r(x, y)$, respectivamente. As funções l e r se combinam multiplicativamente para produzir a função f

$$f(x, y) = l(x, y) \cdot r(x, y) \quad (4.2)$$

onde $0 < l(x, y) < \infty$ e $0 < r(x, y) < 1$.

Devido a características intrínsecas de comportamento do objeto e a luz que incide e reflete sobre ele, pode-se adotar que a componente de iluminação varia lentamente no espaço enquanto a reflectância tem variações abruptas neste domínio, e seguindo uma relação entre espaço e frequência [65], tem-se que a componente $r(x, y)$ está concentrada nas altas frequências e a componente $l(x, y)$ localiza-se nas baixas frequências.

Nota-se que as áreas compostas por nuvens esparsas, nevoeiros ou neblinas (etiquetadas como sendo do tipo 2) possuem grande contribuição da componente de iluminação (localizada nas baixas frequências por serem mais homogêneas), dessa forma, para que se reduza a aparência de nuvem em uma cena é preciso que reduza-se essa contribuição. Para isso, utilizam-se os filtros homomórficos [40, 65, 66].

Pela natureza multiplicativa da Equação 4.2, não é possível filtrar a função $f(x, y)$ no domínio da frequência de maneira trivial, que aconteceria se $f(x, y)$ fosse aditivo. Para

contornar esse problema, é aplicado o logaritmo natural (base e) em ambos os lados da Equação 4.2:

$$z(x, y) = \ln [f(x, y)] = \ln [l(x, y)] + \ln [r(x, y)] \quad (4.3)$$

Utilizando a transformada de *Fourier* na Equação 4.3, então

$$\begin{aligned} \mathcal{F}\{z(x, y)\} &= \mathcal{F}\{\ln [f(x, y)]\} = \mathcal{F}\{\ln [l(x, y)]\} + \mathcal{F}\{\ln [r(x, y)]\} \\ Z(u, v) &= L(u, v) + R(u, v) \end{aligned} \quad (4.4)$$

onde $Z(u, v)$, $L(u, v)$ e $R(u, v)$ são as transformadas de *Fourier* de z , $\ln l$ e $\ln r$, respectivamente.

A função $Z(u, v)$ representa então, a soma de uma imagem de baixa frequência (iluminação) e uma imagem de alta frequência (reflectância). Com base nessa informação, se for aplicado um filtro com função de transferência que suprima as componentes de baixa frequência e realce as componentes de alta frequência, ou seja, um filtro passa-altas (FPA), pode-se reduzir a aparência de nuvens suaves nas áreas atingidas. Assim, a função de transferência $H(u, v)$, exemplificada na Figura 4.2, aplicada à função $Z(u, v)$ assume a forma:

$$S(u, v) = H(u, v) \cdot Z(u, v) = H(u, v) \cdot L(u, v) + H(u, v) \cdot R(u, v). \quad (4.5)$$

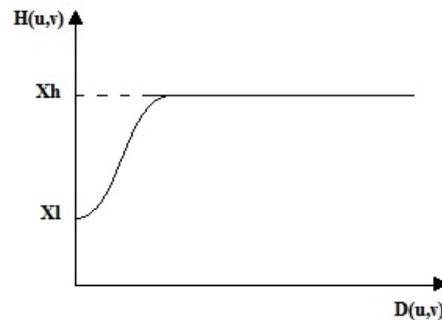


Figura 4.2: Função de transferência $H(u, v)$ de um filtro homomórfico esfericamente simétrico, onde $D(u, v)$ é a distância da origem [66].

Tomando-se a transformada inversa de *Fourier* da Equação 4.5:

$$s(x, y) = \mathcal{F}^{-1}\{S(u, v)\} = \mathcal{F}^{-1}\{H(u, v) \cdot L(u, v)\} + \mathcal{F}^{-1}\{H(u, v) \cdot R(u, v)\} \quad (4.6)$$

Finalmente, como $z(x, y)$ foi obtida usando o logaritmo da imagem original $f(x, y)$, a operação exponencial produz a imagem desejada.

$$g(x, y) = e^{s(x, y)}$$

Assemelhando-se aos testes executados por [38], vários tipos de FPA foram testados nas imagens contaminadas por nuvens esparsas disponíveis. Dentre todos, e principalmente entre os mais populares, os filtros *Butterworth* e *High-boost*, concluiu-se que o FPA *High-boost* mostra-se mais indicado para tratamento das imagens disponíveis. Ao contrário do que fazem muitos filtros passa-alta (FPA), o filtro *High-boost* não elimina totalmente as componentes de baixa frequência, permitindo que estas auxiliem na interpretação da imagem.

Nesse tipo de filtragem, uma pequena parcela de componentes de baixa frequência é mantida pela aplicação de um coeficiente multiplicativo (A) sobre a imagem original. Assim, o filtro FPA que originalmente excluiria todas as componentes até o corte, é definido por

$$FPA = Img_original - FPB$$

passa por uma adaptação, adquirindo a forma

$$\begin{aligned} High - boost &= (A)(Img_original) - FPB \\ &= (A - 1)(Img_original) + Img_original - FPB \\ &= (A - 1)(Img_original) + FPA \end{aligned} \tag{4.7}$$

onde que para $A = 1$, tem-se um FPA comum, enquanto que quando $A > 1$, informações uma parte da imagem original é mantida na saída [67]. O filtro passa-altas *Butterworth* é usado na implementação do filtro *High-boost* usado neste trabalho.

4.1.3 *Inpainting*

As áreas rotuladas com o valor 3 e 0 durante a separação de regiões (Equação 4.1), representam nuvens mais densas e sombras, os quais sofrem muito pouco, ou mesmo nenhum efeito da filtragem homomórfica. A alternativa apresentada para essas áreas é chamada de *Inpainting* [68, 35, 69].

Esse método tem o objetivo de preencher lacunas, regiões indefinidas ou deterioradas de uma imagem, de forma a restaurar ou tornar mais visível a mesma. O algoritmo *Inpainting* usado por [35] é aplicado somente sobre áreas de nuvens densas, porém a presença de nuvens está geralmente associada à existência de sombras, por isso, neste trabalho, a ocorrência de sombras também é prevista e tratada.

O uso de interpolação pode estar associado ao melhoramento de imagens, convertendo uma imagem em outra sem perda de informação visual [70]. Interpolação não-adaptativa ou linear é aplicada para recuperar componentes ausentes e são mais simples que métodos

adaptativos. Interpolação por vizinho próximo, bilinear e bicúbica são alguns métodos de interpolação, que dependendo de sua complexidade, pode envolver centenas de *pixels* adjacentes no processo.

Neste trabalho, o algoritmo *Inpainting* utilizará interpolação por vizinho mais próximo, onde o *pixel* sofre um efeito alargador pela atribuição de seu valor ao *pixel* indefinido. O *pixel* indefinido chamado de x possui um conjunto de amostras de *pixels* vizinhos f_k , formado pelos x_k pontos próximos. A decisão sobre qual valor será atribuído é tomado com base na distância Euclidiana entre x e x_k ($dE_k = ||x - x_k||$). A menor distância dE_k indicará o valor de *pixel* próximo x_k a ser atribuído ao *pixel* indefinido x [71].

4.1.4 Realce não linear no domínio da frequência

A imagem resultante da filtragem homomórfica e da interpolação, descritas nas Subseções 4.1.2 e 4.1.3, apresenta reduzida informação de brilho e para contornar tal situação sugere-se o uso de um método de realce.

O algoritmo de realce empregado, adota a possibilidade de representação de uma imagem como sendo a soma de componentes de baixas e altas frequências, assim como interpretado na Subseção 4.1.2. A obtenção das componentes de frequência é feita pela Representação Piramidal Multiresolução (RPM), que utiliza a Pirâmide Gaussiana como um filtro passa-baixas e a Pirâmide Laplaciana como um filtro passa-altas [72].

A RPM é uma ferramenta para manipulação de informação de escala que visa obter características dos objetos em diferentes escalas, formando pirâmides de imagens. São usadas para aplicações de visão computacional e compressão de imagens, formando um conjunto de imagens com resolução decrescente, onde a base da pirâmide possui a imagem de mais alta resolução enquanto a ponta contém a de mais baixa [73]. Uma representação piramidal pode ser vista na Figura 4.3.¹

A Pirâmide Gaussiana é formada por vários estágios, onde cada nível corresponde à suavização por um núcleo Gaussiano simétrico e subamostrado. A Pirâmide Laplaciana consiste em um filtro passa-banda, onde cada camada da pirâmide é formada pela subtração dos dois níveis adjacentes da Pirâmide Gaussiana (Equação 4.8)

$$La_n = G_n - G_{n+1} \quad (4.8)$$

onde La_n é a camada n da Pirâmide Laplaciana, sendo $n = 0 \dots (N - 1)$ e N o número de

¹Fonte: http://star.pst.qub.ac.uk/idl/Image_Tiling.html

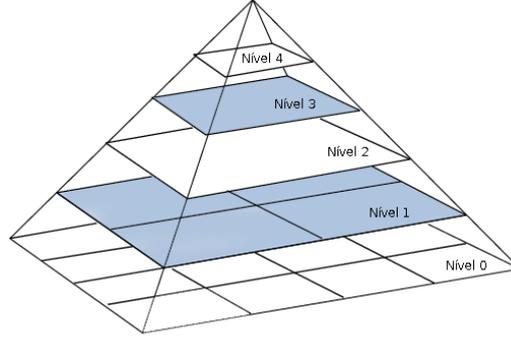


Figura 4.3: Exemplo de RPM com $N = 4$, onde cada nível é subamostrado por um fator de 2. [74]

camadas da pirâmide e G_n e G_{n+1} são camadas da Pirâmide Gaussiana. Com a imagem do nível G_0 , base de uma Pirâmide Gaussiana, pode-se gerar as componentes La_0 e G_1 de altas e baixas frequências, respectivamente.

O realce utilizando pirâmides visa encontrar a mais alta harmônica para o sinal de entrada (imagem) mantendo a informação de fase e a forma do sinal, como descrito por [72]. O método busca prever uma imagem G_{-1} mantendo o formato de G_0 , pela estipulação da componente de alta frequência La_{-1} , a qual terá a mesma forma e será coerente em fase com a componente de alta frequência La_0 . A predição de La_{-1} é feita através de um operador não linear que delimita (ou recorta) a resposta La_0 , multiplicando a forma de onda resultante por um fator de amplificação e remove as baixas frequências presentes, a fim de extrair uma resposta de alta frequência. A equação a seguir formaliza a geração de La_{-1} :

$$La_{-1} = FPA(s \times (BOUND(La_0))) \quad (4.9)$$

onde s é o fator de amplificação e $BOUND(\cdot)$ é definido pela expressão:

$$BOUND(x) = \begin{cases} T, & x > T \\ x, & -T \leq x \leq T \\ -T, & x < -T \end{cases} \quad (4.10)$$

e $T = (1 - c)La_{0_{max}}$. A constante de corte c varia entre 0 e 1 e $La_{0_{max}}$ denota o maior valor de La_0 . O resultado deste cálculo é passado por um filtro passa-alta, de modo a manter apenas as componentes de alta frequência da imagem. Então, a imagem realçada, G_{-1} , é gerada pela equação 4.11:

$$G_{-1} = La_{-1} + G_0 \quad (4.11)$$

A imagem realçada G_{-1} , possui dimensões maiores pois é resultado da soma entre a imagem sobreamostrada da base da pirâmide Laplaciana (L_{-1}) mais a imagem localizada na base da pirâmide Gaussiana G_0 . Por esse motivo, esta imagem é manualmente redimensionada para seu valor original.

O Diagrama 4.4 mostra a sequência de aplicação das técnicas expostas nessa Seção. Neste, o processo chamado *Fusão* corresponde à ação de união de duas imagens para a formação de uma imagem melhor, do ponto de vista da remoção de nuvens. A primeira fusão irá agregar as regiões filtradas pelo algoritmo de *Inpainting* (regiões 3 e 0) com a imagem original, enquanto que a segunda fusão será a união da imagem resultante da primeira fusão realçada e a imagem resultante da filtragem homomórfica (região 1) realçada, produzindo assim, a imagem que será usada no restante do algoritmo.

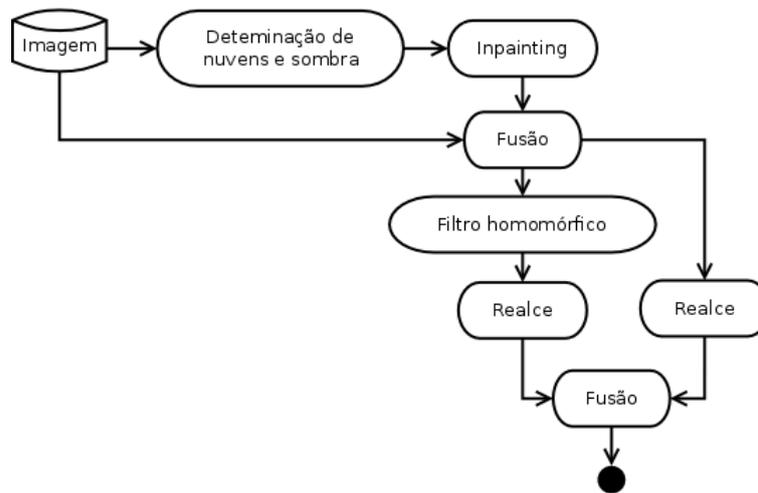


Figura 4.4: Diagrama do algoritmo de remoção de nuvens

4.2 Segmentação

A região destacada na Figura 4.5 mostra as ações a serem realizadas para a segmentação da imagem desejada.

O termo segmentar uma imagem significa agrupar *pixels* com características similares em regiões de homogeneidade. Nesse trabalho, serão utilizadas técnicas de limiarização de histogramas com busca por picos significativos (Subseção 4.2.2), Transformada Cosseno Discreta (DCT, Subseção 4.2.1) e filtros morfológicos (Subseção 4.2.3).

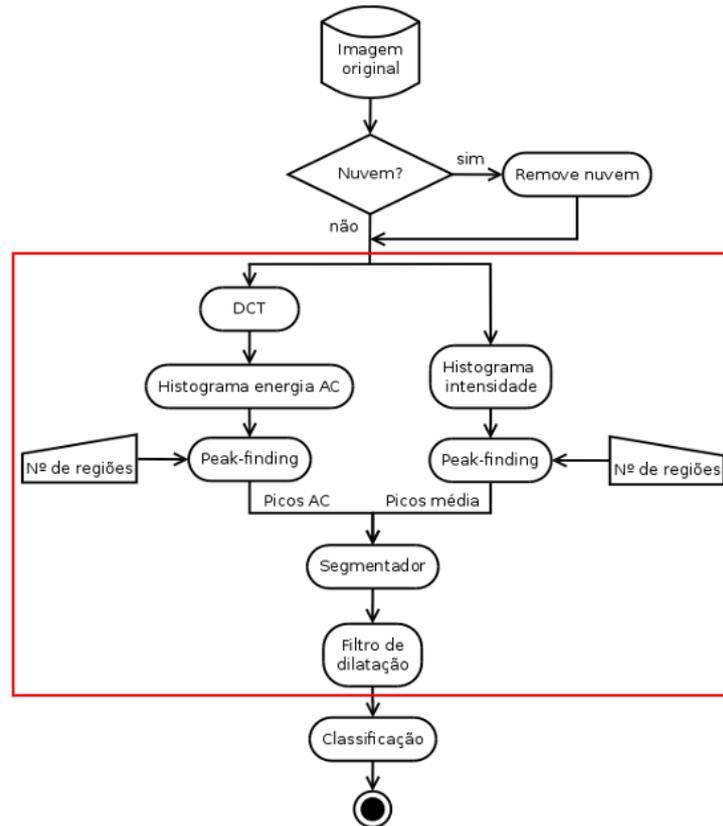


Figura 4.5: Destaque das ações de segmentação do algoritmo proposto.

4.2.1 Transformada Cosseno Discreta

A Transformada de Cosseno Discreta (DCT) é uma transformada unitária, real e separável, cuja operação é similar à Transformada de *Fourier* Discreta, que converte um sinal em forma de soma de cossenos em diferentes frequências usando números reais com simetria [75, 76].

A DCT unidimensional (1D-DCT), $X[k]$, de uma sequência $x[n]$ de N números é definida pela relação

$$X[k] = \alpha[k] \sum_{n=0}^{N-1} \left[x[n] \cos \left(\frac{\pi(2n+1)k}{2N} \right) \right], \quad (4.12)$$

para $0 \leq k \leq N$, onde $\alpha[k]$ é definido como

$$\alpha[k] = \begin{cases} \sqrt{\frac{1}{N}} & \text{para } k = 0; \\ \sqrt{\frac{2}{N}} & \text{para } k \neq 0; \end{cases}$$

Dessa forma, cada coeficiente $X[k]$ é expresso pelo produto interno da sequência $x[n]$ e um

vetor base ortogonal e normalizado.

O processamento de sinais bidimensionais, como as imagens, pode então, ser feito por uma transformada igualmente bidimensional, nesse caso a 2D-DCT, dada por

$$X[k_1, k_2] = \frac{2}{\sqrt{N_1 N_2}} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x[n_1, n_2] \cos \frac{\pi}{N_1} k_1 \left(n_1 + \frac{1}{2} \right) \cos \frac{\pi}{N_2} k_2 \left(n_2 + \frac{1}{2} \right), \quad (4.13)$$

onde N_1 e N_2 são as dimensões da matriz processada e as funções base são matrizes de dimensão $N_1 \times N_2$, formadas pelo produto externo de dois vetores base.

Como a DCT é uma operação separável, a operação descrita na Equação 4.13 pode ser substituída pela aplicação individual de 1D-DCT em linhas e colunas. Quando usada, a 2D-DCT pode ser aplicada separadamente em blocos de dimensão $n \times n$. Quando aplicada de forma individual, o esquema em blocos pode ser mantido, considerando que cada bloco é um vetor (linha ou coluna) de dimensão n [76].

Observando uma matriz transformada $X[k_1, k_2]$, o elemento $X[0, 0]$ corresponde ao coeficiente DC ou coeficiente de frequência zero e representa o valor médio dos *pixels* em um bloco. Os elementos restantes são chamados de coeficientes AC.

A energia média de uma sequência de dimensão $N_1 \times N_2$ é dada por

$$E = \frac{1}{N_1 N_2} \sum_{n_1, n_2 \neq 0} X^2[n_1, n_2]. \quad (4.14)$$

As componentes de energia AC representam a atividade do bloco no domínio espacial, ou seja, o quão rápido ou lento é a variação dos *pixels* em uma sequência. Essa energia pode ser calculada pela Equação 4.15

$$E_{AC} = \frac{1}{N_1 N_2} \sum_{n_1, n_2 \neq 0} X^2[n_1, n_2] - E_{DC} \quad (4.15)$$

onde $E_{DC} = \frac{X^2(0,0)}{N_1 N_2}$.

Em [75], a DCT é usada para a segmentação de imagens pela extração de características de textura. Processando a imagem em blocos de tamanho inicial de 4×4 , calculou-se a distância entre blocos vizinhos usando os valores de mais alta energia de cada bloco a fim de unir blocos com distâncias pequenas, até alcançar o número de regiões pré-determinadas. Os resultados apresentaram performance superior a métodos baseados em Transformada *Wavelet* Discreta e na técnica de corte normalizado, inclusive no aspecto da eficiência.

Atendendo a propósitos diversos, a DCT foi usada em [77] para tratar imagens de tomografia computadorizada de alta resolução. Com o não conhecimento sobre o número de

classes dessas imagens (a aparência dos padrões depende da condição patológica do objeto), foi desenvolvido um algoritmo não supervisionado de extração de características de textura baseado na variação dos coeficientes da transformada. A adaptabilidade do algoritmo foi feita ao longo das iterações pela técnica *k-means* associada ao discriminante linear de classes *Fisher*. O método mostrou-se adequado para tratar imagens que possuem 3 classes, contribuindo para o diagnóstico médico.

Em [78], é apresentado um método baseado em descritores visuais para segmentação de imagens. Os descritores DCT de situação são obtidos para cada ponto p da imagem, que estava localizado no centro do bloco da DCT, para formar um longo vetor que será quantizado para refletir o modelo discriminatório de detalhes frequenciais dos humanos. A redução dos dados oriunda da quantização e o uso do algoritmo de adaptação baseado em *k-means* demonstrou boa performance nos resultados.

Justificado pelo fato que as componentes AC representam o comportamento dos *pixels* em um bloco e sua usabilidade para segmentação de imagens, este trabalho usa a 2D-DCT para transformar blocos onde $n = 4$. A matriz resultante, excluída a componente DC, serviu como fonte de informação para a realização de uma análise de histograma nos moldes realizados por [52]. Nesse trabalho a análise de histograma é feita com base em informações de homogeneidade da imagem. Essa análise faz uso do algoritmo de localização de picos (descrito na Subseção 4.2.2) para a seccionar o histograma em regiões de homogeneidade.

4.2.2 Análise de Histograma

Neste trabalho, são gerados dois histogramas: um baseado nos coeficientes AC (Subseção 4.2.1) e um baseado na informação de média da imagem transformada.

Uma vez que o histograma de uma imagem é a representação da frequência de ocorrência de níveis de cinza, podemos então, com base nessa informação, encontrar regiões de homogeneidade que nos permitam seccionar uma figura em modos. Para tanto, foi estabelecido que uma região é caracterizada pelos *pixels* que estão ao redor de um pico, onde um pico representa um *pixel* cuja frequência é mais alta que seus vizinhos. Essas regiões são delimitadas por vales, ou um valor de *pixel* cuja frequência é muita baixa ou nula. Em imagens que apresentam duas regiões de intensidade distinguíveis tem seu histograma chamado de bimodal (Figura 4.6(a)), enquanto que imagens como várias regiões com distribuição de intensidade distintas tem seus histogramas chamados de multimodais (Figura 4.6(b)).

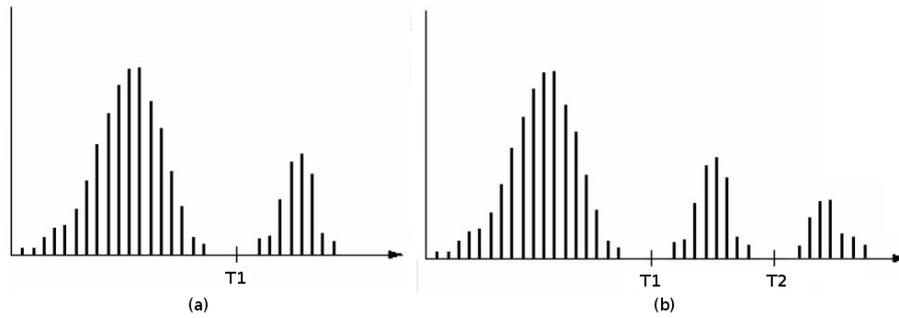


Figura 4.6: Histogramas (a) bimodal (b) multimodal. Adaptado de [6].

4.2.2.1 Algoritmo de localização de picos

A busca de picos em um histograma geralmente não é um trabalho trivial, principalmente quando há bastante heterogeneidade na imagem, como em imagens tomadas de regiões urbanas. A definição dos modos será adaptada do algoritmo de [52] e seguirá três etapas:

1. Busca indiscriminada por picos;
2. Seleção dos picos mais importantes;
3. Determinação de vales entre picos.

Na primeira etapa, são encontrados os máximos locais do histograma, excluindo assim, picos tomados como insignificantes. Para ser considerado insignificante, o máximo local da intensidade c deve atender à seguinte relação

$$\frac{pcs[c]}{pcs_{max}} < 0.05 \quad (4.16)$$

onde $pcs[\cdot]$ é um vetor de intensidades, $1 \leq c \leq 256$ e pcs_{max} é o valor máximo entre os picos (se o vetor é normalizado $pcs_{max} = 1$). São excluídos da busca por picos significativos o primeiro e o último picos que atenderem à relação 4.16.

O segundo passo para determinação de picos importantes é a remoção de picos muito próximos. Aqueles picos que estiverem localizados a menos de 15 posições de distância serão comparados e o menor será excluído. O pico escolhido será escrito em um outro vetor que acomodará somente os valores de picos definitivos (pks).

A seleção de picos importantes continua com a remoção daqueles cujos vales que os separam não são profundos o suficiente. Será considerado um vale insuficiente, aquele cujo valor médio é muito alto em relação aos picos que ele separa. A suficiência é medida pelo

limiar estabelecido em $ts = 0,75$, assim, o valor médio das intensidades entre os picos

$$pks_{av} = \frac{\sum_{p_j=p_1}^{p_2} pks[p_j]}{p_2 - p_1 + 1}$$

onde p_1 e p_2 são posições dos picos consecutivos no vetor pks .

O valor médio encontrado (pks_{av}) será usado na relação 4.17, que se for verdadeira, ou seja, se a porção esquerda for maior que ts , os picos p_1 e p_2 não são suficientemente separáveis, portanto o menor deles é removido e o maior é escrito no vetor pks .

$$\frac{pks_{av}}{(pks(p_1) - pks(p_2))/2} > 0,75 \quad (4.17)$$

Após o segundo passo, todos os picos estarão definidos e restará a busca pelos vales que os separam. Os vales serão os valores cuja frequência seja a menor entre os picos escritos em pks . Esses valores, então, são escritos em um vetor de vales chamado vls .

4.2.3 Filtros morfológicos

Operações morfológicas estão ligadas à extração de componentes úteis na representação e desenho da imagem, como limites e contornos, usadas no pré ou pós-processamento [12].

O filtro de dilatação funciona como um alargador de objetos, controlado por um elemento estruturador. Esses elementos são representados como matrizes binárias (mostrada abaixo) que serão responsáveis por preencher lacunas ou expandir uma classe.

$$\begin{array}{ccc} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{array}$$

O filtro de erosão é usado para afastar grupos de *pixels* que são menores que a máscara, afinando os objetos. Também é controlado por um elemento estruturador.

A dilatação da imagem I pelo elemento E é denotada por $I \oplus E = \{z | (\hat{E})_z \cap I \neq \emptyset\}$, enquanto que a erosão da imagem I pelo elemento E é dada por $I \ominus E = \{z | (\hat{E})_z \cap I^c \neq \emptyset\}$ onde \hat{E} é a reflexão do elemento estruturador e $(\hat{E})_z$ é a translação do elemento pelo ponto $z = (z_1, z_2)$

A combinação dos filtros de erosão e dilatação pode se dar de várias formas. As mais comuns são abertura, fechamento e transformação *hit-or-miss* [12].

A abertura morfológica de I por E ($I \circ E$) é dada pela erosão de I por E , seguida de uma dilatação do resultado pelo mesmo elemento estruturador. Essa operação causa uma

suavização nos contornos e elimina pequenos grupos de *pixels*. O fechamento ($I \cdot E$) é dado pela sequência de dilatação e erosão da imagem I pelo elemento E , causando a suavização dos contornos da imagem, assim como a abertura.

A transformação *hit-or-miss* ($I \otimes E$) é usada para identificar configurações específicas de *pixels*, como aqueles isolados em uma região. Nessa transformação, o elemento E é um par $E = (E_1, E_2)$, dessa forma, sua definição é dada por $I \otimes E = (I \ominus E_1) \cap (I^c \ominus E_2)$.

De forma geral, o pós-processamento visa excluir pequenos e isolados pontos que reduzem a precisão da segmentação e que poderia também contribuir para a queda de precisão na classificação. Dessa forma, foi implementado um filtro transformador morfológico de abertura com o elemento

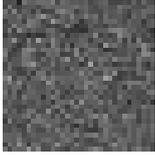
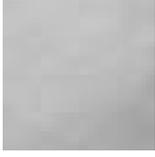
$$E = \begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{matrix}$$

4.3 Classificação

Haja vista que a imagem foi segmentada e suas regiões de homogeneidade foram definidas, resta a atribuição em classes dessas regiões. Essa atribuição de etiquetas é feita mediante o conhecimento prévio sobre o comportamento frequencial das regiões presentes nas imagens. Dessa forma, a classificação apresentada neste trabalho é do tipo supervisionada.

É sabido que regiões urbanas possuem alta heterogeneidade que reflete diretamente no seu comportamento na frequência. A depender da região estudada, a ocorrência de corpos de água apresenta uma aparência de certa heterogeneidade. Isso se deve às variações na profundidade dessas regiões, às variações na qualidade da água, como mais ou menos poluída, mais ou menos turva, entre outras tantas. Na Tabela 4.1 são apresentados alguns exemplos do comportamento frequencial e espacial das classes.

Tabela 4.1: Valores de energia AC para diferentes regiões de uma imagem de satélite.

Classe	Visão espacial (32 × 32)	Energia AC dos coeficientes da DCT em blocos (4 × 4)
Água		<pre> energiaAC = 0.5625 1.5625 1.8398 1.0000 1.6875 0.8711 7.8750 2.9336 0.5000 0.5586 1.4648 2.5586 0.5586 3.2500 9.6211 3.3398 1.3750 0.6836 0.9844 3.9375 0.7148 2.8086 1.4961 2.7461 1.0586 0.8086 0.6094 1.2773 1.1211 6.3125 4.1523 4.5625 0.4336 1.3594 2.6836 1.3086 3.1250 1.5586 2.8711 4.9844 0.4648 1.3398 8.4844 0.6094 7.3086 1.6836 3.8711 5.2500 1.0898 1.1211 2.9844 7.8711 1.9648 0.7148 1.4023 1.7461 0.6523 3.2344 6.8594 1.4336 0.8594 1.2500 0.9023 6.8125 Média = 2.5479 </pre>
Floresta		<pre> energiaAC = 8.6094 10.9023 15.1094 16.5586 30.6250 9.8086 20.2461 43.8398 7.0898 63.2148 34.6094 40.0273 21.5586 29.9375 23.1875 85.5273 18.5625 21.6836 28.2773 18.7344 4.5898 17.3125 48.3398 27.5000 11.3750 40.0000 46.6094 19.0273 16.6836 41.2461 31.0273 13.8125 30.5625 114.0586 31.8398 23.6094 22.0586 47.9648 42.9961 25.1211 26.7773 25.2344 10.3398 18.6836 23.7148 14.7773 48.7344 24.7344 12.6250 12.7461 18.2461 43.3594 41.6523 28.5273 22.6523 53.7344 60.6523 45.8594 43.4375 32.5273 20.9336 193.1211 22.7344 49.1523 Média = 32.7938 </pre>
Área urbana		<pre> energiaAC = 381.1094 143.9375 360.3086 195.1250 151.0273 104.7461 217.4844 280.3125 416.5273 681.6875 418.1875 308.9023 464.2773 295.4023 320.7773 181.1094 447.5898 584.1094 263.5273 452.8086 258.9844 406.2773 240.5625 652.8398 128.8594 483.8594 207.1836 82.5586 546.4844 201.0273 102.2148 254.6094 430.2344 82.9844 382.0586 142.8594 235.1523 557.9023 257.0273 371.2148 174.2500 113.3594 225.3711 349.3750 198.8125 231.1250 221.1094 226.2461 97.4961 366.9648 187.8750 220.1523 441.3125 180.2773 178.8750 224.9336 148.0000 322.5898 202.6836 263.2148 240.1250 322.3125 140.9336 172.4961 Média = 283.4960 </pre>
Nuvens		<pre> energiaAC = 2.4336 1.2148 2.9844 2.1523 2.0000 1.9023 5.0000 2.3594 2.2773 3.5898 0.6875 0.2500 2.7773 4.3594 3.0625 6.6836 6.8125 2.9375 2.9336 1.1094 4.3086 1.2500 0.8594 13.4961 0.6523 0.8750 1.8711 1.1836 0.8086 3.5273 1.9844 10.4648 3.0625 2.0898 1.6250 1.9648 0.7148 3.0586 2.0273 25.1836 5.3086 2.8711 0.8594 4.1875 5.6523 4.3594 10.2344 3.2344 4.4648 11.5586 11.1836 3.1094 3.6094 3.1211 1.9336 3.1211 3.5273 8.3086 1.3398 3.1094 3.9961 2.1523 6.6523 5.2461 Média = 3.9324 </pre>

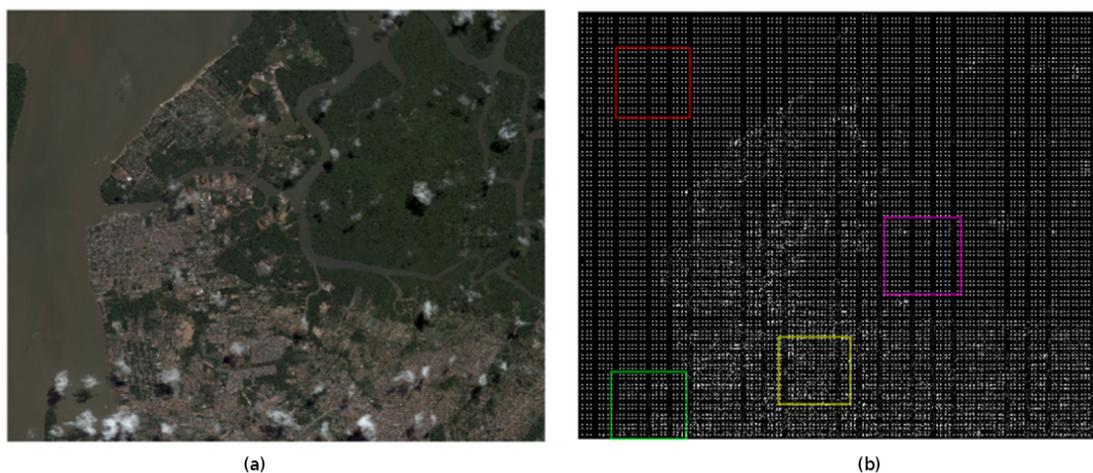


Figura 4.7: Exemplo de imagem usada nos testes: (a) Imagem Original; (b) DCT da imagem original com regiões em destaque.

As amostras frequenciais da Tabela 4.1 foram retiradas na Figura 4.7(b), que é a resposta à transformada DCT da Figura 4.7(a), uma das imagens usadas no estudo. A transformada DCT dessa imagem foi feita com blocos de dimensão 4×4 . A Figura 4.8(b) representa os valores de energia dos blocos sob a forma de uma imagem RGB. Como explicado nas seções anteriores, esses valores estão diretamente ligados à segmentação e classificação das imagens, bem como os valores da média dos blocos mostrados na Figura 4.8(b). Nessa imagem, é possível constatar a justificativa do uso adicional de outras componentes para segmentação e classificação das imagens, além da média.

Em destaque na Figura 4.9, está a região inferior esquerda das imagens, em que é apresentado um exemplo de confusão entre água e área urbana sob a ótica das médias, enquanto que as regiões são completamente distinguíveis sob a ótica das energias.

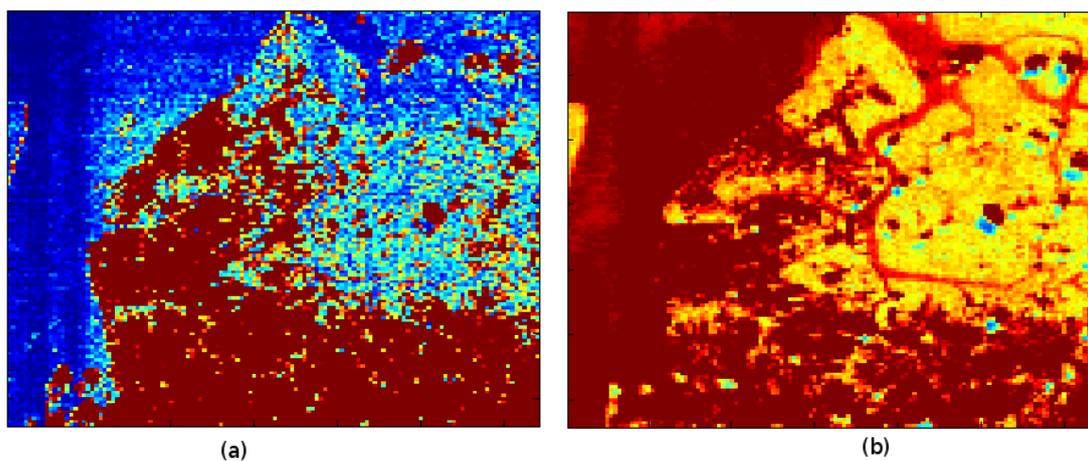


Figura 4.8: Valores de (a) Energia e (b) Média dos blocos da DCT (coeficientes DC)

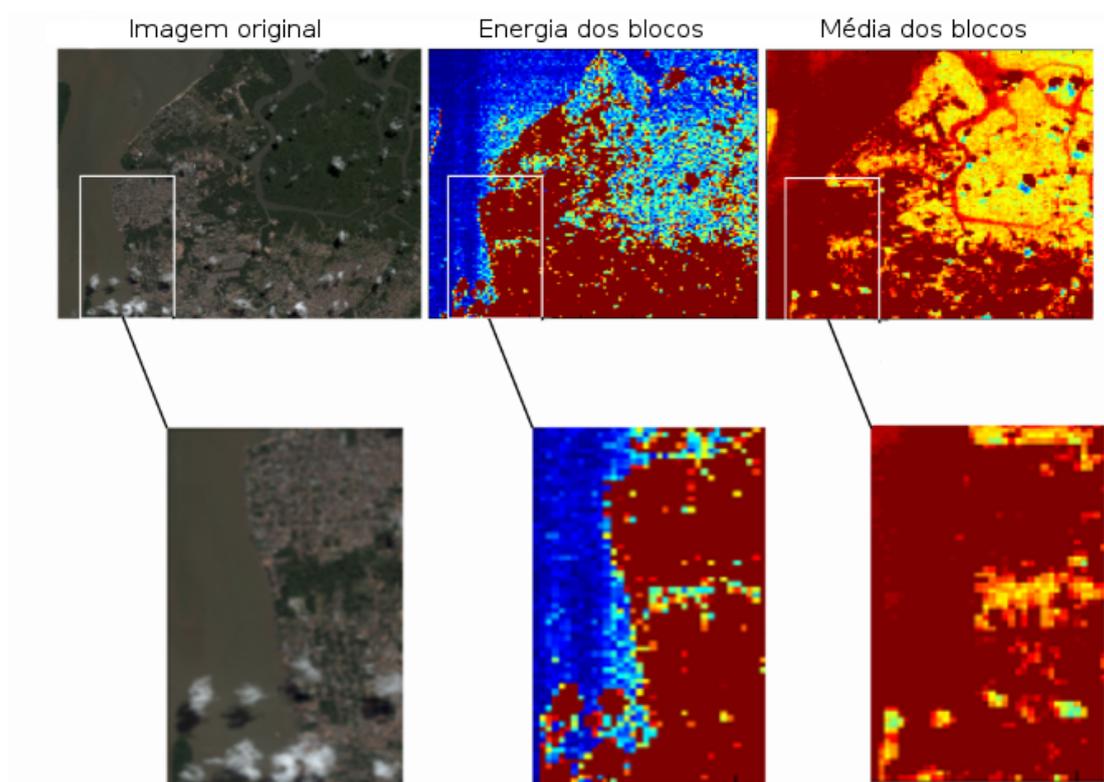


Figura 4.9: Detalhe de uma região de confusão na imagem.

Capítulo 5

Testes e resultados

Neste capítulo serão mostrados os testes executados sobre o banco de imagens disponíveis. Os experimentos serão apresentados passo a passo, de acordo com a ordem de execução do algoritmo.

O banco de imagens testadas é composto de 14 imagens captadas por diversos sensores e com os mais diversos tipos de obstrução sobre elas. Em algumas há a presença somente de nuvens suaves, enquanto outras possuem sombras, nuvens densas e suaves. A cobertura de nuvens é uma ocorrência naturalmente aleatória, portanto a porcentagem de sua ocorrência não tem qualquer influência sobre o algoritmo. Também estão presentes imagens sem obstrução por nuvens.

5.1 Remoção de nuvens

Para exemplificar o processo de remoção de nuvens é apresentada uma das imagens utilizadas nos testes. A Figura 5.1 (a) apresenta uma cena captada da região de Roma, na Itália. A imagem captada pelo satélite *WORLDVIEW-2* em 2009, é obstruída por nuvens suaves e densas, além de sombras, que se sobrepõem às regiões urbanas e de cobertura vegetal.

A Figura 5.1 (b) ilustra a separação da imagem de teste em regiões. Conforme descrito na Subseção 4.1.1, a imagem é dividida em quatro regiões que representam as obstruções por nuvens densas e suaves, sombras e o restante da imagem sem distinção. Segundo a sequência do algoritmo, regiões de nuvem densa e sombra¹ são tratadas pelo mesmo algoritmo de interpolação. Na Figura 5.1(b) essas regiões são destacadas pelas cores preto e vermelho, respectivamente. A região interpretada como contendo nuvens suaves é representada pelos

¹Segundo o algoritmo de detecção, regiões 0 e 3.

pixels verdes na imagem e etiquetada como região do tipo 2, na Subseção 4.1.1. As regiões rosa são regiões etiquetadas como sendo do tipo 1, ou seja, não correspondem a uma obstrução ou não foram identificadas como tal.

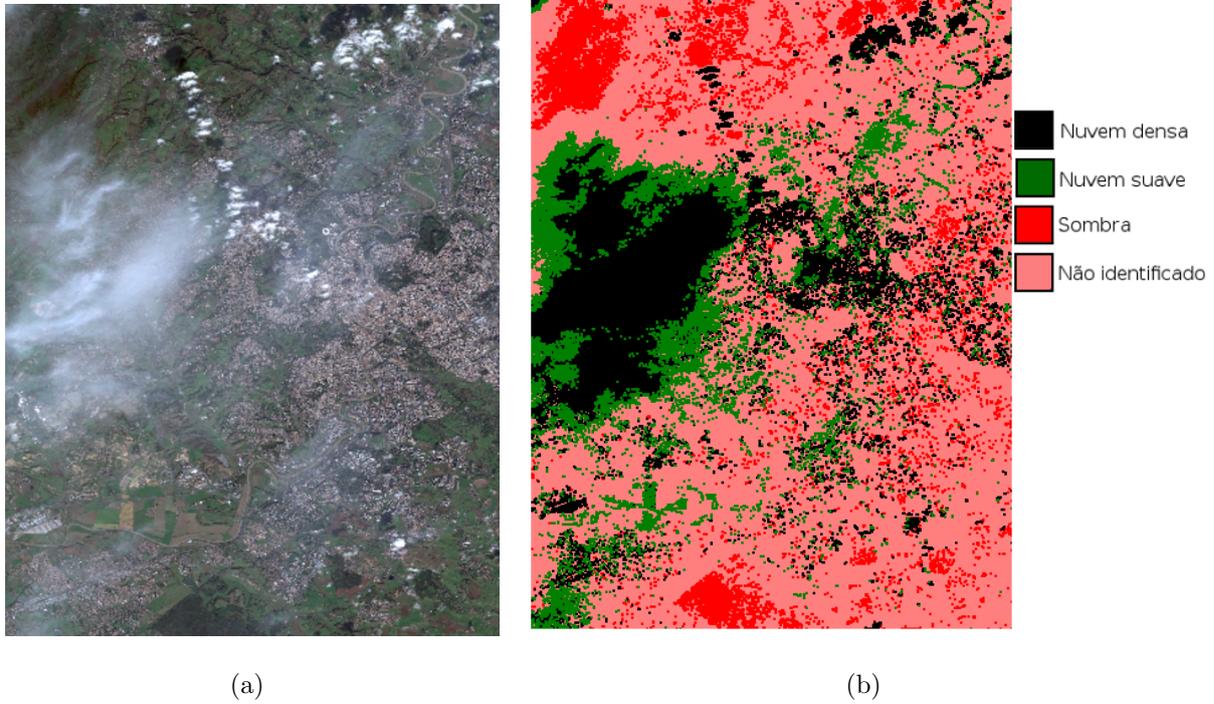


Figura 5.1: Detecção de nuvens e sombras: (a) Figura original retirada do repositório disponível. (b) Figura com as regiões destacadas.

As regiões de nuvem densa e sombra foram tratadas pelo algoritmo *Inpainting* descrito na Subseção 4.1.3. O resultado da interpolação pelo vizinho mais próximo, feita nessa fase, é mostrada na Figura 5.2(a). A filtragem homomórfica, descrita na Subseção 4.1.2, é feita sobre os *pixels* etiquetados como do tipo 2 produzindo resultado como o mostrado na Figura 5.2(b).

Durante a fase de filtragem homomórfica, foram testados 6 filtros para avaliar a relação entre o melhoramento visual provido e a quantidade de informação perdida. Conforme consta na Subseção 4.1.2, este trabalho usou o filtro *High-boost* para compor o processo da filtragem homomórfica. A vantagem desse filtro sobre os filtros *Hanning*, *Ideal*, *Butterworth*, *Bartlett* e *Gaussiano* se sustenta principalmente no fato de que este filtro não exclui toda a informação contida nas frequências mais baixas.

A conservação de parte da informação de baixa frequência aliada à ênfase aplicada nas altas frequências ocasiona menor perda de informação o que geralmente melhora a precisão da classificação. Além disso, também foi avaliada a inserção de ruído na imagem. Na Subseção 5.1.1 serão apresentadas as avaliações de cada filtro e então a escolha pelo filtro *High-boost* será plenamente justificada.

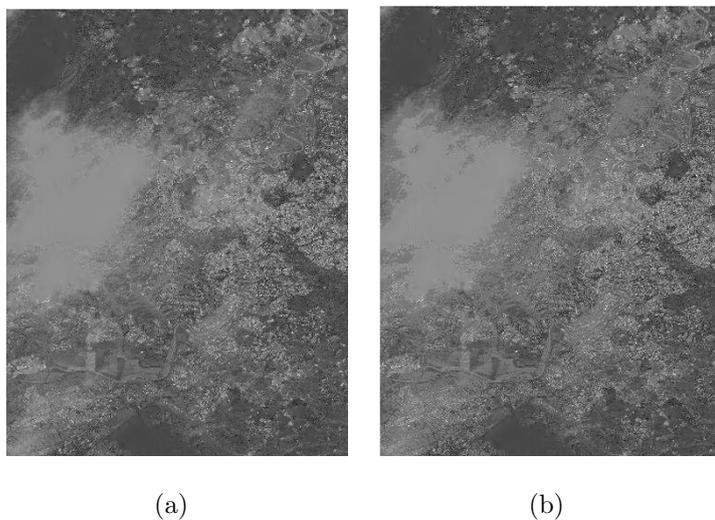


Figura 5.2: Primeira fase: (a) Figura interpolada pelo algoritmo *Inpainting*. (b) Figura tratada pelo filtro Homomórfico.

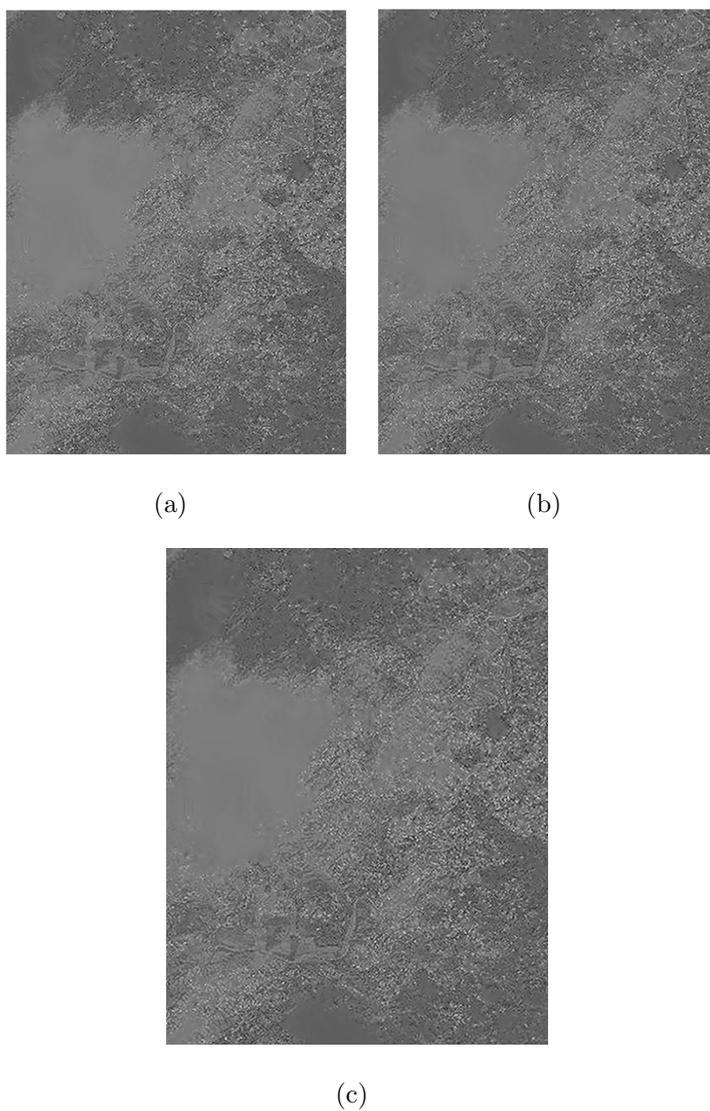


Figura 5.3: Segunda fase: (a) Melhoramento não-linear da Figura 5.2(a). (b) Melhoramento não-linear da Figura 5.2(b). (c) Figura resultante do processo de filtragem de nuvens.

As Figuras 5.3(a) e 5.3(b) mostram o resultado da filtragem de melhoramento não-linear descrito na Subseção 4.1.4. Pode-se notar que essas imagens são mais homogêneas pois houve uma suavização das transições bruscas entre *pixels* dentro de uma mesma região. As informações contidas nas duas imagens geram então, a imagem apresentada na Figura 5.3(c).

5.1.1 Avaliação

O processo de remoção de nuvens foi avaliado conforme seu desempenho na classificação das imagens do repositório, para tal mensuração foram avaliados os valores de coeficiente *Kappa* e pela razão sinal-ruído (PSNR). Usado em [38], o coeficiente *Kappa* é calculado de acordo com a Equação 5.1 e é utilizado para avaliar a concordância entre observações de uma mesma unidade amostral [15]

$$k = \frac{\sum_{j=1}^r s_{jj} - \sum_{j=1}^r s_{j+} s_{+j}}{1 - \sum_{j=1}^r s_{j+} s_{+j}} \quad (5.1)$$

onde r é o número de linhas da matriz de confusão, s_{ii} observação na linha j e coluna j , s_{j+} e s_{+j} são os totais marginais da linha e coluna j , respectivamente. A matriz de confusão, também chamada matriz de erro, é uma matriz quadrada que associa a categoria atribuída aos *pixels* durante a classificação à categoria real a que eles pertencem [79]

Tabela 5.1: Matriz de confusão para a imagem de Roma-ITA original (com nuvens) e limpa manualmente.

	Sombra	Vegetação	Urbana	Nuvem	Total marginal
Sombra	53086	18612	13607	377	85682
Vegetação	25988	17841	25264	321	69414
Urbana	9490	7203	6418	136	23247
Nuvem	7	7	5	0	19
Total	88571	43663	45294	834	178362

Para a produção dessa tabela foram criadas, manualmente, imagens sem interferência de nuvens ou sombras com o auxílio do programa Gimp.² Observa-se na Tabela 5.2, obtida após a remoção de nuvens, que *pixels* antes etiquetados como nuvem e sombra (vide Tabela 5.1) foram redefinidos para a representação urbana e vegetação, efeito este que corrobora a eficácia do algoritmo de remoção apresentado.

²<http://www.gimp.org/>

Tabela 5.2: Matriz de confusão para a imagem de Roma-ITA filtrada de nuvens e limpa manualmente.

	Sombra	Vegetação	Urbana	Nuvem	Total marginal
Sombra	65031	15911	4700	40	85682
Vegetação	23346	30966	15071	31	69414
Urbana	3447	10152	9628	20	23247
Nuvem	1	0	16	2	19
Total	91825	57029	29415	93	178362

Os valores de PSNR foram calculados de acordo com a Equação 5.3. O uso do PSNR vem da necessidade de avaliar o processo de reconstrução de imagens em comparação com a imagem original, onde quanto melhor é a reconstrução maior é o valor dessa métrica, não implicando, no entanto, na garantia de que a imagem seja visualmente aceitável [80].

A formulação denota que Q representa a imagem reconstruída (Q_i é o *pixel* dessa imagem) e P é a imagem original. Assim, o erro médio quadrático (MSE) entre as imagens, ou seja a média da diferença quadrática entre *pixels*, é dado pela Equação 5.2

$$MSE = \frac{1}{n} \sum_{i=1}^n (P_i - Q_i)^2 \quad (5.2)$$

onde $1 \leq i \leq n$. Na Equação 5.3, RMSE é definida como sendo a raiz quadrada do valor MSE.

$$PSNR = 20 \log_{10} \frac{\max_i |P_i|}{RMSE} \quad (5.3)$$

O valor PSNR não possui dimensão, porém a presença do logaritmo faz com que a interpretação seja dada em *decibéis* (dB), e essa presença implica também na redução de sensibilidade para mudanças no RMSE. Por não ter significado absoluto, os valores de PSNR apenas comparam desempenhos de métodos diferentes.

Durante a elaboração do trabalho, foram testados seis filtros passa-altas para a execução da filtragem homomórfica: *Ideal*, *Bartlett*, *Butterworth*, *Gaussiano*, *Hanning* e *High-boost*. Os resultados mostrados nas Figuras 5.4 e 5.5 mostram o desempenho de cada filtro de remoção de nuvens suaves em comparação com a imagem manualmente manipulada para não conter interferências causadas por nuvens e/ou sombras.

A Figura 5.4 apresenta o desempenho, em termos de valores PSNR para cada uma das 12 imagens do repositório acometidas por nuvens. Percebe-se que os filtros *High-boost* e *Ideal* apresentam os melhores resultados com relação à esta métrica. O filtro *Butterworth* apresentou resultados razoavelmente superiores aos filtros *Hanning*, *Gaussiano* e *Bartlett*.

A interpretação da Figura 5.5, dos valores Kappa, confirma a tendência de bons resultados dos filtros Ideal e *High-boost*, além do desempenho pobre do filtro *Butterworth* em relação à esta métrica.

De uma forma geral, o filtro *High-boost* apresentou melhor desempenho em ambas métricas, além do resultado visual ter sido mais aceitável. O filtro Ideal, apesar do bom desempenho nas métricas de avaliação, apresentou resultado visual ruim, com bastante confusão entre regiões devido a inserção de ruído. Vale destacar que durante os testes os parâmetros dos filtros, como a frequência de corte, foram mantidos em equivalência.

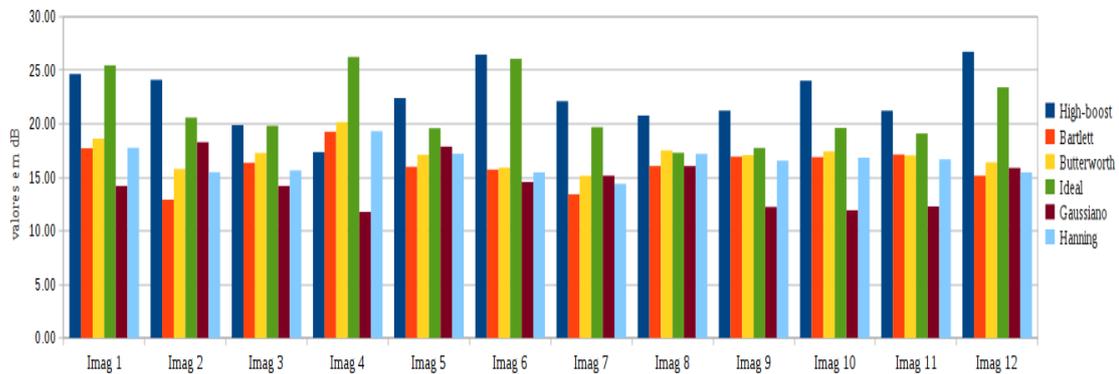


Figura 5.4: Valores PSNR para imagens filtradas por filtros homomórficos.

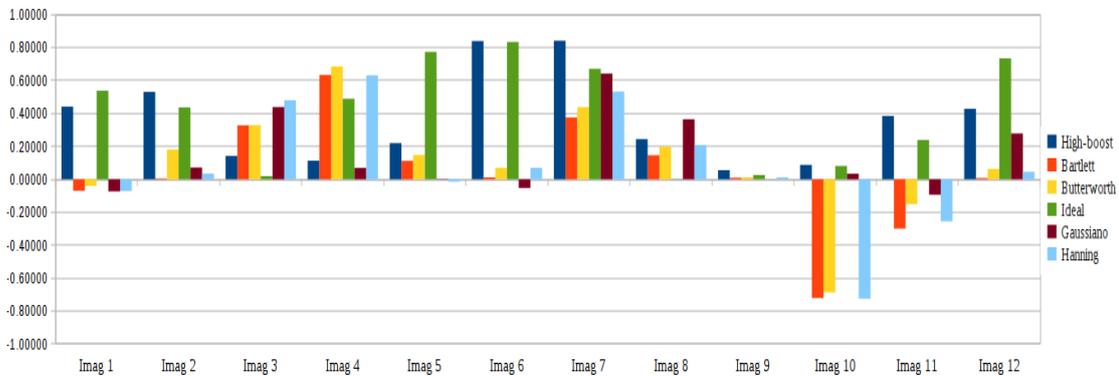


Figura 5.5: Valores Kappa para imagens filtradas por filtros homomórficos.

Os resultados mostrados nas Figuras 5.6 e 5.7 referem-se ao uso da filtragem homomórfica e do algoritmo *Inpainting*, isso é, da execução completa do algoritmo descrito na Seção 4.1 para remoção de nuvens densas, suaves e sombras.

A Figura 5.6 mostra a redução dos valores de PSNR para todos os filtros. Esse fenômeno, também observado por [38], é explicado pela tendência de diminuição da precisão da classificação após uma filtragem. Essa operação, que neste caso é sucedida por uma operação de substituição de *pixels* (*Inpainting*), provoca uma grande redefinição de *pixels* de forma que, ao mesmo tempo que visualmente produz resultados satisfatórios, produz valores de *pixels*

mais distantes daqueles que originaram a imagem tratada.

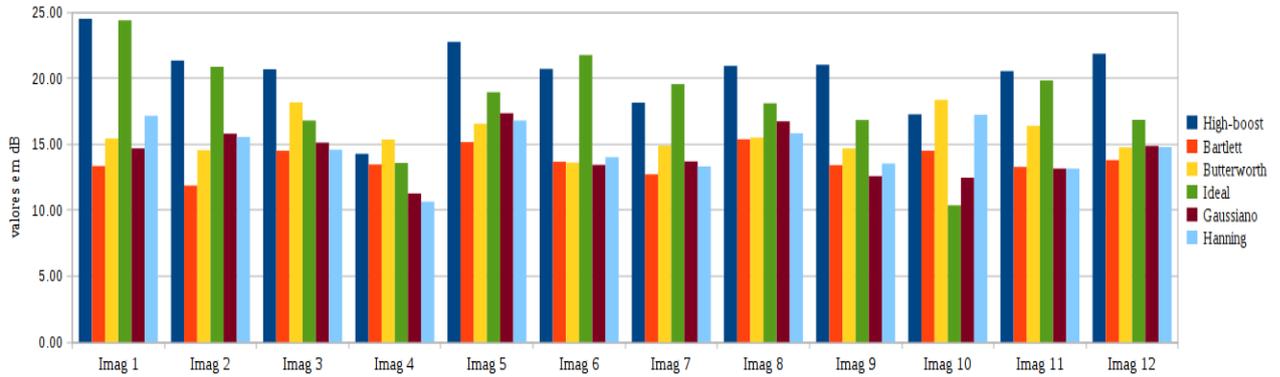


Figura 5.6: Valores PSNR para imagens tratadas por filtragem homomórfica e *Inpainting*.

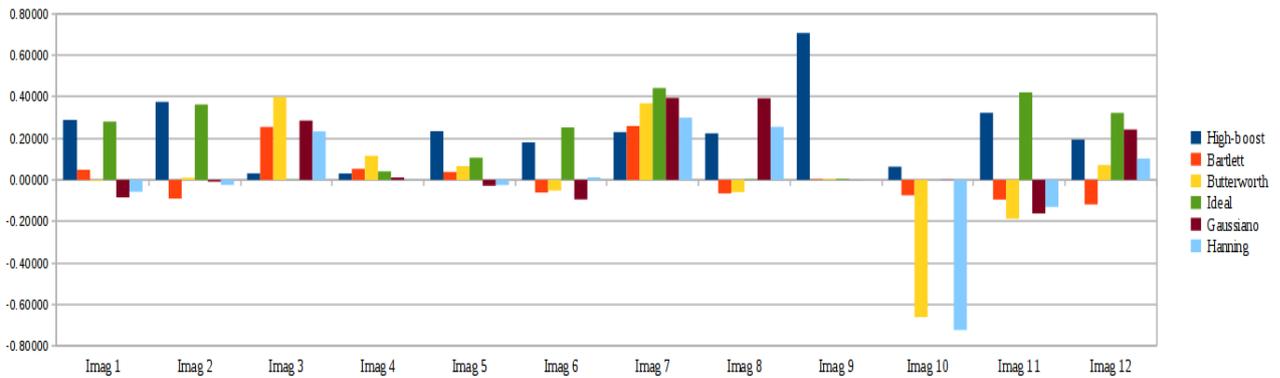


Figura 5.7: Valores Kappa para imagens tratadas por filtragem homomórfica e *Inpainting*.

O uso de filtros para remoção de nuvens deve ser cuidadosamente avaliado, uma vez que a ocorrência de nuvens em uma imagem é fato aleatório e não padronizado, por isso, seu uso indiscriminado pode piorar o desempenho das fases de segmentação e classificação pela supressão de detalhes da figura. Os valores Kappa também sofreram redução, como observa-se na Figura 5.7, pela mesma causa.

Como forma de ilustrar o fator de eficiência visual do método mostrado nesta seção, é apresentada a imagem resultante para os filtros *High-boost* e *Ideal* (vide Figura 5.1.1). De forma geral, como o filtro *High-boost* apresentou melhores resultados para a tarefa de remoção de nuvens, com relação satisfatória entre o efeito visual, coeficiente Kappa e valores PSNR, este filtro será usado para os testes seguintes.

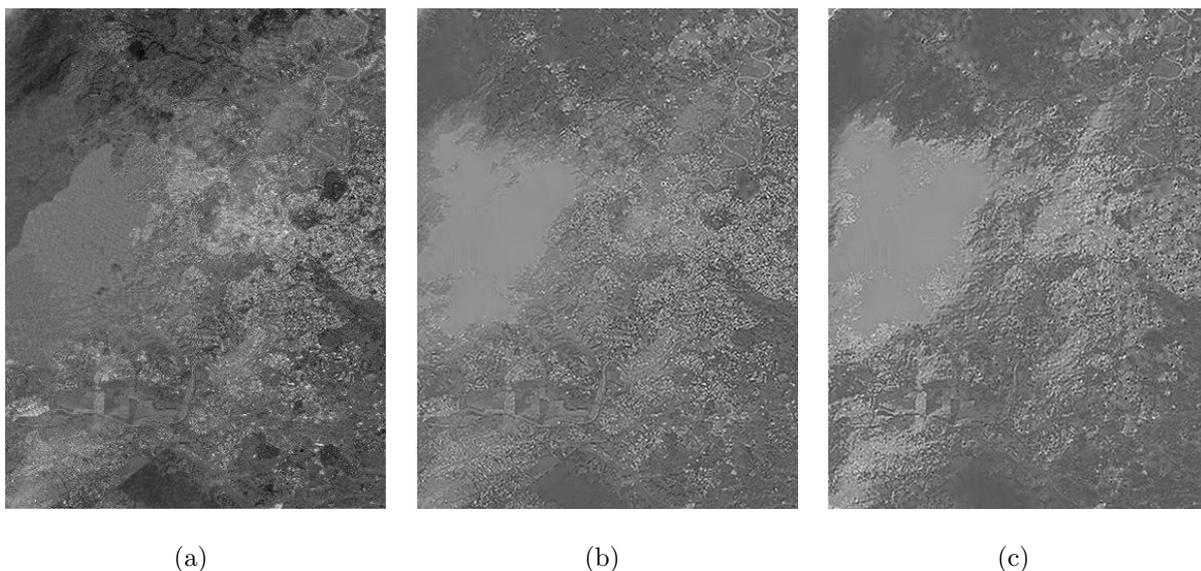


Figura 5.8: Efeito visual: (a) Imagem de Roma-ITA (Imagem 8) limpa de presença de nuvens. (b) Figura tratada por *Inpainting* e filtro homomórfico *High-boost*. (c) Figura tratada por *Inpainting* e filtro homomórfico Ideal.

5.2 Segmentação

O método de segmentação proposto, conforme descrito na Seção 4.2, aplicou a 1D-DCT nas linhas e colunas dos blocos de tamanho 2×2 em todas as imagens do repositório. Foram testados blocos de tamanho 4×4 , 8×8 e 16×16 , porém, constatou-se que com essas configurações o desempenho do algoritmo atinge patamar inaceitável para comparação e avaliação.

As imagens do repositório disponível foram captadas por diversos satélites, como *Ikonos*, *QuickBird02* e *WORLDVIEW-2*. Suas dimensões variam de 2048×2048 (Imagem 1) à 500×252 (Imagem 5), constituindo um banco bastante heterogêneo. A resolução espacial das imagens varia de acordo com o sensor que a captou, mas é sabido que a resolução das imagens 1 e 2, captadas pelo satélite *IKONOS*, possuem as maiores resoluções, 2,4 e 2,88 metros, respectivamente. As imagens foram extraídas de bancos amostrais de empresas que comercializam imagens de sensores remotos e, por esse motivo, os cabeçalhos das imagens não são disponibilizados, bem como informações sobre correções efetuadas sobre as imagens.

5.2.1 Avaliação

A segmentação de imagens pode ser realizada de diversos modos, principalmente pelo fato de que não há padronização de métodos segmentadores, conforme dito anteriormente. Diante disso, será apresentada uma avaliação de segmentação baseada no algoritmo EBC (*Edge-Border Coincidence*, coincidência entre bordas) [81].

A medida de coincidência entre bordas faz uma sobreposição das regiões de borda da imagem segmentada e da imagem original. As bordas relativas à imagem original são encontradas mediante um algoritmo de detecção de bordas, que neste trabalho é baseado no operador Sobel [82]. Essa medida não depende de informação sobre o grau de verdade dos objetos delimitados, ou seja, sua avaliação não leva em consideração a formação de objetos.

O algoritmo EBC adota um conjunto O de *pixels* (x_{pi}, y_{pi}) extraído pelo operador de bordas Sobel da imagem original e o conjunto S de *pixels* (x_{qi}, y_{qi}) encontrados no limiar das regiões da imagem segmentada. Assim,

$$O = p_1, p_2, \dots, p_O = (x_{p1}, y_{p1}), (x_{p1}, y_{p1}), \dots, (x_{pO}, y_{pO})$$

e

$$S = q_1, q_2, \dots, q_O = (x_{q1}, y_{q1}), (x_{q1}, y_{q1}), \dots, (x_{qS}, y_{qS}),$$

definem a expressão

$$EBC = \frac{n(O \cap S)}{n(O)},$$

onde $n(O)$ é o número de elementos do conjunto O e $O \cap S = \{(x, y) | (x, y) \in O \text{ e } (x, y) \in S\}$. Os valores EBC estão localizados no intervalo $[0, 1]$, onde quanto mais próximo de 1 maior a confiança entre as bordas da imagem original e da imagem segmentada.

Essa medida indica, indiretamente, a qualidade da segmentação, por isso, é possível que o valor EBC seja alto enquanto a confiança entre os conjuntos seja baixa e vice-versa. Para o processo de avaliação de segmentação, o EBC será usado como um indicador de eficácia, representando um reforço ao fator visual dos resultados.

Assim, a Tabela 5.3 mostra a comparação dos resultados apontados pelo algoritmo de segmentação proposto (Seção 4.2) em comparação ao *software Spring*³. O *Spring* é um Sistema de Informações Geográficas (SIG) desenvolvido pelo Instituto Nacional de Pesquisas Espaciais (INPE) em parceria com outros órgãos de pesquisa. Nesse sistema, a segmentação das imagens foi feita pelo método de crescimento por regiões, com parâmetro de similaridade igual a 10 e área por *pixel* igual a 10 para todas as imagens.

³<http://www.dpi.inpe.br/spring/>

Tabela 5.3: Valores EBC para as figuras segmentadas pelo método de segmentação proposto e pelo *software Spring*.

Imagens	Método proposto	<i>software Spring</i>
Imagem 1	0,4816	0,5839
Imagem 2	0,4889	0,4887
Imagem 3	0,5867	0,5904
Imagem 4	0,4977	0,2232
Imagem 5	0,571	0,6794
Imagem 6	0,3549	0,6244
Imagem 7	0,2128	0,6354
Imagem 8	0,3446	0,7198
Imagem 9	0,5403	0,7573
Imagem 10	0,4117	0,523
Imagem 11	0,6709	0,7282
Imagem 12	0,3997	0,5898
Imagem 13	0,5609	0,8151
Imagem 14	0,6575	0,8392

Uma vez que o *Spring* é um *software* referência no tratamento de imagens de satélite, é natural que seus resultados sejam mais precisos, porém o que se pretende mostrar na Tabela 5.3 são os resultados promissores do método proposto. Por exemplo, para a Imagem 4, o valor EBC do método proposto foi superior ao apresentado pelo *Spring*.

5.3 Classificação

Após a segmentação das imagens, foi feito o processo de atribuição de classes. Nessa fase, também são usados os valores de energia AC e média dos blocos DCT, e utilizando os limiares encontrados pelo algoritmo descrito na Subseção 4.2.2, são atribuídos valores ou etiquetas às regiões de homogeneidade identificadas na segmentação. Para completar o processo, é feita uma operação de filtragem morfológica de abertura, como descrito na Subseção 4.2.3.

O sistema *Spring* será utilizado para comparação, mais uma vez. A classificação por região desse sistema, aplicada a todas as imagens, utiliza tanto a informação espectral dos *pixels* quanto a informação espacial, que envolve a relação entre os *pixels* e seus vizinhos[83]. O classificador *Isoseg* é o algoritmo usado para essa classificação e as regiões são caracterizadas pelos atributos estatísticos de média e matriz de covariância.

5.3.1 Avaliação

A avaliação da classificação é feita pela análise de entropia relativa da imagem e pela medida do erro médio quadrático.

Entropia relativa ou divergência de Kullback-Leibler, é uma medida não simétrica da diferença entre duas distribuições de probabilidade P e Q . A entropia (H) é definida pela expressão:

$$H(P||Q) = - \sum_{i=1}^n p(x_i) \log_b \left(\frac{p(x_i)}{q(x_i)} \right).$$

Comumente, a distribuição P representa a distribuição “verdadeira” dos dados, enquanto que a medida Q representa o modelo de aproximação de P , e a entropia relativa representa a “medida” de quanta informação P e Q tem em comum, sendo que, em geral, $H(P||Q) \neq H(Q||P)$. $H(P||Q) \geq 0$ sendo nula se e somente se $P = Q$ e atingindo seu valor máximo quando P e Q são independentes.

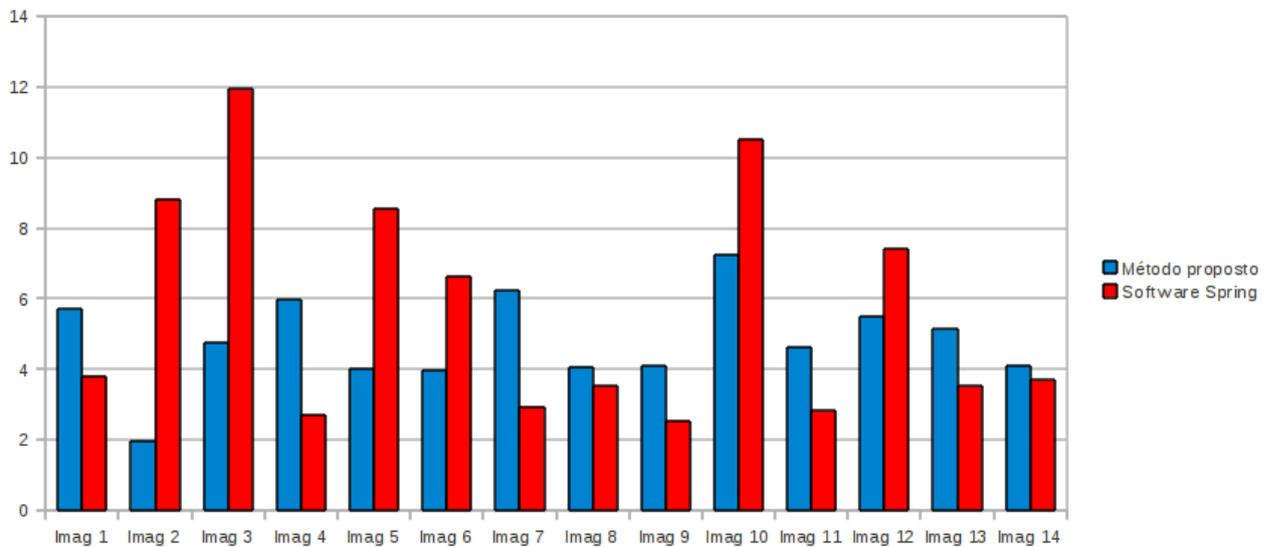


Figura 5.9: Relação de valores de entropia relativa entre as imagens classificadas pelo método proposto e classificadas pelo *software Spring*.

De posse das imagens classificadas pelo método proposto e pelo sistema *Spring*, foram computados os valores de entropia relativa, ilustrados na Figura 5.9. Como é possível observar, os valores de entropia das imagens resultantes do método de classificação proposto tem a tendência de apresentar valores levemente superiores quando comparado com os valores apresentados nas imagens obtidas pela classificação *Isoseg* do *Spring*. Quando o método

proposto obteve valores menores de divergência, podendo ser citadas as imagens 2, 3, 5 e 10, a diferença entre a entropia das imagens chega à 7,17 (Imagem 3). Quando as imagens classificadas pelo método proposto apresentaram entropia maior que as classificadas pelo *Spring*, a maior diferença entre esses valores não foi superior à 3,27, ocorrida na Imagem 7, o que mostra o potencial do método proposto.

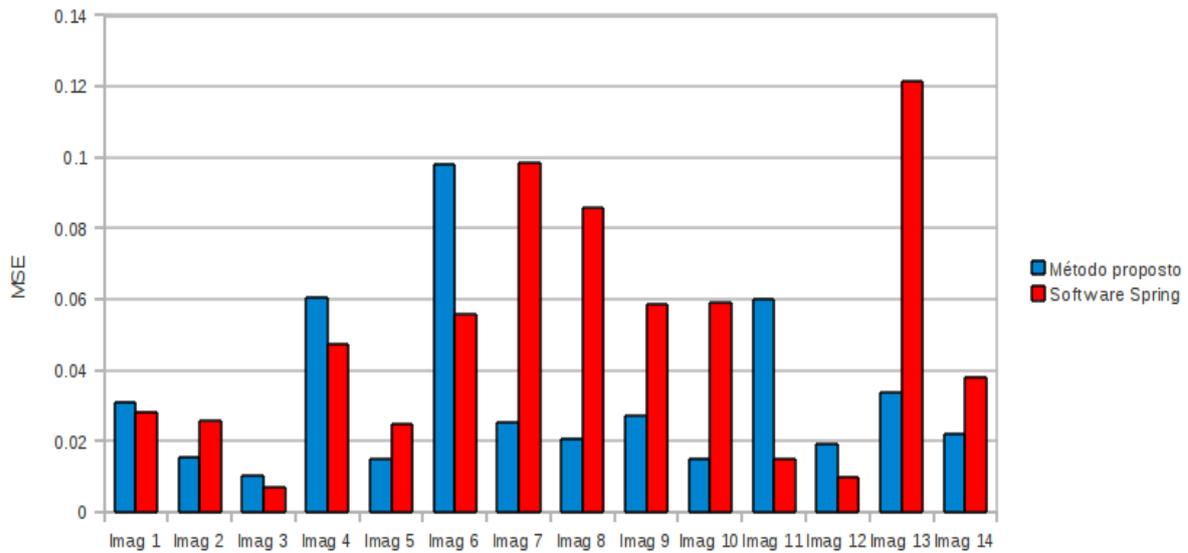


Figura 5.10: Relação de valores de MSE entre as imagens classificadas pelo método proposto e classificadas pelo *software Spring*.

Como a medida de entropia relativa não é uma métrica, pois não se trata do cálculo de distância, e sim da medida de similaridade, a avaliação da classificação será complementada pela apresentação dos valores de MSE das imagens classificadas.

Na Figura 5.10, são mostrados os valores de erro de acordo com a Equação 5.2. Observa-se que em oito ocasiões o erro obtido pela imagem classificada pelo método proposto foi menor em comparação ao obtido pelo classificador *Isoseg*. A maior discrepância entre os valores foi encontrada na Imagem 13, onde o MSE da classificação proposta foi 0,0878 menor que a comparada.

Em paralelo aos valores de entropia relativa e de MSE, as imagens subsequentes visam enfatizar o fator visual dessas classificações. Da Figura 5.11 até a Figura 5.24 são apresentadas as imagens originais presentes no banco, as imagens resultantes da filtragem de remoção de nuvens ou sombras (quando houverem as obstruções), as imagens classificadas pelo método proposto e pelo sistema *Spring*. A legenda oferecida pelo *software Spring* não é apresentada em classes definidas, como água, solo ou vegetação, e sim em termos de classes (ou temas) numerados. As imagens são apresentadas seguindo a ordem adotada em todas as figuras de resultado, na forma “Imagem n ”, onde $n = 1, 2, \dots, 14$.

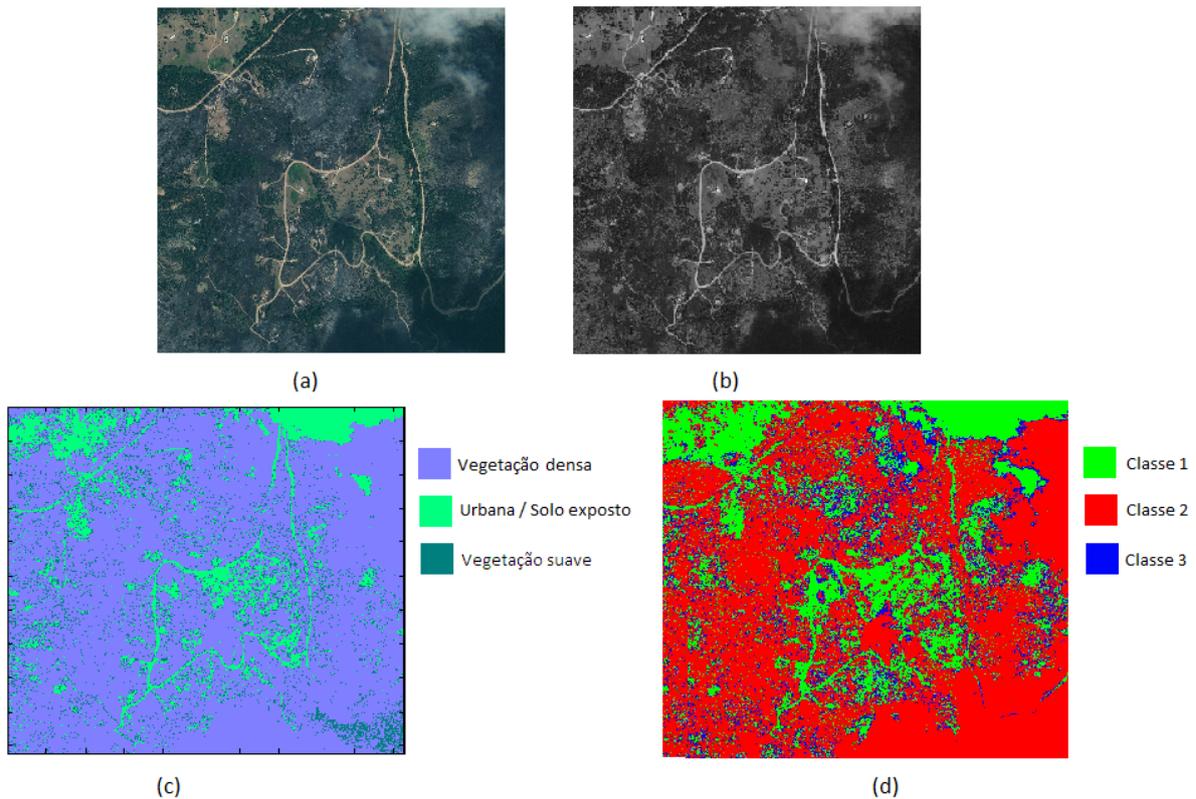


Figura 5.11: Imagem 1 - Cena capturada pelo satélite Ikonos 4: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

A Figura 5.11(c), mostrou boa precisão na identificação das áreas de solo exposto quando comparado ao resultado do *software Spring*, Figura 5.11(d). Seguiu também, a identificação da classe de vegetação densa, porém, a identificação do que se pode considerar como vegetação suave, não foi feita a contento ao final da execução do algoritmo proposto.

A Figura 5.12(c), classificada pelo método proposto, apresenta definição levemente superior à Figura 5.12(d). As áreas identificadas como região urbana possuem maior confiança e menor confusão com as áreas identificadas como solo exposto. Essas classes possuem comportamento similar, por isso são mais sensíveis à limiarização por histograma. Percebe-se que o algoritmo de crescimento de regiões usada pelo sistema *Spring* também não distingue facilmente as regiões.

A Figura 5.13 mostra a clara distinção entre as classes água e solo exposto, detalhe não distinguido pelo *software* usado na comparação. Ainda sim, a definição das bordas mostra-se bastante similar entre as diferentes segmentações e os valores de erro são próximos, tendo o *software* comparado um valor de MSE levemente inferior.

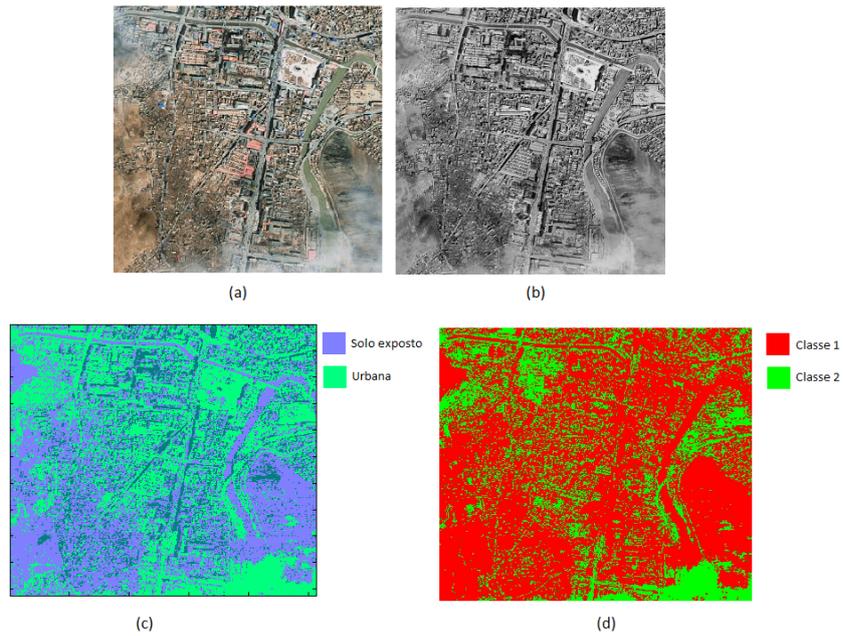


Figura 5.12: Imagem 2 - Cena capturada pelo satélite Ikonos 5: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

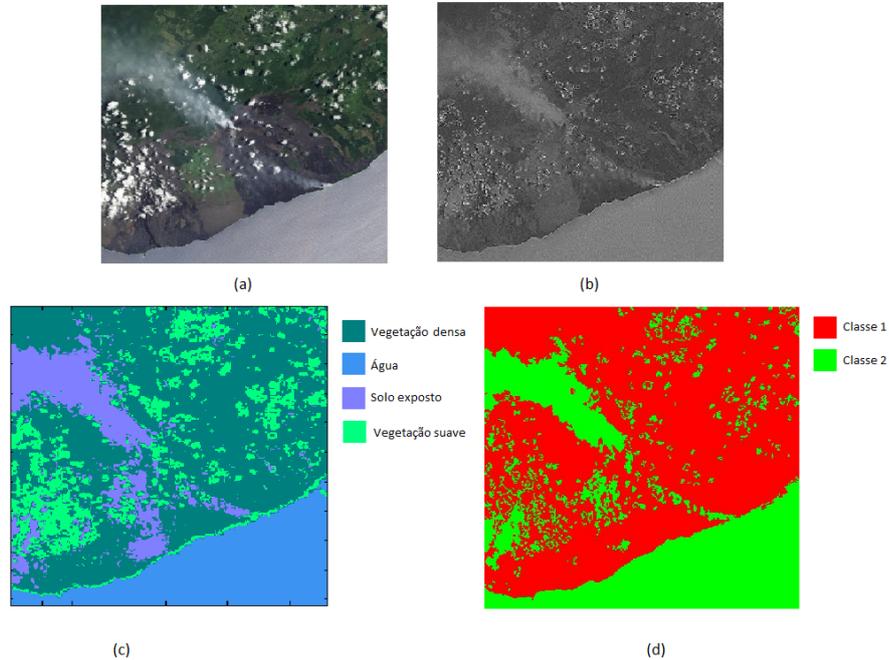


Figura 5.13: Imagem 3 - Cena da região do Hawaii-EUA: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

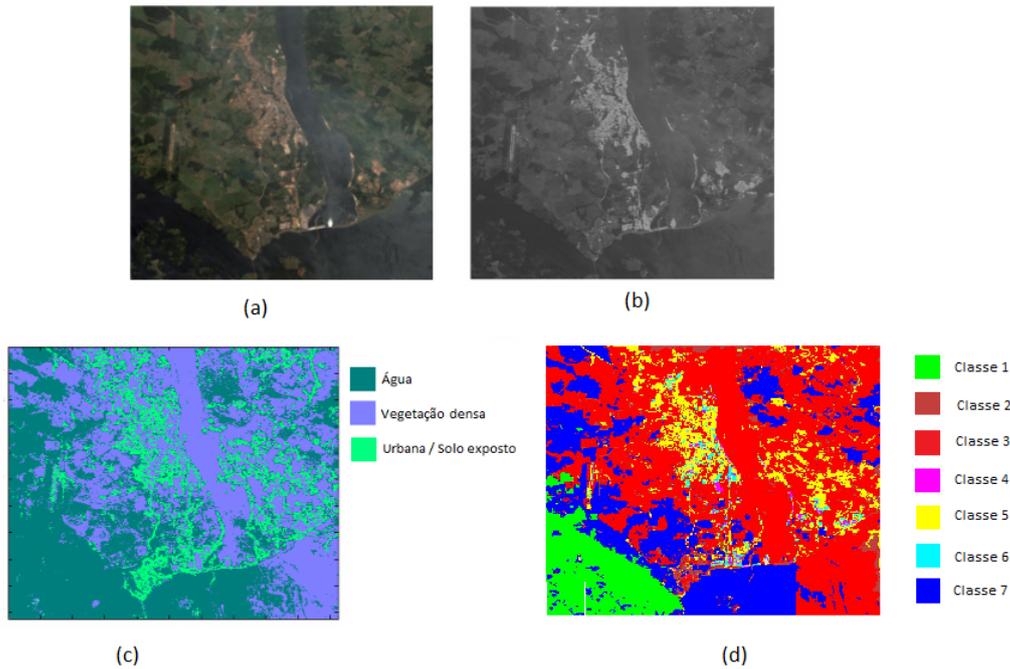


Figura 5.14: Imagem 4 - Cena da região de Tucuruí-Pa: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

A grande presença de nuvem suave na Figura 5.14(a) provocou uma queda na segmentação pelo *software Spring*, fato pouco observado no resultado do método de segmentação proposto (Imagem 4 na Tabela 5.3). A coincidência entre bordas do sistema usado para comparação apresentou um valor baixo e a entropia relativa aponta para uma maior relação com a imagem original.

A Figura 5.15 mostra um resultado visual bastante parecido, com valores EBC similares e com erro levemente superior para a imagem *Spring*. Já a cena imageada da região do lago *Leech*, na Figura 5.16, apresenta resultado visual e EBC inferiores para o método proposto, desempenho este acompanhado pelo valor de erro apresentado, que foi o pior valor de MSE entre todas as imagens do banco para este classificador.

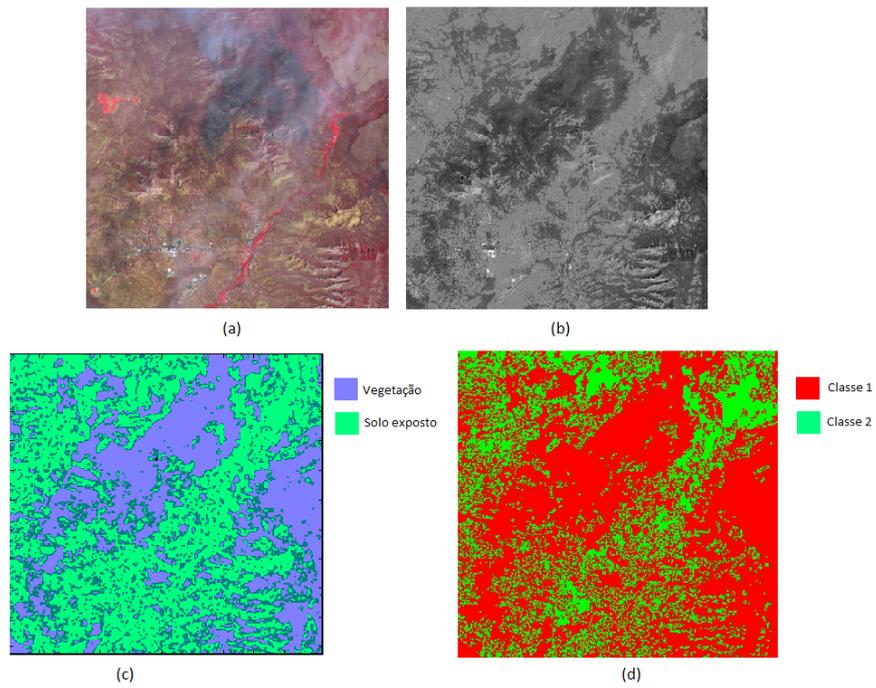


Figura 5.15: Imagem 5 - Cena da região de Sedona, Arizona-EUA: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

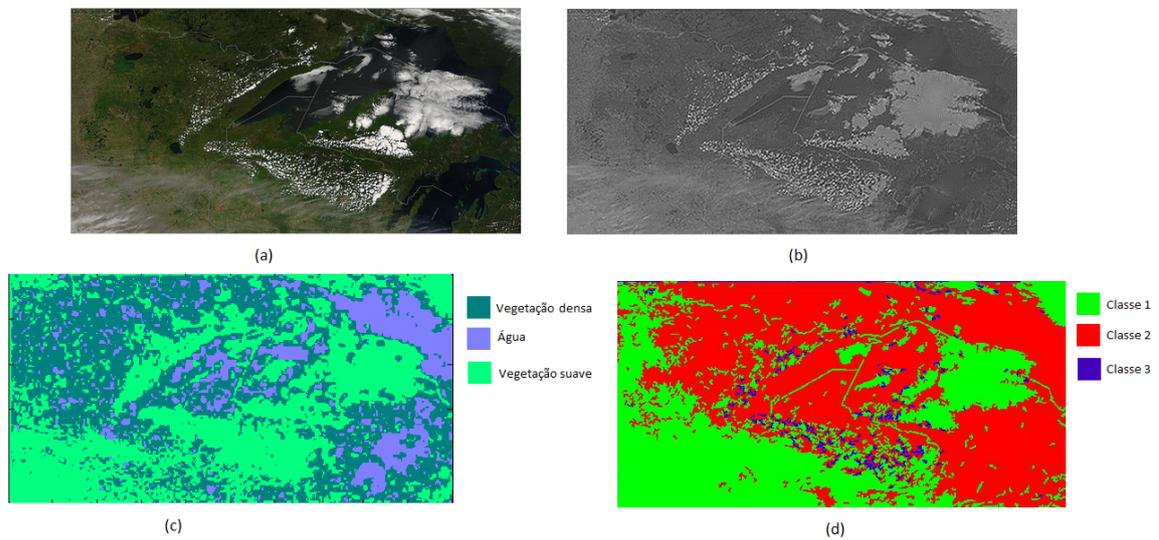


Figura 5.16: Imagem 6 - Cena da região do lago *Leech* em Minnesota-EUA: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

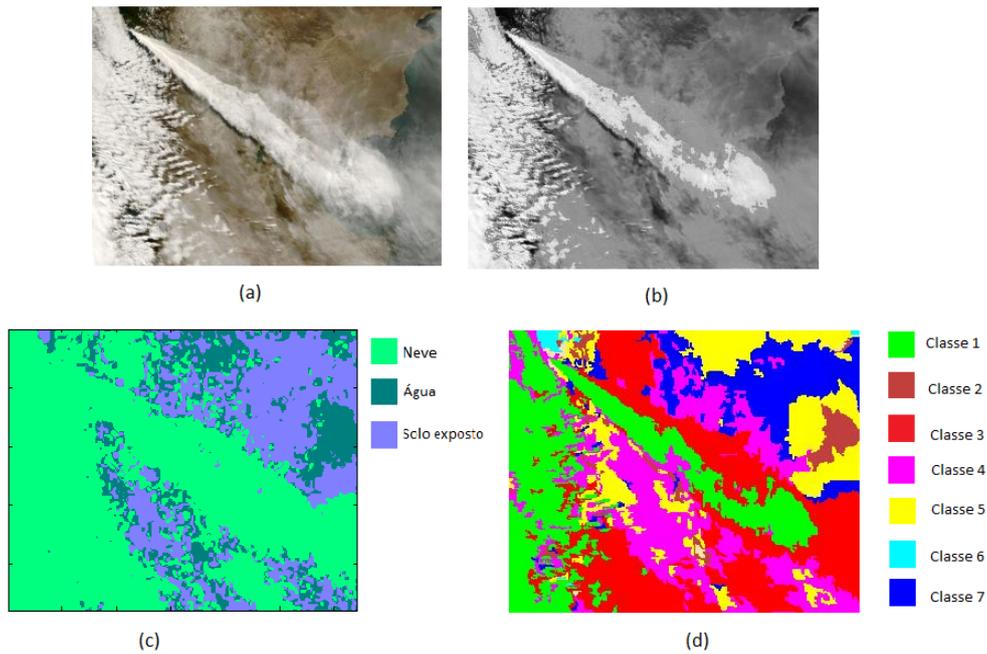


Figura 5.17: Imagem 7 - Cena da região Andina: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

Observando atentamente o resultado da remoção de nuvens na Figura 5.17(b), percebe-se que feições de solo exposto e água, vistas na Figura 5.17(a), perderam sua distinção. Diante disso, percebe-se que o resultado visual da Figura 5.17(c) foi mais homogêneo. Deve-se destacar que a região branca nessa cena representa depósitos acumulados de neve, que foram mantidas na imagem filtrada de nuvens. A classificação por crescimento de regiões (*Spring*) foi capaz de distinguir esses depósitos com mais eficiência que o método proposto.

A Figura 5.18(d), mostrou maior confiança entre bordas, e mais alta heterogeneidade dos dados. Por outro lado, a Figura 5.18(c) mostra maior capacidade de aglutinação de regiões homogêneas. A granularidade deste resultado é menor e, visualmente, o resultado oferecido pelo método proposto é mais homogêneo e mais confiante ao aspecto visual da imagem filtrada.

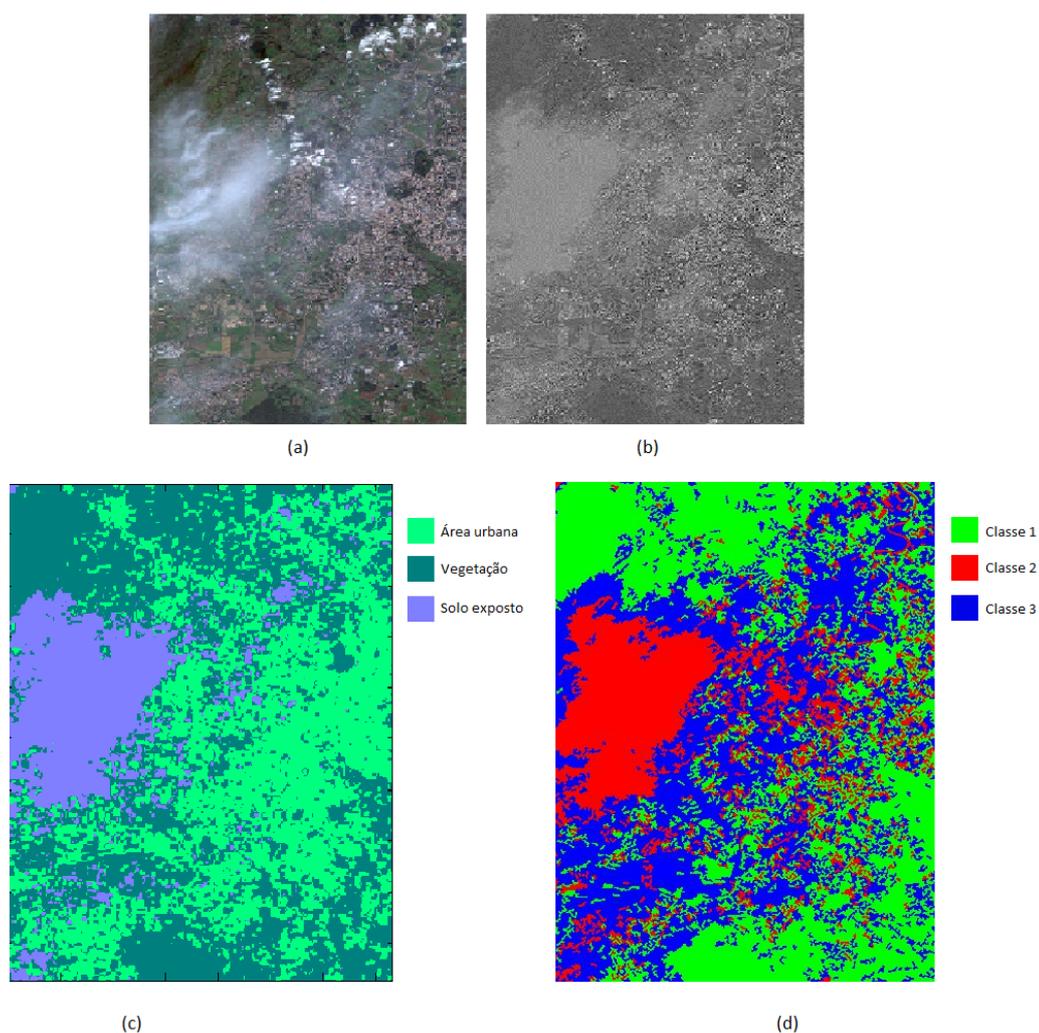


Figura 5.18: Imagem 8 - Imagem de Roma-ITA: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

As feições de área urbana e água na Figura 5.19(a) são facilmente confundidas. A presença de nuvens na Figura 5.19(b) conseguiu ser, em sua maioria, suavizada e aglutinada às regiões em seu entorno. Pela grande confusão entre as áreas urbana e água na imagem original, a Figura 5.19(d), apresentou maior heterogeneidade com a presença de regiões vermelhas entre as regiões azuis. Por outro lado, a Figura 5.19(c), obtida pelo método proposto conseguiu distinguir essas regiões com maior eficiência e alcançou maior homogeneidade nas regiões.

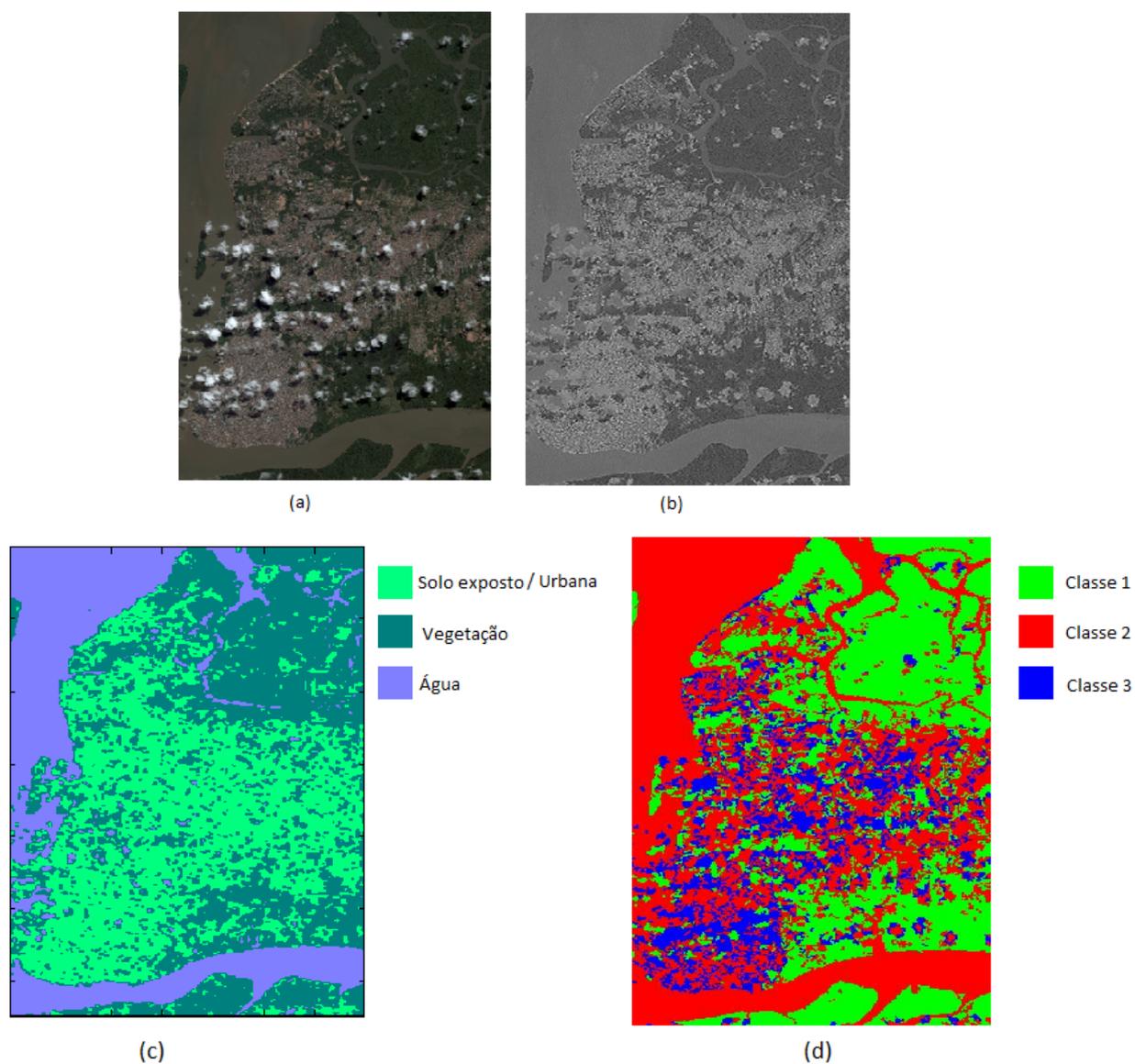


Figura 5.19: Imagem 9 - Imagem da região metropolitana de Belém: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

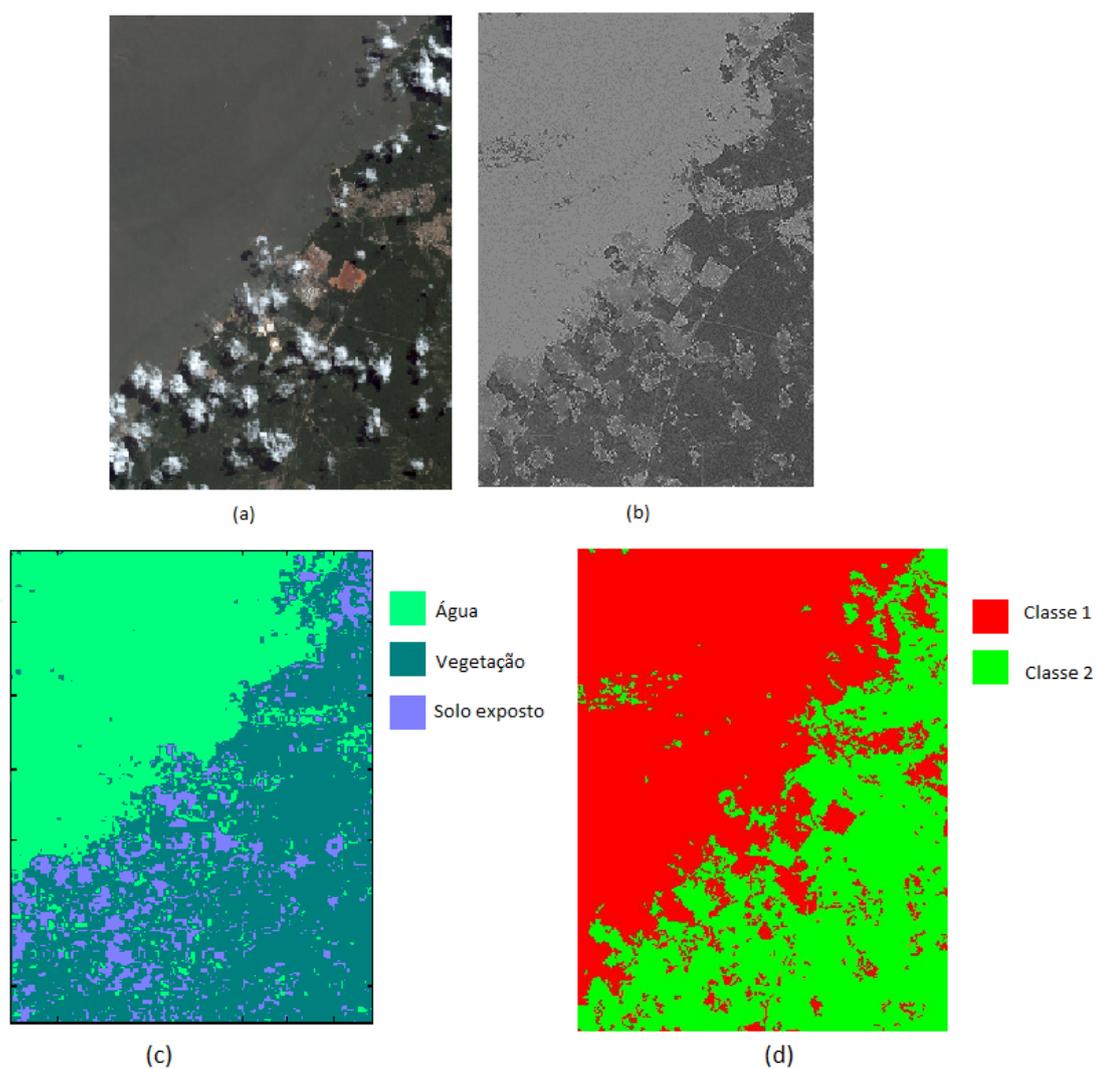


Figura 5.20: Imagem 10 - Cena da região da Baía do rio Capim-Pa: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

Na Figura 5.20(d) observa-se claramente a não distinção entre região de solo exposto e água, distinção feita em partes pelo método proposto e ilustrada na Figura 5.20(c). Esta que apresentou menor coincidência entre bordas e apresentou MSE consideravelmente menor que a comparada.

O maior valor de entropia relativa apresentado pela Figura 5.21(c) deve-se à distinção entre as regiões de solo exposto e água, o que não acontece na figura referente ao *software Spring* (Figura 5.21(d)). Percebe-se ainda que a maior parte das nuvens presentes na imagem foram filtradas pelo método de remoção de nuvens (Figura 5.21(b)).

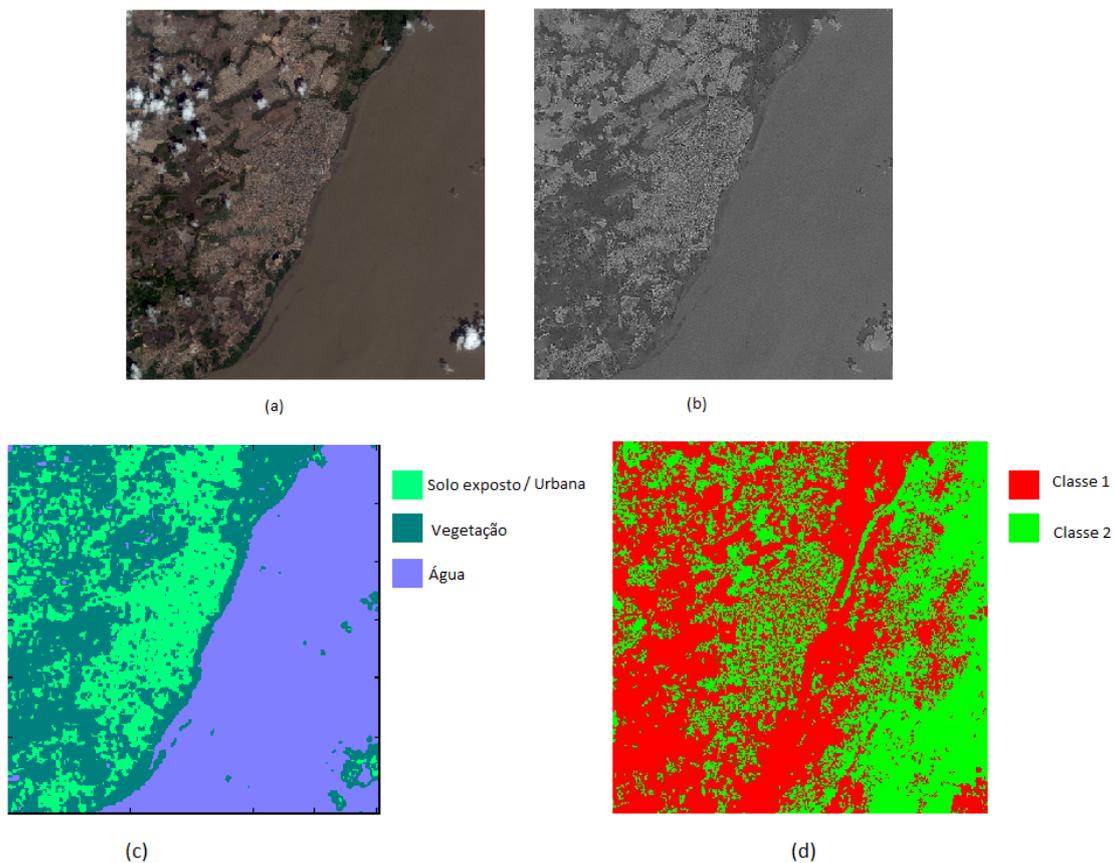


Figura 5.21: Imagem 11 - Cena da cidade de Macapá-Ap: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

A classificação pelo *software Spring* produziu, para a Figura 5.22(a), a identificação de várias classes que diferenciam diversas qualidades de solo exposto e vegetação, como observado na Figura 5.22(d). Segundo o método proposto, ilustrado na Figura 5.22(c), essas qualidades não são diferenciadas, ainda sim, mantendo a interpretação do que essas regiões representam. Por manter a interpretação do que as regiões representam, mas sem diferenciar

essas informações, seu valor de entropia relativa é menor que o valor apresentado para a Figura 5.22(d).

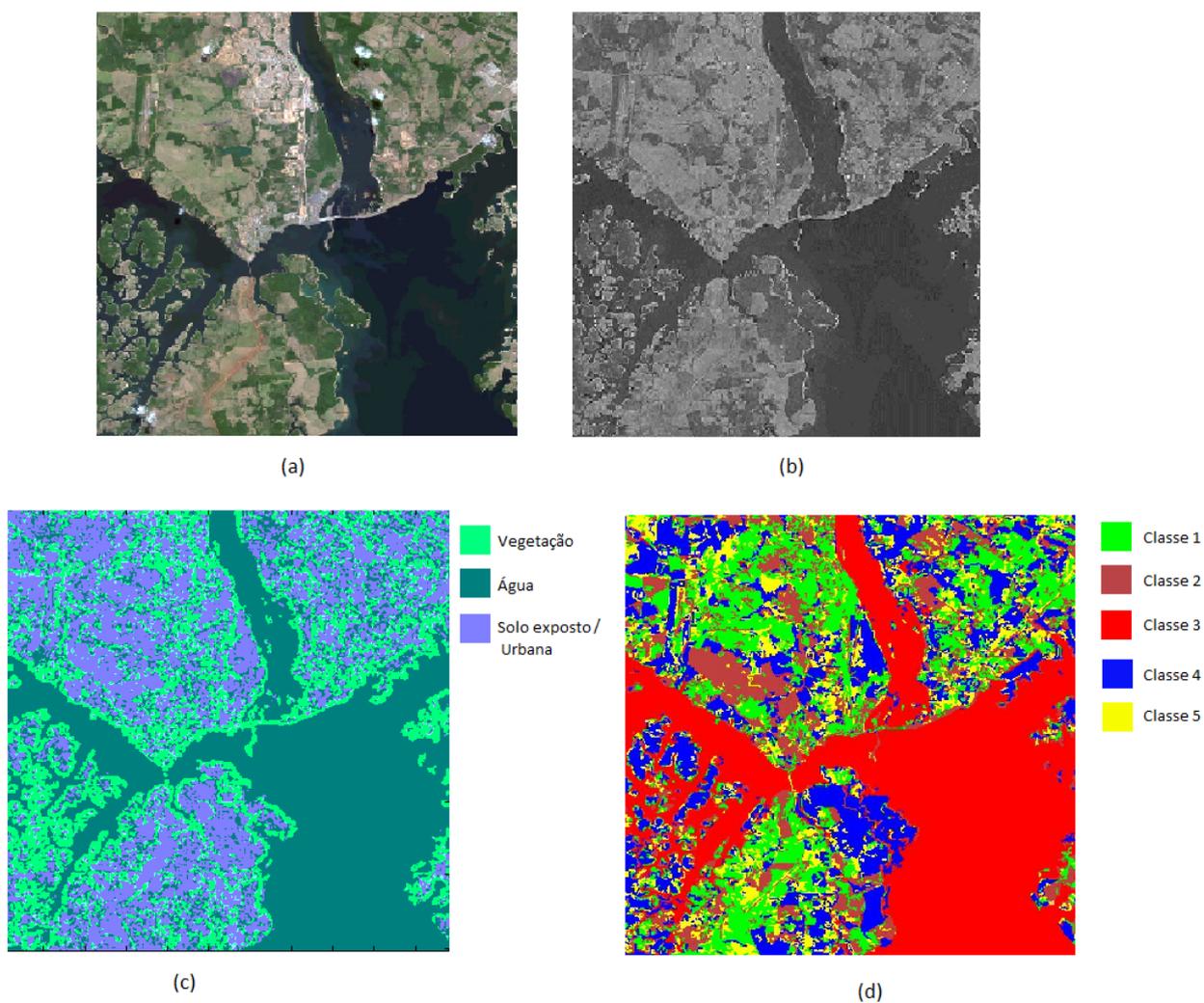


Figura 5.22: Imagem 12 - Cena da região de Tucuruí-Pa: (a) Imagem original (b) Imagem filtrada para remoção de nuvens (c) Imagem classificada pelo método proposto (d) Imagem classificada pelo *software Spring*.

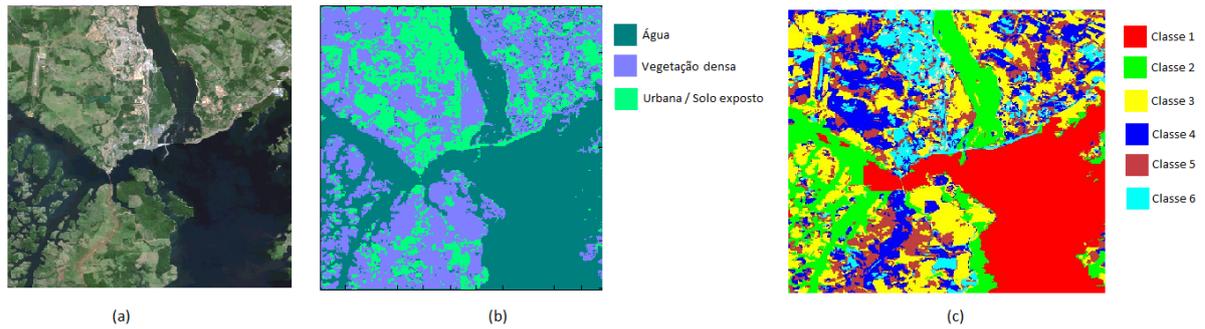


Figura 5.23: Imagem 13 - Cena da cidade de Tucuruí-Pa: (a) Imagem original (b) Imagem classificada pelo método proposto (c) Imagem classificada pelo *software Spring*.

As cenas mostradas na Figura 5.23 e na Figura 5.24 não apresentam obstrução por nuvens ou sombras. A Figura 5.23(b), obteve resultado similar à Figura 5.22(c), porém, sem a distinção entre diferentes qualidades de solo e vegetação encontradas pelo sistema *Spring*, sistema este, que atribuiu diferentes qualidades também para a região de água na figura.

A Figura 5.24(a) apresenta uma cena bastante complexa, pois apresenta níveis de vegetação e água diversos e de valores bastante similares. Diante dessa situação, os dois métodos apresentaram resultados de pouca homogeneidade, diante da imagem original. Com isso, as imagens classificadas pelo método proposto, Figura 5.24(b), e pelo *software Spring*, Figura 5.24(a), apresentaram resultado visual bastante parecido.

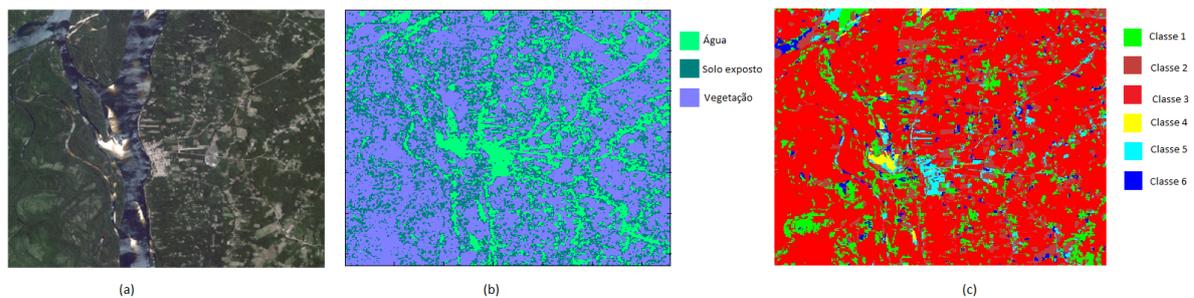


Figura 5.24: Imagem 14 - Cena da região do rio Tocantins-Pa: (a) Imagem original (b) Imagem classificada pelo método proposto (c) Imagem classificada pelo *software Spring*.

Capítulo 6

Conclusão

Das diversas abordagens possíveis para implementação da tarefa de classificação, foi proposto um algoritmo de classificação usando Transformada Cosseno Discreta. O artifício da limiarização de histograma das respostas de energia AC e média, providas por esta transformada, foi empregado para a divisão da imagem em regiões homogêneas e para sua posterior etiquetagem.

Diante da escassez de imagens livres de obstrução por nuvens e sombras, adotou-se duas abordagens para remoção ou suavização dessas interferências: filtragem homomórfica e *Inpainting*. A fase de pré-processamento usou a filtragem homomórfica para tratar a ocorrência de nuvens suaves e o método *Inpainting*, com interpolação por vizinho próximo, foi usado para remoção ou suavização de nuvens densas e sombras.

Após a fase de pré-processamento, foi executada a etapa de segmentação, que usou os valores encontrados por meio da limiarização de histograma, para separar as imagens em regiões de homogeneidade, que posteriormente foram classificadas de acordo com suas características de textura. Os valores de energia AC e média da DCT são usados para dividir a imagem de acordo com as informações texturais dos blocos da transformada. Foram usados blocos quadrados de lado $n = 2$ e a transformada 1D-DCT foi aplicada nas linhas e colunas desses blocos. A classificação seguinte atribuiu etiquetas às regiões homogêneas com base nos mesmos valores de energia e média.

O método de remoção de nuvens foi avaliado com base nas métricas de razão sinal-ruído de pico (PSNR) e coeficiente Kappa. Os testes realizados em vários filtros passa-alta constataram a eficácia superior do filtro *High-boost*. Em particular, esse filtro não suprime todas as frequências abaixo da frequência de corte, e atribui-se a esse fato seu desempenho geralmente superior a outros filtros como *Bartlett* e *Butterworth*.

A segmentação foi avaliada pelo método da coincidência entre bordas (EBC) e comparada ao resultado da segmentação efetuada pelo *software Spring*. Por essa métrica, observou-se que o método de segmentação proposto apresentou resultados promissores. Por fim, a classificação foi avaliada pela medida similaridade entre imagens, a entropia relativa, além da avaliação pela observação do erro médio quadrático e do aspecto visual produzido pelos métodos. O algoritmo de classificação proposto mostrou-se similar aos resultados obtidos pelo sistema *Spring*.

Foi usado um banco de imagens constituído por 14 imagens, sendo 12 delas obstruídas por nuvens e/ou sombras. As imagens foram captadas por diversos sensores e possuem dimensões diversas, constituindo um repositório bastante heterogêneo. Geralmente, imagens de satélite possuem dimensões maiores que as disponíveis no banco atual, porém, espera-se que a execução do algoritmo não se torne custosa computacionalmente quando aplicado à essas novas imagens.

O algoritmo pode ser aprimorado uma vez que seus resultados foram encorajadores. Seu desempenho foi bastante semelhante ao do sistema *Spring*, que é fruto de anos de pesquisa e de colaboração multi-disciplinar. Para tanto, novas ações são sugeridas na Seção 6.1.

6.1 Trabalhos futuros

Visando a melhoria do algoritmo sugere-se a implementação das seguintes ações:

- Inserir parâmetros que aumentem a precisão da fase de detecção de nuvens e sombras.
- Usar um algoritmo evolucionário ou de inteligência artificial para automatizar a escolha dos parâmetros citados no item anterior.
- Avaliar o uso de diferentes implementações do método *Inpainting*.
- Avaliar a inserção de parâmetros adicionais para aumentar a precisão das tarefas de segmentação e classificação, tais quais mediana e variância.
- Expandir o banco de imagens pela aquisição de novas cenas.

6.2 Trabalhos publicados

- A. Siravenha e E. G. Pelaes, “The use of discrete cosine transform for satellite images segmentation and comparison to statistical metrics” - XV Simpósio Brasileiro de Sensoriamento Remoto, 2011.
- A. C. Siravenha, D. Sousa, A. Bispo e E. Pelaes, “The use of high-pass filters and the inpainting method to clouds removal and their impact on satellite images classification” - 16th International Conference on Image Analysis and Processing - ICIAP, 2011.

Referências Bibliográficas

- [1] K. McGuinness, *Image Segmentation, Evaluation, and Applications*, Ph.D. thesis, School of Electronic Engineering, 2009.
- [2] F. Sabine and W.H. Freeman, *Remote sensing, Principles and Interpretation*, Waveland Pr Inc; 3 edition, 2007.
- [3] M. A. T. Elmiro, “Técnicas digitais para produção de carta imagem,” in *VII Simpósio Brasileiro de Sensoriamento Remoto*, 1993, pp. 351–358.
- [4] C. A. Z. Toneli, L. M. T. Carvalho, and F. W. A. Junior, “Utilização das imagens Landsat TM nas épocas de verão e inverno para classificação de diferentes formações vegetais,” in *XIV Simpósio Brasileiro de Sensoriamento Remoto*, 2009, pp. 3127–3134.
- [5] A. Criminisi, P. Pérez, and K. Toyama, “Region filling and object removal by exemplar-based image inpainting,” in *IEEE Transactions On Image Processing*, 2004, vol. 13(9), pp. 1200–1212.
- [6] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley Publishing Company, 2008.
- [7] T. McInerney and D. Terzopoulos, “Deformable models in medical image analysis: A survey,” in *Medical Image Analysis*, 1996, vol. 1, pp. 91–108.
- [8] Y. Q. Chen, M. S. Nixon, and D.W. Thomas, “Texture classification using statistical geometrical features,” in *Pattern Recognition*, 1995, pp. 537–552.
- [9] C.-M Pun and H.-M Zhu, “Image segmentation using discrete cosine texture feature,” in *International Journal of Computers*, 2010, vol. 4(1), pp. 19–26.
- [10] A. T. Sapkal, C. Bokhare, and N. Z. Tarapore, “Satellite image classification using the Back Propagation algorithm of Artificial Neural Network,” in *Geomatrix- Conference on Geospatial Technologies and Applications*, 2009.

-
- [11] S. Yang, J. Wang, C.-I. Chang, J. L. Jensen, and J. O. Jensen, “Unsupervised image classification for remotely sensed imagery,” in *Proceedings of the SPIE - Society of Photo-Optical Instrumentation Engineers*, 2004, vol. 5546, pp. 354–365.
- [12] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing using Matlab*, Gatesmark Publishing., 2009.
- [13] R. P. Souto, “Segmentação de imagem multiespectral utilizando-se o atributo matiz. disponível em: www.obt.inpe.br/pgsere/souto-r-p-2000/publicacao.pdf,” Acessado em: 20 de Janeiro de 2011.
- [14] A. Richards, *Alien vision: exploring the electromagnetic spectrum with imaging technology*, SPIE Press, 2001.
- [15] B. Tso and P. Mather, *Classification methods for remotely sensed data*, CRC Press, 2009.
- [16] R. M. Castro, R. C. Moreira, E. S. C. Espósito, and E. V. D. Lucca, “Sensoriamento remoto como ferramenta para o ensino de Física,” in *XVI Simpósio Nacional de Ensino de Física*, 2005.
- [17] R. Peroni, “Princípios de sensoriamento remoto,” Tech. Rep., Faculdade do Noroeste de Minas, 2004.
- [18] L. E. Keiner, “Remote sensing figures. Disponível em: <http://kingfish.coastal.edu/physics/lkeiner/>,” Acessado em: 25 de Janeiro de 2011.
- [19] DPI/INPE, “SPRING: Tutorial de geoprocessamento,” Tech. Rep., INPE - Instituto Nacional de Pesquisas Espaciais, 2006.
- [20] Imagem, “Conceitos de imagens de satélite e sensoriamento remoto,” Tech. Rep., Imagem - Soluções de engenharia geográfica, 2011.
- [21] S. Amaral and A. M. V. Monteiro, “Abordagens espaciais em estudos de população: métodos analíticos e técnicas de representação,” Tech. Rep., Instituto Nacional de Pesquisas Espaciais - INPE, 2010.
- [22] J.R. Jensen, *Remote sensing of the environment: An Earth resource perspective*, Prentice Hall, 2006.
- [23] M. L. Latorre, “Utilização de um método de correção atmosférica para o processamento de dados hiperespectrais do sensor AVIRIS em regiões tropicais,” M.S. thesis, INPE - Instituto Nacional de Pesquisas Espaciais, 1998.

-
- [24] M. A. Moreira, “Sensoriamento remoto aplicado a estudos ambientais,” Tech. Rep., Instituto Nacional de Pesquisas Espaciais - INPE, 2003.
- [25] H. J. Kramer, *Observation of the earth and its environment: survey of missions and sensors*, Springer, 2002.
- [26] T. A. Warner, G. M. Foody, and M. D. Nellis, *The SAGE handbook of remote sensing*, SAGE Publications Ltd., 2009.
- [27] EURIMAGE, “Eurimage - multimission satellite data. Disponível em: <http://www.eurimage.com>,” Acessado em: 03 de fevereiro de 2011.
- [28] P.K. Varshney and M.K. Arora, *Advanced image processing techniques for remotely sensed hyperspectral data*, Springer, 2004.
- [29] J.A. Richards and X. Jia, *Remote sensing digital image analysis: an introduction*, Springer, 2006.
- [30] P.M. Mather, *Computer processing of remotely sensed images: an introduction*, John Wiley & Sons, 2004.
- [31] J. Gao, *Digital Analysis of Remotely Sensed Imagery*, McGraw Hill, 2009.
- [32] Fan W.L., Du H.Q., Zhou G.M., Xu X.J., Cui R.R., and Dong D.J., “Effects of atmospheric calibration on remote sensing estimation of Moso bamboo forest biomass,” *The journal of applied ecology*, vol. 21(1), pp. 1–8, 2010.
- [33] C. Song, C. E. Woodcock, K. C. Seto, M. Pax-Lenney, and S. A. Macomber, “Classification and change detection using Landsat TM data: When and how to correct atmospheric effects,” *Remote Sensing of Environment*, vol. 75, pp. 230–244, 2001.
- [34] X. Zhang, F. Qin, and Y. Qin, “Study on the thick cloud removal method based on multi-temporal remote sensing images,” in *International Conference on Multimedia Technology (ICMT)*, 2010, pp. 1–3.
- [35] A. Maalouf, P. Carre, B. Augereau, and C. Fernandez Maloigne, “A bandelet-based inpainting technique for clouds removal from remotely sensed images,” *IEEE Transactions On Geoscience and Remote Sensing*, vol. 47, no. 7, pp. 2363–2371, 2009.
- [36] N. T. Hoan and R. Tateishi, “Cloud removal of optical image using SAR data for ALOS applications. Experimenting on simulated ALOS data,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2008.

-
- [37] J. Feng, L. Bai, H. Tang, S. Liu, Q. Zhou, and J. Liu, "A new context-based procedure for the detection and removal of cloud shadow from moderate-and-high resolution satellite data over land," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2010, pp. 1718–1721.
- [38] C. Y. Hau, C. H. Liu, T. Y. Chou, and L. S. Yang, "The efficacy of semi-automatic classification result by using different cloud detection and diminution method," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2008.
- [39] S. Jayaraman, S. Esakkirajan, and T. Veerakumar, *Digital Image Processing*, McGraw-Hill Education (India) Pvt Ltd, 2009.
- [40] C. Loizou, C.S. Pattichis, and C. Pattichis, *Despeckle Filtering Algorithms and Software for Ultrasound Imaging*, Synthesis lectures on algorithms and software in engineering. Morgan & Claypool, 2008.
- [41] H.D. Cheng, X.H. Jiang, Y. Sun, and J. Wang, "Color image segmentation: advances and prospects," *Pattern Recognition*, vol. 34, no. 12, pp. 2259–2281, 2001.
- [42] R. Unnikrishnan and M. Hebert, "Measures of similarity," in *Proceedings of the Seventh IEEE Workshops on Application of Computer Vision*. 2005, vol. 01, IEEE Computer Society.
- [43] K. Bhoyar and O. Kakde, "Color image segmentation based on JND color histogram," in *International Journal of Image Processing (IJIP)*, 2010, vol. 3, pp. 282–293.
- [44] R. Unnikrishnan, C. Pantofaru, and M. Hebert, "A measure for objective evaluation of image segmentation algorithms," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR Workshops.*, June 2005.
- [45] K. Bhoyar and O. Kakde, "Colour image segmentation using fast Fuzzy C-means algorithm," in *Electronic Letters on Computer Vision and Image Analysis*, June 2010, vol. 1, pp. 18–31.
- [46] F. Y. Shih and S. Cheng, "Automatic seeded region growing for color image segmentation," in *Image and Vision Computing*, 2005, vol. 23, pp. 877–886.
- [47] Y.-L. Chen, T.-W. Chen, and S.-Y. Chien, "Fast texture feature extraction method based on segmentation for image retrieval," in *IEEE 13th International Symposium on Consumer Electronics*, 2009, pp. 941–942.

-
- [48] M. Mohamed and P. Gader, “Handwritten word recognition using segmentation-free Hidden Markov Modeling and segmentation-based dynamic programming techniques,” in *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1996, vol. 18, pp. 548–554.
- [49] L. Silva, D. Weingaertner, and O. R. P. Bellon, “Estudo sobre detecção de bordas em imagens de profundidade,” in *Brazilian Symposium on Computer Graphics and Image Processing*, 1997.
- [50] R. Gamanya, P. De Maeyer, and M. De Dapper, “An automated satellite image classification design using object-oriented segmentation algorithms: A move towards standardization,” in *Expert Systems with Applications*, 2007, vol. 32, pp. 616–624.
- [51] K. Johansen, N. C. Coops, S. E. Gergel, and Y. Stange, “Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification,” in *Remote Sensing of Environment*, 2007, vol. 110, pp. 29–44.
- [52] H. D. Cheng and Y. Sun, “A hierarchical approach to color image segmentation using homogeneity,” in *IEEE Transactions on Image Processing*, 2000.
- [53] M. Sezgin and B. Sankur, “Survey over image thresholding techniques and quantitative performance evaluation,” in *Journal of Electronic Imaging*, 2004, vol. 13(1), pp. 146–165.
- [54] S. Mansor, T. C. Chuang, A. R. M. Shariff, and N. Ahmad, “Multi-threshold technique for segmentation of satellite imagery for feature extraction,” Acessado em: 15 de fevereiro de 2011.
- [55] L. Rokach, *Pattern classification using ensemble methods*, Series in machine perception and artificial intelligence. World Scientific, 2010.
- [56] R. L. Santos, A. T. Ohata, N. Crusco, and J. A. Quintanilha, “A teoria da incerteza aplicada a classificação de imagens de satélites,” in *Congresso Brasileiro de Cartografia*, 2003.
- [57] F. A. Prado and M. L. B. T. Galo, “Classificação de imagens multiespectrais baseada em abordagem Fuzzy,” *Revista Brasileira de Cartografia*, 2009.
- [58] L. Samaniego and K. Schulz, “Supervised classification of agricultural land cover using a modified k-NN technique (MNN) and Landsat remote sensing imagery,” *Remote Sensing*, vol. 1, no. 4, pp. 875–895, 2009.

-
- [59] Y. F. Yang, P. Lohmann, and C. Heipke, "Genetic algorithms for the unsupervised classification of satellite images," in *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2011.
- [60] A. Mukhopadhyay, S. Bandyopadhyay, and U. Maulik, "Combining multiobjective Fuzzy clustering and probabilistic ANN classifier for unsupervised pattern classification: Application to satellite image segmentation," in *IEEE Congress on Evolutionary Computation*, 2008, pp. 877–883.
- [61] A. Mukhopadhyay and U. Maulik, "Unsupervised pixel classification in satellite imagery using multiobjective Fuzzy clustering combined with SVM classifier," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 4, pp. 1132–1138, 2009.
- [62] H. B. Kekre and S. Gharge, "SAR image segmentation using co-occurrence matrix and slope magnitude," in *Communication and Control International Conference on Advances in Computing*, 2009, pp. 368–372.
- [63] E. V. D. Lucca, C. C. Freitas, A. C. Frery, and S. J. S. Sant'Anna, "Comparison of SAR segmentation algorithms," in *Anais IX Simpósio Brasileiro de Sensoriamento Remoto*, 1998, pp. 1097–1108.
- [64] S. R. A. Neta, C. C. Freitas, and L. V. Dutra, "Uso de imagens ALOS/PALSAR multipolarizadas para detecção de incremento de desflorestamento na Amazônia," in *Revista Brasileira de Cartografia*, 2010, pp. 417–431.
- [65] K. Delac, M. Grgic, and T. Kos, "Sub-image Homomorphic filtering technique for improving facial identification under difficult illumination conditions," in *International Conference on Systems, Signals and Image Processing*, 2006, pp. 95–98.
- [66] O. Marques and H. Vieira, *Processamento Digital de Imagens*, BRASPORT, 1999.
- [67] H. B. Kekre, A. Athawale, and P. N. Halarnkar, "High payload using High Boost filtering in Kekre's Multiple LSB's algorithm," in *2nd International Conference on Advances in Computer Vision and Information Technology*, 2009.
- [68] H. Liu, W. Wang, and X. Bi, "Study of image Inpainting based on learning," in *Proceedings of The International MultiConference of Engineers and Computer Scientists*, 2010, pp. 1442–1445.
- [69] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 2000, pp. 417–424.

-
- [70] A. N. Htwe, “Image interpolation framework using non-adaptive approach and nl means,” *International Journal of Network and Mobile Technologies*, vol. 1, 2010.
- [71] D. Hale, “Image-guided blended neighbor interpolation of scattered data,” *79th Annual International Meeting, Society of Exploration Geophysicists*, vol. 28, pp. 1127–1131, 2009.
- [72] H. Greenspan, C. H. Anderson, and S. Akber, “Image enhancement by nonlinear extrapolation in frequency space,” *IEEE Transactions on Image Processing*, vol. 9, pp. 1035–1048, 2000.
- [73] P. J. Burt and E. H. Adelson, “The laplacian pyramid as a compact image code,” *IEEE Transactions On Communications*, vol. COM-31, pp. 1035–1048, 1983.
- [74] Astrophysics Research Centre, “Object programming - working with image objects,” Acessado em: 15 de fevereiro de 2011.
- [75] C.-M. Pun and H.-M. Zhu, “Textural image segmentation using Discrete Cosine Transform,” *International Conference On Communications And Information Technology*, pp. 54–58, 2009.
- [76] A. B. Watson, “Image compressing using the discrete cosine transform,” in *Mathematica Journal*, 1994, vol. 1, pp. 81–88.
- [77] P. K. Singh, “Unsupervised segmentation of medical images using DCT coefficients,” *Proceedings of the Pan-Sydney area workshop on Visual information processing*, pp. 75–81, 2005.
- [78] J. Wei, “Image segmentation based on situational DCT descriptors,” *Pattern Recognition Letters*, vol. 23, pp. 295–302, 2002.
- [79] J. A. C. Mangabeira, E. C. Azevedo, and R. A. C. Lamparelli, “Avaliação do levantamento do uso das terras por imagens de satélite de alta e média resolução espacial,” Tech. Rep., Empresa Brasileira de Pesquisa Agropecuária - Embrapa, 2003.
- [80] D. Salomon and G. Motta, *Handbook Of Data Compression*, Springer, 2009.
- [81] B. Bhanu and J. Peng, “Adaptive integrated image segmentation and object recognition,” *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and reviews*, vol. 30(4), pp. 427–441, 2000.
- [82] J.C. Russ, *The image processing handbook*, Image Processing Handbook Series. CRC/Taylor and Francis, 2007.

- [83] E. S. S. Lopes and H. S. Ferreira, “Tutorial SPRING 5.1,” in *Instituto Nacional de Pesquisas Espaciais*, 2009, vol. São José dos Campos, INPE.