

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

Uma abordagem híbrida e semiautomática para
estimativa de regiões cobertas por nuvens e sombras em
imagens de satélite: análise e avaliação

NOME DO AUTOR

Danilo Frazão Sousa

DM 08/2014

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2014

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

NOME DO AUTOR

Danilo Frazão Sousa

**Uma abordagem híbrida e semiautomática para
estimativa de regiões cobertas por nuvens e sombras em
imagens de satélite: análise e avaliação**

DM 08/2014

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2014

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

NOME DO AUTOR

Danilo Frazão Sousa

**Uma abordagem híbrida e semiautomática para
estimativa de regiões cobertas por nuvens e sombras em
imagens de satélite: análise e avaliação**

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 Computação Aplicada.

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
2014

Uma abordagem híbrida e semiautomática para estimativa de regiões cobertas por nuvens e sombras em imagens de satélite: análise e avaliação

Sousa, Danilo Frazão, 1990-

Uma abordagem híbrida e semiautomática para estimativa de regiões cobertas por nuvens e sombras em imagens de satélite: análise e avaliação;

Orientador: Evaldo Gonçalves Pelaes.-2014.

Dissertação (Mestrado) - Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, Belém, 2014.

1. Processamento de imagens - técnicas digitais. 2. Satélites artificiais em sensoriamento remoto.

I. Título.

CDD 22. ed. 621.367

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

Uma abordagem híbrida e semiautomática para estimativa de regiões cobertas por nuvens e sombras em imagens de satélite: análise e avaliação

AUTOR: DANILO FRAZÃO SOUSA

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 31/03/2014

BANCA EXAMINADORA:

.....
Prof. Dr. Evaldo Gonçalves Pelaes (ORIENTADOR - UFPA)
.....
Prof. Dr. Ronaldo Freitas Zampolo (MEMBRO - UFPA)
.....
Prof. Dr. Alessandra Rodrigues Gomes (MEMBRO - INPE)
.....
Prof. Dr. Johelden Campos Bezerra (MEMBRO - IFPA/IESAM)

VISTO:

.....
Prof. Dr. Evaldo Gonçalves Pelaes
COORDENADOR DO PPGEE/ITEC/UFPA

A Deus,
e à minha família.

Agradecimentos

Agradeço em primeiro lugar a Deus por ter me dado tudo o que tenho até hoje e iluminado meu caminho, mesmo nas horas mais difíceis.

Ao meu pai, minhas irmãs e, especialmente, à minha mãe, que sempre me guiou ao caminho certo, mesmo contra a minha vontade em algumas situações. A eles, todo o meu agradecimento por estarem ao meu lado em todos os momentos da vida, sempre me dando amor, apoio e forças para sempre seguir em frente.

Ao restante de minha família além de tios, tias, primos e primas pelo apoio sempre a mim dirigido.

À minha namorada, Isadora, pelo companheirismo, respeito, amor e carinho que me dá força extra para encarar o dia-a-dia, tornando, assim, a minha vida mais especial.

A todos os amigos, companheiros de casa e professores que estiveram ao meu lado nessa jornada, e que ainda estarão nas próximas etapas da vida.

A todos que diretamente ou indiretamente contribuíram para a finalização desta fase da minha vida, o meu imenso muito obrigado!

Resumo

Os principais objetivos deste trabalho são propor um algoritmo eficiente e o mais automático possível para estimar o que está coberto por regiões de nuvens e sombras em imagens de satélite; e um índice de confiabilidade, que seja aplicado previamente à imagem, visando medir a viabilidade da estimação das regiões cobertas pelos componentes atmosféricos usando tal algoritmo. A motivação vem dos problemas causados por esses elementos, entre eles: dificultam a identificação de objetos de imagem, prejudicam o monitoramento urbano e ambiental, e desfavorecem etapas cruciais do processamento digital de imagens para extrair informações ao usuário, como segmentação e classificação. Através de uma abordagem híbrida, é proposto um método para decompor regiões usando um filtro passa-baixas não-linear de mediana, a fim de mapear as regiões de estrutura (homogêneas), como vegetação, e de textura (heterogêneas), como áreas urbanas, na imagem. Nessas áreas, foram aplicados os métodos de restauração *Inpainting* por suavização baseado em Transformada Cosseno Discreta (DCT), e Síntese de Textura baseada em modelos, respectivamente. É importante salientar que as técnicas foram modificadas para serem capazes de trabalhar com imagens de características peculiares que são obtidas por meio de sensores de satélite, como por exemplo, as grandes dimensões e a alta variação espectral. Já o índice de confiabilidade, tem como objetivo analisar a imagem que contém as interferências atmosféricas e daí estimar o quão confiável será a redefinição com base no percentual de cobertura de nuvens sobre as regiões de textura e estrutura. Tal índice é composto pela combinação do resultado de algoritmos supervisionados e não-supervisionados envolvendo 3 métricas: Exatidão Global Média (EGM), Medida De Similaridade Estrutural (SSIM) e Confiança Média Dos Pixels (CM). Finalmente, verificou-se a eficácia destas metodologias através de uma avaliação quantitativa (proporcionada pelo índice) e qualitativa (pelas imagens resultantes do processamento), mostrando ser possível a aplicação das técnicas para solucionar os problemas que motivaram a realização deste trabalho.

PALAVRAS-CHAVE: Processamento de imagens, Sensoriamento Remoto, Remoção de nuvens e sombras, Síntese de Textura, *Inpainting*.

Abstract

The main goals of this work are to propose a more automatic and efficient algorithm to replace regions of clouds and shadows in satellite images as well as an index of reliability that is previously applied to each image, in order to measure the feasibility of the estimation of the regions covered by atmospheric components using that algorithm. The motivation comes from the problems caused by these atmospheric elements, among them: to impede the identification of objects of the image, to make the urban and environmental monitoring more difficult, and to interfere in crucial stages of digital image processing to extract information for the user, such as segmentation and classification. Through a hybrid approach is proposed a method for decomposing regions using a median non-linear low-pass filter, in order to map the regions of structure (homogeneous) and texture (heterogeneous) in the image. In these areas was applied restoration methods Inpainting by Smoothing based on Discrete Cosine Transform (DCT), and Exemplar-Based Texture Synthesis, respectively. It's important to note that the techniques have been modified to be able to work with images obtained through of satellite sensors with peculiar features such as large size and/or high spectral variation. Regarding to the reliability index, it aims to analyze the image that contains atmospheric interference and hence estimate how much reliable will be the redefinition, based on the percentage of cloud cover over the regions of texture and structure. This index is composed by combining the result of supervised and unsupervised algorithms involving three metrics: Average of Accuracy Global, Measure Of Structural Similarity (SSIM) and Average of Pixels Confidence. Finally, it was verified the effectiveness of these methods through a quantitative assessment (provided by the index) and qualitative (the images resulting from processing), showing the possible application of the techniques to solve the problems that motivated this work.

KEYWORDS:Image Processing, Remote Sensing, Clouds and Shadows Removal, Texture Synthesis, Inpainting.

Sumário

Lista de Figuras	ii
Lista de Tabelas	iii
Abreviações e Acrônimos	iv
1 Introdução	1
1.1 Contribuição e Organização do Trabalho	4
2 Fundamentos	6
2.1 Processamento Digital de Imagens	6
2.2 Textura	8
2.3 Componentes Estruturados da Imagem	9
2.4 Sensoriamento remoto	10
2.5 Transformada Cosseno Discreta	12
2.6 Abordagem Supervisionada e Não-supervisionada	13
2.7 Segmentador SRM	13
2.8 Média, Desvio Padrão e Coeficiente de Variação	14
3 Técnicas	16
3.1 Detecção de Nuvens e Sombras	16
3.2 Inpainting por suavização baseado em DCT	17
3.3 Síntese de Textura	18
3.4 Decomposição de Regiões	19
3.5 Mapeamento das regiões de estrutura e de textura	20
3.5.1 Adequação das técnicas para imagens de satélite	21

3.6	Abordagem Híbrida	22
3.7	Índice de Confiabilidade	23
3.7.1	Contaminação Artificial	24
3.7.2	Tamanho relativo das nuvens e sombras	24
3.7.3	Separação por classe	25
3.7.4	Avaliando pela Exatidão Global	27
3.7.5	Avaliando pelo SSIM	28
3.7.6	Avaliando pela Confiança média dos pixels	33
3.7.7	O Índice	34
4	Resultados e Discussão	35
4.1	Materiais	35
4.2	EGM para cada classe	36
4.3	SSIM para cada classe	37
4.4	CM para cada classe	38
4.5	IC médio para cada classe	39
4.6	Amostra das etapas de processamento e dos índices	40
5	Conclusões	44
5.1	Conclusão	44
5.2	Trabalhos Publicados	45
5.3	Trabalhos Futuros	45
6	Apêndice: Detecção de Nuvens e Sombras em Imagens de Satélite usando Algoritmo Genético	46
6.1	Detecção de Regiões	46
6.2	Aplicação do AG	47
6.3	Avaliação usando a métrica EBC	48
6.4	Resultados e Conclusões	49
	Bibliografia	58

Lista de Figuras

2.1	Representação de uma imagem no seu eixo de coordenadas espaciais. Fonte: [16].	7
2.2	Os componentes iluminação (I) e reflectância (R) em relação a uma imagem. Fonte: [16].	7
2.3	Imagem contendo textura lisa e rugosa. Fonte: http://www.satimagingcorp.com/gallery/geoeye-1-kutztown.html	9
2.4	Processo de aquisição das imagens de satélite: (A) fonte de energia ou iluminação, (B) radiação da atmosfera, (C) interação com o alvo, (D) gravação de energia pelo sensor, (E) transmissão, recepção e processamento, (F) interpretação e análise e (G) aplicação. Fonte: CCRS/CCT.	11
2.5	Espectro eletromagnético. Fonte: [20]	11
2.6	Sensibilidade à detecção de determinados objetos, como rochas e vegetação, em relação a um comprimento de onda específico. Fonte: [20]	12
3.1	Síntese de textura baseada em modelos: (a) Imagem original. (b) Fragmento Ψ_p centralizado em $p \in \Phi$. (c) Candidatos mais prováveis $\Psi_{q'}$ e $\Psi_{q''}$. (d) Candidato mais provável é propagado para o fragmento alvo.	18
3.2	Fluxograma do algoritmo híbrido de estimação de nuvens e sombras de imagens de satélite proposto neste trabalho.	23
3.3	(a) Ilustração da forma que o algoritmo varre a vizinhança de um <i>pixel</i> . (b) Exemplo da forma que o algoritmo varre uma imagem binária em busca de <i>pixels</i> ruído.	25
3.4	Matriz de Confusão.	27
3.5	Fluxograma do algoritmo de treinamento para avaliação pela EG.	28
3.6	Diagrama do sistema de medição de similaridade estrutural (SSIM). Fonte: [33].	29

3.7	Fluxograma do algoritmo de treinamento para avaliação pelo SSIM.	32
4.1	Etapas do processamento para a Imagem 1 na Classe 1 (Textura e Estrutura excelentes), que obteve os índices de EG=0.9130, SSIM=0.8615, CM=0.9830 e IC=0.9095. (a) Imagem original; (b) Imagem decomposta; (c) Imagem Segmentada antes do processamento; (d) Imagem resultante da contaminação artificial da classe 1; (e) Resultado da abordagem híbrida; (f) Imagem segmentada após o processamento.	41
4.2	Etapas do processamento para a Imagem 3 na Classe 5, que obteve os índices de EG=0.7329, SSIM=0.8297, CM=0.9204 e IC=0.8605. (a) Imagem original; (b) Imagem decomposta; (c) Imagem Segmentada antes do processamento; (d) Imagem resultante da contaminação artificial da Classe 5; (e) Resultado da abordagem híbrida; (f) Imagem segmentada após o processamento.	42
4.3	Etapas do processamento para a Imagem 7 na Classe 9, que obteve os índices de EG=0.7935, SSIM=0.9101, CM=0.8544 e IC=0.8150. (a) Imagem original; (b) Imagem decomposta; (c) Imagem Segmentada antes do processamento; (d) Imagem resultante da contaminação artificial da Classe 9; (e) Resultado da abordagem híbrida; (f) Imagem segmentada após o processamento.	43
6.1	Imagem “Belém 1”:Parâmetros ótimos $fn = 0.68$ e $fs = 0.79$ que levam à $EBC = 0.74$; (a) Imagem Original; (b) Melhores, piores e média dos indivíduos de cada geração; (c) Máscara; (d) Regiões de nuvens e sombras detectadas na imagem.	50
6.2	Imagem “Ilhas de Belém”:Parâmetros ótimos $fn = 0.94$ e $fs = 4.27$ que levam à $EBC = 0.86$; (a) Imagem Original; (b) Melhores, piores e média dos indivíduos de cada geração; (c) Máscara; (d) Regiões de nuvens e sombras detectadas na imagem.	51
6.3	Imagem “Belém 2”:Parâmetros ótimos $fn = 0.82$ e $fs = 0.82$ que levam à $EBC = 0.9$; (a) Imagem Original; (b) Melhores, piores e média dos indivíduos de cada geração; (c) Máscara; (d) Regiões de nuvens e sombras detectadas na imagem.	52

6.4	Imagem “Mina”:Parâmetros ótimos $fn = 1.26$ e $fs = 0.48$ que levam à $EBC = 0.84$; (a) Imagem Original; (b) Melhores, piores e média dos indivíduos de cada geração; (c) Máscara; (d) Regiões de nuvens e sombras detectadas na imagem.	53
6.5	Imagem “São Paulo”:Parâmetros ótimos $fn = 0.46$ e $fs = 3.56$ que levam à $EBC = 0.82$; (a) Imagem Original; (b) Melhores, piores e média dos indivíduos de cada geração; (c) Máscara; (d) Regiões de nuvens e sombras detectadas na imagem.	54

Lista de Tabelas

3.1	Relação entre Porcentagem de cobertura e Qualidade da Região.	26
3.2	Relação entre as Classes e Qualidade das Regiões de Estrutura e Textura. . . .	26
4.1	Imagens originais.	36
4.2	Média, desvio padrão (DP) e coeficiente de variação (CV) da EG para cada classe.	37
4.3	Média, Desvio Padrão (DP) e coeficiente de variação (CV) do SSIM para cada classe.	38
4.4	Média da CM para cada classe.	39
4.5	Média do IC para cada classe.	39

Abreviações e Acrônimos

AG	-	<i>Algoritmo Genético</i>
CV	-	<i>Coefficiente de Variação</i>
CM	-	<i>Confiança Média</i>
DP	-	<i>Desvio Padrão</i>
DCT	-	<i>Discrete Cosine Transform</i>
EBC	-	<i>Edge-border coincidence - coincidência entre bordas</i>
fn	-	<i>fator de nuvem</i>
fs	-	<i>fator de sombra</i>
IC	-	<i>Índice de Confiabilidade</i>
IDCT	-	<i>Inverse Discrete Cosine Transform</i>
EGM	-	<i>Exatidão Global Média</i>
MS1	-	<i>Multi-espectral 1</i>
MS2	-	<i>Multi-espectral 2</i>
MSE	-	<i>Mean Square Error - erro médio quadrático</i>
PSNR	-	<i>Peak Signal-to-noise ratio - razão sinal-ruído de pico</i>
QB02	-	<i>QuickBird 02</i>
REM	-	<i>Radiação Eletromagnética</i>
RGB	-	<i>Espaço de cores Red (vermelho), Green (verde) e Blue (azul)</i>
SAR	-	<i>Synthetic-aperture radar</i>
SRM	-	<i>Statistical Region Merging</i>
SSIM	-	<i>Medida de Similaridade Estrutural (Measure of Structural Similarity)</i>
WV02	-	<i>World View 02</i>

Capítulo 1

Introdução

Em sensoriamento remoto óptico é comum a presença de interferências atmosféricas, tais como nuvens e neblinas, ou mesmo sombras consequentes destes elementos, durante a captação de uma cena por um sensor. Essas ocorrências estão presentes em várias regiões de clima equatorial e tropical quente e semi-úmido, e podem dificultar a identificação de objetos da superfície da terra através de imagens, afetando atividades como o monitoramento ambiental e urbano e as etapas subsequentes do processamento digital de imagens, como a segmentação e classificação, responsáveis pela extração de informações da cena.

Especialistas em todo o mundo têm realizado diversos estudos com o objetivo de encontrar uma solução eficaz para reduzir ou eliminar o impacto causado pelas nuvens e sombras. É importante perceber que as técnicas usadas para o processamento de imagens de satélite devem ser, na mesma medida, robustas e computacionalmente viáveis, devido às grandes dimensões e variações espectrais desse tipo de imagem. Isso tem incentivado o desenvolvimento de alternativas para melhorar as técnicas existentes nesta área.

A remoção e a redefinição das regiões de nuvens e sombras são feitas basicamente através de duas abordagens: utilizando imagens de referência, com conhecimento *a priori*, em que se tem as informações reais das regiões cobertas, podendo substituí-las pelas áreas correspondentes; ou estimando as regiões ocultadas por estes elementos, quando não se tem nenhuma outra informação *a priori* daquela cena, através de algoritmos automáticos ou semi-automáticos.

A primeira estratégia pode ser realizada fazendo uma análise multi-temporal [1], ou usando uma cena captada por um sensor diferente [2], que usou uma imagem SAR (*Synthetic-aperture radar*). Esta abordagem é mais confiável devido à ideia de que é provável que as regiões da imagem não tenham mudado significativamente entre as datas de imageamento, mas ainda assim haverá sempre uma probabilidade de erro principalmente

se as datas forem muito distantes. Por outro lado, muitas vezes esta operação não pode ser realizada, pelo fato de não se ter em mãos imagens multitemporais da mesma cena, problema recorrente de usuários que acabam abandonando estudos sobre determinadas regiões.

Já para estimar e redefinir as regiões usando apenas informações da cena atual e a “inteligência” de procedimentos computacionais, pode ser usado um método de interpolação como o *inpainting* [3], que tem por objetivo preencher (ou restaurar) regiões danificadas e remover objetos, sem que a operação seja perceptível, o qual funciona bem para regiões homogêneas. Uma variação desta técnica, usando a transformada *Bandelet*, para remover nuvens em imagens de satélite é proposta em [4]. Outro método utilizado para estimar regiões é a síntese de textura, que geralmente trabalha em blocos e alcança bons resultados para áreas que contém algum padrão textural ou que apresentam heterogeneidade. Em [5], é usada uma abordagem de campos hierárquicos de árvores estruturadas (*Tree-Structured Hierarchical Fields*) para modelar o comportamento complexo de dados operacionais de imagens *RADARSAT SAR* de mares e geleiras. Este tipo de abordagem é mais propensa a erros, no entanto, com a utilização de algoritmos eficazes e dependendo do contexto de aplicação, pode redefinir uma região com baixa probabilidade de erro.

Alguns estudos têm explorado soluções híbridas que preservam a eficácia da síntese de textura e do *inpainting* sobre o mesmo processamento. Nesse contexto, cada abordagem é responsável por redefinir as regiões de textura (heterogênea) e estrutura (homogênea) separadamente, garantindo imagens sem componentes atmosféricos e com texturas reais.

Para efeito de conhecimento, os dois objetivos de mais alto nível que se tem interesse com a aplicação de técnicas de mineração de dados são a predição e a descrição [6]. Os padrões do tipo preditivo são os que procuram predizer valores futuros de um ou mais atributos variáveis do banco de dados com base em valores conhecidos de outros atributos. Os do tipo descritivo, por sua vez, procuram descrever padrões encontrados nos dados em um formato que seja interpretável pelo homem [7].

A área de predição em mineração de dados é consolidada, tendo inúmeros trabalhos publicados ao redor do mundo. Além disso, a maioria dos livros sobre o assunto trazem capítulos abordando diferentes técnicas para realizar este tipo de operação. Dado isto, qual a grande diferença entre a predição de dados de imagens para qualquer outro tipo de dado? Como qualquer outra aplicação, pode-se extrair diversas características, variáveis e padrões; e aplicar-se algoritmos de inteligência computacional para predizer elementos ainda desconhecidos, com uma dada probabilidade de erro, podendo ou não se tornar válido o uso desta abordagem.

De fato, instituições como o INPE (Instituto de Pesquisas Espaciais) têm interesse

maior em melhorar a detecção de nuvens, já que a remoção é feita de forma multitemporal usando uma grande base de dados existente. Porém, nem toda instituição tem acesso a dados multitemporais de uma cena. Logo, se existem nuvens e sombras, já são desconsideradas parte das informações. Por isso, na literatura já foram apresentados alguns trabalhos usando diferentes técnicas de *inpainting* com o objetivo de remover e substituir nuvem com as informações da própria imagem processada.

A aplicação deste tipo de método está relacionada também ao risco que pode ser assumido em um determinado trabalho. Por exemplo, para processar imagens médicas talvez esta não fosse uma boa solução, assim como para a detecção de objetos muito pequenos e peculiares proporcionalmente ao tamanho da imagem. No entanto, para extração de características mais gerais, como a porcentagem de vegetação de uma floresta, talvez pudesse ser usado com baixa probabilidade de erro.

No que diz respeito à avaliação e validação, geralmente os artigos que tratam da remoção automática de componentes atmosféricos não entram no mérito de que essa abordagem é sim possível de ser usada em determinados casos ou mesmo impedindo o uso das mesmas em aplicações reais. Esses trabalhos apenas apresentam suas técnicas de redefinição de regiões e em alguns casos comparam com a de outros autores.

Na literatura ainda não existem técnicas em estado de arte no que diz respeito à estimação de erros, medição da qualidade ou de confiabilidade de operações de redefinição de regiões, onde aqui generaliza-se como *inpainting*.

Em [8] foi feito um estudo de erro em relação ao *inpainting*. Intuitivamente, se mais pixels estão ausentes, pode-se esperar um erro maior. No entanto, esta análise mostra que o erro depende mais da geometria e da forma, do que o tamanho ou área total do domínio do *inpainting*.

Em [9] afirma-se que textura, estrutura, movimento (para vídeo), o tamanho da região indefinida e outros atributos geométricos, possuem um papel crucial na determinação da eficácia de um algoritmo de *inpainting*. Ainda segundo o autor, a maioria dessas técnicas comparam o desempenho de seu algoritmo com outros métodos existentes, destacando a eficácia em termos de sua capacidade de lidar com grandes áreas de preenchimento, de redefinir estruturas curvilíneas e regiões com alta frequência de bordas, de replicação de textura e do tempo necessário para concluir a operação. Apesar da natureza inerentemente subjetiva dos resultados deste tipo de método, eles geralmente têm avaliações feitas por métricas objetivas simples, como a relação sinal-ruído de pico (PSNR), o erro médio quadrático (MSE) ou apenas pela comparação visual. Ainda segundo [9], é de amplo conhecimento que esses tipos de avaliação não são adequados para caracterizar a qualidade perceptual da imagem. Portanto,

em tal trabalho foi proposta uma técnica de avaliação baseada em testes feitos levando em consideração a análise de indivíduos acerca dos resultados.

Na tentativa de resolver a situação de não existir métricas quantitativas para avaliação de *inpainting*, em [10] são propostas duas novas métricas para capturar as noções de informação e intenção visual, a fim de avaliar as redefinições de regiões. As métricas propostas usam medida quantitativa da importância visual baseada em um modelo computacional da atenção visual humana.

Em [11] é proposto um método para prever a resposta humana a um processamento de imagem onde os blocos contíguos de pixels foram substituídos. Estimando alocação da atenção humana para imagens pré- e pós-editadas, foram calculadas pontuações quantitativas que resumem o aparecimento de artefatos visuais incoerentes. As experiências indicaram que esses índices se correlacionam fortemente com a opinião qualitativa humana e, portanto, servem como uma base suficiente para a formação de algoritmos de classificação automáticos que imitam a avaliação humana.

Em [12] são apresentadas duas novas métricas de qualidade de *inpainting*, que juntamente com outras métricas de qualidade (inclusive as de [10] e [11]), foram avaliadas em relação à observadores humanos. Dois tipos de avaliação foram efetuados: investigando o desempenho das métricas sobre toda a base de dados de imagens, e outra avaliando a correlação das métricas para imagens individuais. Os resultados mostram que nenhuma das métricas consideradas podem prever adequadamente a qualidade do *inpainting* sobre toda a base de dados, e que a performance das métricas é dependente da imagem.

1.1 Contribuição e Organização do Trabalho

Este trabalho apresenta duas contribuições principais:

1. Apresenta uma nova abordagem híbrida usando *inpainting* por suavização baseado em DCT, proposta por [13], e síntese de textura baseada em exemplos, proposta por [14], ambas com modificações necessárias para o processamento de imagens de satélite. A união das técnicas é permitida por um processo de mapeamento de textura e estrutura baseado em [15], aliado a filtragem não-linear, que permite definir o contexto em que se encontram os elementos a serem removidos.
2. Propor um índice de confiabilidade para um algoritmo de remoção e redefinição de regiões de nuvens e sombras em imagens de satélite, tendo como objetivo analisar a imagem que contém as interferências atmosféricas e daí estimar o quanto confiável será a redefinição

para cada imagem isoladamente. Tal índice é composto pelo resultado de algoritmos supervisionados e não-supervisionados envolvendo 3 métricas: EG, SSIM e confiança média dos pixels.

Este trabalho está organizado da seguinte forma. No Capítulo 2 são descritos e explanados termos comuns ao contexto do processamento digital de imagens e sensoriamento remoto, assim como fundamentos essenciais para a compreensão das metodologias utilizadas neste trabalho. No Capítulo 3 é apresentada a teoria das principais técnicas de processamento utilizadas; onde são abordados os processos de detecção de nuvens e sombras, técnicas de *inpainting*, síntese de textura, decomposição de regiões, o algoritmo híbrido que tratará de estimar as áreas de nuvens densas e sombras das imagens de satélite e o índice de confiabilidade da redefinição dessas regiões. No Capítulo 4 são mostrados os resultados, e uma breve discussão acerca dos resultados obtidos a partir da aplicação dos métodos propostos neste trabalho. Finalmente, no Capítulo 5 é mostrada uma visão geral do trabalho, das abordagens propostas, uma síntese com resultados alcançados e explicações acerca das metodologias empregadas, bem como propostas de trabalhos futuros.

Capítulo 2

Fundamentos

Neste capítulo, são apresentadas a descrição e explanação de termos comuns no contexto do processamento digital de imagens e sensoriamento remoto, e os fundamentos para a compreensão das técnicas utilizadas neste trabalho.

2.1 Processamento Digital de Imagens

A área de processamento de imagens vem sendo objeto de crescente interesse, e tem se tornado mais popular nas últimas décadas, por permitir viabilizar um grande número de aplicações em duas categorias distintas: o aprimoramento de informações pictóricas para interpretação humana, e a análise automática de informações extraídas de uma cena [16].

O processamento digital de imagens pode ser caracterizado pela realização de diversas tarefas, dentre elas: (1) aquisição da imagem, que compreende a recepção (através de sensores, desde simples máquinas digitais até sensores de satélites), e o armazenamento da mesma no formato de um arquivo digital; (2) pré-processamento, que trata de formatar a imagem, corrigir imperfeições ocasionadas durante a formação da cena (tal como remoção de ruído); (3) processamento digital, que inclui a gama de técnicas atualmente disponibilizadas pela literatura, como a segmentação e a classificação, empregadas para que finalmente seja realizada a (4) extração de informações relevantes ao interesse do usuário.

No entanto, não se pode entender o processamento digital sem antes ter conhecimento do que são as imagens digitais em si, e, para isso, tem-se de ter em mente o conceito de *pixel*, que é um número inteiro que representa (em uma dada escala de níveis de cinza) uma determinada área na superfície. As imagens digitalizadas, portanto, consistem basicamente num arquivo digital com os dados captados pelo sensor de uma certa cena da vida real. Deste

modo, estes dados registrados da imagem são resumidos no espaço tridimensional formado pelo plano (x,y) e pelo brilho ou radiância do *pixel* $f(x,y)$, como mostrado na Figura 2.1.

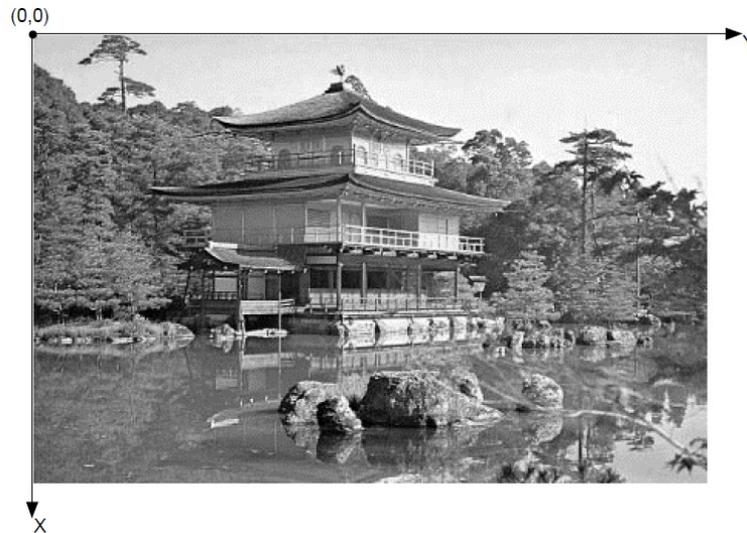


Figura 2.1: Representação de uma imagem no seu eixo de coordenadas espaciais. Fonte: [16].

A função $f(x,y)$ representa o produto da interação entre a iluminação $i(x,y)$ - que exprime a quantidade de luz que incide sobre o objeto - e as propriedades de reflectância ou de transmitância próprias do objeto, que podem ser representadas pela função $r(x,y)$, cujo valor exprime a fração de luz incidente que o objeto vai transmitir ou refletir no ponto (x,y) . Estes conceitos são ilustrados por [16], e podem ser visualizados na Figura 2.2 e expresso como na Equação 2.1.

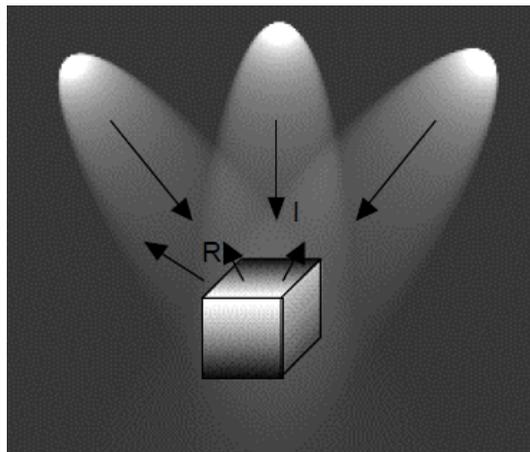


Figura 2.2: Os componentes iluminação (I) e reflectância (R) em relação a uma imagem. Fonte: [16].

$$f(x,y) = i(x,y)r(x,y), \quad (2.1)$$

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

No caso de uma imagem que possui informações em intervalos ou bandas distintas de frequência, é necessária uma função $f(x, y)$ para cada banda. É o caso de imagens coloridas padrão *RGB*, que são formadas pela informação de cores primárias aditivas: vermelho (*R - Red*), verde (*G - Green*) e azul (*B - Blue*).

2.2 Textura

Muitos algoritmos de processamento de imagens utilizam a informação sobre a uniformidade de intensidade dos pixels para proceder as operações desejadas. No entanto, imagens de objetos reais muitas vezes não apresentam regiões de intensidades uniformes. Por exemplo, a imagem de uma superfície de madeira não é uniforme, mas contém variações de intensidades que formam certos padrões repetidos chamados de textura visual. Os padrões podem ser o resultado de propriedades da superfície física, tais como rugosidade ou filamentos orientados, que muitas vezes têm uma qualidade tátil, ou eles podem ser o resultado de diferenças de reflectância, como a cor sobre uma superfície [17].

Embora se possa associar intuitivamente várias propriedades da imagem, tais como suavidade, aspereza, profundidade, regularidade, com textura, não há uma definição formal ou completa de textura [18]. Muitos pesquisadores têm descrito textura usando várias definições, em geral, a textura é definida como o arranjo dos objetos, bem como a frequência da sua variação de tons que se verifica em certas áreas da imagem. Quando essas variações são pequenas, tem-se uma textura lisa, como nos gramados, solo nu, água, etc. Já quando as variações são abruptas, tem-se a textura dita rugosa, como acontecem com a copa das árvores, áreas urbanas, por exemplo. Na Figura 2.3, são perceptíveis os dois tipos de textura citados: rugosa, quando são observadas as regiões de copas das árvores e telhados de construções, e lisa, quando são observadas as regiões de grama e asfalto.

Portanto, a textura da imagem, definida como uma função da variação espacial em intensidades de *pixel* (níveis de cinza) é útil em uma variedade de aplicações e sua análise tem sido objeto de intenso estudo por muitos pesquisadores [17].

Dentre as aplicações da análise de textura da imagem, tem-se o reconhecimento de regiões da imagem usando as propriedades de textura, chamada de classificação de textura, em que o objetivo é produzir um mapa de classificação da imagem de entrada, onde cada região uniforme de textura é identificada com uma classe de textura a qual pertence. Pode-se também encontrar os limites de textura mesmo sem classificar estas superfícies texturizadas,



Figura 2.3: Imagem contendo textura lisa e rugosa. Fonte: <http://www.satimagingcorp.com/gallery/geoeye-1-kutztown.html>

a isso se dá o nome de segmentação de textura, que tem como objetivo principal obter um mapa de bordas de alta frequência da imagem. Da mesma forma, uma aplicação bastante empregada é a síntese de textura, aplicada principalmente na compressão de imagens, e que neste trabalho é utilizada para redefinição/restauração de regiões deterioradas da imagem texturizada (descrito na seção 3.3). Uma descrição mais detalhada de técnicas relacionadas à análise de textura é dada em [17].

2.3 Componentes Estruturados da Imagem

De um modo geral, podem-se distinguir as descrições de uma imagem entre quantitativas e qualitativas. A descrição quantitativa descreve uma imagem por meio de uma função sobre algum domínio, por exemplo, valores de *pixel*, coeficientes de Fourier e *pixels* que definem a textura da imagem. Já a descrição qualitativa, ou estrutural, identifica primitivas da imagem firmes e estáveis, e inter-relacionamentos entre elas, por relações como “é parte de” e “é adjacente a”.

Uma característica das partes estruturadas da imagem é o elevado nível alcançado de invariância às transformações não muito importantes na imagem. Por exemplo, as relações de aninhamento dos contornos de brilhos iguais em uma imagem (isófotos) mantêm-se inalteradas por deformações arbitrárias de suavização no espaço da imagem e transformações monotônicas de luminância [19], ou seja, após o processamento, o agrupamento das bordas deve permanecer o mesmo. Este fator torna os componentes estruturados mais conservadores em relação às descrições quantitativas e qualitativas.

2.4 Sensoriamento remoto

Por sensoriamento remoto entende-se o processo de captação de informações dos fenômenos e feições terrestres, permitindo uma visão sinóptica das unidades de paisagem existentes na superfície, possibilitando interpretação e classificação sem que haja um contato físico com as mesmas, diminuindo significativamente o tempo de pesquisa e o custo de sua aplicação. Por este motivo o sensoriamento remoto tem sido largamente usado para a análise geo-ambiental, auxiliando no monitoramento multitemporal das transformações impostas a essas áreas devido ações antrópicas e/ou naturais [20]. Como tratado pela literatura, sensoriamento remoto é equivalente à aquisição de cenas através sensores acoplados a satélites orbitais.

De um modo geral, a radiação eletromagnética (REM) da superfície da terra é captada e registrada por sensores ópticos instalados a bordo de satélites artificiais (com sensores que podem ser ativos, como os radares, ou passivos, que não emitem radiação, só as recebe) e transmitida digitalmente em tempo real para instalações no solo [21] (um exemplo é uma estação situada em Cachoeira Paulista-SP, onde fica o centro de processamento e distribuição das imagens fornecidas pelo *INPE*¹), que a transformam em uma imagem digital de satélite, pronta para um posterior processamento e adequada extração de informações. A Figura 2.4 descreve o processo de aquisição das imagens de satélite (que estão disponíveis gratuitamente ou comercialmente na internet) pelas seguintes etapas: (A) fonte de energia ou iluminação, (B) radiação da atmosfera, (C) interação com o alvo, (D) gravação de energia pelo sensor, (E) transmissão, recepção e processamento, (F) interpretação e análise e (G) aplicação.

É importante ter em mente também os conceitos de resolução radiométrica, resolução temporal, e resolução espectral da imagem. A primeira determina quantos bits serão usados para representar um *pixel*, por exemplo, se usados 8 bits, então cada *pixel* poderá ser representado por 256 cores, variando de 0 a 255; a segunda corresponde a quanto tempo (contado em dias) que o satélite leva para dar uma volta completa na terra e passar pelo mesmo ponto da superfície.

A definição de resolução espectral de uma imagem varia de acordo com a resposta dos objetos à REM. Como explicado em [21], a radiação eletromagnética é aquela que se propaga na forma de ondas eletromagnéticas através do espaço. Estas ondas são perturbações periódicas de campos elétricos e magnéticos originados por uma fonte energética, sendo a fonte mais conhecida, o Sol. Todas as ondas eletromagnéticas viajam através do espaço na mesma velocidade, 300.000 km/s (velocidade da luz no vácuo), e são caracterizadas pelo seu

¹Instituto Nacional de Pesquisas Espaciais url: <http://www.inpe.br>

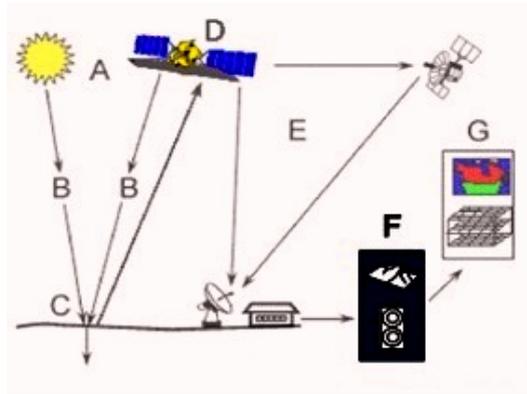


Figura 2.4: Processo de aquisição das imagens de satélite: (A) fonte de energia ou iluminação, (B) radiação da atmosfera, (C) interação com o alvo, (D) gravação de energia pelo sensor, (E) transmissão, recepção e processamento, (F) interpretação e análise e (G) aplicação. Fonte: CCRS/CCT.

comprimento de onda e energia associada. As diferenças entre as ondas eletromagnéticas dependem basicamente da fonte energética que originou a onda. O espectro eletromagnético é visto na Figura 2.5.

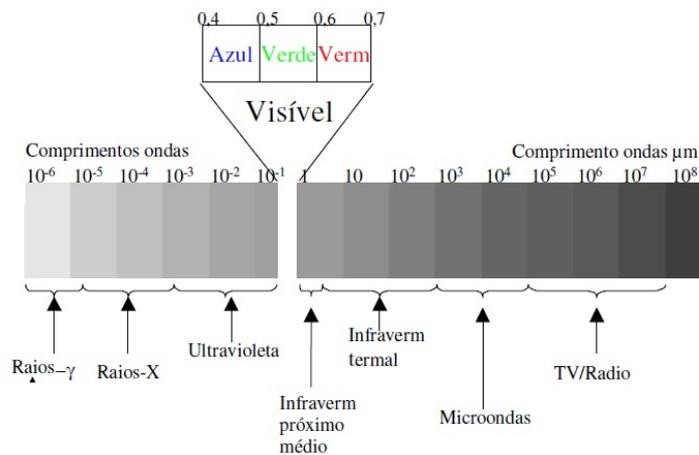


Figura 2.5: Espectro eletromagnético. Fonte: [20]

O comportamento espectral ou a assinatura espectral de um alvo está relacionado ao processo de interação entre os objetos e a REM incidente, e dependem da estrutura atômica e molecular dos alvos. Os elétrons dos materiais estão distribuídos em diferentes níveis energéticos, em torno dos núcleos de seus átomos. Estes níveis eletrônicos podem absorver maior ou menor quantidade da energia da REM. Esta absorção implica na diminuição da quantidade de energia da REM refletida pela matéria, em certas faixas do espectro eletromagnético, faixas estas denominadas bandas de absorção. Estes fatores fazem com que cada alvo terrestre tenha sua própria assinatura espectral [20]. Em outras palavras, cada alvo

absorve ou reflete de modo diferente cada uma das faixas do espectro da luz incidente, como na Figura 2.6, em que é mostrada a porcentagem de sensibilidade à detecção de determinados objetos, como rochas e vegetação, em relação a um comprimento de onda específico.

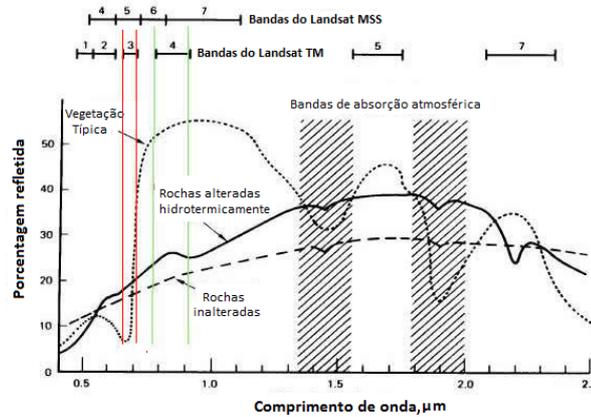


Figura 2.6: Sensibilidade à detecção de determinados objetos, como rochas e vegetação, em relação a um comprimento de onda específico. Fonte: [20]

Portanto, a resolução espectral de um sensor indica a quantidade de regiões do espectro eletromagnético nas quais o sensor é capaz de gerar uma imagem em níveis de cinza, onde cada uma destas imagens representa a energia registrada pelos detectores do sensor numa determinada região do espectro eletromagnético. Assim, quanto maior a quantidade de bandas ou imagens geradas, maior será a resolução espectral sensor.

2.5 Transformada Cosseno Discreta

Como mostrado em [22], a Transformada Cosseno Discreta (*Discrete Cosine Transform-DCT*) tem sido amplamente utilizada no processamento de imagens, especialmente na codificação para compressão de imagens, para um desempenho quase ótimo. A DCT ajuda a separar a imagem em várias partes (ou sub-bandas espectrais) de importâncias diferentes (em relação à qualidade visual da imagem). A DCT transforma um sinal ou uma imagem do domínio espacial para o domínio da frequência.

A DCT no espaço bidimensional $2-D$, é definida de acordo com a Equação 2.2.

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right], \quad (2.2)$$

onde $f(x, y)$ é uma matriz de dimensões $(N \times N)$, $u, v = 0, 1, 2, \dots, N - 1$ e $\alpha(\cdot)$ é definido na equação 2.3.

$$\alpha(x) = \begin{cases} \sqrt{\frac{1}{N}}, & \text{se } x = 0; \\ \sqrt{\frac{2}{N}}, & \text{se } x \neq 0; \end{cases} \quad (2.3)$$

A *IDCT*, ou a DCT inversa, é dada na Equação 2.4.

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u, v) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right], \quad (2.4)$$

onde $x, y = 0, 1, 2, \dots, N - 1$.

Para a maioria das imagens, a maior parte da energia do sinal encontra-se nas baixas frequências, estas aparecem no canto superior esquerdo da compressão da DCT de uma imagem (o primeiro coeficiente da transformada é o valor médio da seqüência de amostras, na literatura este valor é referido como o coeficiente DC, enquanto todos os outros são chamados a coeficientes AC), e são alcançadas a partir dos valores do canto inferior direito, que representa as altas frequências, que muitas vezes são pequenas o suficiente para serem negligenciadas, devido terem pouca distorção visível.

2.6 Abordagem Supervisionada e Não-supervisionada

Através de métodos supervisionados, primeiramente são identificados (para treinamento) exemplos de classes de informações (por exemplo: tipo de cobertura do solo, informações estatísticas) de interesse na imagem, chamadas de *training sites*. Daí então, um algoritmo é usado para identificar uma padronização de cada uma dessas classes de informação, baseado em determinada(s) característica(s). Por conseguinte, a imagem é classificada com base nas informações e padrões resultantes do treinamento [23].

Algoritmos não-supervisionados analisam um grande número de dados e informações de uma única amostra, sintetizando e gerando uma resposta ao usuário. Ao contrário de classificação supervisionada, aqui não requer dados de treinamento fornecidos *à priori* [23].

2.7 Segmentador SRM

A segmentação de imagens é o processo de particionamento de uma imagem digital em vários segmentos (conjuntos de pixels). O objetivo da segmentação é simplificar e/ou alterar a representação de uma imagem em algo que é mais significativo e mais fácil de analisar.

A segmentação é normalmente usada para localizar objetos e limites (linhas, curvas, etc) em imagens. Mais precisamente, a segmentação de imagens é o processo de atribuição de um rótulo para cada pixel de uma imagem de tal forma que pixels com as mesmos rótulos compartilham certas características visuais [18].

O resultado da segmentação de imagens é um conjunto de segmentos que cobrem coletivamente toda a imagem, ou um conjunto de contornos extraídos da imagem. Cada um dos pixels de uma região são semelhantes com relação a alguma característica ou propriedade computada, tais como cor, intensidade, ou textura. Regiões adjacentes são significativamente diferentes em relação à algumas características. Mais informações à respeito de segmentação podem ser encontradas em [18].

O algoritmo de segmentação SRM (*Statistical Region Merging*) é baseado no crescimento e fusão de regiões, capaz de capturar os principais componentes estruturais de uma imagem digital, utilizando uma análise estatística simples, mas eficaz. O algoritmo é baseado em um modelo de geração de imagem, em que o agrupamento é um problema de inferência. Isso proporciona um simples predicado e uma ordenação na fusão, que com alta probabilidade, sofre apenas com uma fonte de erro: *overmerging*. Alcançando com alta probabilidade um baixo erro de segmentação. As vantagens deste método são a sua simplicidade, eficiência computacional, e um excelente desempenho sem a utilização de quantização ou transformações de espaço de cores [24].

2.8 Média, Desvio Padrão e Coeficiente de Variação

Estatística é um conjunto de procedimentos para a coleta, medição, classificação, cálculo, descrição, síntese, análise e interpretação de dados quantitativos sistematicamente adquiridos. A estatística tem dois componentes principais: a estatística descritiva e a estatística inferencial. A estatística descritiva implica em procedimentos numéricos e gráficos para resumir um conjunto de dados de uma forma clara e compreensível para que a estatística inferencial forneça os procedimentos para tirar conclusões sobre a população a partir de uma amostra.

A estatística descritiva ajuda a simplificar grandes quantidades de dados de uma forma sensata. A estatística descritiva reduz a quantidade de dados em um resumo simples. Existem dois métodos básicos: numéricos e gráficos. Usando a abordagem numérica pode-se calcular medidas como a média e o desvio-padrão. Uma vez que as aproximações numéricas e gráficas complementam-se, é aconselhável usar ambas. Há três características principais de uma única variável cuja tendência deve-se considerar: distribuição, tendência central e dispersão.

A distribuição é uma síntese da frequência dos valores individuais ou intervalos de valores para uma variável. A distribuição mais simples seria listar cada valor de uma variável e o número de vezes que cada valor ocorre. Uma das maneiras mais comuns para descrever uma única variável é com uma distribuição de frequência. Distribuição de frequência organiza dados brutos ou observações que foram coletadas em dados agrupados ou não. Os dados não agrupados fornecem uma listagem de todos os resultados possíveis que ocorrem na distribuição e, em seguida, indicam quantas vezes cada um deles ocorre. Dados Agrupados combinam todos os resultados possíveis em classes e, em seguida, indicam quantas vezes cada um deles ocorre dentro de cada classe. Neste último é mais fácil detectar os padrões nos dados, mas as informações sobre os resultados individuais são perdidas.

A tendência central de uma distribuição é uma estimativa do “centro” de uma distribuição de valores. A média é provavelmente o método mais utilizado para descrever a tendência central. Para calcular a média, todos os valores são somados e divididos pelo número de amostras.

Uma medida da tendência central por si só não é suficiente para descrever uma distribuição de frequências. Além disso, deve-se ter uma medida de dispersão das observações. O desvio padrão (DP) é uma estimativa precisa e detalhada de dispersão, mostrando a relação que um conjunto de valores (x_i) em relação à média da amostra (\bar{x}). O DP é a raiz quadrada da soma dos desvios a partir da média ao quadrado, dividido pelo número de amostras (n) menos um, como mostra a equação 2.5.

$$DP = \sqrt{\sum_i \frac{(x_i - \bar{x})^2}{n - 1}} \quad (2.5)$$

Se duas distribuições a serem comparadas são expressas nas mesmas unidades e as suas médias são muito similares, então a sua variabilidade pode ser comparada diretamente usando seus desvios-padrão. Caso contrário, não se pode usar os desvios-padrão para fazer este tipo de comparação. Deve-se usar as medidas relativas de dispersão em tais situações. O coeficiente de variação (CV) é uma unidade de medida livre, expressa em percentagem. O CV vai ser pequeno, se a variação for pequena. Entre dois grupos, aquele com menor CV será o mais consistente. O coeficiente de variação não é confiável se a média é próxima de zero. O CV é calculado dividindo o desvio padrão pela média [25].

Capítulo 3

Técnicas

Este capítulo descreve a teoria por trás das principais técnicas utilizadas neste trabalho. Tanto as utilizadas no algoritmo híbrido de estimação de regiões contaminadas de imagens de satélite, como a detecção de regiões, *inpainting*, síntese de textura e decomposição de regiões; como todo o algoritmo de treinamento para chegar ao índice de confiabilidade de estimação proposto.

3.1 Detecção de Nuvens e Sombras

Em primeiro lugar, serão definidas as regiões da imagem que serão processadas, usando o algoritmo descrito em [26]. Nesta etapa, as características da imagem são identificadas e separadas em quatro classes: nuvem densa, nuvem suave, sombra e área não afetada. Para isso, são calculadas medidas estatísticas da cena, como a média e o desvio padrão da distribuição dos valores digitais dos *pixels*. A equação 3.1 descreve essa operação, que é realizada para as bandas correspondentes à cor verdadeira (*true color*) do sensor.

$$m(x, y) = \begin{cases} f(x, y) < fs \times f_{m-dp}, & f(x, y) \in 0, \\ fs \times f_{m-dp} < f(x, y) < f_m, & f(x, y) \in 1, \\ f_m < f(x, y) < fn \times f_{m+dp}, & f(x, y) \in 2, \\ f(x, y) > fn \times f_{m+dp}, & f(x, y) \in 3, \end{cases} \quad (3.1)$$

onde $f(x, y)$ é o valor digital do pixel, f_m é o valor médio dos pixels da imagem, f_{m+dp} e f_{m-dp} são a soma e a subtração do valor médio com o desvio padrão da imagem, respectivamente. Também há a presença dos fatores fn (fator de nuvem) e fs (fator de sombra) que servem

para regular a detecção das regiões, modificando os limiares de cada classe. Tais fatores são definidos empiricamente para cada imagem, buscando sempre a melhor separação de regiões aos olhos do usuário. Cada região foi rotulada com um respectivo número. Regiões definidas como 0 são as regiões de sombra da imagem, as definidas como 1 são as livres de qualquer tipo de ruído, já as definidas como 2 são rotuladas como sendo nuvens esparsas, e finalmente as definidas como 3 são detectadas como nuvens densas.

Originalmente, essa fórmula utiliza a média aritmética das médias de cada banda $((m_{b1} + m_{b2} + m_{b3})/3)$, porém como mostrado em [27], para aumentar o desempenho do método proposto e aproveitando as particularidades de cada banda, as médias serão utilizadas separadamente. O mesmo acontece para os valores de desvio padrão que também passam a ser usados separadamente.

Para finalizar o processo, é aplicada uma operação de abertura morfológica para remover pequenos objetos, detectados por engano, que podem provocar confusões nos passos seguintes.

3.2 Inpainting por suavização baseado em DCT

Este método foi proposto por [13], e assim como em [3], é baseado na propagação de informação através da suavização. A especificidade desta abordagem está relacionada com a utilização da Transformada Cosseno Discreta (DCT - *Discrete Cosine Transform*) para simplificar e resolver sistemas de equações lineares, gerando resultados eficientes. Em estatística e análise de dados, a suavização é utilizada para reduzir informação ou ruído experimental, e manter as características mais importantes do conjunto de dados.

O processo é aplicado para dados multidimensionais conforme a Eq. 3.2.

$$\hat{y}_{k+1} = IDCTN(\Gamma^N \circ DCTN(y_k)). \quad (3.2)$$

onde y é um *pixel*, \hat{y} é o *pixel* suavizado, $DCTN$ e $IDCTN$ referem-se à DCT N-dimensional e sua inversa, respectivamente; k é o número de iterações, N é o número de dimensões, \circ é o produto Schur e Γ^N representa um tensor de rank N . Para mais informações sobre a definição desses dois últimos, consultar [13].

É importante observar que, quando há valores indefinidos na imagem, a suavização atua como um interpolador de dados, funcionando como um método de *inpainting*. A fim de acelerar a convergência, o processo inicia com uma iteração de interpolação do vizinho mais próximo sobre a imagem a ser restaurada.

3.3 Síntese de Textura

A síntese de textura tem sido um campo intenso de estudos, devido a sua variedade de propósitos. Pode ser aplicada em tarefas de preenchimento de objetos, recuperação de imagens, compressão de vídeos, remoção de plano de fundos e etc.

Textura pode ser definida como um padrão visual num plano infinito 2D que, em alguma escala, tem uma distribuição estável. Então, naturalmente, pode-se obter uma amostra finita de texturas presentes neste plano, a fim de sintetizar outras amostras a partir das mesmas. Esta amostra finita pode ser extraída de inúmeras texturas distintas, o que pode se tornar um problema. Para contorná-lo, assume-se que a amostra deve ser grande o suficiente para capturar a distribuição textural da imagem. Além disso, a síntese de textura é responsável pela fusão de regiões contínuas com mínima imperfeição e percepção da operação, garantindo a qualidade visual [28].

A abordagem proposta por [14] visa eliminar e redefinir objetos em uma imagem digital com informações de regiões vizinhas. Este método utiliza a síntese de textura para preencher as regiões que contém padrões texturais bidimensionais com estocasticidade moderada. Para isso, gera amostragem de novas texturas a partir de áreas fonte da imagem e faz uma cópia para as áreas alvo.

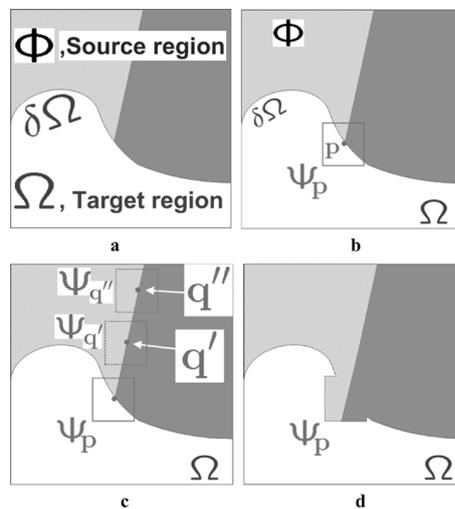


Figura 3.1: Síntese de textura baseada em modelos: (a) Imagem original. (b) Fragmento Ψ_p centralizado em $p \in \Phi$. (c) Candidatos mais prováveis $\Psi_{q'}$ e $\Psi_{q''}$. (d) Candidato mais provável é propagado para o fragmento alvo.

Na Fig. 3.1, é apresentado o algoritmo de [14], no qual se supõe que uma imagem em que existe uma área fonte Φ e uma área alvo Ω , contornada por $\delta\Omega$, claramente distinguível (Fig. 3.1 (a)). O objetivo é sintetizar a área delimitada pelo fragmento Ψ_p , o qual é centralizado

no ponto $p \in \Phi$, ilustrado na Fig. 3.1 (b). Então, são computados os candidatos mais prováveis para preencher Ψ_p apresentados em $\delta\Omega$, por exemplo, $\Psi_{q'}$ e $\Psi_{q''}$ na Fig. 3.1 (c). Entre os candidatos, existe um que mais se aproxima do fragmento alvo, tendo seus pixels correspondentes copiados no fragmento Ψ_p . Este processo é repetido até o preenchimento completo de Ω . Na Fig. 3.1 (d) pode-se notar que a textura, bem como a estrutura (a linha que separa as regiões cinzas, clara e escura), são propagadas sobre o fragmento Ψ_p .

No algoritmo, cada pixel mantém um valor de cor (que é nulo nos *pixels* a serem preenchidos) e um valor de confiança, que reflete a confiança no valor de cor com base no momento em que este pixel foi definido. Durante a execução do algoritmo, os fragmentos localizados no contorno $\delta\Omega$ recebem um valor de prioridade temporário, definindo a ordem de preenchimento. Assim, um processo iterativo é executado na seguinte sequência:

1. Computando prioridades dos fragmentos: o algoritmo realiza a tarefa de síntese através de uma estratégia de preenchimento *best-first*, que depende inteiramente dos valores de prioridade que são atribuídos a cada fragmento na frente de preenchimento. O cálculo da prioridade é tendenciosa para aquelas regiões que: a) estão sobre a continuação do bordas fortes e b) são cercadas por *pixels* de alta confiança.
2. Propagação das informações de estrutura e de textura: uma vez que todas as prioridades na frente de preenchimento são computadas, o fragmento $\Psi_{\hat{p}}$ de maior prioridade é encontrado, sendo depois preenchido com dados extraídos da região fonte Φ .
3. Atualização de valores de confiança: depois que o fragmento $\Psi_{\hat{p}}$ é redefinido com novos valores de *pixel*, o valor de confiança é atualizado na área delimitada por $\Psi_{\hat{p}}$.

3.4 Decomposição de Regiões

O algoritmo *cartoon+texture* (*cartoon* é uma outra denominação de estrutura bastante usada na literatura) decompõe qualquer imagem f na soma de uma parte de estrutura, u , em que aparecem apenas as formas grosseiras da imagem, e uma parte de textura v contendo padrões de oscilação. Tal decomposição $f = u + v$ tem comportamento análogo aos clássicos filtros passa-altas e passa-baixas do processamento de sinais. A parte de estrutura de uma imagem contém bordas fortes, e portanto baixas e médias frequências, enquanto que a textura tem frequências altas, podendo ainda conter médias. Desta forma, os algoritmos de decomposição linear não conseguem fazer uma separação clara entre estrutura e textura, uma vez que, geralmente eles apenas suavizam as bordas, extraem as altas frequências e as

introduzem na imagem de textura. Frequentemente, o filtro passa-baixas acaba não atenuando algumas componentes de textura, que permanecem erroneamente na parte de estrutura.

Em [15] é proposta uma solução usando um problema variacional aproximado e rápido, obtido pela aplicação de um par de filtros não-lineares passa-baixas e passa-altas. Neste algoritmo, uma decisão é tomada para cada ponto da imagem, definindo se o mesmo pertence à parte de estrutura ou à parte de textura. Este processo é feito pelo cálculo de uma variação total local da imagem em torno do ponto, e comparando-a com a variação total local depois da aplicação de um filtro passa-baixas. Quando a imagem é convoluída por um filtro passa-baixas, pontos de estrutura de uma imagem tendem a ter uma variação total local oscilando lentamente. Por outro lado, pontos de textura mostram forte decaimento de variação total local após esta operação.

A parte de *cartoon* mantém os valores originais da imagem em pontos denominados como não-texturais. Em pontos identificados como de textura, esta parte assume o valor filtrado. Já nos pontos em que a decisão é confusa, uma média ponderada entre os mesmos é feita. A parte de textura v é simplesmente a diferença entre a imagem original f e sua parte de estrutura u .

Outro ponto importante é o parâmetro de escala da textura, para especificar o grau de decomposição. Devido a tal parâmetro, não há uma decomposição única de uma imagem em textura e *cartoon*. Uma textura pode ser mantida na parte de estrutura para baixos valores do parâmetro de escala. O parâmetro de escala no algoritmo é crucial, e deve ser escolhido empiricamente.

Tal como indicado no algoritmo, a decomposição *cartoon+texture* apenas requer a aplicação de dois filtros passa-baixas sobre o gradiente da imagem, que são executados diretamente por uma convolução discreta. Para informações mais detalhadas sobre esta técnica, consultar [15].

3.5 Mapeamento das regiões de estrutura e de textura

Uma estratégia foi proposta para mapear áreas de estrutura e textura de uma imagem de satélite com base na decomposição descrita na Seção 3.4. Este processo começa transformando a componente que contém a informação de textura, ou seja, a imagem v , em uma imagem binária com valores 1 para áreas heterogêneas de textura e 0 para áreas de estrutura. Em seguida, um filtro não-linear de mediana [18] é aplicado para tornar homogêneas (suavizadas), as áreas onde pequenas lacunas de uma determinada classe são cercadas por regiões onde predomina uma outra característica.

Essa etapa é crítica devido à presença de nuvens e sombras na imagem, e é realizada de modo a definir corretamente as técnicas a serem utilizadas para cada região. Isso acontece porque as nuvens e as sombras sempre serão componentes da estrutura, então, para definir qual técnica usar para removê-las, deve-se observar as regiões circundantes à esta, isto é, seu contexto. Portanto, como resultado da aplicação do filtro, as regiões a serem redefinidas são mapeadas na imagem binária para regiões de estrutura e textura, para que sejam aplicados o *inpainting* e a síntese de textura na imagem de entrada, respectivamente.

3.5.1 Adequação das técnicas para imagens de satélite

Os métodos de restauração de regiões descritos até aqui, foram projetados e testados especificamente para imagens sintéticas de cenas comuns. Em se tratando de processamento de imagens obtidas por sensores orbitais, devem ser levadas em consideração algumas características peculiares, como o tamanho e alto grau de heterogeneidade entre os *pixels*, resultantes do grande campo de visada do sensor e da alta variação de reflectância dos objetos presentes na superfície terrestre, respectivamente.

O consumo de memória exigidos pelas técnicas está intrinsecamente ligado ao número de operações e ao tamanho da imagem. A máquina usada para testes contém um processador de dois núcleos operando à 1.8 GHz, e memória RAM de 3 GB, sendo considerado um equipamento *desktop* usual. Para este computador, ocorria estouro de memória para a operação de *inpainting* e processamento “sem fim” para a síntese de textura.

Com o objetivo de contornar este problema, optou-se por fazer um processamento por blocos na imagem, ou seja, em cada operação dividiu-se para cada banda a imagem em blocos de tamanhos iguais e equivalentes ao tamanho de imagens sintéticas, permitindo mais economia de memória. Por exemplo, na operação de *inpainting*, ao dividir a imagem em 16 blocos, é feito o processamento para as três bandas do primeiro bloco (canto superior esquerdo), e o resultado é armazenado na mesma posição do bloco na imagem resultante, e assim sucessivamente para todos os blocos até formar a imagem resultante. É importante ter em mente que quanto maior a imagem, maior será a quantidade de memória exigida no processamento, logo, maior deverá ser a divisão em blocos. Desta forma, a imagem se adequa ao limite de memória da máquina, e não contrário.

No entanto, a busca por fragmentos candidatos, relativa à síntese de textura, ainda era um processo muito lento. Para acelerá-lo, adotou-se uma estratégia local para a busca por pequenos blocos. Supondo que a continuação das áreas cobertas por nuvens e sombras estão nos seus arredores, então não faz sentido procurar por substitutos em regiões muito distantes

(como é feito no método original, que percorre a imagem inteira). Logo, é buscado na k -ésima vizinhança, a partir do pixel central $f(i, j)$, o fragmento fonte que mais se aproxima da região alvo. O conceito de vizinhança é dado em [18], sendo k a camada que determina a distância de $f(i, j)$ ao vizinho. Por padrão, k foi convencionado ser 10 vezes o tamanho do fragmento alvo, mas deve ser manipulado de acordo o resultado do processamento, ou seja, se 10 não for o suficiente deve-se aumentar este parâmetro. Portanto, aplicada esta modificação, a síntese de textura reduziu consideravelmente seu tempo de processamento, mantendo a mesma qualidade visual que apresentava sem esta modificação.

3.6 Abordagem Híbrida

A Figura 3.2 mostra o fluxograma do algoritmo híbrido de estimação de nuvens e sombras de imagens de satélite proposto neste trabalho. Inicialmente uma Imagem tem suas regiões contaminadas detectadas e passadas por um operação morfológica de abertura (Seção 3.1), que gera a *Imagem Mapeada* com as regiões detectadas. A mesma imagem inicial passa por um processo de decomposição de regiões e mapeamento de textura e estrutura com o auxílio de um filtro não-linear de mediana (Seção 3.4), gerando a *Imagem Decomposta*, que subdivide-se na *Imagem de Estrutura* e *Imagem de Textura*. O algoritmo de *inpainting* (Seção 3.2) é aplicado na *Imagem de Estrutura* (apenas sobre as regiões da *Imagem Mapeada*), enquanto que a síntese de textura (Seção 3.3) é aplicada na *Imagem de Textura* (novamente, apenas sobre as regiões da *Imagem Mapeada*), resultando na *Imagem Resultante de Estrutura* e na *Imagem Resultante de Textura*. Por fim, ambas as imagens são fusionadas gerando a *Imagem Resultante Híbrida*.

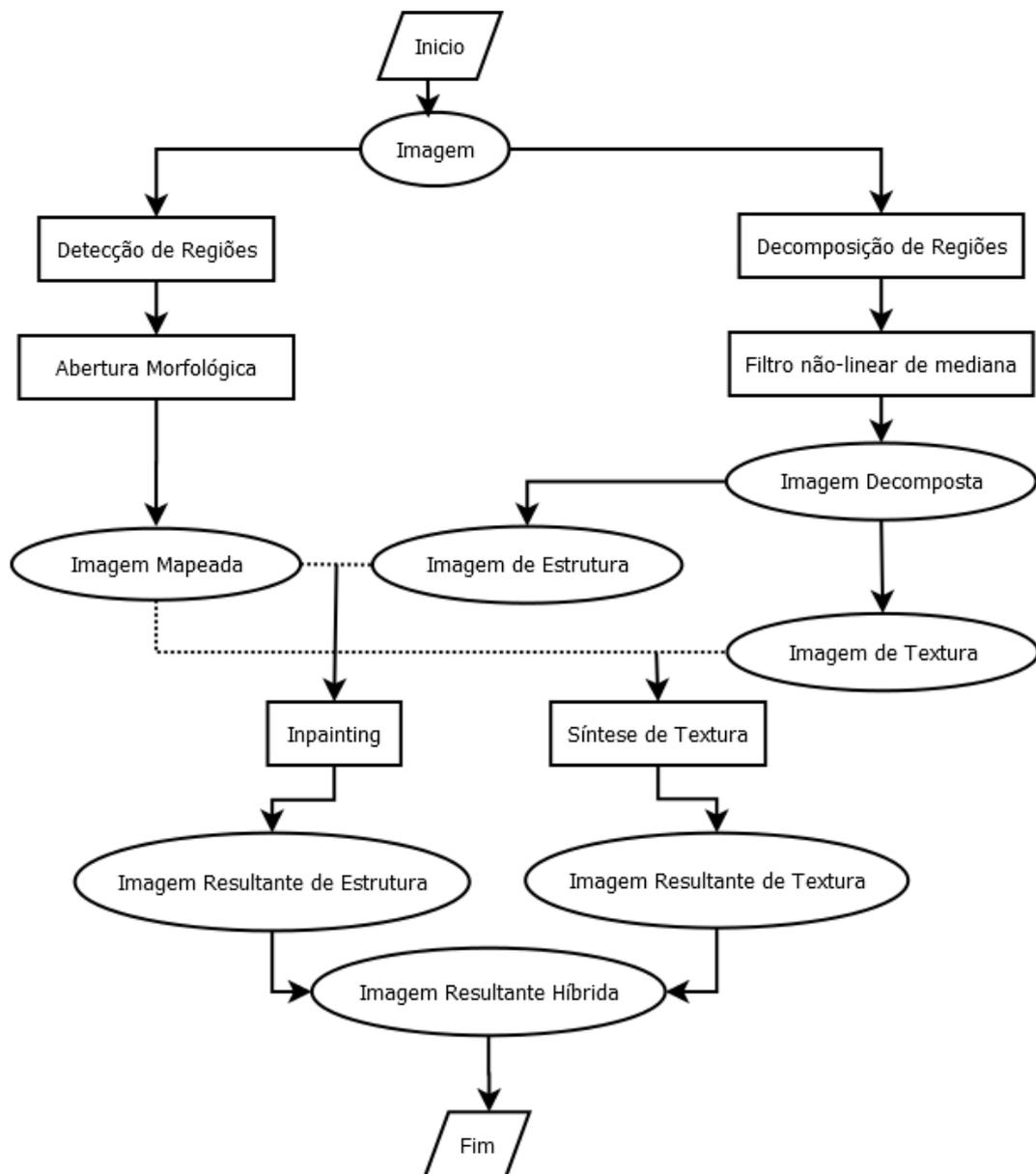


Figura 3.2: Fluxograma do algoritmo híbrido de estimação de nuvens e sombras de imagens de satélite proposto neste trabalho.

3.7 Índice de Confiabilidade

Esta Seção descreve as teorias, artifícios e etapas aplicadas para que se pudesse chegar ao Índice de Confiabilidade de estimação de regiões contaminadas de imagens de satélites proposto neste trabalho.

3.7.1 Contaminação Artificial

Uma vez que não se tem disponível imagens de uma mesma cena com e sem componentes atmosféricos em um ambiente controlado, e principalmente por melhor se encaixar no contexto do trabalho, foi dada preferência à simulação da ocorrência de nuvens e sombras em imagens limpas. Aplicando os algoritmos nestas imagens simuladas, pode então ser estimado o quão eficiente foi a remoção e redefinição de regiões. Mas aqui a questão principal é: como aplicar coberturas atmosféricas fiéis à situações reais?

Em [29] a avaliação quantitativa foi realizada através da repetição de experiências múltiplas vezes para diferentes tamanhos de regiões indefinidas introduzidas aleatoriamente ao longo das imagens. Estes diferentes tamanhos foram escolhidos para verificar a resposta dos mais variados algoritmos.

A fim de tornar possível quantificar a precisão da reconstrução da técnica proposta, [4] adotou um procedimento experimental para as regiões correspondentes de sombras e nuvens. Zonas de inpainting foram criadas manualmente em formatos de disco, a fim de simular áreas indefinidas de contaminação atmosférica. O tamanho das zonas são grandes o suficiente para testar a capacidade da técnica proposta no que diz respeito à reconstrução das áreas.

Com base nestes trabalhos, a simulação de componentes atmosféricas foi feita nas imagens de forma estocástica tanto no que diz respeito ao tamanho, quanto na quantidade e formato das interferências. Exemplos podem ser visto nas Figuras 4.1(d), 4.2(d) e 4.3(d).

3.7.2 Tamanho relativo das nuvens e sombras

Para obter o tamanho das nuvens e sombras mapeadas, e conseqüentemente achar sua proporção em relação à imagem, que será o principal critério para dividir as classes de imagens, o algoritmo primeiramente percorre *pixel a pixel* da imagem em busca do valor numérico que representa este ruído (3 no caso de nuvens densas, 2 nuvens esparsas e 0 sombras). Após localizar o primeiro pixel que corresponde a essa característica, o algoritmo varre recursivamente sua vizinhança buscando por pixels de valores iguais. Esse processo acaba apenas quando não há mais pixels de valor igual adjacentes ao ruído.

A Figura 3.3(a) ilustra a forma que o algoritmo varre a vizinhança de cada pixel. Sabe-se que cada elemento possui 8 vizinhos adjacentes, logo um por um destes é verificado em busca de valores iguais ao núcleo (representado pelo número 1). A seta preenchida indica o início da busca. Caso este valor seja encontrado a busca é interrompida momentaneamente e um novo caso é tratado, onde o mais novo vizinho adjacente encontrado passa a ser o núcleo

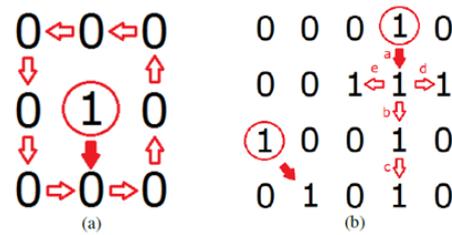


Figura 3.3: (a) Ilustração da forma que o algoritmo varre a vizinhança de um *pixel*. (b) Exemplo da forma que o algoritmo varre uma imagem binária em busca de *pixels* ruído.

da busca.

Esta busca só é interrompida se nenhum vizinho adjacente ao núcleo possui o mesmo valor numérico que ele. Quando isso ocorre, as operações interrompidas anteriormente são retomadas uma a uma até que cada núcleo seja avaliado. Não encontrando mais nenhum pixel ruído adjacente aos vizinhos do núcleo, a busca é finalizada e a varredura principal é retomada para procurar outros possíveis ruídos.

A Figura 3.3(b) exemplifica a maneira que o algoritmo percorre uma imagem I . Nesta imagem o mapeamento só possui dois valores (0's e 1's) onde os 1's representam ruído. Os 1's dentro de cada círculo correspondem aos núcleos principais da busca - àqueles identificados pela varredura pixel a pixel da imagem. As setas preenchidas apontam para o primeiro vizinho encontrado para cada núcleo principal, e as letras a, b, c, d e e indicam a ordem pela qual cada vizinho é encontrado (onde a letra a seria o primeiro vizinho encontrado, ou seja, aqueles apontados pelas setas preenchidas). Após varrer toda a imagem I , dois ruídos são encontrados. O primeiro de tamanho $T = 6$ e o segundo de tamanho $T = 2$ (dado em pixels).

3.7.3 Separação por classe

Como a proposta do índice está atrelada à análise da imagem de entrada, então é importante que características relacionadas às interferências atmosféricas sejam usadas em um treinamento para se ter idéia do quanto elas influenciam o resultado da remoção e redefinição das regiões. Nesta abordagem supervisionada, foi analisado levar em consideração o percentual de cobertura de nuvens e sombras sobre regiões de textura e estrutura, o tamanho das áreas a serem removidas e a forma das mesmas. A última opção foi descartada uma vez que como se trata de objetos específicos para remoção (diferentemente de técnicas de inpainting de propósito geral, em que qualquer coisa pode ser o objeto-alvo), terão sempre formatos semelhantes.

É de praxe na literatura usar a convenção de [30] para avaliação da usabilidade

de imagens para aplicações glaciológicas, geológicas e cartográficas. A mesma define uma excelente cobertura de nuvem em até 5% em relação à cena inteira, boa até 10%, ruim até 15% e imagem não-usável a partir deste valor. Ao mesmo tempo, [31] só considera usáveis as imagens com cobertura de nuvem menor que 15%. Este trabalho usa como base os dois trabalhos citados, de acordo com a Tabela 3.1.

Tabela 3.1: Relação entre Porcentagem de cobertura e Qualidade da Região.

Percentual de Cobertura de nuvem	Classificação da imagem
0 – 5%	Excelente
5 – 10%	Boa
10 – 15%	Ruim
$\geq 15\%$	Não-usável

A diferença é que aqui este percentual é analisado separadamente para regiões de estrutura e de textura, e também além das nuvens serão contabilizadas todas as interferências atmosféricas. Como não serão usadas imagens classificadas como não-usáveis, são obtidas 9 classes de imagens, determinadas conforme Tabela 3.2, onde “Estrutura” e “Textura” correspondem à Cobertura de componentes atmosféricos sobre região de Estrutura e Textura, respectivamente.

Tabela 3.2: Relação entre as Classes e Qualidade das Regiões de Estrutura e Textura.

Classes de imagens	Estrutura	Textura
Classe 1	Excelente	Excelente
Classe 2	Excelente	Boa
Classe 3	Excelente	Ruim
Classe 4	Boa	Excelente
Classe 5	Boa	Boa
Classe 6	Boa	Ruim
Classe 7	Ruim	Excelente
Classe 8	Ruim	Boa
Classe 9	Ruim	Ruim

Quanto ao tamanho das nuvens e sombras, foi contabilizado o percentual do tamanho das mesmas em relação ao tamanho da imagem (somando todas), e também do percentual do tamanho da maior nuvem da imagem em relação ao tamanho da imagem.

3.7.4 Avaliando pela Exatidão Global

Ao referir-se ao desempenho de um modelo de classificação, deve-se levar em consideração a capacidade do modelo para prever ou separar corretamente as classes. Ao olhar para os erros cometidos por um modelo de classificação, a matriz de confusão dá a situação completa.

A matriz de confusão da Figura 3.4 mostra como as previsões são feitas pelo modelo. As colunas correspondem às classe conhecidas de dados, isto é, os dados rotulados. As linhas correspondem às previsões feitas pelo modelo. O valor de cada elemento na matriz é o número de previsões feitas com a classe correspondente à coluna de exemplos com o valor correto conforme representado pela linha. Assim, os elementos da diagonal mostram o número de classificações corretas feitas para cada classe, e os elementos fora da diagonal mostram os erros cometidos.

		+i = colunas (referência)			Total da linha
		1	2	k	x_{i+}
i+ = linhas (classificação)	1	x_{11}	x_{12}	x_{1k}	x_{1+}
	2	x_{21}	x_{22}	x_{2k}	x_{2+}
	k	x_{k1}	x_{k2}	x_{kk}	x_{k+}
Total da coluna	x_{+i}	x_{+1}	x_{+2}	x_{+k}	N

Figura 3.4: Matriz de Confusão.

Foi decidido não usar o índice Kappa [32] por tratar-se de uma abordagem não supervisionada, e uma vez que aqui as classes são formadas dinamicamente de acordo com o resultado da segmentação SRM, não se tem como construir toda a matriz de confusão. Ao invés disso, usa-se apenas a informação contida na diagonal principal, relacionada às classificações corretamente previstas, que levam à exatidão global.

A Exatidão Global (EG) [32] é a razão entre a soma de todos os elementos classificados corretamente pelo numero total de elementos. Esta medida é calculada pela expressão:

$$EG = \frac{\sum_{i=1}^r x_{ii}}{N}, \quad (3.3)$$

onde x_{ii} são os elementos da diagonal principal, r é o número de classes e N o número total de amostras. A Exatidão Global Média (EGM) é a média da EG levando em consideração todas as classes.

O algoritmo de treinamento para avaliação pela EG é realizado para cada uma das classes (N_j) definidas na Seção 3.7.3. Uma Imagem I_{ij} (dentre todas as N_i imagens) da

classe j é a única entrada necessária. O primeiro passo é segmentar a imagem de entrada (Usando o segmentador SRM da Seção 2.7), com o objetivo de obter a Imagem Segmentada de Referência (ISR_{ij}). Paralelamente, são adicionadas contaminações atmosféricas artificiais na imagem I_{ij} (Seção 3.7.1), simulando uma imagem com ruído e gerando a Imagem Contaminada IC_{ij} . É nesta última que é aplicado o algoritmo híbrido mostrado nesse trabalho, a imagem reconstituída é chamada de Imagem Híbrida (IH_{ij}). Tal como no início, a imagem IH_{ij} é segmentada, formando a Imagem Segmentada Híbrida ISH_{ij} . A exatidão global deste processo é obtida com base na diagonal principal da matriz de confusão de ISR em relação a ISH . Quando encontradas todas as EGs da classe, são obtidas as medidas estatísticas de média, desvio padrão e coeficiente de variação da mesma, passando para a próxima classe. O fluxograma do algoritmo é mostrado na Figura 3.5.

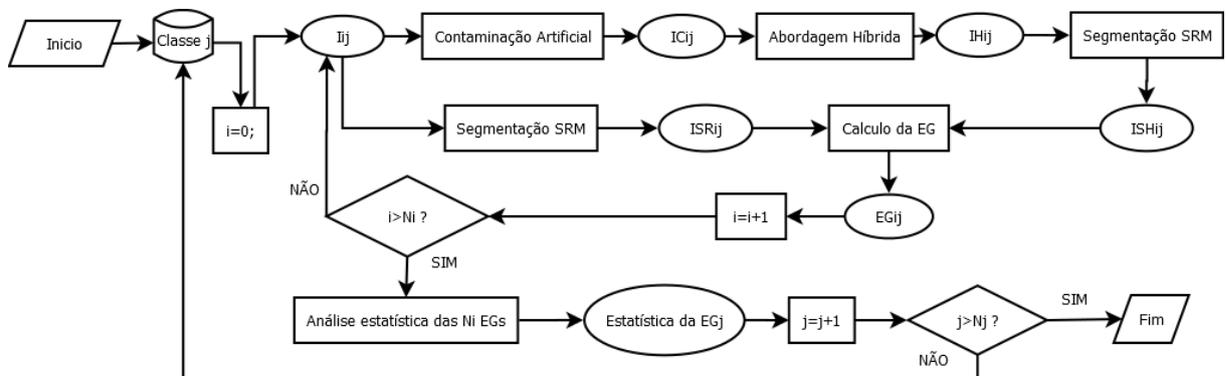


Figura 3.5: Fluxograma do algoritmo de treinamento para avaliação pela EG.

3.7.5 Avaliando pelo SSIM

Sinais de imagem naturais são altamente estruturados: os pixels apresentam fortes dependências, especialmente quando eles são espacialmente próximos, e essas dependências carregam informações importantes sobre a estrutura dos objetos na cena visual. A motivação para a Medida de Similaridade Estrutural (*Measure of Structural Similarity - SSIM*) é encontrar uma forma mais direta para comparar as estruturas de referência com os sinais distorcidos, e também que seja baseada na perspectiva de formação da imagem.

A luminância da superfície de um objeto observado é o produto da iluminação e da reflectância, mas as estruturas dos objetos na cena são independentes da iluminação. Conseqüentemente, para explorar a informação estrutural de uma imagem, a influência da iluminação foi separada. A informação estrutural em uma imagem é definida como os atributos que representam a estrutura de objetos na cena, independente da luminância média e do contraste. Uma vez que luminosidade e contraste podem variar em uma cena, usa-se a

luminosidade local e contraste para esta definição.

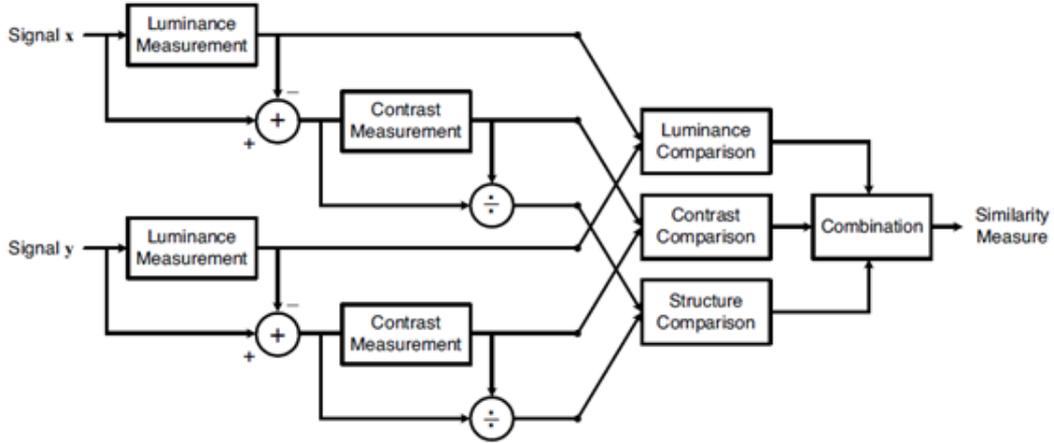


Figura 3.6: Diagrama do sistema de medição de similaridade estrutural (SSIM). Fonte: [33].

O diagrama do sistema de avaliação de qualidade proposto é mostrado na Figura 3.6. Supondo que dois sinais não negativos de imagem tenham sido alinhados um com o outro (por exemplo, fragmentos espaciais extraídos de cada imagem), e que um dos sinais têm uma qualidade perfeita, então a medida de semelhança pode servir como uma medida quantitativa de qualidade do segundo sinal. O sistema separa a tarefa de medida de similaridade em três comparações: luminância, contraste e estrutura.

Primeiramente, a luminância de cada sinal é comparada. Assumindo sinais discretos, isto é estimada como a intensidade média

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i. \quad (3.4)$$

A função de comparação de luminância $l(x, y)$ é então uma função de μ_x e μ_y .

Em seguida, a intensidade média do sinal é removida. Na forma discreta, o sinal resultante $x - \mu_x$ corresponde à projeção do vetor x sobre o hiperplano definido por

$$\frac{1}{N} \sum_{i=1}^N x_i = 0. \quad (3.5)$$

O desvio-padrão é usado como uma estimativa do contraste. Uma estimativa *unbiased* na forma discreta é dada por

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2}. \quad (3.6)$$

A comparação de contraste $c(x, y)$ é então uma comparação de σ_x e σ_y .

Em terceiro lugar, o sinal é normalizado pelo seu próprio desvio padrão, de modo que os dois sinais sendo comparados apresentam desvio padrão unitários. A comparação de estrutura $s(x, y)$ é realizada nestes sinais normalizados $(x - \mu_x)/\sigma_x$ e $(y - \mu_y)/\sigma_y$.

Finalmente, os três componentes são combinados para obter uma medida global de similaridade

$$S(x, y) = f(l(x, y), c(x, y), s(x, y)). \quad (3.7)$$

Um ponto importante é que os três componentes são relativamente independentes. Por exemplo, a mudança de luminância e/ou contraste não afetará as estruturas de imagens.

A fim de completar a definição da medida de similaridade, em 3.7, é necessário definir a três funções $l(x, y)$, $C(x, y)$ e $S(x, y)$, bem como a função de combinação $f(\cdot)$. O autor ressalta também que a medida de similaridade deve satisfazer as seguintes condições.

- Simetria: $S(x, y) = S(y, x)$;
- *Boundedness*: $S(x, y) \leq 1$;
- Máximo único: $S(x, y) = 1$, se e somente se $x = y$ (nas representações discretas $x_i = y_i$, para todo $i = 1, 2, \dots, N$).

Para comparação da luminância, é definido:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (3.8)$$

onde a constante C_1 é incluída para evitar a instabilidade quando $\mu_x^2 + \mu_y^2$ está muito próximo de zero. Especificamente, é escolhido

$$C_1 = (K_1 L)^2 \quad (3.9)$$

Em que L é a faixa dinâmica dos valores de pixel (255 para 8 bits em tons de cinza), e $K_1 \ll 1$ é uma pequena constante. Considerações semelhantes também se aplicam ao comparador de contrastes e comparador de estrutura descrita mais tarde. A equação 3.8 obedece as três propriedades listadas anteriormente.

A equação 3.8 é também qualitativamente consistente com a lei de Weber, que tem sido amplamente utilizada para modelar a adaptação de luz (também chamado de mascaramento

de luminância) no HVS. De acordo com a lei de Weber, a magnitude de uma mudança apenas perceptível de luminância é ΔI aproximadamente proporcional à luminância de fundo para uma ampla gama de valores de luminância. Em outras palavras, o HVS é sensível à mudança de luminância relativa, e não a mudança de luminância absoluto. Considerando que R seja o tamanho da mudança de luminosidade em relação ao fundo de luminosidade, reescrevemos a luminância do sinal distorcido como . Substituindo isso em 3.8 tem-se

$$l(x, y) = \frac{2(1 + R)}{1 + (1 + R)^2 + C_1/\mu_x^2} \quad (3.10)$$

Se C_1 for suficientemente pequeno (em relação a μ_x^2) para ser ignorado, então $l(x, y)$ é uma função apenas de R , qualitativamente consistente com a lei de Weber

A função de comparação de contraste assume uma forma semelhante

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (3.11)$$

onde $C_2 = (K_2L)^2$ e $K_2 \ll 1$. Esta definição de novo satisfaz as três propriedades listadas anteriormente. Uma característica importante desta função é que, com a mesma quantidade de mudança de contraste, esta medida é menos sensível ao caso de contraste de base alta σ_x do que o contraste de base baixa. Isto é consistente com a característica de mascaramento de contraste de mascaramento do HVS.

A comparação de estrutura é conduzida após a subtração de luminância e normalização da variância. Especificamente, são associados dois vetores de unidade $(x - \mu_x)/\sigma_x$ e $(y - \mu_y)/\sigma_y$, cada um encontra-se no hiperplano definido em 3.5, com a estrutura das duas imagens. A correlação (produto interno) entre eles é uma medida simples e eficaz para quantificar a similaridade estrutural. Note-se que a correlação entre $(x - \mu_x)/\sigma_x$ e $(y - \mu_y)/\sigma_y$ é equivalente ao coeficiente de correlação entre x e y . Assim, define-se a função de comparação de estrutura como se segue:

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (3.12)$$

Como nas medidas de luminância e de contraste, foi introduzida uma pequena constante em ambos, denominador e numerador. Na forma discreta, σ_{xy} pode ser estimada com

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \quad (3.13)$$

Geometricamente, o coeficiente de correlação corresponde ao cosseno do ângulo entre os vetores $x - \mu_x$ e $y - \mu_y$. Note também que $s(x, y)$ pode assumir i 's negativos.

Finalmente, podem-se combinar as três comparações de 3.8, 3.11 e 3.12 formando o índice de medida de similaridade SSIM entre os sinais x e y .

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (3.14)$$

onde $\alpha > 0$, $\beta > 0$ e $\gamma > 0$ são parâmetros usados para ajustar a importância relativa dos três componentes. É fácil verificar que esta definição satisfaz as três condições anteriormente indicadas. A fim de simplificar a expressão, define-se $\alpha = \beta = \gamma = 1$, e $C_3 = C_2/2$ no artigo. Isto resulta numa forma específica do índice SSIM

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_1)} \quad (3.15)$$

Em [33], o índice SSIM foi implementado utilizando uma janela deslizante 11×11 ao longo de todo o espaço de imagem. Em cada coordenada da imagem, o índice SSIM é calculado localmente dentro da janela, o mapa de índice SSIM resultante pode ser utilizado para visualizar o mapa de qualidade das imagens distorcidas. Finalmente, dos valores do índice SSIM obtidos ao longo da imagem, podem ser sintetizados fazendo uma média que representa a qualidade da imagem.

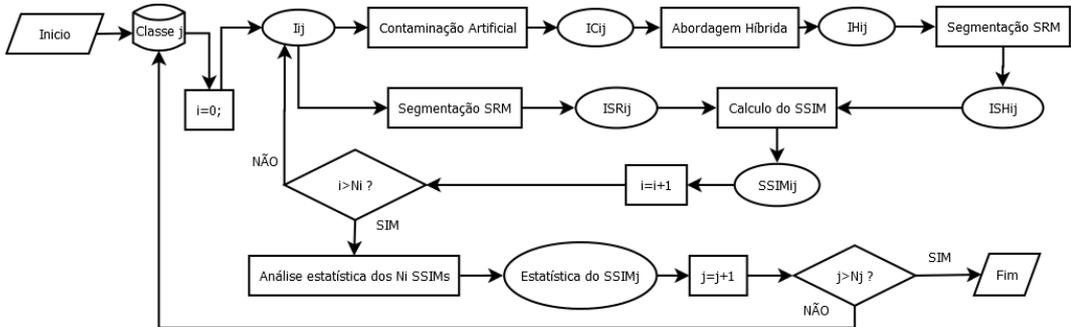


Figura 3.7: Fluxograma do algoritmo de treinamento para avaliação pelo SSIM.

O algoritmo de treinamento para avaliação utilizando SSIM é semelhante ao visto na seção anterior, que usava EG, sendo realizado para cada uma das N_j classes definidas na seção 3.7.3. Novamente, uma Imagem I_{ij} (dentre todas as N_i imagens) da classe j é a única entrada necessária. O primeiro passo é segmentar a imagem de entrada (Usando o segmentador SRM da Seção 2.7), com o objetivo de obter a Imagem Segmentada de Referência (ISR_{ij}). Paralelamente, são adicionadas contaminações atmosféricas artificiais na imagem I_{ij} (Seção 3.7.1), simulando uma imagem com ruído e gerando a Imagem Contaminada

ICij. É nesta última que é aplicado o algoritmo híbrido mostrado nesse trabalho, a imagem reconstituída é chamada de Imagem Híbrida (*IHij*). Tal como no início, a imagem *IHij* é segmentada, formando a Imagem Segmentada Híbrida *ISHij*. A exatidão global deste processo é obtida com base na diagonal principal da matriz de confusão de *ISR* em relação a *ISH*. Quando encontrados todos os SSIMs médios da classe, são obtidas as medidas estatísticas de média, desvio padrão e coeficiente de variação da mesma, passando para a próxima classe. O fluxograma do algoritmo é mostrado na Figura 3.7.

3.7.6 Avaliando pela Confiança média dos pixels

A idéia aqui é aproveitar uma etapa da síntese de textura [14] para avaliar como está a confiança média em torno de todos os pixels que estão nas bordas das regiões indefinidas da imagem.

Dado um fragmento Ψ_p centrado no ponto p para algum $p \in \delta\Omega$, tem-se o termo de confiança $C(p)$, definido como:

$$C(p) = \frac{\sum_{q \in \Psi_p \cap (f - \Omega)} C(q)}{|\Psi_p|}. \quad (3.16)$$

onde $|\Psi_p|$ é a área de Ψ_p . Durante a inicialização, a função $C(p)$ é definida como $C(p) = 0 \forall p \in \Omega$ e $C(p) = 1 \forall p \in f - \Omega$. O termo de confiança $C(p)$ pode ser pensado como uma medida da quantidade de informações confiáveis em torno do pixel p .

De forma geral, o termo $C(p)$ impõe aproximadamente a ordem de preenchimento desejável. Ao longo do processamento, pixels nas camadas exteriores da região alvo tendem a ser caracterizados por altos valores de confiança, portanto são preenchidos antes; pixels no centro da região alvo terão valores menores de confiança.

O termo confiança resume a percentagem de pixels úteis em uma janela. A informação estrutural não é incluída. Por exemplo, uma linha vertical e de uma linha horizontal em duas manchas diferentes (com o mesmo número de pixels úteis) podem ter o mesmo termo de confiança. Assim, o uso do termo confiança apenas indica a quantidade de informação numa região alvo que pode ser utilizada como uma referência, enquanto se olha para fragmentos semelhantes na área fonte.

Aqui não há fluxograma, porque a Confiança Média (CM) é obtida de forma não-supervisionada, sendo calculada para cada imagem e por consequência obtendo um valor diferente para cada uma. A CM portanto contribui para que o índice final de confiabilidade apresente um valor diferente para cada imagem, independente de qual classe a imagem de

entrada pertença (ainda que os valores de cada classe sejam próximos como será visto adiante).

3.7.7 O Índice

Os algoritmos supervisionados buscam fazer uma relação entre as características de contaminação exploradas e os índices EGM e SSIM através de medidas estatísticas, buscando mapear um índice compatível com a região averiguada. Deste modo, ao ser analisada a imagem, são atribuídos índices EG e SSIM de acordo com as Tabelas 4.2 e 4.3, respectivamente, e caso não se enquadrem nas condições mostradas nas tabelas tem atribuído valor de 0. Por outro lado a CM dos pixels é verificada isoladamente por cada imagem. Finalmente, cada um desses fatores é combinado com pesos proporcionais à estatística obtida, que nada mais são do que a média do coeficiente de variação obtida durante o treinamento de cada medida, para encontrar o então índice de confiabilidade IC daquela imagem, determinado pela equação:

$$IC = \frac{k_1 \times EG + k_2 \times SSIM + CM}{3} \quad (3.17)$$

Onde k_1 e k_2 são constantes que representam pesos da EG e do SSIM, respectivamente, e o IC varia de 0 a 1, sendo que 1 representa a confiança máxima na redefinição de regiões (imagem sem ruído).

Os pesos são encontrados levando em consideração o Coeficiente de Variação (CV) do treinamento da EGM e do SSIM, que são respectivamente, 0.12 e 0.0761. Como a CM é única para cada imagem, foi decidido usá-la em sua integralidade. No entanto, EG e SSIM foram obtidas de um treinamento sobre um conjunto de imagens, portanto mais peso será dado ao treinamento que teve o menor CV, pois se julga que este teve mais estabilidade na geração de seus resultados.

Levando em consideração que $k_1 + k_2 = 2$, os CV médios foram somados e foi percebido que o CV médio da EGM corresponde a 61,19% do total. Logo, 61,22% de 2 é igual a 1.2238 é o k_2 dado ao SSIM, pois quanto maior o CV menor o peso. O mesmo procedimento é feito para a EG, obtendo k_1 igual a 0.7762. Deste modo a equação final do IC é:

$$IC = \frac{0.7762 \times EG + 1.2238 \times SSIM + CM}{3} \quad (3.18)$$

Capítulo 4

Resultados e Discussão

Este capítulo apresenta os materiais e os resultados obtidos a partir da aplicação das técnicas apresentadas neste trabalho, além de uma breve descrição e discussão sobre os mesmos.

4.1 Materiais

Na Tabela 4.1 são descritas as 20 imagens originais usadas para a geração das imagens contaminadas por componentes atmosféricos (Seção 3.7.1) e separadas em 9 classes de acordo com a Seção 3.7.3, totalizando 180 cenas, em que os métodos e algoritmos usados neste trabalho foram aplicados. As imagens são dos sensores QuickBird 02 (QB02) [34], combinando as bandas PAN e Multi-espectral 1 (MS1) [35], e World View 02 (WV02) [36], combinando as bandas PAN, Multi-espectral 1 (MS1) e Multi-espectral 2 (MS2) [37]; e foram obtidas gratuitamente pela *Digital Globe*¹, com cenas dos estados do Pará, Maranhão e Rio de Janeiro, obtidas entre 2005 e 2012.

¹<http://www.digitalglobe.com>

Tabela 4.1: Imagens originais.

Imagem	Local	Sensor	Bandas
Imagem 1	Belém-PA	WV02	Pan-MS1-MS2
Imagem 2	Região do Salgado-PA	QB02	Pan-MS1
Imagem 3	Região do Salgado-PA	QB02	Pan-MS1
Imagem 4	São Luís-MA	WV02	Pan-MS1-MS2
Imagem 5	São Luís-MA	QB02	Pan-MS1
Imagem 6	Tucuruí-PA	QB02	Pan-MS1
Imagem 7	Tucuruí-PA	WV02	Pan-MS1-MS2
Imagem 8	Marabá-PA	QB02	Pan-MS1
Imagem 9	Marabá-PA	QB02	Pan-MS1
Imagem 10	Altamira-PA	QB02	Pan-MS1
Imagem 11	Altamira-PA	WV02	Pan-MS1-MS2
Imagem 12	Rio de Janeiro-RJ	WV02	Pan-MS1-MS2
Imagem 13	Rio de Janeiro-RJ	WV02	Pan-MS1-MS2
Imagem 14	Rio de Janeiro-RJ	WV02	Pan-MS1-MS2
Imagem 15	Rio de Janeiro-RJ	WV02	Pan-MS1-MS2
Imagem 16	Rio de Janeiro-RJ	QB02	Pan-MS1
Imagem 17	Rio de Janeiro-RJ	QB02	Pan-MS1
Imagem 18	Rio de Janeiro-RJ	WV02	Pan-MS1-MS2
Imagem 19	Rio de Janeiro-RJ	QB02	Pan-MS1
Imagem 20	Rio de Janeiro-RJ	QB02	Pan-MS1

4.2 EGM para cada classe

A EGM foi obtida para cada classe seguindo o treinamento mostrado na subseção 3.7.4, de acordo com o fluxograma apresentado na Figura 3.5. Na Tabela 4.2 são mostradas a Média, Desvio Padrão (DP) e coeficiente de variação (CV) da EG para cada classe. A partir desta tabela são tirados os argumentos necessários relacionado à EGM, no momento do cálculo do IC, dependendo de qual classe pertença a imagem, conforme apresentado na subseção 3.7.7.

Tabela 4.2: Média, desvio padrão (DP) e coeficiente de variação (CV) da EG para cada classe.

Classes	Estrutura	Textura	Média	DP	CV
Classe 1	Excelente	Excelente	0.83	0.089	0.1072
Classe 2	Excelente	Boa	0.76	0.087	0.1144
Classe 3	Excelente	Ruim	0.78	0.077	0.0987
Classe 4	Boa	Excelente	0.81	0.113	0.1395
Classe 5	Boa	Boa	0.78	0.082	0.1051
Classe 6	Boa	Ruim	0.77	0.083	0.1077
Classe 7	Ruim	Excelente	0.85	0.102	0.12
Classe 8	Ruim	Boa	0.79	0.105	0.1329
Classe 9	Ruim	Ruim	0.76	0.118	0.1552

Dos valores da Tabela 4.2 é possível perceber que a EG está intimamente ligada à quantidade de cobertura de nuvens e sombras sobre regiões de textura, ocorrendo variação incoerente da EGM quando se leva em consideração a qualidade da área de estrutura. Por outro lado o DP e o CV se tornam maiores proporcionalmente à quanto pior (mais contaminada) está a região de estrutura, logo nessas regiões há maior variação sobre valores de EG. Daí pode-se inferir que quanto melhores são as condições de cobertura sobre regiões de textura, melhor o valor da EGM, bem como, em proporções menores, quanto melhores as condições de cobertura sobre regiões de estrutura, mais fiel é o treinamento.

4.3 SSIM para cada classe

O SSIM foi obtido para cada classe seguindo o treinamento mostrado na subseção 3.7.5, de acordo com o fluxograma apresentado na Figura 3.7. Na Tabela 4.3 são mostradas a média, DP e CV da SSIM para cada classe. A partir desta Tabela são tirados os argumentos necessários relacionado ao SSIM, no momento do cálculo do IC, dependendo de qual classe pertença a imagem, conforme apresentado na subseção 3.7.7.

Uma vez que a EGM dá menos importância à região de estrutura, o SSIM vem complementar a avaliação sendo sensível à análise de estrutura da imagem, mas que também não deixa de analisar regiões de textura, uma vez que também leva em consideração informações de luminância e reflectância. O resultado mostrado na Tabela 4.3 mostra que a qualidade da redefinição das regiões combina a qualidade das regiões de textura e estrutura

Tabela 4.3: Média, Desvio Padrão (DP) e coeficiente de variação (CV) do SSIM para cada classe.

Classes	Estrutura	Textura	Média	DP	CV
Classe 1	Excelente	Excelente	0.8998	0.0622	0.0691
Classe 2	Excelente	Boa	0.8696	0.0611	0.0702
Classe 3	Excelente	Ruim	0.8414	0.0623	0.0740
Classe 4	Boa	Excelente	0.8855	0.0687	0.0775
Classe 5	Boa	Boa	0.8626	0.0638	0.0739
Classe 6	Boa	Ruim	0.8300	0.0637	0.0767
Classe 7	Ruim	Excelente	0.8764	0.0688	0.0785
Classe 8	Ruim	Boa	0.8562	0.0670	0.0782
Classe 9	Ruim	Ruim	0.8177	0.0713	0.0871

de forma simultânea. Ao mesmo tempo todas as classes tem uma variação semelhante de SSIM para cada classe, mostrando um resultado de treinamento igualmente confiável em todas as classes.

4.4 CM para cada classe

A algoritmo para encontrar a CM (mostrado na subseção 3.7.6 e baseado na Equação 3.16) foi aplicado à cada imagem, classe por classe. Na Tabela 4.4 é mostrada a média da confiança média para cada classe.

A média da confiança calculada inicialmente por [14] considera as regiões circundantes de cada pixel, quanto pior a qualidade da região, pior o índice inicial conforme mostra a Tabela 4.4. No entanto esse índice é calculado de imagem para imagem, servindo como parâmetro de como se encontra a cena avaliada em termos de contaminação atmosférica.

Tabela 4.4: Média da CM para cada classe.

Classes	Estrutura	Textura	Média
Classe 1	Excelente	Excelente	0.9651
Classe 2	Excelente	Boa	0.9333
Classe 3	Excelente	Ruim	0.8931
Classe 4	Boa	Excelente	0.9345
Classe 5	Boa	Boa	0.9163
Classe 6	Boa	Ruim	0.8681
Classe 7	Ruim	Excelente	0.9125
Classe 8	Ruim	Boa	0.8837
Classe 9	Ruim	Ruim	0.8438

4.5 IC médio para cada classe

Achados os argumentos que serão usados para preencher os parâmetros da EGM e do SSIM para cada classe. O IC é calculado imagem à imagem das nossas classes, a fim de verificar o comportamento do mesmo frente aos seus índices base, assim como sua coerência. Na Tabela 4.5 é mostrada a média do Índice de Confiabilidade para cada classe.

Tabela 4.5: Média do IC para cada classe.

Classes	Estrutura	Textura	Média
Classe 1	Excelente	Excelente	0.903525
Classe 2	Excelente	Boa	0.865177
Classe 3	Excelente	Ruim	0.84449
Classe 4	Boa	Excelente	0.88295
Classe 5	Boa	Boa	0.85934
Classe 6	Boa	Ruim	0.83013
Classe 7	Ruim	Excelente	0.88106
Classe 8	Ruim	Boa	0.8497
Classe 9	Ruim	Ruim	0.81433

Semelhante ao resultado obtido por [38], um limite em torno 0.85 foi identificado como nível aceitável de confiança, neste caso para que uma região contaminada por elementos atmosféricos seja restaurada, e que a área sem acesso seja estimada. Além do limite, é importante ter em mente que quanto maior o IC, maior a confiança sobre a estimação das regiões.

4.6 Amostra das etapas de processamento e dos índices

A Figura 4.1 mostra um exemplo de processamento conforme as etapas descritas neste trabalho, assim como os índices resultantes deste processamento. A cena trabalhada é da Imagem 1 na Classe 1. Segundo o índice IC de 0.9095, a imagem representa o melhor caso, e é adequada para que se possa estimar suas regiões indeterminadas, fato que é constatado quando se observa o sucesso da estimação visto na Figura 4.1(d) e da segmentação vista na Figura 4.1(e).

Já a Figura 4.2 mostra outro exemplo, da Imagem 3 na Classe 5 (Textura e Estrutura boas). Apresentando o índice IC de 0.8605, a imagem representa um caso médio, e ainda é adequada para que se possa estimar suas regiões indeterminadas, o que é constatado quando se observa o sucesso (não tão aparente quanto no exemplo da Figura 4.1, mas ainda assim eficiente) da estimação visto na Figura 4.2(d) e da segmentação vista na Figura 4.2(e).

Finalmente, na Figura 4.3 tem-se um exemplo da Imagem 7 na Classe 9 (Textura e Estrutura ruins). Através do índice IC de 0.8150 obtido, e principalmente da estimação vista na Figura 4.3(d), em que constam borrões próximo do centro, além de fragmentos incoerentes no lugar das regiões indeterminadas, a imagem resultante representa um dos piores casos, sendo inadequada para que se possa estimar suas áreas.

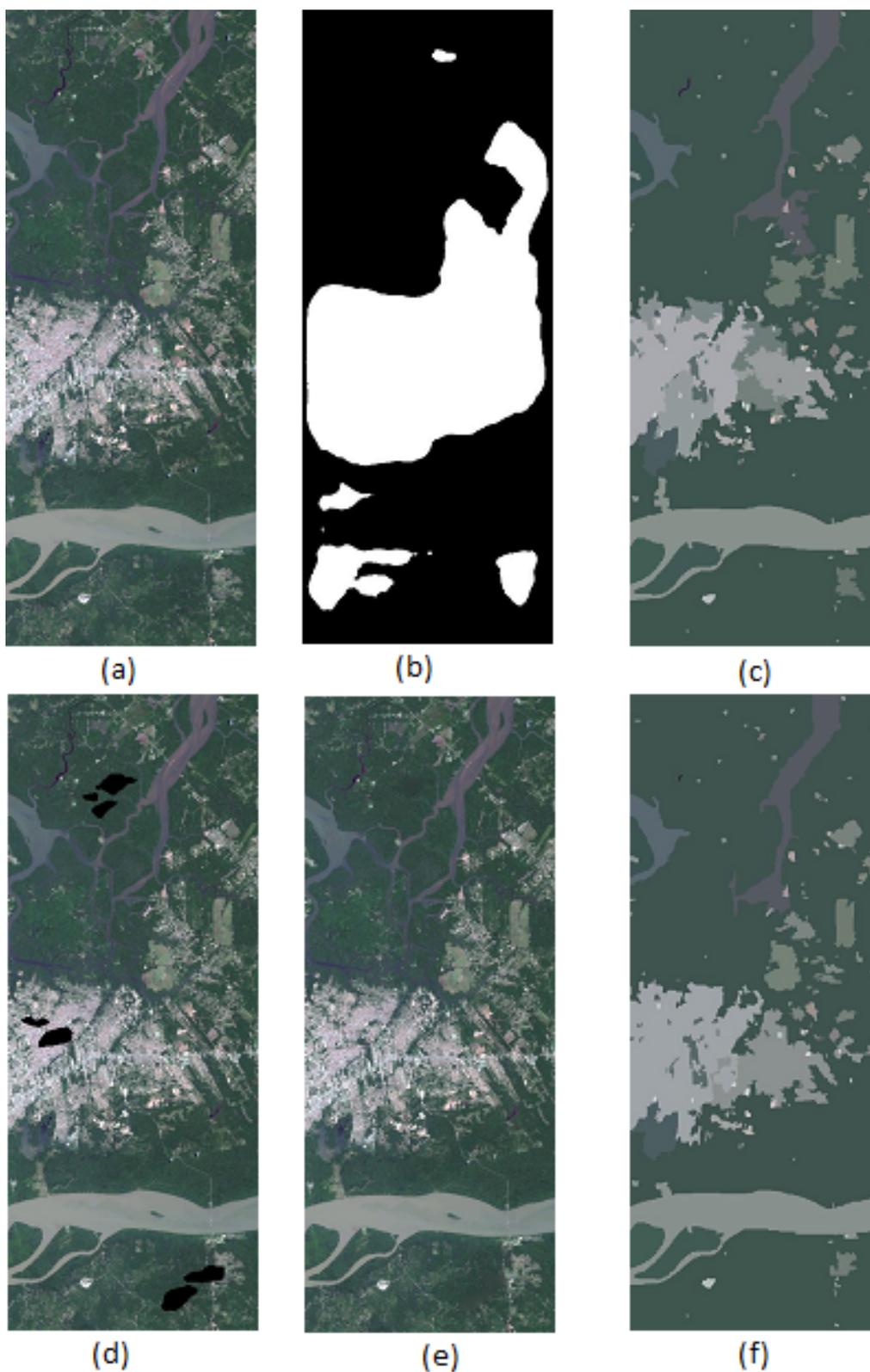


Figura 4.1: Etapas do processamento para a Imagem 1 na Classe 1 (Textura e Estrutura excelentes), que obteve os índices de $EG=0.9130$, $SSIM=0.8615$, $CM=0.9830$ e $IC=0.9095$. (a) Imagem original; (b) Imagem decomposta; (c) Imagem Segmentada antes do processamento; (d) Imagem resultante da contaminação artificial da classe 1; (e) Resultado da abordagem híbrida; (f) Imagem segmentada após o processamento.

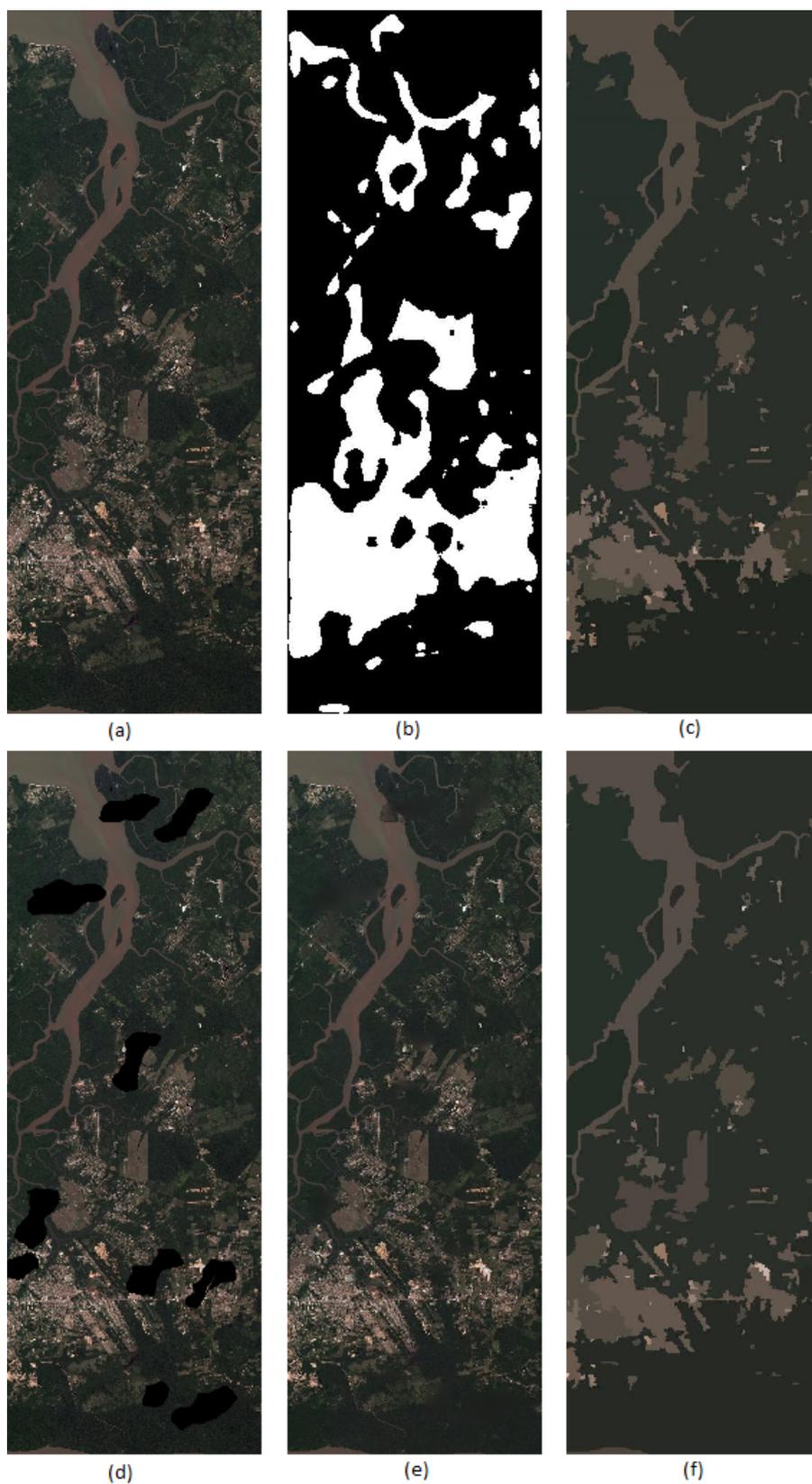


Figura 4.2: Etapas do processamento para a Imagem 3 na Classe 5, que obteve os índices de $EG=0.7329$, $SSIM=0.8297$, $CM=0.9204$ e $IC=0.8605$. (a) Imagem original; (b) Imagem decomposta; (c) Imagem Segmentada antes do processamento; (d) Imagem resultante da contaminação artificial da Classe 5; (e) Resultado da abordagem híbrida; (f) Imagem segmentada após o processamento.

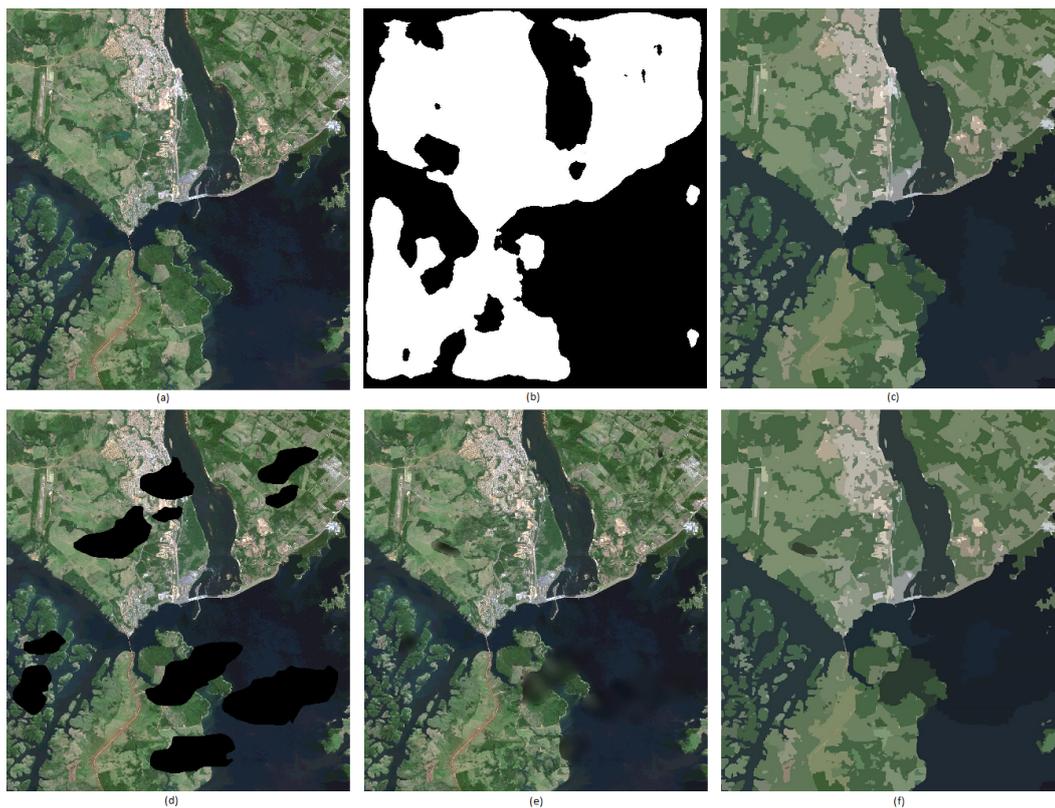


Figura 4.3: Etapas do processamento para a Imagem 7 na Classe 9, que obteve os índices de $EG=0.7935$, $SSIM=0.9101$, $CM=0.8544$ e $IC=0.8150$. (a) Imagem original; (b) Imagem decomposta; (c) Imagem Segmentada antes do processamento; (d) Imagem resultante da contaminação artificial da Classe 9; (e) Resultado da abordagem híbrida; (f) Imagem segmentada após o processamento.

Capítulo 5

Conclusões

Este capítulo mostra uma visão geral do trabalho, abordagens propostas e uma síntese com resultados alcançados e explicações acerca das metodologias empregadas. Por fim, são indicados os trabalhos publicados, assim como são apresentadas propostas de trabalhos futuros.

5.1 Conclusão

Este trabalho apresentou uma abordagem híbrida para estimação e redefinição de regiões de nuvens e sombras em imagens de satélite, além de uma metodologia de avaliação que culminou em um índice de confiabilidade para tal processamento. A ideia é primeiramente aplicar o índice, e com base na proporção de cobertura de ruídos sobre regiões de textura e estrutura, checar se as regiões contaminadas de uma determinada imagem são estimáveis, e se são, qual o grau de confiabilidade desta estimação. Dependendo do IC calculado, o algoritmo de redefinição de regiões híbrido é então aplicado à imagem em questão, obtendo-se desta forma uma nova imagem sem interferências atmosféricas.

Analisando os resultados mostrados no capítulo anterior, principalmente da Tabela 4.5, que mostra uma média da aplicação do IC sobre cada classe de imagens usadas no trabalho (como um processo de validação); e da Seção 4.6, onde são apresentadas as imagens resultantes das principais etapas apresentadas neste trabalho; é constatado que a relação entre as condições de contaminação de uma imagem (quanto menor for a contaminação sobre estrutura e textura, esta última principalmente, melhor são as condições), os valores de índice e a qualidade das imagens resultantes são diretamente proporcionais. Portanto, as metodologias descritas aqui estão aptas a serem usadas em condições reais de cobertura de nuvens e sombras em imagens de satélite, sendo uma nova opção para a estimação deste tipo de região indefinida.

5.2 Trabalhos Publicados

- D. F. Sousa, A. C. Q. Siravenha, D. C. O. de Resende, R. C. S. Filho, e E. G. Pelaes, Uma abordagem híbrida e semi-automática para a remoção de nuvens e sombras de imagens de satélite, Simpósio Brasileiro de Sensoriamento Remoto, 2013.
- D. C. O. de Resende, R. C. S. Filho, D. F. Sousa, A. C. Q. Siravenha e E. G. Pelaes, A influência do elemento estruturante no processo de detecção de nuvens em imagens de sensoriamento remoto, Simpósio Brasileiro de Sensoriamento Remoto, 2013.
- A. C. Q. Siravenha, D. F. Sousa, D. C. O. de Resende, R. C. S. Filho, e E. G. Pelaes, Uso de índices de diferença normalizada na detecção de nuvens e sombras em imagens landsat-5 tm, Simpósio Brasileiro de Sensoriamento Remoto, 2013.
- A. C. Q. Siravenha, D. F. Sousa and E. G. Pelaes, The Development of a Hybrid Solution to Replacement of Clouds and Shadows in Remote Sensing Images, Computational Modeling of Objects Presented in Images Lecture Notes in Computational Vision and Biomechanics Volume 15, pp 269-283, 2014.

5.3 Trabalhos Futuros

- Otimização dos algoritmos automáticos de detecção e estimação de regiões de nuvens e sombras;
- Aprimoramento do índice de confiabilidade, utilizando outras informações relacionadas à cena analisada;
- Criação de um software de fácil uso com interface gráfica para o usuário utilizar os algoritmos definidos neste trabalho a fim de resolver seus problemas relacionados à contaminação atmosférica em imagens de satélite, se possível disponível para web.

Capítulo 6

Apêndice: Detecção de Nuvens e Sombras em Imagens de Satélite usando Algoritmo Genético

Este apêndice apresenta uma forma de detecção genérica de nuvens e sombras em imagens de satélite baseada em [39], usando as bandas RGB (*Red, Green, Blue*) e operações de morfologia matemática, com automatização na escolha dos parâmetros de nuvem e de sombra propostos em [40, 26]. Para tal, será usado um algoritmo genético (AG) que aperfeiçoe a escolha de tais parâmetros da detecção, de modo que os erros de classificação sejam minimizados, por exemplo, que regiões de vegetação densa não sejam confundidas com vegetação.

6.1 Detecção de Regiões

A função que representa este processo de separação de regiões foi mostrada na Equação 3.1, vista na Seção 3.1 deste trabalho, sendo novamente mostrada na Equação 6.1. É importante ter em mente que esta equação é aplicada para todas as bandas da imagem, por exemplo, um *pixel* em D_m dimensões é considerado como região de sombra se em todas as bandas D_m da imagem tem seu valor de nível de cinza menor que a multiplicação de fn pela subtração do valor médio dos *pixels* pelo desvio padrão.

$$m(x, y) = \begin{cases} f(x, y) < fs \times f_{m-dp}, & f(x, y) \in 0, \\ f_{m-dp} < f(x, y) < f_m, & f(x, y) \in 1, \\ f_m < f(x, y) < fn \times f_{m+dp}, & f(x, y) \in 2, \\ f(x, y) > fn \times f_{m+dp}, & f(x, y) \in 3, \end{cases} \quad (6.1)$$

onde $f(x, y)$ é o valor digital do pixel, f_m é o valor médio dos pixels da imagem, f_{m+dp} e f_{m-dp} são a soma e a subtração do valor médio com o desvio padrão da imagem, respectivamente. Também há a presença dos fatores fn (fator de nuvem) e fs (fator de sombra) que servem para regular a detecção das regiões, modificando os limiares de cada classe.

É importante de perceber que se fn está entre 0 e 1, o algoritmo fica mais propenso à detectar nuvem, e vice-versa, se $fn > 1$. Já o parâmetro fs tem comportamento contrário ao de fn , durante a detecção de sombras.

Cada região foi rotulada com um respectivo número. Regiões definidas como 0 são as regiões de sombra da imagem, as definidas como 1 são as livres de qualquer tipo de ruído, já as definidas como 2 são rotuladas como sendo nuvens esparsas, e finalmente as definidas como 3 são detectadas como nuvens densas.

6.2 Aplicação do AG

O AG foi setado com as seguinte características:

- Representação Binária com 50 bits, sendo 25 destinados à cada variável (parâmetro de nuvem ou de sombra);
- Os valores reais das variáveis variam de 0 à 5, de acordo com a experiência dos fóto-intérpretes;
- População de 60 indivíduos;
- Apenas um ponto de corte;
- Seleção por roleta ponderada;
- Com e sem Elitismo;
- Taxa de cruzamento = 0.7;
- Taxa de mutação = 0.01;

- Critério de parada: 400 à 600 gerações, ou se nenhum indivíduo mais apto aparecer depois de 200 gerações;
- Avaliação das aptidões dos indivíduos pela métrica EBC.

6.3 Avaliação usando a métrica EBC

O EBC (*Edge-Border Coincidence*) é uma métrica que mede a qualidade de segmentação de uma determinada imagem, sobrepondo as regiões de borda na imagem segmentada, em relação às áreas encontradas pelo detector de bordas (no nosso caso é utilizado um filtro passa-altas Sobel) sobre a imagem original, sendo definido da seguinte forma: dado E , um conjunto formado por pixels (x_{pi}, y_{pi}) , extraídos após a limiarização da imagem, e o conjunto S , formados pelos pixels (x_{qi}, y_{qi}) , encontrados após a segmentação das regiões desta imagem por um algoritmo. Assim, temos:

$$E = \{p_1, p_2, \dots, p_E\} = \{(x_{P1}, y_{P1}), (x_{P2}, y_{P2}), \dots, (x_{PE}, y_{PE})\}$$

$$S = \{q_1, q_2, \dots, q_S\} = \{(x_{q1}, y_{q1}), (x_{q2}, y_{q2}), \dots, (x_{qS}, y_{qS})\}$$

Dessa forma o EBC é definido como:

$$EBC = n(E \cap S) / n(E) \quad (6.2)$$

Onde $n(\cdot)$ computa o número de elementos do seu argumento e $E \cap S = \{(x, y) | (x, y) \in E \text{ e } (x, y) \in S\}$.

É importante ter em mente que o valor de EBC varia entre 0 e 1, e quanto maior este valor, melhor é a segmentação da imagem.

A separação de regiões de uma imagem, nada mais é do que uma segmentação das áreas em nuvens densas e esparsas, sombras e locais livres de contaminação. Portanto, as bordas da imagem resultante da identificação das regiões deve ter o maior número de coincidências possíveis com as bordas da imagem original, mesmo que nunca sejam iguais ($EBC = 1$), os melhores indivíduos devem maximizar o valor de EBC entre estas 2 imagens.

6.4 Resultados e Conclusões

Esta seção apresenta os resultados do algoritmo de detecção de regiões usando um AG para automatizar as escolhas dos melhores parâmetros de nuvens e de sombras. Para cada experimento é variado o critério de parada e mostradas as regiões identificadas na imagem e os melhores, piores e média das aptidões dos indivíduos, bem como o indivíduo ótimo e sua aptidão.

O primeiro experimento foi feito com a imagem “Belém 1”, o critério de parada adotado foi o não aparecimento de um indivíduo melhor que os anteriores em um período de 200 gerações, usando elitismo. Desta forma, o AG parou na geração 320, obtendo os parâmetros ótimos $fn = 0.68$ e $fs = 0.79$ que levam à $EBC = 0.74$. A curva da média das aptidões está próxima da dos melhores indivíduos, mas não igual, o que mostra que o AG está convergindo, porém ainda mantém uma diversidade. São mostradas a imagem original, comportamento do AG, máscara e imagem com regiões identificadas na Fig. 6.1.

O experimento realizado com a imagem “Ilhas de Belém”, o critério de parada adotado foi o término de 400 gerações, usando elitismo. Desta forma, o AG obteve os parâmetros ótimos $fn = 0.94$ e $fs = 4.27$ que levam à $EBC = 0.86$. A curva da média das aptidões está próxima da dos melhores indivíduos, mas não igual, o que mostra que o AG está convergindo, porém ainda mantém uma diversidade. São mostradas a imagem original, comportamento do AG, máscara e imagem com regiões identificadas na Fig. 6.2

Já no experimento realizado com a imagem “Belém 2”, o critério de parada adotado foi o término de 500 gerações, usando elitismo. Desta forma, o AG obteve os parâmetros ótimos $fn = 0.82$ e $fs = 0.82$ que levam à $EBC = 0.9$. A curva da média das aptidões está próxima da dos melhores indivíduos, mas não igual, o que mostra que o AG está convergindo, porém ainda mantém uma diversidade. São mostradas a imagem original, comportamento do AG, máscara e imagem com regiões identificadas na Fig. 6.3

Em relação ao experimento feito com a imagem “Mina”, o critério de parada adotado foi o término de 500 gerações, porém sem elitismo. Desta forma, o AG obteve os parâmetros ótimos $fn = 1.26$ e $fs = 0.48$ que levam à $EBC = 0.84$. A curva da média das aptidões está próxima da dos melhores indivíduos, mas não igual, o que mostra que o AG está convergindo, porém ainda mantém uma diversidade. São mostradas a imagem original, comportamento do AG, máscara e imagem com regiões identificadas na Fig. 6.4

Finalmente, no experimento feito com a imagem “São Paulo”, o critério de parada adotado foi o término de 800 gerações, usando elitismo. Desta forma, o AG obteve os parâmetros ótimos $fn = 0.46$ e $fs = 3.56$ que levam à $EBC = 0.82$. A curva da média

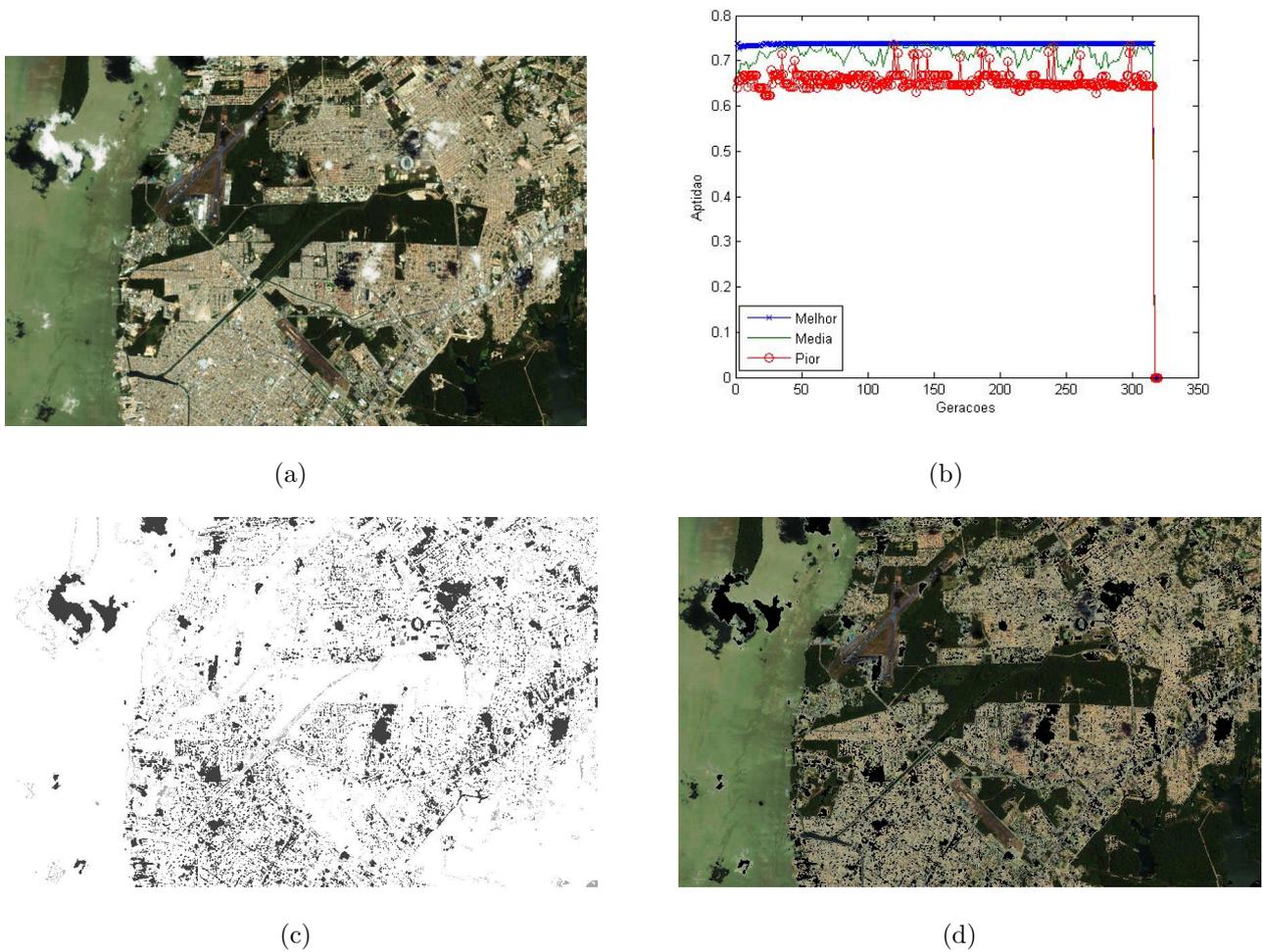


Figura 6.1: Imagem “Belém 1”: Parâmetros ótimos $fn = 0.68$ e $fs = 0.79$ que levam à $EBC = 0.74$; (a) Imagem Original; (b) Melhores, piores e média dos indivíduos de cada geração; (c) Máscara; (d) Regiões de nuvens e sombras detectadas na imagem.

das aptidões está próxima da dos melhores indivíduos, mas não igual, o que mostra que o AG está convergindo, porém ainda mantém uma diversidade. São mostradas a imagem original, comportamento do AG, máscara e imagem com regiões identificadas na Fig. 6.5

Portanto, foi constatado que, para gerar melhores resultados e não perder tempo em “processamento inútil”, o AG deve usar elitismo e o critério de parada deve ser o não aparecimento de um indivíduo melhor que os anteriores em um determinado período de gerações.

O AG demonstrou ser uma ótima ferramenta para automatizar a escolha dos parâmetros de nuvem e de sombra, uma vez que as regiões foram bem detectadas. No entanto, em alguns casos, as nuvens confundiram-se com pequenos fragmentos da área urbana e sombras deixaram de ser detectadas por serem mal classificadas como vegetação densa. Como trabalho futuro,

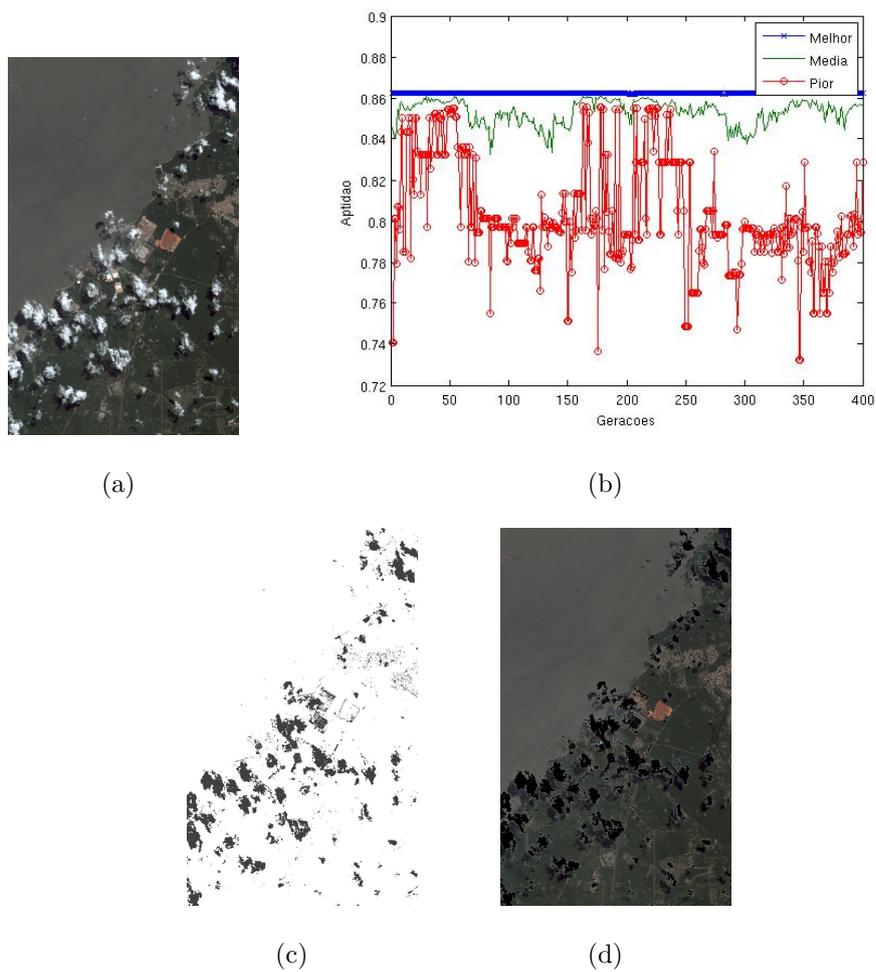


Figura 6.2: Imagem “Ilhas de Belém”: Parâmetros ótimos $fn = 0.94$ e $fs = 4.27$ que levam à $EBC = 0.86$; (a) Imagem Original; (b) Melhores, piores e média dos indivíduos de cada geração; (c) Máscara; (d) Regiões de nuvens e sombras detectadas na imagem.

pretende-se eliminar ou ao menos diminuir estes erros.

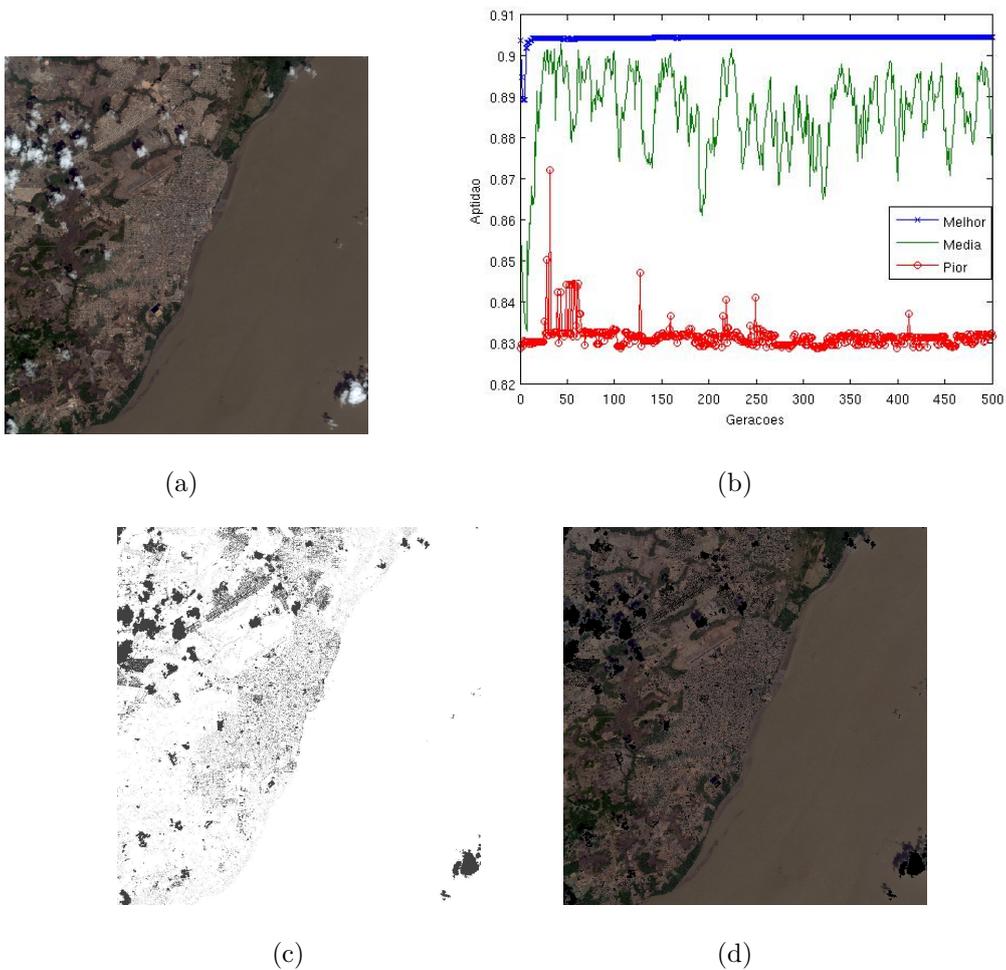
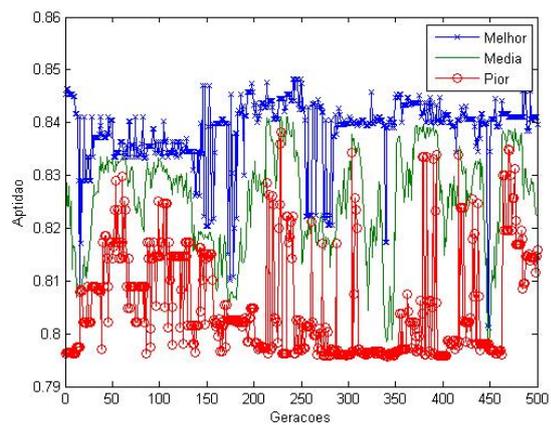


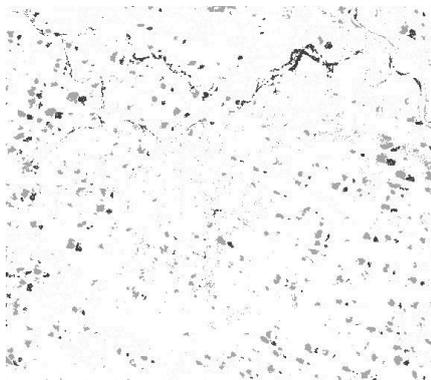
Figura 6.3: Imagem “Belém 2”:Parâmetros ótimos $fn = 0.82$ e $fs = 0.82$ que levam à $EBC = 0.9$; (a) Imagem Original; (b) Melhores, piores e média dos indivíduos de cada geração; (c) Máscara; (d) Regiões de nuvens e sombras detectadas na imagem.



(a)



(b)



(c)



(d)

Figura 6.4: Imagem “Mina”: Parâmetros ótimos $fn = 1.26$ e $fs = 0.48$ que levam à $EBC = 0.84$; (a) Imagem Original; (b) Melhores, piores e média dos indivíduos de cada geração; (c) Máscara; (d) Regiões de nuvens e sombras detectadas na imagem.

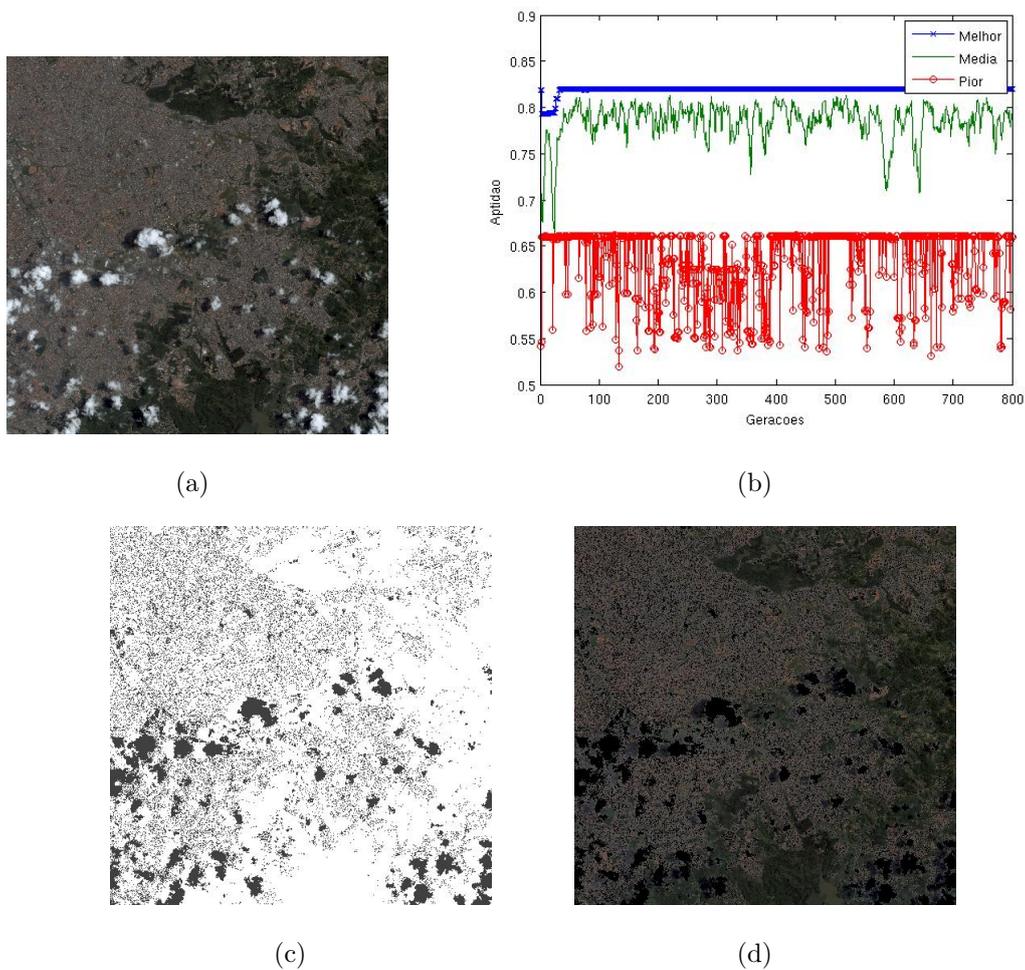


Figura 6.5: Imagem “São Paulo”: Parâmetros ótimos $fn = 0.46$ e $fs = 3.56$ que levam à $EBC = 0.82$; (a) Imagem Original; (b) Melhores, piores e média dos indivíduos de cada geração; (c) Máscara; (d) Regiões de nuvens e sombras detectadas na imagem.

Referências Bibliográficas

- [1] 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.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] Y. Liu, A. Wong, and P. Fieguth, “Remote sensing image synthesis,” in *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, 2010, pp. 2467–2470.
- [6] U. Fayyad, G. Piatetsky-shapiro, and P. Smyth, “From data mining to knowledge discovery in databases,” *AI Magazine*, vol. 17, pp. 37–54, 1996.
- [7] R. M. PITONE, “Mineração de regras de associação nos canais de informação do direto,” Master’s thesis, Universidade Federal do Rio Grande do Sul, 2002.
- [8] T. F. Chan and S. H. Kang, “Error analysis for image inpainting.” *Journal of Mathematical Imaging and Vision*, vol. 26, pp. 85–103, 2006.
- [9] V. V. Mahalingam, “Digital inpainting algorithms and evaluation,” Ph.D. dissertation, University of Kentucky, 2010.
- [10] P. A. Ardis and A. Singhal, “Visual salience metrics for image inpainting,” vol. 7257, 2009, pp. 72 571W–72 571W–9.

-
- [11] P. A. Ardis, C. M. Brown, and A. Singhal, “Inpainting quality assessment.” *J. Electronic Imaging*, vol. 19, p. 011002, 2010.
- [12] A. I. O. Feier, “Digital inpainting for artwork restoration: Algorithms and evaluation,” Master’s thesis, Gjøvik University College, 2012.
- [13] D. Garcia, “Robust smoothing of gridded data in one and higher dimensions with missing values,” *Computational Statistics & Data Analysis*, vol. 54, no. 4, pp. 1167 – 1178, 2010.
- [14] A. Criminisi, P. Perez, and K. Toyama, “Region filling and object removal by exemplar-based image Inpainting,” in *IEEE Transactions On Image Processing*, vol. 13(9), 2004, pp. 1200–1212.
- [15] A. Buades, T. Le, J.-M. Morel, and L. Vese, “Cartoon+texture image decomposition,” 2011. [Online]. Available: <http://lx.pe/b707>
- [16] O. Marques and H. Vieira, *Processamento Digital de Imagens*. BRASPORT, 1999.
- [17] M. Tuceryan and A. K. Jain, *Texture Analysis, In The Handbook of Pattern Recognition and Computer Vision (2nd Edition)*. World Scientific Publishing Co., 1998, pp. 207–248.
- [18] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Addison-Wesley Publishing Company, 2008.
- [19] L. D. Griffin and A. C. F. Colchester, “Superficial and deep structure in linear diffusion scale space: isophotes, critical points and separatrices,” *Image and Vision Computing*, vol. 13, 7, pp. 543–557, 1995.
- [20] D. FIGUEIREDO, “Conceitos básicos de sensoriamento remoto,” http://www.conab.gov.br/conabweb/download/SIGABRASIL/manuais/conceitos_sm.pdf, 2005.
- [21] J. Campbell, *Introduction to Remote Sensing*. New York: Guilford. (eds. 1-4), (1987, 1996, 2002, 2007).
- [22] A. Buz, *Image Compression Using Discrete Cosine Transform*. Çankaya University. The Department of Computer Engineering, 2002. [Online]. Available: <http://books.google.com.br/books?id=S-5TNwAACAAJ>
- [23] Eastman, *IDRIS for Windows, User’s Guide*. Clark University, 1995.
- [24] R. Nock and F. Nielsen, “Statistical region merging,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 1452–1458, 2004.

-
- [25] S. Jaggi, *Descriptive Statistics and Exploratory Data Analysis*, <http://www.iasri.res.in/ebook/EB.SMAR/index.htm>, 2007, ch. 2.
- [26] D. Sousa, A. C. Siravenha, and E. Pelaes, “A hybrid approach to clouds and shadows removal in satellite images,” in *Computational Modelling of Objects Represented in Images III*. CRC Press, August 28, 2012 2012, pp. 153–158.
- [27] D. e. R. D. e. F. R. e. P. E. SIRAVENTHA, Ana Carolina e SOUSA, “Uso de índices de diferença normalizada na detecção de nuvens e sombras em imagens landsat-5 tm,” in *Anais XVI Simpósio Brasileiro de Sensoriamento Remoto*, 2013, pp. 7942–7949.
- [28] A. Efros and T. Leung, “Texture synthesis by non-parametric sampling,” in *In International Conference on Computer Vision*, 1999, pp. 1033–1038.
- [29] E. B. K. Sangeetha, P.Sengottuvelan, “Comparative analysis and evaluation of image inpainting algorithms,” *Journal of Information Engineering and Applications*, 2011.
- [30] V. SCHYTT, “Glaciers of europe-glaciers of sweden,” in *U.S. GEOLOGICAL SURVEY PROFESSIONAL PAPER 1386-E-4*. U.S. Government Printing Office, 1993, pp. E111–E125.
- [31] D. J. Reeves, Matthew; Bedunah, “A comparison of low cost satellite imagery for pastoral planning projects in central asia,” in *Rangelands of Central Asia: Proceedings of the Conference on Transformations, Issues, and Future Challenges*, 2006, pp. 120–127.
- [32] C. M. D. d. Pinho, “Análise orientada a objetos de imagens de satélites de alta resolução espacial aplicada à classificação de cobertura do solo no espaço intra-urbano: o caso de são José dos campos-sp,” Master’s thesis, Instituto Nacional de Pesquisas Espaciais, 2006.
- [33] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 13, no. 4, pp. 600–612, 2004.
- [34] D. Globe, “Quickbird,” 2013. [Online]. Available: <http://www.digitalglobe.com/sites/default/files/QuickBird-DS-QB-PROD.pdf>
- [35] A. S. T. R. Anne Puissant, C. Weber, “Fusion of quickbird ms and pan data for urban studies,” *23rd Symposium of the European Association of Remote sensing Laboratories (EARSeL)*, 2003.
- [36] D. Globe, “Worldview-2,” 2013. [Online]. Available: <http://www.digitalglobe.com/downloads/WorldView2-DS-WV2-Web.pdf>

- [37] F. P. C. Padwick, M. Deskevich and S. Smallwood, "Worldview-2 pan-sharpening," ASPRS, San Diego, California, Tech. Rep., 2010.
- [38] G. Foody, "Status of land cover classification accuracy assessment," *Remote Sensing of Environment*, vol. 80, no. 1, pp. 185–201, 2002. [Online]. Available: <http://eprints.soton.ac.uk/15173/>
- [39] 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.
- [40] A. Siravenha, "Um método para classificação de imagens de satélite usando transformada cosseno discreta com detecção e remoção de nuvens e sombras," 2011.